Studies of Lagrangian Modeling Techniques with Applications to Deep Ocean Carbon Sequestration

by

Peter Hampus Israelsson

Submitted to the Department of Civil and Environmental Engineering on August 21, 2008, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the field of Environmental Fluid Mechanics

Abstract

This thesis consists of four separate studies which explore the use of random walk particle tracking (RWPT) in simulating environmental mass transport. Two of the studies also consider the efficacy and marine impact of ocean CO$_2$ sequestration by direct injection.

The first study compares RWPT to two other simple Lagrangian techniques (forward and backward Gaussian puff tracking) to simulate mixing beyond the near field of a pollutant discharge. RWPT is found to be more accurate, but also more computationally expensive, thus motivating hybrid approaches where Lagrangian calculations transition to Eulerian schemes in the far field.

The second study considers 1D RWPT when strong gradients in ambient diffusivity exist. For step profiles, the work of past investigators is unified and extended, and the Thomson et al. (1997) particle reflection approach is recommended. For piecewise linear profiles, a novel and efficient particle reflection with probability translation approach is proposed.

The third study implements RWPT to emulate the tracer transport of an ocean general circulation model (OGCM) using the OGCM’s flow and diffusivity fields. A high level of agreement between RWPT and OGCM results is achieved. Particle reflection with probability translation successfully handled sharply varying vertical diffusivities. However, precisely mimicking OGCM calculations proved difficult due to complications in specifying the sub-grid scale variation of isopycnal slope and diffusivity in steeply sloped or convectively unstable regions, and in accurately implementing the Gent-McWilliams eddy-induced transport. Further development is recommended to resolve spurious upwelling occurring mainly in the Southern Ocean. The utility of RWPT is demonstrated by calculating domain-wide CO$_2$ sequestration efficiencies using a novel book-keeping method. Additional RWPT benefits to oceanographic investigation are also proposed.

The fourth study updates the Auerbach et al. (1997) and Caulfield et al. (1997) joint assessment of acute environmental impact to zooplankton resulting from ocean CO$_2$ discharges. Acute toxicity data are used to estimate the cumulative harm accrued by passive organisms drifting through idealized CO$_2$ plumes generated by three promising discharge
methods. Results suggest that discharges can be engineered to largely avoid acute impacts, and that ocean carbon sequestration should not be dismissed on the basis of environmental impact alone.

Thesis Supervisor: E. Eric Adams
Title: Senior Research Engineer of Civil and Environmental Engineering
Acknowledgments

As my own “sequestration” at MIT comes to a close and I reflect upon my time here, I realize that I have many people to thank on many different levels, both within and outside of the Institute. Since it is a practical impossibility for me to adequately thank everyone in the text below, let me begin by expressing my gratitude to all who have guided and supported me, and enriched my life over the years. I have benefited greatly from your collective generosity, and consider myself very fortunate in this regard.

MIT is by its very nature an interesting and stimulating place to be, and I thank the funding sources that have made my time here possible: the MIT Presidential Graduate Fellowship Program, the Martin Family Society of Fellows for Sustainability, the Ocean Carbon Sequestration Program, Biological and Environmental Research (BER), US Department of Energy (grant number DE-FG02-01ER63078), and the National Energy Technology Laboratory, US Department of Energy (grant number DE-FG26-98FT40334).

My technical knowledge and research skills have been greatly enhanced by my thesis supervisor, Dr. E. Eric Adams, whom I thank for his guidance, support, patience, and friendship as I have pursued this Ph.D. along with a concurrent S.M. in Technology and Policy. He has been instrumental in each part of my research, providing many creative ideas and solutions to roadblocks along the way. Hopefully I have gained some of Eric’s enviable ability to make complex topics approachable, and to assimilate disparate research efforts into a coherent story with an appropriate level of detail.

I would also like to thank the other members of my committee, Professor Heidi Nepf (MIT CEE), Dr. Mick Follows (MIT EAPS), and Dr. Ken Caldeira (Carnegie Institution’s Department of Global Ecology at Stanford University) for their valuable insights into the research process and vibrant exchanges throughout. Perhaps most importantly, together with Eric they helped me see the forest at times when I was too focused on the trees.

Special thanks to Ken Caldeira for our collaboration over the years, which I hope will continue. Ken has been extremely generous throughout, hosting me on multiple trips to the Lawrence Livermore National Laboratory (LLNL) and to the Carnegie Institution at Stanford University. In addition, he granted me significant computational resources across multiple computer clusters throughout my research, which would have been infeasible without his generosity. Beyond logistics, I have found our collaboration highly stimulating, due in large part to Ken’s creativity, diverse interests, and infectious love of science.

I wish to acknowledge several people for specific help on some parts of the thesis. For Chapter 2, I thank my coauthor Young Do Kim for our collaboration. For Chapter 3, I thank Marco Ghisalberti for the original test case scenarios and for help in model testing (and for fun along the way). For Chapter 4, I thank Ken for his constant insights, ideas, and advice; Michael Wickett (LLNL) and Daniel Enderton (MIT EAPS) for early help with the LLNL MOM and post-processing software; Ken’s research groups and colleagues at LLNL and Carnegie for patiently sharing computational resources with me; and the many LLNL and Carnegie system administrators who have helped me solve all kinds of computing problems over the years. For Chapter 5, I thank Costas Tsouris and his colleagues at ORNL for their work on developing the co-flowing jet hydrate reactor which forms the basis of most discharge scenarios considered herein, as well as T. Kikkawa (MERI) and Y. Watanabe (GETC) for
sharing their acute CO₂ toxicity data. Lastly, I thank my close friend and colleague Aaron Chow for his technical help throughout (e.g., plume simulations and figures in Chapter 5), many insightful discussions, and most importantly his friendship as we trekked through our research side-by-side.

I have enjoyed my time as a Parsons Lab resident and I thank the Parsons community, including the faculty who educated me and the people who keep the Lab running (for my part of it, that is mainly Sheila Frankel, Gayle Sherman, Jim Long, and Vicki Murphy), as well as the many friends that I have made among Parsons students and post-docs. From this latter group, I think my memories of the Lab will be dominated by Aaron Chow, Jeremy Pal, Rolf Reichle, Marco Ghisalberti and Leah Kaplan, Blake Landry, Yoshi and Kayo Tajima, Matt Hancock, Meng-Yi Chen, Marine Herrmann, Geraldine Grataloup, Brian White, Anne Thompson, Loretta Fernandez, Jingfeng Wang, Ramahi Sarma-Rupavtarm, Frederic Chagnon, Homero and Xanat Flores-Cervantes, Dana Hunt, Matt Orosz, Hanan Karam, Jean Fitzmaurice, Anne Lightbody, David Gonzalez-Rodriguez, and Yukie Tanino.

Beyond Parsons, I thank Cynthia Stewart, Kris Kipp, Jeanette Marchocki and their colleagues within CEE administration for much assistance over the years; Sydney Miller for helping me navigate TPP; the good friends I made along the way among fellow MEng and TPP students; Dave Marks of CEE/LFEE for periodic but steady mentorship; and Audra Bartz for her invaluable guidance.

I also thank my many friends outside of MIT for giving me much-needed reprieves over these years. From this group, special thanks go to Jeff Anderson, Nirosh and Narissa Ramachandran, Janira Arocho, Michael and Tanya Augustine, Kanishka and Rochelle Wijayasiri, Luis and Cholo Kålfors, Uffe and Lars Galsgaard, Maury Kimball, Katrina Zafiriadis, Katrina Schwartz, and Vanessa and Robert Breese.

In addition, I thank my friends and colleagues at QEA for their past mentorship and support, especially that of Kirk Ziegler and John Connolly.

Finally, I reserve my deepest thanks for my family. I thank my dear siblings and their families: Monica (+ Nabil, Caroline, Robbie, Michelle), Lena (+ Roger, Kristina, Julia), and Johan (+ Amy, Maria, Susanna, Daniel), as well as my extended family back in Sweden, for their collective support in what to them must have seemed a very strange process. I thank P.K., Kala, and Lakshman Nathan, as well as their extended family, for encouraging me and always making me feel a true part of the family. Most of all, I thank my parents Bo Hampus and Karin, my wife Janani, and my daughter Maya; I am at a loss for words to describe how important your unwavering love and support have been to me. All I can say is that you are constant reminders of that which is most important to me.

I dedicate this work to the memory of my father.
Contents

1 Introduction and Overview ........................................... 23
   1.1 Lagrangian modeling ........................................... 23
   1.2 Deep ocean carbon sequestration ............................ 25
Bibliography .............................................................. 27

2 A Comparison of Three Lagrangian Approaches for Extending Near Field
   Mixing Calculations .................................................. 29

3 Random Walk Particle Tracking in Step and Piecewise Linear Diffusivity 49
   3.1 Introduction ...................................................... 49
   3.2 Preliminary Considerations .................................... 52
      3.2.1 Alternate forms of the random walk equation .......... 52
      3.2.2 Variable timestepping ................................... 55
      3.2.3 Boundary handling ....................................... 56
   3.3 Step Diffusivity .................................................. 56
      3.3.1 Illustration of the problem ............................... 57
      3.3.2 Particle statistics at a boundary ........................ 59
      3.3.3 Analysis of discontinuity error .......................... 72
      3.3.4 Correction techniques ................................... 74
      3.3.5 Validation and comparison of corrections .............. 84
   3.4 Piecewise Linear Diffusivity .................................. 89
      3.4.1 Illustration of the problem ............................... 89
      3.4.2 Correction methodology ................................... 90
4.A.1 Analysis of bias in RWPT implementation ............................. 297
4.A.2 Paths for future refinement ............................................. 305
4.B Appendix: additional sequestration figures ............................. 322

Bibliography ................................................................. 329

5 Evaluation of the Environmental Viability of Direct Injection Schemes for
Ocean Carbon Sequestration ............................................. 337
5.1 Introduction ..................................................................... 337
5.2 CO₂ Toxicity Studies ....................................................... 341
  5.2.1 pH toxicity studies ....................................................... 342
  5.2.2 CO₂ toxicity studies of fish .......................................... 345
  5.2.3 CO₂ toxicity studies of pelagic copepods .......................... 359
  5.2.4 CO₂ toxicity studies of benthic organisms ......................... 363
  5.2.5 Combined mortality dataset for all species ......................... 368
  5.2.6 Sub-lethal impacts ....................................................... 369
5.3 Modeling Approach and Scenarios ..................................... 372
  5.3.1 Previous studies ......................................................... 372
  5.3.2 CO₂ discharge scenarios & plume modeling ....................... 382
  5.3.3 Biological impact analysis approach ................................. 398
5.4 Results ........................................................................... 415
  5.4.1 Water quality impacts of discharges .................................. 415
  5.4.2 Biological impacts of discharges ...................................... 417
  5.4.3 Sensitivity analysis on isomortality function ....................... 423
  5.4.4 Discussion of results .................................................... 428
5.5 Summary, Conclusions, and Policy Implications ..................... 439

Bibliography ................................................................. 447

6 Summary and Major Findings ........................................... 465

Bibliography ................................................................. 469
List of Figures

3-1 Two-layer test domain with a discontinuous diffusivity profile. .......................... 57
3-2 Predicted time-averaged concentration and ratio of mass below to mass above $z = 100 \text{ cm}$ for the test domain using an initially uniform particle density and $\Delta t = 1 \text{ s}$ and $\Delta t = 0.1 \text{ s}$. .................................................................................. 58
3-3 Number of particles predicted and observed to arrive at $z' = 0$ in one $\Delta t$ for simulations of various $D$ and $\Delta t$ values in an unbounded domain with an initially uniform particle distribution. ................................................................. 63
3-4 Number of particles predicted and observed to arrive at $z' = 0$ in one $\Delta t$ for simulations of various $D$ and $\Delta t$ values in a bounded domain ($0 \leq z' \leq \sqrt{2D\Delta t}$) with an initially uniform particle distribution and an absorbing boundary at $z' = \sqrt{2D\Delta t}$. ................................................................. 64
3-5 Predicted and observed cumulative distribution function ($P(0 \leq z'_i \leq Z_i)$) and probability density function ($p(Z_i)$) (unbounded domain). ........................ 66
3-6 Predicted and observed cumulative distribution function ($P(0 \leq z'_i \leq Z_i)$) and probability density function ($p(Z_i)$) (bounded domain). .......................... 66
3-7 Predicted and observed cumulative distribution function ($P(0 \leq z'_f \leq Z_f)$) and probability density function ($p(z'_f = Z_f)$). ............................................. 69
3-8 Predicted and observed cumulative distribution function ($P(0 \leq f_b \leq F_b)$) and probability density function ($p(F_b)$). ......................................................... 72
3-9 Predicted and observed cumulative distribution function ($P(0 \leq f_b \leq F_b)$) and probability density function ($p(F_b)$). ......................................................... 73
3-10 Probability density for particle displacements beyond the test domain discontinuity in one $\Delta t$ of 0.5 s & 0.05 s for an initially uniform particle density .......................... 75
3-11 Predicted time-averaged concentration \((C)\) and ratio of mass below to mass above \(z = 100\) cm \((R)\) for the test domain with an initially uniform particle density and \(\Delta t = 1\) s using particle reflection with midstep adjustment.

3-12 Predicted time-averaged concentration and ratio of mass below to mass above \(z = 100\) cm for the test domain with an initially uniform particle density and \(\Delta t = 1\) s using two different correction methods.

3-13 Observed cumulative distribution function \((P(0 \leq z'_f \leq Z_f))\) for the final position \(z'_f\) of particles crossing from a \(D = 10\) cm\(^2\)/s region (initialized with a uniform particle) into a \(D = 0.1\) cm\(^2\)/s region using step shortening coupled with the particle reflection flux correction \((\Delta t = 1\) s everywhere).

3-14 Predicted concentration at end of the simulation and ratio of mass below to mass above \(z = 100\) cm for the test domain with an initially uniform particle density using timestep balancing with step shortening for two large timesteps.

3-15 Predicted concentration and ratio \(R\) of mass below to mass above \(z = 100\) cm for the unsteady case for \(\Delta t = 1\) s.

3-16 Predicted concentration and ratio of mass below to mass above \(z = 100\) cm for the unsteady case for \(\Delta t = 100\) s.

3-17 Two-layer test domain with a discontinuous diffusivity gradient.

3-18 Predicted time-averaged concentration and mass ratio for the test domain with an initially uniform particle density using constant \(\Delta t\) values of 1 s, 0.1 s, and 0.01 s.

3-19 Predicted time-averaged concentration and mass ratio for the test domain with an initially uniform particle density using the proposed correction methodology with constant \(\Delta t\) values of 1 s, 10 s, and 100 s.

3-20 Predicted time-averaged concentration and mass ratio for the test domain with an initially uniform particle density in the top half of the domain.

3-21 Non-dimensional vertical diffusivity profile from Ghisalberti and Nepf (2005) for flow over a submerged aquatic canopy.

3-22 Piecewise linear approximation of the observed diffusivity profile in Ghisalberti and Nepf (2005).
3-23 Predicted time-averaged concentration and mass ratio for the test domain
with an initially uniform particle density in the top half of the domain. . . . 102
3-24 Sample diﬀusivity proﬁles from an LLNL MOM simulation. . . . . . . . . . . 103
3-25 Predicted time-averaged concentration and mass ratio for the ﬁrst LLNL
MOM sample diﬀusivity proﬁle with an initially uniform particle density. . . 105
3-26 Predicted concentration at 240 days and mass ratio for the ﬁrst LLNL MOM
sample diﬀusivity proﬁle with particles initially in depth range 229 - 309 m. . 106
3-27 Predicted time-averaged concentration and mass ratio for the second LLNL
MOM sample diﬀusivity proﬁle with an initially uniform particle density. . . 107
3-28 Predicted concentration at 20 days and mass ratio for the second LLNL MOM
sample diﬀusivity proﬁle with particles initially in depth range 229 - 309 m. . 108
3-29 Probability density function of reﬂected particle position required to maintain
the well-mixed condition for a D = 0 impermeable boundary at z = 0. . . . . 109
3-30 Predicted time-averaged concentration and mass ratio for a domain with impermeable boundaries at the top and bottom with D = 100 and 0 cm2 /s,
respectively, and an initially uniform particle density. . . . . . . . . . . . . . 110
3-31 Predicted concentration at t = 20 s and mass ratio for a domain with impermeable boundaries at the top and bottom with D = 100 and 0 cm2 /s,
respectively, with particles introduced as a point source at z = 30 cm. . . . . 111
3-32 Predicted concentration at t = 20 s and mass ratio for a domain with impermeable boundaries at the top and bottom with D = 102 and 2 cm2 /s,
respectively, with particles introduced as a point source at at z = 30 cm . . . 112
3-33 The 5 possible shapes of the ξc (zi ) curve in the continuous domain. . . . . . 122
3-34 Comparison of predicted and observed Pnr (0 < zf < Zf ) for “non-reﬂected”
particles for each of the 5 cases presented in Table 3.2 and Figure 3-33. . . . 126
4-1 LLNL MOM computational grid. . . . . . . . . . . . . . . . . . . . . . . . . 141
4-2 LLNL MOM computational grid (plan view). . . . . . . . . . . . . . . . . . . 142
4-3 LLNL MOM computational grid (proﬁle view).

. . . . . . . . . . . . . . . . 142

4-4 Isopycnal diﬀusivity as a function of isopycnal slope in LLNL MOM. . . . . . 145
13


4-5  Advection variables available from standard LLNL MOM output. ........  150
4-6  Diffusion variables available from standard LLNL MOM output. ........  150
4-7  Annual mean LLNL MOM isopycnal slope. ...................................  154
4-8  Average percent of year that isopycnal diffusivity is tapered in LLNL MOM.  155
4-9  Annual mean LLNL MOM vertical diffusivity. .................................  156
4-10 Tracer test release simulation sites. ...........................................  186
4-11 Initial condition for the 5 depth intervals for the tracer test release simulations 186
4-12 LLNL MOM and RWPT 100-year tracer concentrations for Site a1 ..........  193
4-13 LLNL MOM and RWPT 100-year tracer concentrations for Site a2 ..........  194
4-14 LLNL MOM and RWPT 100-year tracer concentrations for Site a3 ..........  195
4-15 LLNL MOM and RWPT 100-year tracer concentrations for Site a4 ..........  196
4-16 LLNL MOM and RWPT 100-year tracer concentrations for Site a5 ..........  197
4-17 LLNL MOM and RWPT 100-year tracer concentrations for Site b1 ..........  198
4-18 LLNL MOM and RWPT 100-year tracer concentrations for Site b2 ..........  199
4-19 LLNL MOM and RWPT 100-year tracer concentrations for Site b3 ..........  200
4-20 LLNL MOM and RWPT 100-year tracer concentrations for Site b4 ..........  201
4-21 LLNL MOM and RWPT 100-year tracer concentrations for Site b5 ..........  202
4-22 LLNL MOM and RWPT 100-year tracer concentrations for Site c1 ..........  203
4-23 LLNL MOM and RWPT 100-year tracer concentrations for Site c2 ..........  204
4-24 LLNL MOM and RWPT 100-year tracer concentrations for Site c3 ..........  205
4-25 LLNL MOM and RWPT 100-year tracer concentrations for Site c4 ..........  206
4-26 LLNL MOM and RWPT 100-year tracer concentrations for Site c5 ..........  207
4-27 LLNL MOM and RWPT 100-year tracer concentrations for Site d1 ..........  208
4-28 LLNL MOM and RWPT 100-year tracer concentrations for Site d2 ..........  209
4-29 LLNL MOM and RWPT 100-year tracer concentrations for Site d3 ..........  210
4-30 LLNL MOM and RWPT 100-year tracer concentrations for Site d4 ..........  211
4-31 LLNL MOM and RWPT 100-year tracer concentrations for Site d5 ..........  212
4-32 LLNL MOM and RWPT 100-year tracer concentrations for Site e1 ..........  213
4-33 LLNL MOM and RWPT 100-year tracer concentrations for Site e2 ..........  214
4-34 LLNL MOM and RWPT 100-year tracer concentrations for Site e3 ..........  215
4-35 LLNL MOM and RWPT 100-year tracer concentrations for Site e4 216
4-36 LLNL MOM and RWPT 100-year tracer concentrations for Site e5 217
4-37 LLNL MOM and RWPT 100-year tracer concentrations for Site f1 218
4-38 LLNL MOM and RWPT 100-year tracer concentrations for Site f2 219
4-39 LLNL MOM and RWPT 100-year tracer concentrations for Site f3 220
4-40 LLNL MOM and RWPT 100-year tracer concentrations for Site g1 221
4-41 LLNL MOM and RWPT 100-year tracer concentrations for Site g2 222
4-42 LLNL MOM and RWPT 100-year tracer concentrations for Site g3 223
4-43 LLNL MOM and RWPT 100-year tracer concentrations for Site g4 224
4-44 LLNL MOM and RWPT 100-year tracer concentrations for Site g5 225
4-45 LLNL MOM and RWPT 100-year tracer concentrations for Site h1 226
4-46 LLNL MOM and RWPT 100-year tracer concentrations for Site h2 227
4-47 LLNL MOM and RWPT 100-year tracer concentrations for Site h3 228
4-48 RWPT uniformity test initial condition 231
4-49 RWPT uniformity test results averaged over interval 50-100 years 232
4-50 RWPT uniformity test results averaged over interval 150-200 years 233
4-51 Illustration of first passage travel time estimation 242
4-52 GOS sites for validating the RWPT sequestration efficiency calculations, which
are approximately colocated with the GOSAC study locations (Bay of Biscay, New York, Rio de Janeiro, San Francisco, Tokyo, Jakarta, and Bombay) 251
4-53 RAN and ADD sites for validating the RWPT sequestration efficiency calculations. 252
4-54 LLNL MOM and RWPT sequestration efficiency for Sites a3 and a4 255
4-55 LLNL MOM and RWPT sequestration efficiency for Sites a5 and b3 256
4-56 LLNL MOM and RWPT sequestration efficiency for Sites b4 and b5 257
4-57 LLNL MOM and RWPT sequestration efficiency for Sites c3 and c4 258
4-58 LLNL MOM and RWPT sequestration efficiency for Sites c5 and d3 259
4-59 LLNL MOM and RWPT sequestration efficiency for Sites d4 and d5 260
4-60 LLNL MOM and RWPT sequestration efficiency for Sites e3 and e4 261
4-61 LLNL MOM and RWPT sequestration efficiency for Sites e5 and f3 262
4-62 LLNL MOM and RWPT sequestration efficiency for Sites g3 and g4 .... 263
4-63 LLNL MOM and RWPT sequestration efficiency for Sites g5 and h3 .... 264
4-64 LLNL MOM and RWPT sequestration efficiency for Sites GOS1 (Bay of Biscay) and GOS2 (New York) .................................................. 267
4-65 LLNL MOM and RWPT sequestration efficiency for Sites GOS3 (Rio de Janeiro) and GOS4 (San Francisco) .......................... 268
4-66 LLNL MOM and RWPT sequestration efficiency for Sites GOS5 (Tokyo) and GOS6 (Jakarta) .................................................. 269
4-67 LLNL MOM and RWPT sequestration efficiency for Site GOS7 (Bombay) .................. 270
4-68 Error summary for sequestration efficiency calculations ............... 272
4-69 Spatial distribution of relative errors of RWPT sequestration efficiency ... 274
4-70 Number of particle states used for sequestration efficiency in Figure 4-73 .... 275
4-71 Temporal trends in mean predicted sequestration efficiency ............ 277
4-72 Probability density function for predicted sequestration efficiencies ....... 278
4-73 Global sequestration efficiency for an instantaneous source at approximate depths of 800 m, 1,500 m, and 3,000 m .......................... 279
4-74 Global relative sequestration efficiency for an instantaneous source at approximate depths of 800 m, 1,500 m, and 3,000 m ............. 281
4-75 Hill et al. (2004) global injection efficiency at 935 m .......................... 283
4-76 Global relative sequestration efficiency with a reduced surface decay $k = 0.5 \text{ yr}^{-1}$ .................................................. 286
4-77 RWPT uniformity test results averaged over interval 150-200 years, with zero isopycnal slope .................................................. 298
4-78 LLNL MOM and RWPT 100-year tracer concentrations for Site a3, with zero isopycnal slope .................................................. 299
4-79 LLNL MOM and RWPT 100-year tracer concentrations for Site g4, with zero isopycnal slope .................................................. 300
4-80 RWPT uniformity test results averaged over interval 150-200 years, excluding the Gent-Mcwilliams eddy-induced transport terms ......... 302
4-81 LLNL MOM and RWPT 100-year tracer concentrations for Site a3, excluding the Gent-Mcwilliams eddy-induced transport terms. ......................... 303
4-82 LLNL MOM and RWPT 100-year tracer concentrations for Site g4, excluding the Gent-Mcwilliams eddy-induced transport terms. ......................... 304
4-83 LLNL MOM and RWPT 100-year tracer concentrations for Site g3, using an alternate treatment of the Gent-Mcwilliams eddy-induced transport terms. . 309
4-84 LLNL MOM and RWPT 100-year tracer concentrations for Site c5, excluding the Gent-Mcwilliams eddy-induced transport terms. ......................... 311
4-85 RWPT uniformity test results averaged over interval 150-200 years, using an older code version. ......................................................... 314
4-86 Error summary for sequestration efficiency calculations using an older code version. ................................................................. 315
4-87 LLNL MOM and RWPT sequestration efficiency for Site d4 using an older code version. ................................................................. 316
4-88 LLNL MOM and RWPT 100-year tracer concentrations for Site d4 using an older code version. ................................................................. 317
4-89 LLNL MOM and RWPT 100-year tracer concentrations for Site g5, using upwind vertical advection in the OGCM. ................................. 320
4-90 LLNL MOM and RWPT sequestration efficiency for sites RAN1-6 .......... 323
4-91 LLNL MOM and RWPT sequestration efficiency for sites RAN6-12 ....... 324
4-92 LLNL MOM and RWPT sequestration efficiency for sites RAN12-18 ...... 325
4-93 LLNL MOM and RWPT sequestration efficiency for sites RAN19-20 ...... 326
4-94 LLNL MOM and RWPT sequestration efficiency for site ADD1 .......... 327
4-95 LLNL MOM and RWPT sequestration efficiency for site ADD3 .......... 328

5-1 pH mortality data for the combined dataset of Auerbach et al. (1997) for a variety of marine zooplankton and benthic species. ......................... 343
5-3 Mortality data for adult fish as a function of pCO₂ and pH. ................ 349
5-4 Reported CO$_2$ LC$_{50}$ data from Kikkawa et al. (2004) for developmental stages of red sea bream, Japanese sillago, Japanese flounder, and eastern tuna.

5-5 CO$_2$ LC$_{50}$ data by development stage for red sea bream, Japanese sillago, Japanese flounder, and eastern tuna.

5-6 Combined CO$_2$ LC$_{50}$ dataset for developmental fish for red sea bream, Japanese sillago, Japanese flounder, and eastern tuna cleavage-stage eggs with tomato clownfish embryos.

5-7 Combined CO$_2$ LC$_{50}$ dataset for adult and developmental fish for red sea bream and Japanese flounder.

5-8 Reported LT$_{50}$ values for copepods in the Western North Pacific from Watanabe et al. (2006).

5-9 Combined LC$_{50}$ dataset for mortality due to pH depression by CO$_2$ and other acids.

5-10 Combined LC$_{50}$ dataset for mortality due to CO$_2$ exposure.

5-11 Centerline dilution over time for the base case scenarios considered by Caulfield (1996).

5-12 Compiled pH mortality dataset and derived isomortality curves from Auerbach (1996).

5-13 Descent depth as a function of particle diameter and conversion efficiency for 30 cm long cylindrical CO$_2$ hydrate particles released at 1,500 m depth into a typical ocean stratification, from Chow (2008).

5-14 Schematic illustration of the stationary CO$_2$ hydrate plume discharge, adapted from Adams and Wannamaker (2005).

5-15 Schematic illustration of the towed pipe CO$_2$ hydrate plume discharge in a crossflow, from Chow (2008).

5-16 Schematic illustration of the source due to the towed pipe CO$_2$ hydrate plume discharge into a cross flow. Figure adapted from Chow (2008).

5-17 Illustration of towed pipe moving reference frame.

5-18 Schematic illustration of the bottom manifold CO$_2$ discharge from Adams and Wannamaker (2005).
5-19 Schematic of the simulation of organisms drifting through a CO$_2$ plume. ... 399
5-20 Reported LT$_{50}$ values for Western North Pacific copepods from Watanabe et al. (2006). ........................................... 408
5-21 Isomortality function derived from shallow copepod samples from the subarctic and transitional regions of the Western North Pacific of Watanabe et al. (2006). ........................................... 409
5-22 Isomortality function derived from shallow copepod samples from the subtropical region of the Western North Pacific of Watanabe et al. (2006). ... 410
5-23 LC$_{50}$ values for developmental fish from Kikkawa et al. (2003). ........ 413
5-24 Isomortality function derived from the selected subset of developmental fish data, plotted together with the main copepod isomortality function and the selected subset of copepod data. ......................... 414
5-25 Impact volumes for discharge scenarios in terms of pH and pCO$_2$ for mass loadings of 100 kg/s and 1,000 kg/s. ........................ 416
5-26 Centerline variation in pH and pCO$_2$ as a function of time from the source for discharge scenarios with mass loadings of 100 kg/s and 1,000 kg/s. ... 418
5-27 Centerline variation in pH and pCO$_2$ as a function of distance from the source for discharge scenarios with mass loadings of 100 kg/s and 1,000 kg/s. ... 419
5-28 Simulated trajectory of an organism cluster traveling along the centerline of the stationary hydrate plume for a 100 kg/s discharge. ............... 420
5-29 Fractional mortality incurred by an organism traveling down the plume centerline for all scenarios and loadings considered. ......................... 421
5-30 Simulated trajectory of an organism cluster traveling along the centerline of the bottom manifold plume for a 100 kg/s discharge. ............... 422
5-31 Simulated trajectory of an organism cluster traveling along the centerline of the towed pipe (2.5 cm diameter) plume for a 100 kg/s discharge. ... 422
5-32 Integrated mortality of organisms encountering each discharge plume. ...... 424
5-33 Average mortality of organisms encountering each discharge plume. ...... 424
5-34 Simulated trajectory of an organism cluster traveling along the centerline of the stationary hydrate plume for a 100 kg/s discharge, when the dual isomortality functions in Figure 5-24 are used. ........................................ 426
5-35 Fractional mortality incurred by an organism traveling down the plume centerline, when the dual isomortality functions in Figure 5-24 are used. .... 427
5-36 Average mortality of organisms encountering each discharge plume, when the dual isomortality functions in Figure 5-24 are used. ................. 427
5-37 Induced vs natural mortality over various percentages of the ocean volume using the base case isomortality analysis .................................... 436
5-38 Induced vs natural mortality over various percentages of the ocean volume using the base case isomortality analysis ..................... 437
List of Tables

3.1 Timestep sensitivity for step diffusivity correction methods for the well-mixed condition test. ................................................................. 84
3.2 $L(\xi')$ for the 5 possible shapes of the $\xi'_c(z'_i)$ curve. .................... 121
3.3 $D_b$, $\frac{dD'}{dz'}$, and $\Delta t$ values used to verify (3.85) for each case. ... 125
4.1 LLNL MOM vertical discretization used in present study. ....................... 143
4.2 Horizontal extent of tracer test source regions. ................................. 187
4.3 Vertical extent of tracer test source regions. .................................... 187
4.4 Summary of relative error in RWPT sequestration calculations for all 70 comparison sites. ................................................................. 273
4.5 Global statistics for RWPT predicted sequestration efficiencies. ............ 276
4.6 Sensitivity of 1,000-year mean RWPT predicted sequestration efficiency to surface decay rate ($k$). .................................................. 285
5.1 Plume characteristics modeled by Caulfield (1996). ............................ 374
5.2 Acute impact estimates from Caulfield (1996). .................................. 379
5.3 Regression coefficients for various isomortality functions of form (5.40) for time in hours and $\Delta p\text{CO}_2$ in kPa. ................................. 410
5.4 Plume characteristics of the modeled discharge scenarios. .................... 417
5.5 Change in predicted centerline and average mortalities (i.e., fractions of organisms killed) when dual isomortality functions are applied. .... 428
Chapter 1

Introduction and Overview

This thesis consists of four studies which, while largely independent of each other, are connected by two main themes. First, all four studies focus on the use of a Lagrangian framework for simulating mass transport in the environment, with particular emphasis on the application of random walk models in a variety of contexts. Second, two of the studies apply these modeling techniques to the evaluation of deep ocean carbon sequestration by direct injection, i.e., a proposed climate change mitigation strategy in which anthropogenic CO₂ is purposefully injected into the deep ocean in order to at least temporarily “sequester” it from the atmosphere. Organizationally, each study exists as a stand-alone chapter with distinct introductory and concluding sections, and the technical content and major findings of each study are summarized in the final chapter of this thesis (Chapter 6). With the aim of serving as a “roadmap” for navigating through the document as a whole, this first chapter provides an overview¹ of how each of the four studies fit into the two major themes of the thesis.

1.1 Lagrangian modeling

The use of Lagrangian models to simulate tracer (e.g., a pollutant) mass transport in the environment is central to each study, and over the course of the thesis there is a gradual

¹Because this discussion is focused on common themes and synergies between the four studies that comprise the thesis, technical details and specific findings are largely deferred to later chapters. Some readers may therefore find it useful to review the more detailed thesis summary in Chapter 6 before proceeding.
progression from general issues to specific applications which should provide the previously unfamiliar reader with a solid introduction to the topic. Taken as a whole, the four studies cover a wide range of considerations within the field of Lagrangian modeling.

The first study (Chapter 2) gives an overview of Lagrangian techniques as they are commonly applied in environmental contexts, and emphasizes that the choice of technique is first and foremost dictated by the desired representation of turbulent diffusion. Indeed, a central advantage of a Lagrangian framework over an Eulerian framework is the flexibility with which diffusion can be represented; “relative” or “absolute” diffusion can be modeled, and diffusion need not be parameterized by an eddy diffusivity. Within the context of absolute diffusion and surface water tracer transport, this chapter compares three specific Lagrangian approaches: random walk particle tracking (RWPT) and forward/backward gaussian puff tracking (FGPT and BGPT). A detailed discussion of implementation and simulation efficiency issues is provided, ultimately concluding that RWPT, while less computationally efficient, offers greater flexibility in a variety of contexts.

The second study (Chapter 3) focuses exclusively on RWPT, and in particular on the use of the approach in one dimension in the presence of step and piecewise linear diffusivity profiles, both of which can introduce severe inaccuracies to RWPT simulations. A generic toolbox of techniques to handle these situations with substantially better computational efficiency is developed, by combining previously proposed and new methods. Since RWPT simulation is widely applied in a variety of fields, these methods could be useful in many applications beyond environmental transport modeling.

The third study (Chapter 4) takes the RWPT approach and applies it to a specific goal, namely to emulate the tracer transport calculations of an ocean general circulation model (OGCM) such that the advantages of a Lagrangian framework can be exploited as an advanced diagnostic of the OGCM. Although these advantages are illustrated via a specific application (calculating ocean carbon sequestration efficiencies, as described below), the general utility of the RWPT approach to oceanographic investigations is emphasized. The study goal is for the most part successfully achieved, although certain remaining implementation challenges may require additional refinement of the RWPT model. One major challenge, handling the OGCM’s sharply varying vertical diffusivities, is successfully addressed using
the method developed in Chapter 3 of this thesis; in fact, it was the original motivation for that chapter. Most importantly, the third study demonstrates the value of using RWPT to extract tracer transport information from an OGCM that would be difficult or even impossible to extract using the OGCM alone.

The fourth and final study (Chapter 5) demonstrates a very different application of the Lagrangian framework, which is employed in two ways. First, RWPT is used to simulate the diffusion of drifting zooplankton relative to a hypothetical CO$_2$ plume centerline, thus illustrating the applicability of RWPT to relative diffusion problems as discussed in Chapter 2. Second, the acute toxicity of each zooplankter is considered via an “isomorality” method, i.e., a Lagrangian framework is used to diagnose the aggregate harm to a plume of organisms by resolving the impact to individuals sampling different portions of the plume over time.

1.2 Deep ocean carbon sequestration

The last two studies (Chapters 4 and 5) consider the other theme of this thesis, deep ocean carbon sequestration by direct injection. This proposed climate change mitigation strategy involves capturing anthropogenic CO$_2$ from large point sources, compressing it, and injecting it into suitable locations in the oceans where it will remain largely sequestered from the atmosphere for at least several centuries. The logic behind this approach is based on expectations that (a) over long timescales the atmosphere and oceans will equilibrate such that most CO$_2$ will eventually reside in the deep ocean (about 80% at present conditions [5]) regardless of where it is initially discharged; and (b) by discharging the CO$_2$ directly to the deep ocean the peak atmospheric CO$_2$ concentration will be reduced, thus lessening deleterious climate effects such as the warming of the atmosphere and the acidification of the surface oceans$^2$.

Chapter 4 considers the expected efficacy of direct ocean injection as a carbon sequestration strategy. In particular, it uses an RWPT model, specifically designed to mimic the tracer

\footnote{For a more comprehensive overview of ocean carbon sequestration, the reader is referred to the introduction of Chapter 5 and to the Intergovernmental Panel on Climate Change (IPCC) Special Report on Carbon Capture and Sequestration [5]; the latter reference also contains an overview of carbon capture and storage in other receptors.}
calculations of the Lawrence Livermore National Laboratory version of the Modular Ocean Model (LLNL MOM), to map out the global distribution of CO$_2$ sequestration efficiencies as predicted by this OGCM. A variety of past studies have considered this same problem, either by using OGCMs to simulate sequestration at specific sites (most comprehensively by the Global Ocean Storage of Anthropogenic Carbon (GOSAC) study [4]), or by generating ocean sequestration maps via a very different (non-Lagrangian) approach (Hill et al. [3]). The spatial trends of the results generally support the observations of previous studies, save for some nuances regarding the relative efficiency of certain regions. For example, the strong variation of mean sequestration efficiency with depth noted in past studies is confirmed, and efficiency is found to be less variable at deeper sites.

Chapter 5 addresses the marine impact aspect of ocean sequestration, seeking to update a previous assessment performed a decade ago (Auerbach et al. [1] and Caulfield et al. [2]) with revised discharge scenarios and newer, more relevant (CO$_2$-induced) acute toxicity data. The discharge scenarios considered are thought to be among the more promising in terms of maximizing dilution. Recent toxicity data indicate that marine organisms are much more sensitive to CO$_2$ perturbations than they are to an equivalent pH perturbation by another acid. Idealized modeling analyses were conducted, incorporating the new toxicity data into a refined version of the Auerbach et al. isomortality approach to estimate cumulative toxicity over time. The analysis suggests that despite the higher sensitivity of organisms to a CO$_2$ plume, dilution strategies can be employed which minimize the predicted impact from an acute toxicity standpoint, and perhaps even from an ecosystem standpoint.

The findings from both studies are caveated, Chapter 4 due to the coarse resolution of the parent OGCM and lingering problems in the RWPT implementation, and Chapter 5 due to the idealized nature of the calculations and the use of a small group of target organisms as representative of ocean ecosystems. Nonetheless, together they provide some insight into two overarching technical questions that will need to be addressed should there ever be renewed interest in implementing ocean sequestration as a climate change mitigation strategy.
Bibliography


Chapter 2

A Comparison of Three Lagrangian Approaches for Extending Near Field Mixing Calculations$^{1,2}$

---

$^1$This chapter has been previously copyrighted by Elsevier Ltd. and is therefore excluded from the MIT copyright covering other chapters of this thesis. Citation: P.H. Israelsson, Y.D. Kim, and E.E. Adams. A comparison of three Lagrangian approaches for extending near field mixing calculations. *Environmental Modelling and Software*, 21: 1631-1649, 2005.

$^2$Erratum: Equation (3) is true only if $E$ is a symmetric tensor.
A comparison of three Lagrangian approaches for extending near field mixing calculations

Peter H. Israelsson a,*, Young Do Kim b,1, E. Eric Adams c,2

a Department of Civil & Environmental Engineering, Massachusetts Institute of Technology, 48-216, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
b School of Environmental Science and Engineering, Inje University, 607 Obang-dong, Gimhae, Gyeongnam 621-749, South Korea
c Department of Civil & Environmental Engineering, Massachusetts Institute of Technology, 48-216B, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Received 31 August 2004; received in revised form 30 June 2005; accepted 22 July 2005
Available online 3 October 2005

Abstract

Lagrangian techniques have previously been employed to extend initial mixing calculations beyond the near field, either alone or in combination with Eulerian models. Computational efficiency and accuracy are of prime importance in designing these ‘hybrid’ approaches to simulating a pollutant discharge, and we characterize three relatively simple Lagrangian techniques in this regard: random walk particle tracking (RWPT), forward Gaussian puff tracking (FGPT), and backward Gaussian puff tracking (BGPT). RWPT is generally the most accurate, capable of handling complexities in the flow field and domain geometry. It is also the most computationally expensive, as a large number of particles are generally required to generate a smooth concentration distribution. FGPT and BGPT offer dramatic savings in computational expense, but their applicability is limited by accuracy concerns in the presence of spatially variable flow or diffusivity fields or complex no-flux or open boundary conditions. For long simulations, particle and/or puff methods can transition to an Eulerian model if appropriate, since the relative computational expense of Lagrangian methods increases with time for continuous sources. Although we focus on simple Lagrangian models that are not suitable to all environmental applications, many of the implementation and computational efficiency concerns outlined herein would also be relevant to using higher order particle and puff methods to extend the near field.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Lagrangian; Mass transport simulation; Random walk particle tracking; Gaussian puffs

1. Introduction

Contaminant mixing near pollution sources is usually handled using an initial (or near field) mixing model. Our interest here concerns the use of Lagrangian models that can be used to extend the domain of near field analysis. While similar considerations exist for groundwater and atmospheric plumes, our concern is principally with surface water plumes such as those released to rivers, lakes and coastal waters.

Initial mixing models in common use for surface waters include CORMIX-GI (Jirka and Akar, 1991; Doneker, 2005), RSB (Roberts et al., 1989), Visual Plumes (EPA, 2005), and VISJET (Lee et al., 2000; Lee and Wang, 2005). All are basically steady state...
and focus on active mixing (e.g., plume entrainment) near the source. While ambient mixing processes can be included, they are usually handled in a steady state manner by tacking on a 1-, 2- or 3-D diffusivity calculation. Brooks (1960) provided an early example in simulation of lateral diffusion downstream from a coastal wastewater diffuser.

Yet transient effects are often important, as in discharge to a reversing tidal current where previously discharged mass may “return to the scene”. Lagrangian models, in which mass is book kept as particles or puffs, represent a convenient way to keep track of the resulting build-up of background concentration. Other advantages of Lagrangian (as opposed to Eulerian grid-based methods) include the following: (1) they are inherently mass conservative, (2) they can more easily represent sub-grid scale processes without suffering from numerical dispersion, (3) they concentrate computational effort at the point of interest, (4) they offer advantages with certain types of boundary conditions, and (5) they are useful for visualizing transport pathways. Disadvantages include (1) computational expense, (2) noisy concentration distributions if insufficient particles/puffs are used, and (3) difficulty in simulating chemical kinetic processes that are not first order.

The term “Lagrangian model” can mean many things, and we begin by giving a brief overview of this general class of models. In these models, mass is embodied as particles or puffs that are advected and diffused with a computed flow/diffusivity field. A particle is the simplest form in which mass is represented as a point. Stochastic particle methods differ with respect to the amount of correlation between particles. The more common “one particle” stochastic models simulate the motion of a series of independent particles, whereas “two particle” models track the relative position of particle pairs in order to predict concentration variance and relative diffusion (e.g., see Thomson, 1990; Reynolds, 1999; Liu and Du, 2003). Our focus here is on “one particle models”, which differ primarily in their parameterization of turbulence. The simplest models use an eddy diffusivity to predict the evolution of particle position, while more complex models introduce particle “memory” whereby a particle’s velocity and acceleration may be correlated in time for simulating motions smaller than the Lagrangian integral timescale (e.g., see Wolf-Gladrow, 2000).

As implied above, the models considered herein assume that plume spreading can be described by a diffusivity, which warrants some discussion. The output from a near field model is inherently time-averaged because of the characteristic short time scales of plume-induced turbulence. Beyond the near field, however, a plume is spread by ambient processes which occur over a wide range of length and time scales. A common approximation is to describe horizontal plume spreading in terms of a relative diffusivity, i.e., the rate of growth of the variance relative to the center of mass, using tracer release data from the modeled site or from similar sites. Plume variance is typically observed to grow initially in a nonlinear fashion, resulting in a diffusivity that depends on the scale of the plume. Although the enhanced spreading is caused by the plume intercepting increasingly larger eddies, diffusivity is commonly given as a function of the age of a tracer patch (e.g., Okubo, 1971). After long enough time scales, the diffusivity may be assumed to approach a constant value as the plume becomes large enough to be spread by the largest eddies unresolved by the mean flow. This constant diffusivity is consistent with the notion of absolute diffusion, which describes the spreading of the ensemble average plume relative to a fixed coordinate system.

Once a scale-dependent or constant diffusivity is defined, the three methods considered herein can be applied. Of course, the accuracy of the methods will be limited by the accuracy of describing plume spreading by a diffusivity. Even though this approximation may be poor in some cases (e.g., for oceans, see Davis, 1985 or Berloff et al., 2002), it is commonly used in environmental assessments because of its simplicity. The two puff...
methods, FGPT and BGPT, are naturally suited to using either a scale-dependent (based on puff age) or constant diffusivity since the size of a Gaussian puff is completely described by its variance. The particle method, RWPT, is traditionally only used with a constant diffusivity because it is a "one-particle" model in which particles move independently of each other, i.e., each particle is an individual realization of a trajectory and an ensemble of trajectories yields the ensemble average plume. To be consistent with the notion of a constant diffusivity, this mode of simulation is only valid for particle motions longer than the Lagrangian integral timescale (i.e., the timescale over which a particle’s turbulent velocity becomes decorrelated). However, the RWPT algorithm can also be invoked with a scale-dependent diffusivity (based on particle age) to simulate shorter timescales in simple cases. Here particles of the same age would not be independent, but would rather be considered part of a cluster diffusing in a scale-dependent way (e.g., lateral diffusion in a 2-D shear flow might be modeled like this). Applied in this manner, it is perhaps inaccurate to refer to the model as RWPT; it becomes something analogous to a puff model which allows puff distortion. Nonetheless, both algorithms can be applicable to both scale-dependent and constant diffusivities, and can thus be applicable to extending the near field in simple cases.

It should be noted that RWPT and the puff methods considered herein are only applicable to computing mean concentrations, since they are based on diffusivities which describe the mean spreading of a plume. If a constant diffusivity is used, then the result is the ensemble average concentration of many realizations (absolute diffusion), i.e., the standard interpretation of "one-particle" model results. If a scale-dependent relative diffusivity is applied, then the result is the mean concentration distribution of a typical realization, i.e., one for which plume meandering has been ignored.

In comparing the three models, we use examples consistent with absolute diffusion, motivated by the availability of analytical solutions and by the fact that the comparison of the computational aspects of the methods does not change if scale-dependent diffusivities are used. Three of the four simulations consider "point" sources with constant diffusivity, while the last considers a "point" source with diffusivity that varies longitudinally due to temporal flow averaging. In these cases, mass introduced at a source can spread only by absolute diffusion (not by relative diffusion since the mass occupies the same "point" in time and space). Such a diffusivity would be defined by averaging turbulence over some natural length or time scale used to distinguish turbulence from mean flow. It could be dictated by the presence of physical boundaries or by an averaging interval of particular relevance (e.g., the 1-h average exposure concentrations of an aquatic organism). When using mean flow predictions from an Eulerian model, this scale is dictated by the grid size, i.e., the scale to which velocities are resolved.

We begin by describing the structure of the three models and implementation issues in Section 2, and then present simple calculations to characterize the computational efficiency and accuracy of these methods in Sections 3 and 4. Based on these considerations, we outline recommendations in Section 5 as to how these particle and puff methods might be used in combination with each other and/or Eulerian models to create hybrid approaches for extending the near field (e.g., as in Kim et al., 2002; Zhang and Adams, 1999; Suh, in press).

2. Lagrangian approaches and implementation issues

2.1. Random walk particle tracking

The RWPT method represents tracer mass as discrete particles. Particle advection is usually achieved by integrating velocities interpolated from a parent hydrodynamic model, and diffusion is modeled as a random walk process, i.e., using first order integration

\[ \Delta x = A \Delta t + B \cdot Z \sqrt{\Delta t} \]  

(1)

where \( \Delta x \) is the change in 3-D particle position over timestep \( \Delta t \), \( A \) is a deterministic forcing vector and \( B \) is a deterministic scaling matrix defined by

\[ A = u + \nabla \cdot E \]  

(2)

\[ BB^T = 2E \]  

(3)

Here \( u \) is the 3-D velocity vector, \( E \) is a tensor of diffusivities, and \( Z \) is a vector of three independent random variables with zero mean and unit variance (Tompson and Gelhar, 1990). The timestep should be constrained such that the particle excursion length is less than the scale of spatial gradients in the velocity and diffusivity fields. The above integration is consistent with the Fokker–Planck analog between particle density and mass concentration. As discussed previously, Eq. (1) is but one example of a stochastic particle model; alternate formulations can be found in the sample references provided in Section 1.

An important aspect of the RWPT model is the term which allows for spatial variability in the diffusivity field, \( \nabla \cdot E \), sometimes referred to as the "pseudovelocity" term because of its dimensions (Zhang, 1995). As discussed in Section 4, the absence of an equivalent term in FGPT and BGPT compromises the accuracy of these methods in non-uniform diffusivity.
2.2. Gaussian puff tracking methods

Tracer mass is represented as an ensemble of "puffs", where each puff represents the mass loading over a given timestep, i.e.,

\[ m = \int q(\tau) \, d\tau \quad (4) \]

where \( m \) is the total mass of the puff, and \( q \) is the mass loading rate. The mass of each puff is assumed to be distributed along a Gaussian distribution where the peak is located at the center of mass, \( x_s = (x_c, y_c, z_c) \),

\[ C(x, y, z) = \frac{m}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z} \times \exp \left( \frac{(x-x_c)^2}{2\sigma_x^2} - \frac{(y-y_c)^2}{2\sigma_y^2} - \frac{(z-z_c)^2}{2\sigma_z^2} \right) \quad (5) \]

where \( \sigma \) is the standard deviation about the mean and a useful metric for describing puff size (Adams et al., 1986). The puff center of mass is advected by the ambient flow field (usually interpolated from a parent hydrodynamic grid) and puff dimensions grow in time by a diffusion law (e.g., one based on observations). Thus,

\[ \Delta x = u \Delta t \quad (6) \]

and

\[ \sigma^2 = \sigma^2_0 + \int_0^t 2E \, d\tau \quad (7) \]

where \( \sigma_0 = (\sigma_{0x}, \sigma_{0y}, \sigma_{0z}) \) is the initial puff size. Here \( E \) is a vector of diffusivities \( (E_x, E_y, E_z) \). Because puffs can be generated either backward or forward in time, there are two distinct puff tracking methods as illustrated in Fig. 1.

The FGPT method, shown schematically in Fig. 1(a), uses a constant timestep. Thus, in the case of a steady continuous source of a conservative tracer, each puff has the same mass. At each timestep, a new puff is added at the source and previous puffs are advected by Eq. (6) and grow by Eq. (7). In a uniform current, the degree of puff overlap increases downstream from the source. To illustrate this, we define a puff overlap parameter, \( \alpha \), as the ratio of the puff size in the downstream direction \( (\sigma_z) \) and the distance between the centers of adjacent puffs \( (uDelta t) \):

\[ \alpha = \frac{\sigma_z(t)}{u \Delta t} \quad (8) \]

Small values of \( \alpha \) indicate large puff spacing and reduced simulation accuracy, i.e., there may not be enough overlapping puffs to resolve the plume. Conversely, large values of \( \alpha \) indicate a high degree of puff overlap and better plume representation. However, large values of \( \alpha \) also mean that more puffs are being simulated and thus the computational burden is higher. Because \( \sigma \) grows as the puff becomes older and \( u \Delta t \) is constant in the case shown in Fig. 1, \( \alpha \) increases with puff age and distance from the source in FGPT. This is undesirable because closer puff spacing (high \( \alpha \)) is ideally found in the region of highest concentration gradients, i.e., near the source. Furthermore, an unnecessarily high value of \( \alpha \) is likely found in low concentration regions away from the source, resulting in unnecessary computational burden.

These undesirable characteristics of FGPT are addressed by BGPT, as illustrated in Fig. 1(b). Rather than advancing old puffs at each timestep, the entire plume is reconstructed at each desired output time, thus avoiding the need for simulation between outputs. The BGPT method is like a convolution, where time flows backward. Adequate puff coverage is assured by specifying a constant \( \alpha \) and allowing the puff timestep \( \Delta t \) to vary. Puffs near the source (where \( \sigma \) is small) contain less mass than older puffs. Puff spacing is thereby optimized and computational burden is reduced.

Despite its advantages, BGPT is not always superior to FGPT. For example, FGPT can be made more efficient by aggregating puffs when \( \alpha \) becomes unnecessarily small. Furthermore, considerations such as the number of outputs and input/output burden affect the relative efficiency of the methods (see Section 3.4).

2.3. Computing concentration distributions

The particle and puff representations of mass are shown in Fig. 2. Fig. 2(a) shows the analytical solution of a 2-D plume from a continuous line source in a uniform current with constant lateral and longitudinal diffusivity; the representation of this plume by RWPT, FGPT, and BGPT is illustrated in Fig. 2(b–d), respectively. Here puffs are represented by a characteristic size (one standard deviation). RWPT results are plotted as the positions of individual particles released at the source at fixed intervals throughout the simulation. Recall that the use of a constant diffusivity implies absolute diffusion, meaning that the depicted plume represents an ensemble average of many individual realizations.

In RWPT, particle density is converted to mean concentration by overlaying a "counting" grid of arbitrary spatial resolution and binning particles into grid cells. Because particles are discrete packets of mass, the resulting concentration distribution is not smooth, although the smoothness increases as more particles are used. Smoothness can also be improved by using projection functions, whereby mass from a particle is partially distributed over adjacent grid cells (e.g., Bagtzoglou et al., 1992; Moeller and Adams, 1994). As discussed in Section 3, grid resolution is an important consideration.

Calculating the overall concentration distribution for the puff methods is straightforward; the distributions of
individual puffs are superimposed. In contrast to RWPT or Eulerian models, puff methods generate continuous, smooth concentration fields rather than grid-based mean concentrations. While this is convenient if comparing results to an analytical solution, it presents a minor problem when concentrations are to be projected onto a grid (e.g., in a hybrid scheme combining a puff method with an Eulerian model). Simply using puff concentrations at the grid cell centers is inaccurate as this does not conserve mass. A more sophisticated averaging scheme can be used, e.g., by subdividing the grid into smaller cells and averaging over these cells. Another solution is to use grid cell center concentrations and then adjust all concentrations by a constant factor so that total mass is conserved (Kossik et al., 1987), but this does not preserve higher moments of the distribution. The appropriate level of sophistication is case-specific.

Counting grid resolution has different implications for the RWPT and puff techniques. When projecting RWPT particles onto a grid, the quality of the resulting concentration field depends on the particle density per grid cell. Thus, increasing the resolution of the counting grid requires using a larger numbers of particles. This is not the case when projecting puff results onto a grid since their distributions are continuous. This effect is studied in Section 3.

Fig. 1. Simulation of a continuous source of conservative tracer in a uniform current using (a) FGPT and (b) BGPT.

Fig. 2. (a) Concentration contours for a 2-D continuous line source of conservative tracer in a uniform current with constant diffusivity (absolute diffusion) and the representation of this ensemble average plume by (b) RWPT, (c) FGPT, and (d) BGPT.
2.4. Boundary considerations

Implementation issues for three types of boundary conditions are discussed below: source boundary condition, no-flux boundary condition, and open boundary condition.

2.4.1. Source boundary condition

A source boundary condition is a mass loading to the model domain, either as a point source or a distributed source, and can occur instantaneously or continuously. A point source cannot exactly be represented in an Eulerian framework because tracer mass must be spread over a finite volume to form a grid-based concentration, which requires the use of a mixed boundary condition. Point sources are handled more naturally in a Lagrangian framework, where the loading can be specified directly at the source location without a mixed boundary condition. Particles can exactly represent a point source because they have no physical size, whereas puffs can approximate point sources by using a small initial puff size.

For instantaneous point sources (in non-zero diffusivity), multiple particles or a single puff are released at the start of the simulation. Continuous point sources are simply a series of instantaneous sources spaced in time. For RWPT, the multiple particles for each time interval can be spread out over the interval to reduce downstream concentration gaps. We will refer to the total number of particles or puffs introduced in a simulation as the source resolution, which is dictated by downstream accuracy considerations (i.e., to get smooth concentrations). In general, puff models require much lower source resolution than RWPT because puffs are continuous distributions. Section 3 examines the trade-off between source resolution and accuracy.

Extending near field calculations requires disaggregating the predicted concentration field at the edge of the near field into particles or puffs. Zhang and Adams (1999) and Kim et al. (2002) provide examples of how near field concentrations predicted at the trap height of a buoyant plume can be converted into particles. The near field model output provides a distributed source that can be modeled as a collection of point sources, where the overall source resolution is the product of the number of point sources and the number of particles/puffs per source. Lagrangian models can generally better represent distributed sources than Eulerian models (which can only resolve sources larger than its grid size). In general, particles can better reproduce distributed sources because they have no shape, although small puffs can also be used.

2.4.2. No-flux boundary condition

No-flux boundary conditions are most commonly encountered at solid walls. In RWPT, particles are displaced by a combination of advection and diffusion. If the hydrodynamic field satisfies the no slip or free slip condition at the boundaries and the numerical scheme is perfectly accurate, particles will not cross these boundaries by advection. However, particles may cross the boundaries by diffusion, and the simplest solution is to reflect particles off of the boundaries. Reflection can also be used in the event that particles are advected across these boundaries because of errors in the hydrodynamic field or numerical integration scheme. Any number of boundaries and domain geometry can be handled in this manner. For the two puff methods, the situation is more complicated. The puff center of mass, which is transported by advection alone, should not cross a no-flux boundary if the numerical integration is accurate. The concentration distribution which the puff represents, however, may cross these boundaries. Enforcing the boundary condition requires the use of an image source, where an imaginary puff is placed on the other side of the boundary (equidistant from the boundary as the real puff). The imaginary puff moves and grows as the mirror image of the real puff, and the sum of the real and imaginary puffs remains constant such that total mass within the domain is conserved. The concept is easily applied to the case of a single straight wall, but becomes difficult for more complex boundary geometry. While theoretically feasible, the handling of complex no-flux boundaries is a likely hindrance to using FGPT and BGPT. A more sophisticated puff model which allows for puff distortion (e.g., Sykes and Henn, 1995) might be better suited in this regard, but is beyond the present scope.

2.4.3. Open boundary condition

Imposing a constant concentration (i.e., Dirichlet) condition at an open boundary in Eulerian models is straightforward; a constant concentration at the edge of the domain is specified. In RWPT, the counting grid must first be chosen so that a constant particle density can be specified at the edge of the domain. At each timestep, particles are allowed to pass from the boundary grid cell into the domain based on the velocity and diffusivity at the edge of the domain. Particles crossing into the boundary cell from the domain are ignored and the boundary particle density is reset at the beginning of each timestep.

Conversely, there is no accurate way of specifying a constant concentration condition in the FGPT and BGPT methods. This can be illustrated by the condition that the boundary concentration should equal zero (physically unrealistic in most environmental applications but sometimes imposed in modeling contexts). In RWPT, no particles would enter from the open boundary and particles leaving through the boundary would be removed from simulation, thus satisfying the boundary condition. In FGPT and BGPT, the shape of the
2.5. Modeling non-conservative behavior

For first order processes, e.g., a decay of the form \( C = C_0 e^{-kt} \), the mass of each puff or particle is simply reduced at each timestep. For higher order processes (e.g., when a rate of transformation depends on concentration), the situation is more complex because superposition no longer applies. Concentrations must first be calculated (either by superimposing puffs or by binning particles) before the transformation can be applied. After transformation, the overall concentrations must be redistributed into discrete puffs or particles so that the simulation can proceed. This process can be handled in a number of different ways (e.g., Tompson and Gelhar, 1990) but is likely to homogenize the puff/particle distribution to some extent. For example, in RWPT, spatial concentration gradients beneath the counting grid scale will likely be lost.

3. Efficiency considerations

The relationship between computational efficiency and accuracy is an important consideration in comparing the Lagrangian methods. Two simple test cases (a 1-D case and a 3-D case) for which analytical solutions exist are used below to illustrate the computational efficiency of the approaches. These test cases were selected because all three Lagrangian methods are capable of reproducing the analytical solution if the source is sufficiently resolved, i.e., if enough particles or puffs are used. As noted in Section 1, these test cases use constant rather than scale-dependent diffusivities and are thus implicitly absolute diffusion examples (i.e., predictions of an ensemble average plume). This choice was motivated by the availability of exact analytical solutions and by the fact that the efficiency considerations noted below are unaffected by the use of a scale-dependent diffusivity. Before presenting these simulations, the relationship between computational expense and source resolution is examined.

3.1. Computational expense and source resolution

The increase in computational expense will be a limiting factor in choosing source resolution. The two main factors that control how long a simulation takes are the number of particles/puffs introduced per timestep \( (N_P) \) and the number of total timesteps \( (N_T) \). Computational expense can be expressed as the number of calculations required during the simulation \( (N_{calc}) \), where one “calculation” is needed to advance a single particle or puff by one timestep (i.e., calculation time is assumed to be the same for each method on a per particle or puff basis). Defining computational expense in this manner avoids complicating issues such as computer resources, code efficiency, and I/O overhead costs. For RWPT, \( N_{calc} \) is defined (for large \( N_T \)) by:

\[
N_{calc} = \sum_{n=1}^{N_T} nN_P = N_P \frac{N_T(N_T + 1)}{2} = \frac{1}{2} N_P N_T^2
\]

where \( N_T \) is a function of the timestep \( \Delta t \) and the total simulation time \( T_{tot} \):

\[
N_T = T_{tot} \Delta t
\]

Thus, for a steady continuous source, computational expense varies linearly with \( N_P \) and quadratically with \( N_T \). Halving \( \Delta t \) or doubling \( T_{tot} \) increases computational expense by a factor of four, whereas doubling \( N_P \) doubles the computational expense. For FGPT, the relationship between \( N_{calc} \), \( N_P \), and \( N_T \) is identical to RWPT because FGPT advances each puff with a fixed timestep. Although scaling differences between FGPT and RWPT may arise from practical implementation issues, \( \Delta t \) RWPT will remain significantly more expensive in realistic applications.

For BGPT, the scaling relationship is markedly different. In a 1-D plane source, \( N_P = 1 \) for BGPT. Because the puffs are completely remapped at each output time, \( N_{calc} \) is equal to the number of outputs \( N_{calc} \) proportional to \( N_P \) and to \( \Delta t^{-2} \). FGPT may in practice scale differently than RWPT.
multiplied by the number of timesteps per output. The number of timesteps per output is time variable as it depends on the variables in Eq. (8). The fact that the size of each timestep depends on the previous timestep, coupled with possible spatial and temporal variability in current velocity and diffusivity, prevents the derivation of a useful scaling relationship for $N_{calc}$. Nonetheless, BGPT’s relative efficiency over FGPT (and thus RWPT) in terms of the number of calculations is obvious by considering how BGPT and FGPT would be used to construct a plume at time $t$ using $n$ puffs. FGPT must advance $n$ puffs by one timestep (i.e., all puffs in the system at time $(t - \Delta t)$ plus a new puff added at the source), while BGPT must track a single puff backward in time for $n$ time intervals. Thus, the number of calculations in FGPT for that timestep is roughly the same as the number of calculations in BGPT for that output interval. The difference is that calculations are only performed at the output times for BGPT, while they are performed at each timestep for FGPT. The number of calculations required for the two methods only converges in the limit of a BGPT simulation in which the output interval equals the timestep in a corresponding FGPT simulation. Although this means that BGPT is generally less computationally expensive, other factors such as input/output considerations may significantly affect its computational efficiency (see Section 3.4).

3.2. Efficiency considerations in one-dimensional space

To illustrate the computational superiority of FGPT and BGPT over RWPT, we consider the simple case of a steady continuous loading ($q' = 10^{-3}$ kg/m²/s) in a uniform current ($v = 0.1$ m/s) with homogeneous diffusion ($E = 10$ m²/s). Tracer transport is governed by

$$\frac{\partial C}{\partial t} + u \frac{\partial C}{\partial x} = D \frac{\partial^2 C}{\partial x^2}$$

for which the exact analytical solution is given by Holley and Harleman (1965):

$$C = \frac{q}{2u} \exp \left( \frac{x}{2u} \right) \left[ \text{erf} \left( \frac{x + ut}{\sqrt{4Et}} \right) + 1 \right] \times \exp \left( \frac{-x}{2E} \right) \quad \left[ \text{erf} \left( \frac{x - ut}{\sqrt{4Et}} \right) + 1 \right] \exp \left( \frac{-x}{2E} \right)$$

where the ± sign is negative for $x > 0$ and positive for $x < 0$.

The test case was simulated using different source resolutions, and the resulting concentration profiles after 6 days are plotted together with the exact solution in Fig. 3. Three different grid spacings were used for calculating RWPT concentrations: 0.05, 0.5, and 5 km. The quality of the solutions clearly improves with better source resolution for each method, although the improvement is most dramatic for RWPT. For even a small number of puffs, both puff methods reproduce the analytical solution with fair accuracy. RWPT requires orders of magnitude more particles to achieve results of comparable quality. The RWPT concentration profile is not smooth like the puff results as concentrations are based on particle binning. The contrast in efficiency is more clearly shown in Fig. 4, where mean absolute relative error is plotted against source resolution. The error was calculated as:

$$\text{Error} (\%) = \frac{100}{N_{calc}} \sum_{n=1}^{N_{calc}} \frac{|C_{m,n} - C_{e,n}|}{C_{e,n}}$$

where $C_{m,n}$ and $C_{e,n}$ are the model-predicted and exact analytical concentrations for grid point $n$, and $N_{calc}$ is the number of grid points with analytical concentrations above 1% of the maximum analytical concentration in the domain. Grid points with smaller concentrations were omitted to prevent small errors in low concentrations from dominating the overall relative error. In BGPT, the number of puffs per day is not constant nor is it easily predicted prior to simulation; it was calculated as the total number of puffs needed to construct the concentration field divided by the output time. In Fig. 4 and subsequent error plots, some of the trend lines have positive slopes for low source resolutions, which incorrectly suggests an increase in error due to improved source resolution. These anomalous trends are merely artifacts of the error calculation. For low source resolutions, each particle represents a large amount of mass and many grid cells have no particles, i.e., $C_m = 0$. Grid cells with a single particle can therefore have an error much larger than 100% while cells with no particles have a 100% error. Thus, cells with no particles can actually reduce the overall mean error. As the source resolution increases but remains low, the number of empty cells may decrease more than the number of cells with a single particle, causing the error to increase. Fig. 4 shows that the source resolution required for RWPT greatly exceeds the resolution required for either puff method, meaning that RWPT is much more expensive. Although we might expect BGPT to be more accurate than FGPT because it optimizes puff spacing, this trend is not clear in Fig. 4. This issue is revisited in Section 3.3.

The accuracy of the RWPT results is dependent on the resolution of the grid used to convert particle densities to concentrations. Fig. 3 presents RWPT results for three different grid resolutions ($\Delta x = 5, 0.5,$ and 0.05 km) and three different source resolutions ($N_p = 24,000, 2400,$ and 240 particles/day). As expected, concentrations are most variable when the finest resolved grid ($\Delta x = 0.05$ km) is combined with the lowest source resolution ($N_p = 240$ particles/day).
source resolution is clearly too low for such a fine grid resolution; the number of particles per grid cell is too low and the amount of mass contributed by each particle is too high. The agreement between RWPT and the analytical solution improves as the source resolution increases, or as the resolution of the grid decreases. This effect can be quantified using the error analysis in Fig. 4. For low source resolution, the error is lowest for the coarsest grid resolution; a much larger number of particles are required to achieve a similar error for the finer grid resolutions.

While using a coarser grid resolution lowers the amount of noise in the concentration field, it also decreases the extent to which details of the exact solution can be reproduced. After 6 days of injection, the contaminant front has traveled about 52 km by advection and has spread longitudinally over a 10−15 km region by diffusion. The 5-km grid spacing only uses three points to represent this front and thus the front is not smooth. Even in the limit of an infinite number of particles (i.e., no noise), the quality of the solution cannot improve beyond some minimum error for this reason.

Fig. 3. Concentration profiles for a 1-D simulation of a continuous source in a uniform current using (a) FGPT, (b) BGPT, and RWPT with counting grid resolutions of (c) $\Delta x = 5$ km, (d) $\Delta x = 0.5$ km, and (e) $\Delta x = 0.05$ km. In each panel, simulation results for three different source resolutions are plotted together with the exact analytical solution.
The same is true for any grid resolution, but the minimum error attainable in the limit of infinite source resolution decreases as the grid resolution improves. As shown in Fig. 4, a minimum error for the 5-km grid spacing of approximately 4% is achieved using a source resolution of about 2400 particles per day. Using a larger number of particles has minimal effect on the quality of the concentration field for this grid resolution. In contrast, the 0.5-km resolution results reach a 4% error when about 10,000 particles per day are used, and the error continues to decrease for higher source resolutions. The rate at which the RWPT error decreases with source resolution is initially the same for all grid resolutions; mean absolute error is proportional to $N_p^{1/2}$, where $N_p$ is the number of particles per day. This is consistent with past studies (Tompson and Gelhar, 1990; Bzagzoglou et al., 1992). This rate gradually begins to decrease and the minimum attainable error for each grid resolution is eventually reached. Although the range of source resolutions evaluated is insufficient to identify the minimum error for the 0.5-km grid resolution, it is clearly less than or equal to 0.1%. For the 0.05-km grid spacing, over 100,000 particles per day are required to reach the minimum error of the 0.5-km grid (4%), and extrapolating the slopes of the 0.5-km and 0.05-km grid resolutions beyond the maximum source resolution simulated (24,000,000 particles/day) suggests that the minimum achievable error will be lower for the 0.05-km grid than for the 0.5-km grid.

Source and grid resolutions must be considered jointly in RWPT simulation. For a real environmental problem, the grid resolution will be dictated by the dynamics that need to be resolved. Since computational expense increases linearly with the number of particles used, it is desirable to use the fewest number of particles possible. While the lowest acceptable source resolution depends on the error that the modeler is willing to accept, it is possible to estimate the maximum number of particles that should be used with an error analysis similar to the one presented above. One important modification is necessary, however, since the exact analytical solution is rarely available. Instead of comparing a series of simulations to an exact solution, simulations of increasing source resolutions can be compared to each other. The maximum source resolution that should be used is the one at which the difference between successive runs tends to zero (i.e., where the error plateaus because the solution is no longer noisy). For large domains, the analysis may need to be restricted to the domain subregion that is of most interest.

The above conclusion holds regardless of how the error comparison is performed. In the present analysis, error for RWPT simulations was calculated by comparing a grid cell average value to the analytical solution at the center of the cell. Error could have been reduced by averaging the analytical solution over the grid cell. However, there would still exist a minimum attainable error even in the limit of infinite source resolution. Only in the limit of infinitesimally small grid spacing would RWPT results exactly match the analytical solution, but such a grid resolution is obviously unattainable in practice.

3.3. Efficiency considerations in three-dimensional space

The relative efficiency of the three Lagrangian methods is dependent on the number of dimensions being simulated, and there are nuances of the counting grid resolution impact on RWPT results that are not apparent in the 1-D case. Therefore, a simple 3-D case is presented below. A continuous source loading ($q=1$ kg/s) is placed in a uniform current ($u=0.1$ m/s) with homogeneous, anisotropic diffusion ($E_x=E_y=10$ m$^2$/s, $E_z=10^{-7}$ m$^2$/s), for which the transport is governed by:

$$\frac{\partial c}{\partial t} + u \frac{\partial c}{\partial x} = E_x \frac{\partial^2 c}{\partial x^2} + E_y \frac{\partial^2 c}{\partial y^2} + E_z \frac{\partial^2 c}{\partial z^2}$$  \hspace{1cm} (14)$$

and for which the exact analytical solution is:

$$C = \frac{q \exp\left(\frac{a b}{4}\right)}{16\pi (E_x E_y E_z)^{1/2}} \left[ \exp\left(2\sqrt{ab}\right) \text{erfc}\left(\frac{a}{2\sqrt{b}}\right) + \exp\left(-2\sqrt{ab}\right) \text{erfc}\left(\frac{a}{2\sqrt{b}}\right) \right]$$  \hspace{1cm} (15)$$

where $a = x^2/(4E_x) \left(1 + (E_x y^2/E_y x^2) + (E_z z^2/E_x x^2)\right)$ and $b = a^2/(4E_x)$. 

---

**Fig. 4.** Mean absolute relative error as a function of source resolution for a 1-D simulation of a continuous source in a uniform current. RWPT results have been projected onto three different grid resolutions.
As in the 1-D case, the quality of the RWPT solution is highly dependent on the choice of grid resolution. Fig. 5 shows the centerline ($y=0$) concentration distribution for several RWPT simulations of different source and grid resolutions together with the exact analytical solution after 4 days of injection. Vertical grid spacings of 0.1, 1, and 10 m were evaluated with constant longitudinal and lateral resolution of 1 km (only vertical spacing was varied for computational reasons). As the number of particles increases, the variability of the concentration profile decreases in all cases. For the coarse grid resolution ($\Delta z = 10$ m), the agreement is similar for 240,000 and 2,400,000 particles per day, suggesting that the former is sufficient. This is not true for the finest grid resolution ($\Delta z = 0.1$ m), where the prediction improves slightly when the source resolution is increased from 240,000 to 2,400,000 particles per day.

Although the distribution becomes noisier for finer grid resolutions, a better representation of the distribution is possible with finer grids. This is demonstrated by the degree to which the RWPT simulation can reproduce the high concentration region near the source, which improves as the vertical resolution is increased (see insets in Fig. 5). The 3-D case thus presents an added complexity in that the counting grid resolution in one-dimension affects the concentration distribution in the other two.

The relationship between mean absolute relative error and source resolution is shown in Fig. 6 (error has been calculated in the same way as in the 1-D case). The minimum attainable error for a given grid resolution is also observable here. For the maximum source
resolution simulated (24,000,000 particles per day), this minimum error was reached only for the coarsest resolution, $\Delta z = 10$ m. The $\Delta z = 1$ m line is beginning to flatten out at this point as well, whereas the finest resolution ($\Delta z = 0.1$ m) is still showing reduced noise with increasing source resolution. The relative advantage of the puff methods over RWPT in three-dimensions is clearly greater than in the 1-D case. In the 1-D domain (Fig. 4), approximately four orders of magnitude more particles than puffs are required to achieve similar errors. In the 3-D case, this difference increases to about five or six orders of magnitude. This is due to the fact that particles are spread out over more dimensions in the 3-D case and thus downstream particle dilution occurs more rapidly than in the 1-D case. This dilution effect does not occur for puffs, since modeling more dimensions simply means tracking additional spatial moments of each puff.

The extent of particle dilution depends on the size of the domain and the quality of the solution decreases with increasing distance from the source. Fig. 7 shows the error along three separate transects at 1.5, 5.5, and 30.5 km downstream of the source. For a given grid resolution, the error increases with distance (i.e., the error curves shift toward the right). Furthermore, the source resolution required to reach the minimum achievable error (i.e., where the curves flatten out) for a given grid resolution increases with increasing distance from the source. This effect is due to the fact that the horizontal axis of the error plots is the number of particles being discharged at the source, and the particle density per grid cell achieved for a given source resolution decreases with increasing distance from the source. Plotting error as a function of source resolution is appropriate from a practical standpoint since source resolution is the metric that can be varied (downstream particle density cannot). This effect may prevent the application of RWPT to long-term simulations in large domains; the computational expense required to achieve results of acceptable quality far away from the source may be overwhelming.

In contrast to RWPT, the accuracy of the puff methods improves further away from the source. This is mainly due to two reasons. First, spatial concentration gradients become smaller (see Fig. 5), meaning that there are less sharp features that are poorly resolved by low source resolution. Second, the importance of the initial puff size decreases as the puff grows. Furthermore, FGPT becomes more accurate relative to the BGPT method because there is more puff overlap in FGPT.
with increasing distance from the source (i.e., \( \alpha \) increases), whereas puff overlap remains constant in BGPT.

It should be noted that the reason BGPT appears generally more accurate in Fig. 6 is because the calculation of overall domain error is dominated by the high concentration near source regions (where BGPT has more optimal puff spacing); in Fig. 4 the concentration front is further downstream (where FGPT has higher puff overlap) at the time of the error calculation and thus the overall BGPT and FGPT errors are comparable.

### 3.4. Input/output considerations

Although the previous discussion suggests that BGPT is usually more efficient than FGPT, this is not necessarily the case because BGPT incurs input/output computational expense. All three methods often rely on a parent hydrodynamic model for flow field data, which are generally read from an external data file. RWPT and FGPT only require hydrodynamic values for the current point of the simulation; no knowledge of the past is needed. In contrast, to map the concentration field at any given time, BGPT requires access to the entire preceding hydrodynamic record. Whether stored completely in memory or read piecewise from an external file, BGPT incurs some “overhead” computational cost. Because this process is repeated at each output, computational expense increases with output frequency. The magnitude of the effect is case-specific, since it depends on the simulation details, the model code, and computer resources. Furthermore, the fact that BGPT requires knowledge of the entire hydrodynamic record makes it impractical to incorporate directly into a hydrodynamic model for simultaneous hydrodynamic and mass transport calculations; RWPT and FGPT are more practical in this respect.

### 4. Application to more complex flow fields

Although FGPT and BGPT are more computationally efficient, they are not as generally applicable as RWPT. The more non-uniform the velocity and diffusion fields become, the less suitable become the puff methods and the Gaussian concentration distributions upon which they rely. In this section, the limitations of FGPT and BGPT are demonstrated by test cases that include spatially variable flow and diffusivity.

#### 4.1. Spatially variable velocity field

The transport of an instantaneously released tracer in a steady 2-D shear flow, in which the current velocity varies linearly in the lateral direction (\( u(y) = \alpha y \)), with homogeneous, isotropic diffusivity is governed by:

\[
\frac{\partial C}{\partial t} + (u_0 + \lambda y) \frac{\partial C}{\partial x} - \frac{E}{45} \left( \frac{\partial^2 C}{\partial x^2} + \frac{\partial^2 C}{\partial y^2} \right)
\]

(16)

The analytical solution of Eq. (16) is (Okubo and Karwe, 1969):

\[
C = \frac{m'}{4\pi Et(1 + \lambda^2 t^2/12)} \exp \left( \frac{(x - x_0 - u_0 t - 0.5\lambda(y - y_0) t)^2}{4Et(1 + \lambda^2 t^2/12)} + \frac{(y - y_0)^2}{4Et} \right)
\]

(17)

where \( m' \) is mass per unit depth released at \((x_0, y_0, t_0)\), \( E \) is the horizontal diffusivity, \( u_0 \) is the centerline velocity, and \( \lambda = \partial u / \partial y \) is the gradient of velocity in the lateral direction.

The puff methods are unable to resolve the non-uniform velocity field because the dimensions of the puffs grow by a diffusion law; the velocity field is only used to determine the trajectory of each puff’s center of mass. As the puffs grow larger (i.e., with increasing distance from injection), the accuracy of the solution decreases because the puffs cover a larger range of velocities and experience a greater shear. The magnitude of this effect can be expressed as a (time-dependent) dimensionless shear number \( S(t) \):

\[
S(t) \equiv \frac{\lambda \sqrt{\alpha_0^2 + 2Et}}{u_0}
\]

(18)

It should be noted that the latter equality in Eq. (18) is true only for homogeneous, isotropic diffusion. \( S(t) \) is a measure of the velocity deviation experienced by the puff, normalized by the centerline velocity. RWPT simulation results for three different values of \( \lambda \) are shown in Fig. 8 (\( m' = 10^9 \) kg/m, \( u_0 = 0.1 \) m/s, \( E = 10 \) m²/s) at 6 days after the mass release. Puff results are not included in the figure for two reasons: (1) they look exactly the same in each case (very similar to the analytical solution for \( S = 0 \)), and (2) there is no difference between the FGPT and BGPT methods for an instantaneous release.

Although RWPT is able to reproduce the analytical solution in each case, FGPT and BGPT results can only be considered reasonable for small values of \( S \). Since the puff solution is independent of \( \lambda \) and \( S \), the quality of the puff prediction decreases with increasing \( S \) (i.e., as velocity shear becomes more important). Furthermore, since \( S \) is also a function of time, the quality of the puff solution for a given \( \lambda \) decreases with increasing time. Thus, even small velocity gradients can introduce significant errors if the simulation time is sufficiently long. The increase in error with increasing velocity gradient and/or simulation time is further illustrated in Fig. 9, where mean absolute relative error is...
Fig. 8. Concentration contours (mg/L) of an instantaneous release in 2-D flows with shear numbers of (a,d) 0, (b,e) 0.032, and (c,f) 0.32, as computed by the exact analytical solution (top, a–c) and RWPT (bottom, d–f). The velocity profile is plotted along with the analytical solution.
plotted as a function of $S$. Fig. 9 contains the results of five puff and five RWPT simulations with different $\lambda$ values, each simulation has 12 outputs. The puff results have the lowest error for small $S$, but the error increases sharply with $S$ and eventually grows larger than the error for RWPT. RWPT error also increases with $S$, but at a slower rate than the puff error. The cause of the error increase is different for each method. For FGPT and BGPT results, the solution degrades with increasing $S$ because the velocity shear is not accounted for. For RWPT, the error is caused by the increase in noise associated with particle dilution; as $S$ increases, the number of particles per grid cell decreases. The trends are best illustrated by the clusters of points formed by each simulation in Fig. 9. Each cluster corresponds to a puff or RWPT simulation of a given $\lambda$ value. The lowest and highest errors in each cluster correspond to the first and last outputs of the simulation, and the range of errors increases as $\lambda$ increases. In contrast to the puff results, the RWPT error can be compensated for by using more particles.

In most applications, velocity will be spatially variable to some extent. This means that the puff methods are likely to decrease in accuracy as a simulation progresses to the point at which the puff size is of the same scale as the velocity variations. FGPT and BGPT are therefore in practice probably not suitable over long simulation times and large distances. RWPT would be better suited in such applications. More sophisticated puff models could also be considered in this context if the computational burden of RWPT becomes prohibitive, such as those that allow puff distortion (e.g., Sykes and Henn, 1995) or puff splitting (e.g., de Haan and Rotach, 1998), or that combine puffs and particles (e.g., Hurley, 1994).

4.2. Spatially variable diffusivity field

A continuous source loading is placed in a steady current in a 1-D domain in which the diffusivity increases with $x^2$ from the injection source ($x = 0$) to the end of the domain ($x = L$). This is analogous to a tidal inlet of constant cross-sectional area where the tidal excursion amplitude grows linearly from the head of the inlet ($x = 0$) to the mouth ($x = L$). The resulting time-averaged diffusivity is spatially variable. Steady state tracer distribution is governed by:

$$u \frac{dC}{dx} = \frac{d}{dx} \left( \frac{\partial C}{\partial t} \right)$$

which has the analytical solution (Arons and Stommel, 1951):

$$C = C_0 \left[ 1 - \exp \left( \frac{u}{aL} \left( \frac{x - L}{x} \right) \right) \right]$$

where $C_0 = (q^*/u)$, $u =$ current velocity, $q^* =$ rate of mass injection per unit area; $E = a\alpha^2$ = diffusivity, and $\alpha =$ proportionality coefficient (with units [1/T]). The above solution imposes boundary conditions $C (x = 0) = C_0$ and $C (x = L) = 0$.

Although the diffusivity is scale-dependent from the perspective of the contaminant plume, this is still a case of absolute diffusion since the increase in diffusivity with downstream distance is due to temporal flow averaging (on a fixed coordinate system) rather than the growth of the contaminant plume. Nonetheless, it is analogous to how relative diffusion could be modeled in these methods using a scale-dependent diffusivity of the form $E = a\alpha^2$ (e.g., Okubo, 1971). The growth of a tracer patch undergoing scale-dependent diffusion is given by:

$$\sigma(t)^2 = \sigma_0^2 + \int_0^t 2E(t) \, dt.$$  

In this case, however, diffusivity varies in space rather than time. Using, as the puff models do, $t = x/u$ with $E = a\alpha^2$, leads to

$$\sigma(x)^2 = \sigma_0^2 + \int_0^x 2a\alpha^2 \frac{x}{u} \, dx = \sigma_0^2 + \frac{2a\alpha^3}{3u}$$

but this ignores the variation in diffusivity across the puff.

The spatial variation in diffusivity causes a skewed concentration distribution. Since the gradient in diffusivity increases linearly with $x$ ($dE/dx = 2a\alpha$), the skewness is more pronounced further downstream. Such gradients in diffusivity are accounted for in RWPT by the deterministic ‘pseudovelocity’ term $\nabla E$ in Eq. (2). The FGPT and BGPT methods lack such a correction term and become less accurate with increasing $x$, as a Gaussian puff becomes an increasingly worse approximation of the concentration distribution. The
Fig. 10. Predicted concentration distributions after 160 days for a one-dimensional uniform current with spatially variable diffusivity corresponding to (a,d) $PeLZ_{10}$, (b,e) $PeLZ_{5}$, and (c,f) $PeLZ_{2}$. For further details, see Tables 3-5.
magnitude of this effect depends on the relative strength of advection to diffusion, which can be expressed as a Peclet Number ($Pe$):

$$Pe(x) = \frac{u x}{E(x)}$$

Simulation results for a range of Peclet numbers (evaluated at $x = L = 10^{5}$ m for $u = 10^{-7}$ m/s, $q = 10^{-7}$ kg/m$^2$/s) are presented in Fig. 10 after 160 days (top) and 50 days (bottom) of tracer release. After 160 days, steady state has been reached and the RWPT result matches the analytical solution, while the puff results do not. The failure of the puff methods is, however, not solely due to the diffusivity gradient; it is also due to the difficulty in enforcing $C = 0$ at the downstream open boundary for the puff methods (see Section 2.4.3). To isolate the error introduced by variable diffusivity, puff results are also compared to RWPT results at an earlier simulation time ($t = 50$ days, before steady state and before the problematic open boundary has been reached). Since RWPT has been shown to be accurate in this case, the quality of the puff results can be evaluated by comparing how well they match the RWPT results at this time. As expected, neither puff method is able to reproduce the skewness caused by variable diffusivity; they both become less accurate with decreasing $Pe$ (increasing distance from the source or increasing gradient). This trend is clearly shown in Fig. 11, which shows the mean absolute relative error as a function of $Pe_t$ (i.e., $Pe$ at $x = L$) at steady state (i.e., $t = 160$ days in this case). The puff results are least accurate for low $Pe_t$ and approaches a minimum error of approximately 8% as $Pe_t$ becomes large; this is the error caused by the incorrectly specified open boundary condition. The error in the particle results is roughly constant at about 3–5%, which is due to noise. If source resolution were improved, the error in the particle results would decrease but the puff error would remain unchanged.

The inability of the FGPT and BGPT to handle spatial gradients in diffusion is analogous to the shear flow situation. Up to a certain scale of variability, the error in the puff methods may be acceptable. Once the puffs become larger or equal to this scale, the puff methods become less desirable. Thus, even for relatively small spatial variability, the error in FGPT and BGPT results will continue to grow in time as the puffs grow larger. We realize that the effect is more pronounced in this 1-D case where dispersion is only occurring in one direction; there would likely be less of an effect in 2-D and 3-D situations where lateral diffusion could occur. Nonetheless, for the reasons noted above, the puff tracking methods may not be appropriate for long simulations. This does not necessarily mean that RWPT is the only option; some of the more sophisticated puff models cited in Section 4.1 may provide appropriate alternatives.

It should be noted that the ability of RWPT to handle spatially variable diffusivity is compromised when diffusivity or diffusivity gradient is discontinuous; artificial particle accumulation occurs in low diffusivity regions. Past studies have attempted to address this problem with modifications to RWPT (e.g., Thomson et al., 1997; Hoteit et al., 2002; Ross and Sharples, 2004), but this topic is beyond the scope of the present study.

5. Conclusions

The relative strengths and weaknesses of the three Lagrangian techniques are summarized in Table 1. RWPT is the most generally applicable; it can be applied to complex velocity or diffusivity fields, and is most convenient for representing a distributed source such as a plume predicted by a near field model. In addition, no-flux and open boundary conditions can more easily be specified in RWPT. Provided that spatial velocity and diffusivity variations are not too severe and that the boundary conditions are simple, FGPT and BGPT are attractive alternatives because of their vastly superior computational efficiency. If the entire hydrodynamic record is readily available for interpolation during the simulation, BGPT is more efficient than FGPT and offers improved solution accuracy near the source point.

The observations in Table 1 can be useful in designing efficient hybrid approaches, where Lagrangian techniques can be used (alone or combined with Eulerian models) to extend near field calculations. For example, Moeller and Adams (1994) and Zhang and Adams (1999) demonstrate how BGPT and RWPT, respectively, can be used to interface a near field model with a far field Eulerian model; Kim et al. (2002) use RWPT to extend a near field calculation and show favorable agreement.
between their results and laboratory data; and Suh (in press) transitions from RWPT in the near field to an Eulerian—Lagrangian scheme in the far field. Since the FGPT and BGPT methods are most computationally efficient, they should be considered as a first alternative for extending the near field calculation. Once boundaries are encountered or flow complexities become important, these puff methods could transition to an Eulerian model if the plume spreading is sufficient to be resolved by the model grid (as in Adams et al., 1986). If spreading is insufficient, puff methods can transition to RWPT for continued simulation until the Eulerian grid can resolve the plume. If flow conditions make FGPT and BGPT unsuitable even at the edge of the near field, RWPT can be used alone until it is appropriate to transition to the Eulerian model (as in Kim et al., 2002; Zhang and Adams, 1999).

Alternately, the use of a more sophisticated puff technique may be considered (e.g., Hurley, 1994; Sykes and Henn, 1995; de Haan and Rotach, 1998). In all of these cases, Lagrangian methods are used to provide an Eulerian model with a better source boundary condition for long-term simulations (Eulerian methods are better for long-term simulations because the relative computational expense of Lagrangian models increases with simulation length). Alternately, it may be desirable to use RWPT alone even for long-term simulations because of its ability to help visualize dominant transport pathways in complex flows to an extent not possible with puff or Eulerian techniques.

Lastly, it should be noted that many of the implementation and efficiency considerations in Sections 2 and 3 will also be relevant to the use of more sophisticated particle and puff models in hybrid approaches. While many environmental modeling endeavors may require the use of a higher order model (e.g., the absolute diffusion of an ocean plume may require a particle model in which velocities are correlated in time), the implementation and incorporation of such a model into hybrid approaches will still be subject to considerations discussed herein.

### Acknowledgements

Young Do Kim received support from the Postdoctoral Fellowship Program of Korea Science & Engineering Foundation (KOSEF), and Peter H. Israelsson received support from theMartin Family Society of Fellows for Sustainability. Additional support was provided by the Ocean Carbon Sequestration Program, Biological and Environmental Research (BER), US Department of Energy (grant number DE-FG02-01ER63078) and the National Energy Technology Laboratory, US Department of Energy (grant number DE-FG26-98FT40334).

### References


Chapter 3

Random Walk Particle Tracking in Step and Piecewise Linear Diffusivity

3.1 Introduction

The diffusion of a passive tracer in one dimension is commonly modeled by a random walk (or random displacement) simulation of particles undergoing motion of the following form:

\[ \Delta z = \frac{dD}{dz} \Delta t + \xi \sqrt{2D(z_0)\Delta t}. \]  

(3.1)

where \( \Delta z \) is the particle spatial step from \( z = z_0 \) over timestep \( \Delta t \), \( D(z_0) \) is the diffusivity at \( z_0 \), and \( \xi \) is a random number drawn from a Gaussian distribution with zero mean and unit standard deviation (e.g., as derived in [7, 18, 27] for multiple dimensions). Particle motion is thus taken as the sum of a stochastic term (i.e., the second term on the right side of (3.1)) to represent the diffusion process and a deterministic “drift correction” [21] or “pseudovelocity” [33] term (i.e., the first term on the right side of (3.1)) which accounts for the spatial variability of diffusivity and prevents particle accumulation in regions of low diffusivity. (3.1) is selected as the basis of the present work because of its ease of implementation and widespread use; discussion of its derivation and alternate forms are deferred to Section 3.2. In applying (3.1), the trajectories of many particles are simulated.
and particle density is converted to tracer concentration by gridding or with the aid of a density kernel (e.g., [5]).

An implicit constraint on the application of (3.1) is that the diffusivity profile be differentiable throughout the domain (e.g., [2, 22]). Our concern here is with the application of (3.1) in two situations which violate this constraint: (1) when the diffusivity is discontinuous (step diffusivity), and (2) when the diffusivity is continuous but its gradient is discontinuous (piecewise linear diffusivity). Both situations may be handled by smoothing the discontinuities over some distance and selecting a timestep that is small enough so that particles adequately sample the diffusivity profile (e.g., [2, 19, 22, 26]). Such smoothing is physically justified by the notion that step changes in diffusivity or diffusivity gradient are in most cases not real but rather approximations of the true diffusivity profile. However, this approach requires subjectively choosing a smoothing method and distance; a smaller smoothing distance requires a smaller timestep and is thus more computationally expensive, while a larger smoothing distance permits a larger $\Delta t$ but causes a greater distortion of the true diffusivity profile. Thus, it is useful to develop correction methods which allow (3.1) to be applied despite the presence of discontinuities in diffusivity or diffusivity gradient.

The simulation of (3.1) in step diffusivity has been adequately addressed in two previous studies, Thomson et al. [26] and Hoteit et al. [14]. The former considered a two-layer model of the atmosphere with a step diffusivity at the top of the atmospheric boundary layer, while the latter considered the discontinuity that arises in groundwater aquifers between adjacent layers of vastly different soil properties. The correction methods derived in these studies are similar and can be proven accurate for arbitrary jumps in diffusivity (see Section 3.3).

By contrast, the application of (3.1) to discontinuities in diffusivity gradient is much less developed, presumably because they can be handled by smoothing and/or by the use of a very small timestep. The main goal of this study is to analyze the error introduced by these discontinuities and develop a correction methodology which allows larger timesteps to be used. Although this work has general applicability to random walk modeling, it is specifically motivated by the application of (3.1) to simulate vertical tracer diffusion using the diffusivity profiles computed by an ocean general circulation model (OGCM). Here vertical diffusion coefficients are used to represent a number of lumped processes in both the ocean boundary
layer and the interior ocean (e.g., convective mixing and shear instability; see [20] and [8]), and as a result diffusivity may vary over several orders of magnitude across a single OGCM grid cell. In this case, we seek to avoid smoothing the diffusivity profiles because (1) they are highly variable in time and horizontal position such that arbitrary smoothing rules would have to be specified \textit{a priori}, and (2) smoothing can distort the profile such that physically important interfaces such as the pycnocline are not respected. At the same time, we seek to avoid using an impractically small timestep. To these ends an approximate correction method has been developed which allows improved accuracy for larger timesteps, and which affords a certain level of automation in that it can handle any diffusivity profile which may arise.

The present study begins by examining the step diffusivity case and establishes a framework for analyzing the resulting error. The validity of the correction methods of Thomson et al. [26] and Hoteit et al. [14] is confirmed and an alternative correction method is also developed, thereby unifying and extending previous studies. This framework is then adapted to the more complex case of a discontinuity in diffusivity gradient in Section 3.4, for which a correction methodology is proposed and tested. Overall, the present study presents a set of tools with applicability to a wide range of situations in which diffusivity variability is an important consideration.
3.2 Preliminary Considerations

Before delving into the step and piecewise linear diffusivity cases, some elaboration on the random walk method introduced in Section 3.1 is required. In particular, we briefly discuss alternate forms of (3.1), boundary treatment in applying the random walk, and the issue of variable timestepping.

3.2.1 Alternate forms of the random walk equation

Equation (3.1) is but one possible choice for simulating the random walk in variable diffusivity, albeit one that is commonly used. Its derivation is discussed in a wide variety of studies, and is outlined very briefly below; the reader is referred to Dimou and Adams [7], Thompson and Gelhar [27], Gardiner [12], van Kampen [28] or Rodean [21] for more thorough treatments. Following Dimou and Adams [7] with clarification from Rodean [21] and Gardiner [12], the derivation follows from interpreting the non-linear Langevin equation for the particle position \( z(t) \) using the Ito calculus to yield a stochastic differential equation (SDE) of the form

\[
dz = z(t + dt) - z(t) = A(z(t), t) dt + B(z(t), t) \ dW(t)
\]  

(3.2)

where \( A \) and \( B \) are coefficients (or a vector and a tensor in the 2D/3D cases) which characterize deterministic and random forces on the particle, and \( dW(t) \) is the (random) incremental Weiner process, the distribution of which is Gaussian with \( <dW(t) = 0> \) and \( <dW(t)^2 >= dt \) [7, 21]. As demonstrated in Gardiner [12], (3.2) is a local approximation to the diffusion process described by the 1D Fokker-Planck equation

\[
\frac{\partial p(z, t|z_0, t_0)}{\partial t} = -\frac{\partial}{\partial z} [A(z, t)p(z, t|z_0, t_0)] + \frac{1}{2} \frac{\partial^2}{\partial z^2} [B(z, t)^2 p(z, t|z_0, t_0)]
\]  

(3.3)

for the conditional (or transition) probability density \( p(z, t|z_0, t_0) \) for the change from \( z_0 \) at \( t_0 \) to \( z \) at \( t \), which in turn is an approximation to the differential Chapman-Komolgorov (or Master) equation describing a Markov process [21, 12]. The random walk equation (3.1) is the discrete form of (3.2) in which \( dW(t) \) has been replaced by \( \xi \sqrt{\Delta t} \) and the coefficients \( A \)
and $B$ have been replaced by $\frac{dD}{dz}$ and $\sqrt{2D}$, which follows from writing the 1D advection-diffusion equation in a form identical to (3.3):

$$\frac{\partial c}{\partial t} = -\frac{\partial}{\partial z} \left[ \frac{\partial D}{\partial z} c \right] + \frac{\partial^2}{\partial z^2} [Dc].$$

(3.4)

It is in this last step that the formal connection between concentration $c$ and the conditional probability density $p(z,t|z_0,t_0)$ is established, thus allowing (3.1) to simulate (3.4) in the limit of small $\Delta t$ and many particles.

One variation which is common in the literature concerns the distribution of the random variate $\xi$. Although the Gaussian random forcing is consistent with the underlying theory [21, 24], some studies (e.g., [22, 14] and references therein) draw $\xi$ from a uniform distribution, in which case the $\xi$ in (3.1) is replaced by $\frac{\xi}{\sqrt{r}}$ where $r$ is the variance of the distribution. For example, for the uniform distribution $\xi \in [-1,1]$, the Gaussian $\xi$ in (3.1) would be replaced by $\xi \sqrt{3}$ since in this case $r = \frac{1}{3}$. As discussed in Hunter et al. [15], the central limit theorem dictates that this method converges to the Gaussian forcing result as the number of timesteps increases ([15] examines the rate of this convergence). Nonetheless, the Gaussian forcing is employed in the present study because it is more general. If a uniform forcing is to be used, then the correction method in Section 3.4 (for piecewise linear diffusivity) could be rederived, while those in Section 3.3 (for step diffusivity) could be applied directly without rederivation.

Another variation concerns the value of $D$ to use in the stochastic term of (3.1), and thus only applies to the piecewise linear diffusivity case (where $\frac{dD}{dz}$ is finite and non-zero). Using a derivation technique based on an analysis of the moments of the transport equation rather than the Ito SDE / Fokker-Planck analog described above, Hunter et al. [15] and Visser [29] derive a random walk equation in which the $D(z_0)$ in (3.1) is replaced by $D \left( z_0 + \frac{dD}{dz} \frac{\Delta t}{2} \right)$. This formulation is used in other studies as well (e.g., [22, 23]). The present study does not attempt to reconcile the differences between these two formulations, but looking ahead to the analysis of a discontinuity in $\frac{dD}{dz}$ in Section 3.4, we note that (1) the alternate formulation does not reduce the magnitude of the observed error, and (2) the proposed correction methodology could be rederived for use with this formulation.
The issue of which value of $D$ to use during the particle step is related to a larger question sometimes referred to as the Ito-Stratonovich dilemma. The choice of $D = D(z_0)$ in (3.1) follows from the fact that Ito calculus was used to interpret the Langevin equation, which assumes that the particle’s step is a non-anticipating function, i.e., using only the initial value of $D$. An alternate interpretation is the Stratonovich integral, which uses $D$ at the midpoint of the step [21]. Gard [11] draws the analogy with the numerical schemes commonly used to integrate ordinary differential equations, in which the Ito approach amounts to a (first order) Euler solution of an SDE while the Stratonovich approach is tantamount to a more accurate second order Runge Kutta scheme. While this suggests that a Stratonovich-type random walk could perhaps better simulate the diffusion equation in the presence of spatially variable $D$, no such equations were found in the literature. This is perhaps explained by the discussion in Rodean [21], which points out that only the Ito SDE is consistent with a Markov process and yields the diffusion equation as a solution, and the demonstration by Gardiner [12] that the Stratonovich SDE is the solution of a different Fokker-Planck equation [12, 21]. These observations suggest that the analog between tracer concentration and particle density described above cannot exactly be drawn in the Stratonovich case, which prevents a random walk equivalent to the diffusion equation from being derived. Hoteit et al. [14] mentions the possibility of using the Stratonovich interpretation, but dismisses it on the basis that it is difficult to compute. Weitbrecht et al. [31] states that the Stratonovich approach is simulated by using $D (z_0 + \frac{dD}{dz} \Delta t)$, but this is an over-simplification; a true Stratonovich approach would also need to consider the stochastic part of the step. This issue is not dealt with further in the present work; rather, (3.1) is employed throughout and we note that even if a Stratonovich random walk consistent with the diffusion equation were derived, it would be unlikely to handle the sharp discontinuities in diffusivity gradient considered in Section 3.4.

Lastly, three alternative formulations of the random walk specifically designed to handle inhomogeneous diffusivity are worth mention. The first is the method proposed by Hunter et al. [15] and Scott [23] in which a coordinate transform is used to write the diffusion equation in terms of a constant $D$, thus eliminating the “drift” term in (3.1); however, this approach is not pursued further here because it requires that the diffusivity profile
be analytically specified and it is unclear how to extend this to a piecewise linear profile. The second method is proposed by Ermak and Nasstrom [9], in which (3.1) is modified to use a non-Gaussian forcing with faster convergence of particle simulations in domains with linearly varying diffusivity which vanishes at a boundary. The method is a significant improvement to (3.1) for the test cases considered and appears to be an effective manner to handle inhomogeneous diffusivity profiles that are smoothed. However, the approach is not considered further herein as it does not address sharp discontinuities in diffusivity gradient, and its iterative nature would make it difficult to adapt to these situations. The third is the random walk method constructed by Diehl et al. [6], in which particle displacements are calculated by locally approximating any diffusivity profile as piecewise linear. However, it does not provide any advantage over (3.1) in handling discontinuities (in diffusivity or diffusivity gradient), and thus is not considered further.

### 3.2.2 Variable timestepping

Variable timestepping schemes are sometimes used in the application of (3.1) to simulate inhomogeneous diffusion. These typically constrain the timestep by selecting a characteristic length scale for diffusing particles (i.e., $\Delta t = L/D$) (e.g., in [9]) or by considering the second derivative (i.e., $\Delta t << \frac{D}{dz^2}$ [22]) if the diffusivity profile is smooth. However, if a variable timestepping scheme is to be used, it must be implemented with care as it can lead to biased results [2, 32], e.g., spatial gradients in timesteps can yield unphysical particle accumulations. As a result, some authors recommend avoiding variable timestepping altogether [2, 32].

In the present context, choosing a timestep based on $\frac{D}{dz^2}$ is not applicable since $\frac{D}{dz^2}$ is undefined for step and piecewise linear diffusivity profiles. Choosing a timestep based on a characteristic length scale is also not implemented here because (1) for a step diffusivity this approach is not helpful, and (2) for a piecewise linear diffusivity this approach leads to larger timesteps near discontinuities where diffusivity is small, which is also where the error induced by the discontinuity is usually greatest. Therefore, variable timestepping schemes of the types mentioned above are not considered in the present study. Future studies should investigate how variable timestepping could be combined with the approaches developed herein for maximum computational efficiency.
3.2.3 Boundary handling

Unless boundaries are made “nonattainable” [32] by making the timestep infinitessimally small as boundaries are approached, a reflection scheme must be used in the implementation of stochastic models. Perfect reflection of particles at impermeable boundaries is known to be an exact solution for homogeneous diffusivity but only an approximate solution when diffusivity is inhomogeneous near the boundary (e.g., [9, 22]), as demonstrated by the inability of (3.1) to meet the well-mixed condition [2, 24], i.e., to maintain an initially uniform particle density. Ross and Sharples [22] examine this effect when both $D$ and $\frac{dD}{dz}$ are non-zero at the boundary and propose several approximate approaches, including using a well-mixed boundary layer and smoothing the profile to bring $\frac{dD}{dz}$ to zero at the boundary. For the case of $D = 0$ at the boundary, the error is exactly handled by the method developed by Ermak and Nasstrom [9], which features a non-Gaussian forcing derived by combining (3.1) with the analytical solution for this special diffusivity profile. Their approach does not correct the case when both $D$ and $\frac{dD}{dz}$ are non-zero at the boundary; such cases are handled by bringing $\frac{dD}{dz}$ to zero near the boundary in which case their equations automatically reduce to (3.1).

In the course of developing a correction method for a piecewise linear diffusivity in Section 3.4, a reflection method is derived which allows a uniform particle density to be preserved regardless of the values of $D$ and $\frac{dD}{dz}$. For a non-uniform particle density, the results are only approximate. Nonetheless, a byproduct of the present study is a new approach to treating impermeable boundaries during the application of (3.1) (see Section 3.4.3.4).

3.3 Step Diffusivity

The step diffusivity case is treated first as it is simpler to analyze and has an analytical solution. This section begins by illustrating the error induced by a discontinuous diffusivity, and deriving a set of equations which describe particle statistics at the discontinuity. These equations are then used to develop a set of correction methodologies which allow accurate simulation across the discontinuity. The approaches developed herein are compared to the approaches of Thomson et al. [26] and Hoteit et al. [14].

It should be noted that two of the issues raised in Section 3.2 do not apply in this case.
First, the Ito-Stratonovich dilemma is moot; both interpretations yield the same equation if the diffusivity is constant, as does the alternate formulation used by Hunter et al. [15], Visser [29], and Ross and Sharples [22]. Second, treatment of impermeable boundaries by particle reflection is well-known to be a perfect solution [9, 32].

### 3.3.1 Illustration of the problem

A 1D test domain (Figure 3-1) with impermeable boundaries at \( z = 0 \) and \( z = 200 \) cm will be used to illustrate the problem. There is a discontinuity at \( z = z_b = 100 \) cm, which divides the domain into two layers with constant diffusivities of 0.1 and 10 cm\(^2\)/s. The spatial scales and diffusivities are not based on any physical situation; they were chosen arbitrarily to test the particle method.

The test domain was initialized with \( 10^6 \) particles distributed uniformly (\( \rho = 5000 \) particles/cm). In order to be accurate, the RWPT model (3.1) must maintain the uniform particle distribution, i.e., it must maintain the steady state solution. Thomson [24] argues that this condition alone is sufficient to judge the quality of a one-particle stochastic model such as (3.1).

A simulation was run for \( 10^4 \) s with a \( 10^3 \) s output interval, and concentrations were calculated on a 2-cm counting grid. Particles were reflected off of the bottom and top boundary.
impermeable boundaries as needed, consistent with standard practice and the notion of an image source \( \text{(e.g., [17, 26])} \). Figure 3-2 shows the average concentration profile and the ratio of mass in the two halves of the domain \( (R = \frac{M_{z<100\text{cm}}}{M_{z>100\text{cm}}}) \) using timesteps \( \Delta t = 1 \text{ s} \) and \( \Delta t = 0.1 \text{ s} \). Concentrations are time-averaged over the simulation (to get a smoother profile) and normalized by the initial concentration, so that the steady state solution is \( C = 1 \) and \( R = 1 \). The RWPT model (3.1) fails to maintain the uniform distribution, with severe particle accumulations developing in the low diffusivity region \( (R >> 1) \). This result is true for both timesteps, and additional simulations confirm that the error is not resolved by decreasing \( \Delta t \) further.

The problem noted above has been addressed in two previous studies, Thomson et al. [26] and Hoteit et al. [14]. Although [26] and [14] differ in approach and implementation details, both studies solved the problem by reflecting a fraction of particles at the discontinuity
to prevent accumulation in the low diffusivity region. The present study generalizes and expands upon their findings using a different analytical approach.

### 3.3.2 Particle statistics at a boundary

Before designing a correction method to the problem noted above, it is necessary to understand the error that occurs at the discontinuity in the presence of an initially uniform particle distribution. The uniform particle distribution is convenient because it permits analytical expressions for several quantities that are useful in understanding and correcting the error, each of which will be derived in this section:

- the number of particles crossing the boundary in a single timestep;
- the distributions of initial and final positions of crossing particles;
- the distribution of boundary crossing times.

For analytic convenience in the derivations that follow, a boundary coordinate system is employed to describe particle motion relative to a boundary at $z = z_b$:

\[ z' \equiv |z - z_b| \quad (3.5a) \]

\[ \xi' \equiv \begin{cases} 
-\xi & z < z_b \\
\xi & z > z_b 
\end{cases} \quad (3.5b) \]

Applying (3.5) to (3.1) yields:

\[ \Delta z' = \xi' \sqrt{2D(z)\Delta t} \quad (3.6) \]

where the diffusivity gradient term $\frac{\partial D}{\partial z}$ in (3.1) is omitted since it is zero on either side of the discontinuity considered here. In this coordinate system, a negative $\xi'$ moves the particle toward the boundary ($\Delta z' < 0$) and a positive $\xi'$ moves the particle away from the boundary ($\Delta z' > 0$). This coordinate system is not practical for simulation since one value of $z'$ represents two $z$ points, but is employed in the following derivations because it eliminates the need to specify whether a particle crosses $z_b$ from $z < z_b$ or $z > z_b$. 

59
Also, in the probability distribution derivations below, the following notation will be used. For a positive particle quantity $x$ ($0 \leq x \leq x_{\text{max}}$), the probability that $x$ is in the interval $0 \leq x \leq X$ is given by the cumulative distribution function (cdf) $P(X) \equiv P(0 \leq x \leq X)$. The corresponding probability density function (pdf) is then $p(X) \equiv \frac{\partial P(0 < x < X)}{\partial X}$.

### 3.3.2.1 Particle flux across a boundary

We begin by considering the motion of a particle in a region of constant diffusivity $D$ which ends at a boundary at $z = z_b$. In the coordinate system given by (3.5), the particle is initially at $z'_i$ and the boundary is at $z' = 0$ (by definition). The domain is unbounded in the other direction (extending to $z' = \infty$). The particle’s displacement over a single timestep is stochastic since it depends on $\xi'$, which means that the possible particle displacements are normally distributed with a mean of zero and a standard deviation of $\sqrt{2D\Delta t}$. Thus, the probability that it will reach the boundary in $\Delta t$ is the probability that $\Delta z' \leq -z'_i$, i.e., the probability that $\xi'$ will be less than or equal to a critical $\xi'_c$:

$$-z'_i = \xi'_c \sqrt{2D\Delta t} \quad \Rightarrow \quad \xi'_c(z'_i) \equiv \frac{-z'_i}{\sqrt{2D\Delta t}} \quad (3.7)$$

$$P(\Delta z' \leq -z'_i) = P(\xi' \leq \xi'_c(z'_i)) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\xi'_c(z'_i)} e^{-\xi'^2/2} d\xi'. \quad (3.8)$$

It should be noted that in this coordinate system (3.5), $\xi'_c$ is always a negative number, and thus (3.8) states that $\xi'$ must be more negative than $\xi'_c$ for the particle to reach the boundary in $\Delta t$.

The total flux of particles reaching the boundary from the adjacent region can be calculated by considering the probability given by (3.8) for each particle initially in the region (which are uniformly distributed, as discussed previously). We conceptualize the region as consisting of an infinite number of small slices of thickness $dz'_i$ extending from $(z'_i - dz'_i)$ to $z'_i$. In one timestep $\Delta t$, the number of particles reaching the boundary from a slice ($N_{b,\text{slice}}$) is equal to the number of particles in the slice with $\Delta z' \leq -z'_i$, which equals the number of
particles with $\xi' \leq \xi'_c$:

$$N_{b,\text{slice}} = \rho P(\Delta z'_0 \leq -z'_0) \, dz'_0 = \rho P(\xi' \leq \xi'_c(z'_i)) \, dz'_0$$  \hspace{1cm} (3.9)

where $\rho$ is the spatially uniform particle density and $\xi'_c$ is assumed constant over the infinitesimally thin slice. The total number of particles arriving at the boundary from the region ($N_b$) is the sum of the contributions from all slices. The expected value of this stochastic quantity is thus

$$N_b = \rho \int_0^\infty P(\xi' \leq \xi'_c(z'_i)) \, dz'_i.$$  \hspace{1cm} (3.10)

Combining (3.7), (3.8), and (3.10) yields

$$N_b = \frac{\rho}{\sqrt{2\pi}} \int_0^\infty \int_{\xi'_c(z'_i)}^{\infty} e^{-\xi'^2/2} \, d\xi' \, dz'_i.$$ \hspace{1cm} (3.11)

An alternate way of deriving $N_b$ is to recognize that each value of $\xi'$ has the same probability density $p(\xi')$ everywhere in the domain. Consequently, the infinitesimally small interval from $(\xi' - d\xi')$ to $\xi'$ defines a spatial region from which particles will reach the boundary in one timestep, which from (3.8) includes all $z'_i$ values for which $\xi'_c(z'_i) \geq \xi'$ (i.e., $\xi'_c(z'_i)$ is less negative than $\xi'$). This spatial region is bounded by $z' = 0$ (the boundary) and $z' = -\xi'\sqrt{2D\Delta t}$, giving it length

$$L(\xi') = \begin{cases} -\xi'\sqrt{2D\Delta t} & \xi' < 0 \\ 0 & \xi' \geq 0 \end{cases}$$ \hspace{1cm} (3.12)

Since $L(\xi')$ has probability $p(\xi')d\xi'$, the number of particles reaching the boundary from this region ($N_{b,L}$) is

$$N_{b,L}(\xi') = \rho \, L(\xi') \, p(\xi') \, d\xi'.$$ \hspace{1cm} (3.13)

The total number of particles reaching the boundary in one timestep ($N_b$) is thus

$$N_b = \rho \int_{-\infty}^0 L(\xi') \, p(\xi') \, d\xi' = -\frac{\rho\sqrt{D\Delta t}}{\sqrt{\pi}} \int_{-\infty}^0 \xi' e^{-\xi'^2/2} \, d\xi',$$ \hspace{1cm} (3.14)
where the upper bound of integration is zero because \( L(\xi') = 0 \) for \( \xi' > 0 \) (only \( \xi' < 0 \) moves the particle toward the \( z' = 0 \) boundary). Equation (3.14) amounts to reversing the order of integration in (3.11). This alternate conceptualization, while perhaps slightly less intuitive, turns out to be more useful in addressing the more complex case of a discontinuous diffusivity gradient (see Section 3.4). Both conceptualizations ((3.11) and (3.14)) are employed interchangeably throughout the present study.

The solution to both (3.11) and (3.14) is

\[
N_b = \rho \sqrt{\frac{D \Delta t}{\pi}}, \tag{3.15}
\]

indicating that the flux at the boundary is proportional to the length scale \( \sqrt{D \Delta t} \).

To verify (3.15), single timestep simulations were performed with all 45 combinations of \( D \in [0.01, 0.1, 0.2, 0.5, 1, 2, 5, 10, 15] \text{ cm}^2/\text{s} \) and \( \Delta t \in [0.01, 0.1, 1, 2, 3] \text{ s} \). In each case, \( 10^7 \) particles were distributed uniformly over the interval \( 0 \leq z' \leq 6\sqrt{2D \Delta t}, \) which was selected to represent an unbounded domain (it is extremely unlikely that particles beyond this distance will reach the boundary, \( P(\xi' \leq -6) < 10^{-8} \)). The number of particles reaching the boundary at \( z' = 0 \) in one \( \Delta t \) was counted for each simulation. Figure 3-3 shows the comparison of \( N_{b,\text{obs}} \) (observed from simulation) and \( N_{b,\text{pred}} \) (predicted by (3.15)) for all simulations, where \( N_{b,\text{obs}} \) has been normalized by density \( \rho \) on the horizontal axis because \( \rho = \frac{10^7}{6\sqrt{2D \Delta t}} \) varies between simulations. The excellent agreement (0.998 < \( \frac{N_{b,\text{pred}}}{N_{b,\text{obs}}} < 1.003 \); \( \langle \frac{N_{b,\text{pred}}}{N_{b,\text{obs}}} \rangle = 1.0002 \) between the simulated and predicted results confirms the validity of (3.15).

The quantity \( N_b \) in (3.15) is for an unbounded domain; a similar quantity can be derived for a domain bounded at \( z' = z'_{\text{max}} \). Here we calculate the number of particles arriving at the \( z' = 0 \) boundary (\( N_{b,z'_{\text{max}}} \)) which are initially in the region \( 0 \leq z' \leq z'_{\text{max}} \) if the boundary at \( z' = z'_{\text{max}} \) is treated as an absorbing boundary (i.e., particles are not reflected at \( z' = z'_{\text{max}} \)). The absorbing boundary assumption ensures that \( N_{b,z'_{\text{max}}} \) does not include particles reaching \( z' = 0 \) after reflecting off of the boundary at \( z' = z'_{\text{max}} \). In this case, (3.11) becomes

\[
N_{b,z'_{\text{max}}} = \frac{\rho}{\sqrt{2\pi}} \int_0^{z'_{\text{max}}} \int_{-\infty}^{\xi_{\text{c}}(z_0)} e^{-\frac{\xi'^2}{2}} d\xi' d\xi \tag{3.16}
\]
since the constraint $z_i' < z_{\text{max}}'$ is imposed. Alternately, (3.14) can be used if the contributing length $L(\xi')$ in (3.12) is redefined as

$$L(\xi', z_{\text{max}}') = \begin{cases} 
  z_{\text{max}}' & \xi' \leq \xi_c'(z_{\text{max}}') \\
  -\xi' \sqrt{2D\Delta t} & \xi_c'(z_{\text{max}}') < \xi' \leq 0 \\
  0 & \xi' > 0
\end{cases}$$

(3.17)

which includes the constraint $L(\xi') \leq z_{\text{max}}'$. In either case, the solution is

$$N_{b,z_{\text{max}}'} = \rho \sqrt{\frac{D\Delta t}{\pi}} \left[ 1 + \frac{z_{\text{max}}' \sqrt{\pi}}{2\sqrt{D\Delta t}} \text{erfc} \left( \frac{z_{\text{max}}'}{2\sqrt{D\Delta t}} \right) - \exp \left( - \left( \frac{z_{\text{max}}'}{2\sqrt{D\Delta t}} \right)^2 \right) \right].$$

(3.18)

where $\text{erfc}(x) \equiv 1 - \text{erf}(x)$ is the complementary error function and

$$\text{erf}(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} \, dy$$

(3.19)

is the error function. Equation (3.18) makes intuitive sense since $N_{b,z_{\text{max}}'} \to N_b$ as $z_{\text{max}}' \to \infty$, i.e., (3.15) is recovered.

Test simulations were also run to verify (3.18). The same set of single timestep simulations described above was performed again, this time limiting the domain length to $z_{\text{max}}' = \sqrt{2D\Delta t}$.
Figure 3-4: Number of particles predicted and observed to arrive at \( z' = 0 \) in one \( \Delta t \) for simulations of various \( D \) and \( \Delta t \) values in a bounded domain \( (0 \leq z' \leq \sqrt{2D\Delta t}) \) with an initially uniform particle distribution and an absorbing boundary at \( z' = \sqrt{2D\Delta t} \).

but using the same particle density, \( \rho = \frac{10^7}{6\sqrt{2D\Delta t}} \). Figure 3-4 shows the comparison of \( N_{b,obs} \) (simulated result) and \( N_{b,pred} \) (predicted by (3.18)) for all simulations. The excellent agreement \( (0.997 < \frac{N_{b,pred}}{N_{b,obs}} < 1.002; \langle \frac{N_{b,pred}}{N_{b,obs}} \rangle = 1.0002 \) between the simulated and predicted results confirms the validity of (3.18) for a bounded domain.

### 3.3.2.2 Initial position of crossing particles

The number of particles expected to arrive at \( z' = 0 \) which have initial position \( z'_i \) within a distance \( Z_i \) of the boundary \( (0 \leq z'_i \leq Z_i) \) can immediately be written from (3.18) by replacing \( z'_{\text{max}} \) with \( Z_i \):

\[
N_{b,Z_i} = \rho \sqrt{\frac{D\Delta t}{\pi}} \left[ 1 + \frac{Z_i\sqrt{\pi}}{2\sqrt{D\Delta t}} \text{erfc} \left( \frac{Z_i}{2\sqrt{D\Delta t}} \right) - \exp \left( - \left( \frac{Z_i}{2\sqrt{D\Delta t}} \right)^2 \right) \right]. \tag{3.20}
\]

Equation (3.20) applies to both unbounded and bounded domains, the difference being only that \( Z_i \leq z'_{\text{max}} \) is required in the latter case.

In an unbounded domain \( (z'_{\text{max}} \to \infty) \), the probability that a particle originated within
the $0 \leq z' \leq Z_i$ interval is given by the ratio of (3.20) and (3.15):

$$P(0 \leq z'_i \leq Z_i) \equiv \frac{N_b Z_i}{N_b} = 1 + \frac{Z_i \sqrt{\pi}}{2 \sqrt{D \Delta t}} \text{erfc} \left( \frac{Z_i}{2 \sqrt{D \Delta t}} \right) - \exp \left( - \left( \frac{Z_i}{2 \sqrt{D \Delta t}} \right)^2 \right). \tag{3.21}$$

The corresponding pdf $p(Z_i)$ is thus

$$p(Z_i) \equiv \frac{P(0 \leq z'_i \leq Z_i)}{\partial Z_i} = \frac{\sqrt{\pi}}{2 \sqrt{D \Delta t}} \text{erfc} \left( \frac{Z_i}{2 \sqrt{D \Delta t}} \right). \tag{3.22}$$

If the domain is bounded at $z' = z'_{\text{max}}$ then $P(0 \leq z'_i \leq Z_i)$ is given by

$$P(0 \leq z'_i \leq Z_i) \equiv \frac{N_b Z_i}{N_b z'_{\text{max}}} = 1 + \frac{Z_i \sqrt{\pi}}{2 \sqrt{D \Delta t}} \text{erfc} \left( \frac{Z_i}{2 \sqrt{D \Delta t}} \right) - \exp \left( - \left( \frac{Z_i}{2 \sqrt{D \Delta t}} \right)^2 \right) \frac{1 + \frac{z'_{\text{max}} \sqrt{\pi}}{2 \sqrt{D \Delta t}} \text{erfc} \left( \frac{z'_{\text{max}}}{2 \sqrt{D \Delta t}} \right) - \exp \left( - \left( \frac{z'_{\text{max}}}{2 \sqrt{D \Delta t}} \right)^2 \right)}{1 + \frac{z'_{\text{max}} \sqrt{\pi}}{2 \sqrt{D \Delta t}} \text{erfc} \left( \frac{z'_{\text{max}}}{2 \sqrt{D \Delta t}} \right) - \exp \left( - \left( \frac{z'_{\text{max}}}{2 \sqrt{D \Delta t}} \right)^2 \right)}. \tag{3.23}$$

and the corresponding probability density function $p(Z_i)$ is thus

$$p(Z_i) \equiv \frac{\partial P(0 \leq z'_i \leq Z_i)}{\partial Z_i} = \frac{\sqrt{\pi}}{2 \sqrt{D \Delta t}} \text{erfc} \left( \frac{Z_i}{2 \sqrt{D \Delta t}} \right) \frac{1 + \frac{z'_{\text{max}} \sqrt{\pi}}{2 \sqrt{D \Delta t}} \text{erfc} \left( \frac{z'_{\text{max}}}{2 \sqrt{D \Delta t}} \right) - \exp \left( - \left( \frac{z'_{\text{max}}}{2 \sqrt{D \Delta t}} \right)^2 \right)}{1 + \frac{z'_{\text{max}} \sqrt{\pi}}{2 \sqrt{D \Delta t}} \text{erfc} \left( \frac{z'_{\text{max}}}{2 \sqrt{D \Delta t}} \right) - \exp \left( - \left( \frac{z'_{\text{max}}}{2 \sqrt{D \Delta t}} \right)^2 \right)}. \tag{3.24}$$

It should be noted that (3.23) and (3.24) do not include particles that reached $z' = 0$ after reflecting off of the boundary at $z'_{\text{max}}$, since the derivation of $N_{b, z'_{\text{max}}}$ (3.18) assumed this to be an absorbing boundary. The contribution of reflected particles is discussed in Section 3.3.4 where these equations are applied.

Results of the test simulations previously used to verify $N_b$ were used to verify (3.21) and (3.22), and the simulations previously used to verify $N_{b, z'_{\text{max}}}$ were used to verify (3.23)
Figure 3-5: Predicted (line) and observed (points) cumulative distribution function \( P(0 \leq z'_i \leq Z_i) \), top panel) and probability density function \( p(Z_i) \), bottom panel) where \( z'_i \) is the initial position of particles reaching \( z' = 0 \) in one \( \Delta t \) from an unbounded domain with \( D = 2 \) cm\(^2\)/s, \( \Delta t = 1 \) s, and an initially uniform particle distribution.

Figure 3-6: Predicted (line) and observed (points) cumulative distribution function \( P(0 \leq z'_i \leq Z_i) \), top panel) and probability density function \( p(Z_i) \), bottom panel) where \( z'_i \) is the initial position of particles reaching \( z' = 0 \) in one \( \Delta t \) from a bounded domain \( 0 \leq z' \leq \sqrt{2D\Delta t} \) with \( D = 2 \) cm\(^2\)/s, \( \Delta t = 1 \) s, and an initially uniform particle distribution.

and (3.24). Figures 3-5 and 3-6 compare the observed and predicted distribution of initial positions for the simulations in which \( D = 2 \) cm\(^2\)/s and \( \Delta t = 1 \) s for the unbounded and bounded cases, respectively. The agreement is excellent in both cases, confirming the validity of (3.21), (3.22), (3.23), and (3.24).
3.3.2.3 Final position of crossing particles

Derived below is the distribution of final particle positions \( z'_f \) if particles are allowed to cross the \( z' = 0 \) boundary. This can be deduced from the number of particles expected to reach a certain distance \( Z_f \) beyond the \( z' = 0 \) boundary, \( N_b(Z_f) \).

Before proceeding, it should be noted that in the boundary coordinate system (3.5), a particle at this distance has final position \( z'_f = Z_f \), even though it is actually on the other side of the boundary. For example, using untransformed coordinates, if the boundary is at \( z_b = 0 \) and a particle with initial position \( z_i = 1 \) is displaced \( \Delta z = -3 \), then it travels \( Z_f = 2 \) beyond the boundary to its final position \( z_f = z_i + \Delta z = -Z_f = -2 \); in boundary coordinates, \( z'_i = 1 \) and \( z'_f = 2 \) for this particle. As noted previously, this is impractical for simulation but useful in deriving particle statistic equations.

The equation for \( N_b(Z_f) \) is essentially derived by shifting the boundary by a distance \( Z_f \). To accomplish this, the critical random number \( \xi'_c \) in (3.7) is redefined to be

\[
-z'_i - Z_f = \xi'_c \sqrt{2D\Delta t} \quad \Rightarrow \quad \xi'_c(z'_i, Z_f) \equiv \frac{-z'_i - Z_f}{\sqrt{2D\Delta t}}. \tag{3.25}
\]

For an unbounded domain (\( z'_{\text{max}} \to \infty \)), the number of particles expected to reach a distance \( Z_f \) past the boundary is simply (3.11) using the \( \xi'_c(z'_i, Z_f) \) in (3.25). Alternately, (3.14) can be used if the contributing length \( L(\xi') \) in (3.12) is redefined as

\[
L(\xi', Z_f) = \begin{cases} 
-\xi'\sqrt{2D\Delta t} - Z_f & \xi' \leq \xi'_c(0, Z_f) \\
0 & \xi' > \xi'_c(0, Z_f)
\end{cases} \tag{3.26}
\]

since not even particles at \( z'_i = 0 \) can reach \( Z_f \) beyond the boundary when \( \xi' > \xi'_c(0, Z_f) \). In either case, the solution is

\[
N_b(Z_f) = \frac{\rho \sqrt{D\Delta t}}{\sqrt{\pi}} \exp \left( -\left( \frac{Z_f}{2\sqrt{D\Delta t}} \right)^2 \right) - \frac{\rho Z_f}{2} \operatorname{erfc} \left( \frac{Z_f}{2\sqrt{D\Delta t}} \right). \tag{3.27}
\]

The probability that a particle will have final position \( z'_f \) in the interval \( 0 \leq z' \leq Z_f \) beyond
the boundary is given by

\[ P(0 \leq z_f' \leq Z_f) \equiv 1 - \frac{N_b(Z_f)}{N_b(Z_f = 0)} \]

\[ = 1 + \frac{Z_f\sqrt{\pi}}{2\sqrt{D\Delta t}} \text{erfc} \left( \frac{Z_f}{2\sqrt{D\Delta t}} \right) - \exp \left( -\left( \frac{Z_f}{2\sqrt{D\Delta t}} \right)^2 \right), \quad (3.28) \]

which has the pdf \( p(Z_f) \)

\[ p(Z_f) \equiv \frac{\partial P(0 \leq z_f' \leq Z_f)}{\partial Z_f} \]

\[ = \frac{\sqrt{\pi}}{2\sqrt{D\Delta t}} \text{erfc} \left( \frac{Z_f}{2\sqrt{D\Delta t}} \right). \quad (3.29) \]

Equations (3.28) and (3.29) are identical to (3.21) and (3.22) when \( Z_f \) is substituted for \( Z_i \). This results from the symmetry of the random walk equations when \( \frac{dD}{dz} = 0 \) in that the magnitude of \( \Delta z \) is not dependent on the sign of \( \xi \).

In the bounded case, the distribution of \( z_f' \) is not equivalent to distribution of \( z_i' \). Here \( N_b(Z_f) \) is given by (3.16) using the \( \xi_c(z_i', Z_f) \) in (3.25). Alternately, (3.14) can be used if the contributing length \( L(\xi', z_{\text{max}}') \) in (3.17) is redefined as

\[
L(\xi', z'_{\text{max}}, Z_f) = \begin{cases} 
z'_{\text{max}} & \xi' \leq \xi_c(z_{\text{max}}', Z_f) \\
-\xi'\sqrt{2D\Delta t} - Z_f & \xi' \leq \xi_c(0, Z_f) \\
0 & \xi' > \xi_c(0, Z_f)
\end{cases}
\]

\[ (3.30) \]

In either case, the solution is

\[
N_b(Z_f) = \frac{\rho\sqrt{D\Delta t}}{\sqrt{\pi}} \left[ \exp \left( -\left( \frac{Z_f}{2\sqrt{D\Delta t}} \right)^2 \right) - \exp \left( -\left( \frac{z_{\text{max}} + Z_f}{2\sqrt{D\Delta t}} \right)^2 \right) \right] + \\
\frac{\rho Z_f}{2} \left[ \text{erf} \left( \frac{Z_f}{2\sqrt{D\Delta t}} \right) - \text{erf} \left( \frac{z_{\text{max}} + Z_f}{2\sqrt{D\Delta t}} \right) \right] + \\
\frac{\rho z_{\text{max}}}{2} \text{erfc} \left( \frac{z_{\text{max}} + Z_f}{2\sqrt{D\Delta t}} \right). \quad (3.31) \]
Thus,

\[ P(0 \leq z_f' \leq Z_f) \equiv 1 - \frac{N_b(Z_f)}{N_b(Z_f = 0)}, \] (3.32)

which for brevity will not be expanded, and which has the pdf \( p(Z_f) \)

\[
p(Z_f) \equiv \frac{\partial P(0 \leq z_f' \leq Z_f)}{\partial Z_f} = \frac{\sqrt{\pi}}{2\sqrt{D\Delta t}} \left[ \text{erf} \left( \frac{z_{\text{max}}+Z_f}{2\sqrt{D\Delta t}} \right) - \text{erf} \left( \frac{Z_f}{2\sqrt{D\Delta t}} \right) \right] \]

\[
1 + \frac{z_{\text{max}}\sqrt{\pi}}{2\sqrt{D\Delta t}} \text{erfc} \left( \frac{z_{\text{max}}}{2\sqrt{D\Delta t}} \right) - \exp \left( - \left( \frac{z_{\text{max}}}{2\sqrt{D\Delta t}} \right)^2 \right) \] (3.33)

Equations (3.28), (3.29), (3.32), and (3.33) were verified in the same way as the corresponding \( z_i' \) equations in Section 3.3.2.2. The results for (3.28) and (3.29) (not shown) are identical to those in Figure 3-5 if \( Z_f \) is substituted for \( Z_i \) on the horizontal axis. For the bounded case, (3.32) and (3.33) are verified in Figure 3-7. The excellent agreement in all cases confirms the validity of these equations.
3.3.2.4 Particle crossing times

For reasons that will become clear in Section 3.3.4, it is also useful to know the distribution of particle crossing times. Using the random walk equation (3.1) to analyze particle motion for \( t < \Delta t \) is artificial; (3.1) only predicts the final position of particles and not their path (and for times less than the Lagrangian integral timescale, a particle’s turbulent velocities are correlated so that the random walk and the notion of a diffusivity do not apply [17]). Nonetheless, for the present purpose we define the fractional timestep

\[
fb = \frac{tb}{\Delta t}
\]  

(3.34)

required to reach the boundary for a particle that crosses the boundary in timestep \( \Delta t \), where \( tb \) can be defined in one of two ways:

1. by assuming that the particle travels at a constant velocity; or

2. by calculating the equivalent timestep required to reach the boundary.

The distribution of \( fb \) using each of these definitions is derived below for a uniform particle distribution in an infinite domain. The bounded domain case is not derived below; although its derivation is straightforward, the results are not needed in developing correction methodologies.

In the constant particle velocity approach to calculating \( fb \), the time to reach the boundary \( (tb) \) is taken as

\[
tb = \frac{z^{'i}}{v_i}
\]  

(3.35)

where the particle initially at \( z_i \) is assumed to travel at a constant velocity \( (v_i) \) based on the diffusivity at \( z_i \) \( (D_i) \):

\[
v_i = \frac{\Delta z'}{\Delta t} = \frac{\xi' \sqrt{2D_i}}{\sqrt{\Delta t}}
\]  

(3.36)
The distribution of \( f_b \) is calculated from \( N_b(F_b) \), the number of particles with \( 0 \leq f_b \leq F_b \), which can be calculated using (3.14) if \( L(\xi') \) is replaced by

\[
L(\xi', F_b) = \begin{cases} 
-F_b \xi' \sqrt{2D\Delta t} & \xi' < 0 \\
0 & \xi' \geq 0 
\end{cases} 
\] (3.37)

The solution to (3.14) is then

\[
N_b(F_b) = \rho F_b \sqrt{\frac{D \Delta t}{\pi}}. 
\] (3.38)

Thus, the cumulative distribution function \( P(0 \leq f_b \leq F_b) \) is

\[
P(0 \leq f_b \leq F_b) \equiv \frac{N_b(F_b)}{N_b(F_b = 1)} = F_b, 
\] (3.39)

and the pdf is

\[
p(F_b) \equiv \frac{\partial P(0 \leq f_b \leq F_b)}{\partial F_b} = 1. 
\] (3.40)

In this case, all values of \( f_b \) have the same probability of occurring.

In the equivalent timestep approach to calculating \( f_b \), the time to reach the boundary \( (t_b) \) is

\[
t_b = \frac{1}{2D_i} \left( \frac{z_i'}{\xi'} \right)^2, 
\] (3.41)

i.e., \( t_b \) is here the timestep that would bring a particle at \( z_i' \) to the boundary with random number \( \xi' \). In this case, \( N_b(F_b) \) can be calculated using (3.14) if \( L(\xi') \) is replaced by

\[
L(\xi', F_b) = \begin{cases} 
-\xi' \sqrt{2DF_b \Delta t} & \xi' < 0 \\
0 & \xi' \geq 0 
\end{cases} 
\] (3.42)

The solution to (3.14) is then

\[
N_b(F_b) = \rho \sqrt{\frac{DF_b \Delta t}{\pi}}. 
\] (3.43)
Thus, the probability $P(0 \leq f_b \leq F_b)$ is

$$P(0 \leq f_b \leq F_b) = \frac{N_b(F_b)}{N_b(F_b = 1)} = \sqrt{F_b}, \quad (3.44)$$

and the pdf is

$$p(F_b) = \frac{\partial P(0 \leq f_b \leq F_b)}{\partial F_b} = \frac{1}{2\sqrt{F_b}}. \quad (3.45)$$

The equations derived above were verified in the same way as the infinite domain equations in preceding sections. The results for (3.39) and (3.40) are shown in Figure 3-8, and the results for (3.44) and (3.45) are shown in Figure 3-9. There is excellent agreement between predicted and observed values in all cases.

### 3.3.3 Analysis of discontinuity error

The error observed in Section 3.3.1 can be analyzed using the equations for particle statistics developed above. In particular, two sources of error are identified: (1) a flux error at the discontinuity and (2) a particle placement error when crossing the discontinuity.
Figure 3-9: Predicted (line) and observed (points) cumulative distribution function \( P(0 \leq f_b \leq F_b) \), top panel) and probability density function \( p(F_b) \), bottom panel) where \( f_b \) is the fractional timestep for particles to reach the \( z' = 0 \) boundary based on an equivalent timestep calculation. Particles are initially uniformly distributed in an infinite domain with \( D = 2 \text{ cm}^2/\text{s} \) and \( \Delta t = 1 \text{ s} \).

### 3.3.3.1 Particle flux error

In order for the random walk model to maintain a uniform concentration distribution (i.e., meet the well-mixed condition), the particle flux from either side of the boundary must be equal when the particle density is uniform, i.e.,

\[
\dot{N}_{b,1} = \dot{N}_{b,2} \rightarrow \frac{N_{b,1}}{\Delta t_1} = \frac{N_{b,2}}{\Delta t_2}
\]

where \( \dot{N}_b \) is the particle flux at the boundary and the numeric subscripts indicate regions 1 and 2 which are on opposite sides of the discontinuity. Substituting (3.15) into (3.46) yields

\[
\frac{D_1}{\Delta t_1} = \frac{D_2}{\Delta t_2}.
\]

For the test simulation in Section 3.3.1, \( \dot{N}_b \) with \( \Delta t_1 = \Delta t_2 = 0.5 \text{ s} \) for the regions above and below the boundaries are 1262 and 126.2 particles/s, respectively. Thus, the flux from the high diffusivity \( (z > 100 \text{ cm}) \) region is an order of magnitude larger than the flux from the low diffusivity region \( (z < 100 \text{ cm}) \). This ratio is constant for any choice of \( \Delta t \) (provided \( \Delta t_1 = \Delta t_2 \)), and thus the particle flux error is not affected by decreasing \( \Delta t \) (consistent with
3.3.3.2 Particle placement error

The random walk model (3.1) includes a pseudovelocity or drift term $\frac{\partial D}{\partial z} \Delta t$ which accounts for the spatial variability of diffusivity. For a diffusing particle cloud in linearly varying diffusivity, this term adjusts the center of mass of particles to prevent artificial accumulation of mass in low diffusivity regions. For a step diffusivity, this term is zero everywhere except for the discontinuity where it is infinite, meaning that the correction term is absent from simulation. Nonetheless, particles crossing the boundary experience variable diffusivity, and the result is that particles crossing from high to low diffusivity travel too far into the low diffusivity region, and too short in the opposite case. This is apparent by comparing the distributions of final positions $z'_f$ for particles crossing from each region in one $\Delta t$ if the initial particle distribution is uniform. For the test domain in Section 3.3.1, these distributions (Figure 3-10) are given by (3.29) since the impermeable boundaries are sufficiently far away from the discontinuity that the domain can be considered unbounded. The magnitude of this error is, however, affected by the timestep; as $\Delta t \to 0$, the spatial scale of particle displacements also goes to zero as does the spatial scale of the erroneous particle distribution.

3.3.4 Correction techniques

The following section develops numerous correction techniques to the two sources of error noted above. Two particle flux error corrections are proposed first, and then various particle placement corrections are tested together with the particle flux corrections as appropriate.

3.3.4.1 Particle flux corrections

The particle flux imbalance that occurs across a discontinuity is embodied in (3.46) and (3.47), which suggest two approaches to solving the flux imbalance:

1. Reflect excess particles from the high diffusivity region.

2. Use different timesteps in each region so that (3.47) balances.
3.3.4.1.1 Flux correction by particle reflection

In this approach, a probability of reflection is defined for particles coming from the high diffusivity region:

\[(1 - P_r) \hat{N}_{\text{high}} = \hat{N}_{\text{low}} \quad \rightarrow \quad P_r = 1 - \frac{\hat{N}_{\text{low}}}{\hat{N}_{\text{high}}}\]  

(3.48)

where the subscripts refer to regions of higher and lower diffusivity respectively. Using \( \hat{N} = \frac{N_b}{\Delta t} \) with \( N_b \) given by (3.15) yields

\[P_r = 1 - \frac{\sqrt{D_{\text{low}}}}{\sqrt{D_{\text{high}}}}.\]  

(3.49)

Particles coming from the low diffusivity region are not reflected, i.e., \( P_r = 0 \) for particles in this region. This is accomplished numerically by drawing a random number \( r \) from a [0, 1] uniform distribution and comparing it to \( P_r \). If \( r \leq P_r \), the particle is reflected at the discontinuity. If \( r > P_r \), the particle is transmitted. The success of this approach is illustrated in Section 3.3.4.2, where it is paired with various particle placement techniques.

The correction described above is identical to the one proposed by Thomson et al. [26],
although the derivation and justifications are markedly different. It is similar but not identical to the one proposed by Hoteit et al. [14], which instead defines probabilities for a particle arriving at the boundary to go into the high \((P_{\text{high}})\) and low \((P_{\text{low}})\) diffusivity regions:

\[
P_{\text{high}} = \frac{\sqrt{D_{\text{high}}}}{\sqrt{D_{\text{high}}} + \sqrt{D_{\text{low}}}} \quad (3.50a)
\]

\[
P_{\text{low}} = 1 - P_{\text{high}} = \frac{\sqrt{D_{\text{low}}}}{\sqrt{D_{\text{high}}} + \sqrt{D_{\text{low}}}}. \quad (3.50b)
\]

The Hotiet et al. approach is equivalent to letting \(P_{\text{low}}\) and \(P_{\text{high}}\) be the reflection probabilities for particles coming from the low and high diffusivity regions, respectively. Although this does succeed in balancing the particle flux at the discontinuity, it differs from our approach (and that of [26]) in that particles coming from the low diffusivity region may be reflected, and consequently more particles must be reflected from the high diffusivity region. However, results in Section 3.3.5 suggest that the impact of this difference is small.

### 3.3.4.1.2 Flux correction by timestep balancing

Equation (3.47) indicates that the flux error can be corrected by using a smaller timestep in the region of lower diffusivity:

\[
\Delta t_{\text{low}} = \frac{D_{\text{low}}}{D_{\text{high}}} \Delta t_{\text{high}}. \quad (3.51)
\]

This result is counter-intuitive in that conventional thinking on variable timesteps suggests decreasing the timestep in the high diffusivity region. The practicality of this “timestep balancing” approach is questionable given the increase in computation time. Nonetheless, the success of the technique is demonstrated in Section 3.3.4.2. This approach is somewhat analogous to introducing a stretched coordinate system, as is sometimes done in groundwater transport calculations in layered aquifers with varying hydraulic conductivity (e.g., [10]).

### 3.3.4.2 Particle placement corrections

Two strategies for determining the final location of particles reaching the discontinuity are developed:
• Midstep adjustment

• Step shortening

These corrections alone are insufficient; they must be paired with one of the particle flux corrections in the preceding section.

In developing these approaches, it is helpful to first consider the treatment of impermeable boundaries in random walk modeling. The standard approach taken at these boundaries is to “bounce” particles, e.g., a particle initially at \( z_i \) given a displacement \( \Delta z = -1.5z_i \) toward an impermeable boundary at \( z = 0 \) would have final position \( z_f = 0.5z_i \) (all in untransformed coordinates, not boundary coordinates). The reader should note that for the remainder of this study, we refer to this as bouncing and not reflection as the latter term is used to refer to the particle flux correction method from the previous section; these terms are generally interchangeable in other studies. In constant diffusivity, particle bouncing is an exact solution, consistent with the notion of an image source [32]. In non-constant diffusivity, particle bouncing introduces an error which is discussed in several studies [22, 21, 9], and which is revisited in Section 3.4.3.4.

The accuracy of particle bouncing at impermeable boundaries in a region of constant diffusivity can be understood analytically in terms of the probability distributions derived in Section 3.3.2. In the example that follows, untransformed coordinates are used instead of the boundary coordinate system previously employed. Consider a domain with an impermeable boundary at \( z = 0 \) extending to \( z = \infty \) with a uniform particle distribution. In order for the uniform distribution to be maintained, the particle flux must be equal everywhere. In one timestep, the initial positions of particles reaching the \( z = 0 \) boundary are given by (3.22). Bouncing these particles off of the boundary is equivalent to considering them as originating from image sources on the other side of the boundary with the same initially uniform distribution. As such the final positions of these bounced/image particles is determined by the distribution in (3.29), which as previously noted is identical to the initial position distribution in (3.22). In a sense, particles reaching the boundary are replaced by image particles. Put another way, consider a plane at \( z = z_{plane} \) which is near the \( z = 0 \) impermeable boundary. The number of particles reaching the plane from \( z > z_{plane} \) is given
by (3.15), i.e., the flux from an infinite domain. This is balanced by the sum of three groups of particles:

1. particles initially in the region $0 \leq z_i \leq z_{\text{plane}}$ which reach the plane by positive displacement ($\Delta z > 0$); this quantity is given by (3.18) using $z'_{\text{max}} = z_{\text{plane}}$.

2. particles initially in the region $0 \leq z_i \leq z_{\text{plane}}$ which reach the plane after negative displacement ($\Delta z < 0$) and bouncing off the impermeable boundary at $z = 0$; this quantity is given by (3.31) using $Z_f = z'_{\text{max}} = z_{\text{plane}}$.

3. particles initially in the region $z_i > z_{\text{plane}}$ which crossed $z_{\text{plane}}$, bounced off of the impermeable boundary at $z = 0$, and crossed $z_{\text{plane}}$ again; this quantity is given by (3.27) using $Z_f = 2z_{\text{plane}}$.

It can easily be shown that the sum of these three quantities is exactly equal to the number of particles crossing $z_{\text{plane}}$ from $z > z_{\text{plane}}$ (given by (3.15)). Thus, the region next to the impermeable boundary does not “feel” the boundary in that it receives the same particle flux that it would have if there were no boundary. As a result, the uniform particle distribution is maintained (in a statistical sense).

The goal of the correction methodologies developed below will be to duplicate this statistical harmony in the vicinity of a diffusivity discontinuity. For a uniform particle distribution, the flux of particles should be the same everywhere and equal to the flux in an unbounded domain, i.e., no region of the domain should “feel” the discontinuity just as impermeable boundaries are not “felt” when particle bouncing is employed.

3.3.4.2.1 Midstep adjustment The midstep adjustment technique is the generalization of the particle bouncing described above for the impermeable boundary. It is only compatible with the particle reflection approach to flux correction, and is thus described in this context below (i.e., particles are either transmitted or reflected at the discontinuity and the same timestep is used on either side of the discontinuity).

In this approach, the remaining displacement of the particle is adjusted as it crosses the discontinuity. The time to reach the boundary $t_b$ is calculated using the constant velocity assumption (3.35), which can be expressed as a fractional timestep $f_b$, defined in (3.34). If
the particle is reflected, then its incoming velocity \(v_i\), given by (3.36) is simply reversed for the remainder of the timestep, \(v_{\text{new}} = -v_i\). If the particle is transmitted, then its new velocity is based on the diffusivity in the new region \(D_{\text{new}}\):

\[
v_{\text{new}} = \frac{\xi \sqrt{2D_{\text{new}}}}{\sqrt{\Delta t}} = v_i \sqrt{\frac{D_{\text{new}}}{D_i}}.
\]

(3.52)

The final position of the particle is

\[
z_f = z_b + v_{\text{new}}(1 - f_b)\Delta t.
\]

(3.53)

In the case of reflection, this amounts to bouncing the particle in the same manner as for impermeable boundaries. In the case of transmission, this amounts to adjusting the distance traveled beyond the boundary \(Z_f\) so that the distribution of final positions \(z'_f\) in (3.29) is consistent with the new region rather than the old region (e.g., going from one distribution to the other in Figure 3-10). The result is that neither region “feels” the discontinuity because the particle flux is everywhere the same as it would be if there were no discontinuity, i.e., the particle flux in an unbounded domain.

Simulation results for the combination of particle reflection with midstep adjustment is shown in Figure 3-11 (using the test domain and plotting metrics described in Section 3.3.1). The uniform concentration distribution is in this case maintained, i.e., the error in Figure 3-2 has been corrected.

This combination of approaches is identical to the solution proposed by Thomson et al. [26]. Midstep adjustment is also the approach taken by Hoteit et al. [14], although they combined it with a slightly different flux balancing approach (see Section 3.3.4.1). Nonetheless, the Hoteit et al. [14] approach also maintains the uniform distribution in this manner.

Lastly, it is noted that this approach is incompatible with the timestep balancing flux correction because it results in no velocity adjustment for crossing particles:

\[
v_{\text{new}} = \frac{\xi \sqrt{2D_{\text{new}}}}{\sqrt{\Delta t_{\text{new}}}} = \frac{\xi \sqrt{2D_i}}{\sqrt{\Delta t_i}} = v_i.
\]

(3.54)
Figure 3-11: Predicted time-averaged concentration \( C \) (top panels) and ratio of mass below to mass above \( z = 100 \text{ cm} \) \( R \) (bottom panels) for the test domain with an initially uniform particle density and \( \Delta t = 1 \text{ s} \) using particle reflection with midstep adjustment. The expected values are \( C=1 \) and \( R=1 \).

This does not mean that there is no particle placement error when timestep balancing is used; it is just masked by the fact that more timesteps are required in the lower diffusivity region because of the smaller \( \Delta t \).

3.3.4.2.2 Step shortening In this approach the master timestep \( \Delta t_m \) is split into two separate, independent steps for particles reaching the discontinuity. The first partial step is the timestep that would have been required for the particle to exactly reach the boundary, \( f_b \Delta t_m \), where \( f_b \) is calculated using the \( t_b \) defined in (3.41). The second partial step is the remaining timestep, \( (1 - f_b) \Delta t_m \), and the final particle displacement is calculated using (3.1) with a new random number \( \xi_{\text{new}} \) and the diffusivity in the region the particle enters \( (D_{\text{new}}) \):

\[
z_f = z_b + \xi_{\text{new}} \sqrt{2D_{\text{new}}(1 - f_b)\Delta t_m}.
\tag{3.55}
\]

When used together with the particle reflection flux correction, \( \xi_{\text{new}} \) is drawn until it causes the particle to move in the direction consistent with its determined fate (either reflected or transmitted). When used together with timestep balancing, \( \xi_{\text{new}} \) is drawn until it causes the
particle to be transmitted, and different master timestep $\Delta t_m$ values are used for either side of the discontinuity. Also, when used with timestep balancing, particles that have encountered the discontinuity will be at various time levels since their shortened steps will have combined different master timesteps, so the last timestep before output will vary between particles because it must be the time remaining until the output time (when all particles must be at the same time level). In this last case, the timesteps used should be much smaller than the output interval since particle fluxes will not be in balance for the time immediately preceding output.

The step shortening method also works for impermeable boundaries, with the only difference being that $\xi_{\text{new}}'$ is drawn until it causes the particle to be reflected. Indeed, if the discontinuities are being treated by step shortening, then for consistency impermeable boundaries should also be handled by this approach.

The step shortening approach differs from midstep adjustment in three main ways: (1) it can be used together with either flux correction technique in Section 3.3.4.1; (2) it uses a different time calculation to reach the boundary ((3.41) instead of (3.35)); and (3) a new value of $\xi$ is drawn for the second step.

The success of the step shortening approach is shown in Figure 3-12 using either particle flux correction technique. In both cases, the error in Figure 3-2 has been prevented. Figure 3-13 provides further insight into why the step shortening approach is accurate. Here the observed cdf of final positions $z_f'$ on the low diffusivity side ($D = 0.1 \text{ cm}^2/\text{s}$) of the test domain boundary is shown after a single timestep of a step shortening/particle reflection simulation. It is plotted together with the distribution of $z_f'$ given by (3.28), i.e., for an infinite domain with no discontinuities. The distributions are identical, so step shortening maintains the uniform particle distribution because it places particles such that the discontinuity is not “felt” by neighboring regions. The same accuracy would be observed for $z_f'$ in the high diffusivity region, and also if step shortening were instead combined with the timestep balancing flux correction (in the latter case, the $\Delta t_m$ value would differ on either side of the discontinuity).

The accuracy of step shortening can also be shown analytically. The number of particles
Figure 3-12: Predicted time-averaged concentration (top panels) and ratio of mass below to mass above $z = 100$ cm (bottom panels) for the test domain with an initially uniform particle density and $\Delta t = 1$ s using two different correction methods: particle reflection with step shortening (left panels) and timestep balancing with step shortening (right panels). The expected values are $C=1$ and $R=1$.

reaching a distance $Z_f$ beyond the boundary after step shortening is

$$
N_{b, Z_f} = N_b \int_0^1 p(f_b) P(0 \leq z' \leq Z_f) df_b
$$

$$
= \frac{\sqrt{2}N_b}{\sqrt{\pi}} \int_0^1 \frac{1}{2\sqrt{f_b}} \int_{-\infty}^{-Z_f/\sqrt{2D_{\text{new}}(1-f_b)\Delta t_m}} e^{-\xi^2/2} d\xi d\xi d\xi
$$

$$
= N_b \int_0^1 \frac{1}{2\sqrt{f_b}} \text{erfc} \left( \frac{Z_f}{\sqrt{2D_{\text{new}}(1-f_b)\Delta t_m}} \right) df_b
$$

(3.56)
Figure 3-13: Observed (points) cumulative distribution function \( P(0 \leq z'_f \leq Z_f) \) for the final position \( z'_f \) of particles crossing from a \( D = 10 \text{ cm}^2/\text{s} \) region (initialized with a uniform particle) into a \( D = 0.1 \text{ cm}^2/\text{s} \) region using step shortening coupled with the particle reflection flux correction (\( \Delta t = 1 \text{ s} \) everywhere). The observed distribution is identical to the distribution expected (line) beyond any plane in a continuous infinite domain with \( D = 0.1 \text{ cm}^2/\text{s} \) and \( \Delta t = 1 \) with an initially uniform particle density.

where (3.44) was used. The probability \( P(0 \leq z'_f \leq Z_f) \) is therefore

\[
P(0 \leq z'_f \leq Z_f) \equiv 1 - \frac{N_b Z_f}{N_b} \]

\[
= 1 - \int_0^1 \frac{1}{2\sqrt{f_b}} \text{erfc} \left( \frac{Z_f}{\sqrt{2D_{new}(1-f_b)\Delta t_m}} \right) df_b. \tag{3.57}
\]

Although (3.57) has no closed-form solution, numerical integration proves that it is equivalent to (3.28).

Lastly, it should be noted that two variations on the step shortening method have also been shown to be successful at maintaining a uniform particle distribution: (1) a full \( \Delta t_m \) can instead be used when particles leave the boundary; and (2) when combined with timestep balancing, the first value of \( \xi'_{\text{new}} \) drawn can be used when leaving the boundary, thus causing particle reflection or transmission with equal probability. These options are set aside, however, in favor of the main approach described above since the latter can be shown analytically to reproduce the steady state particle distribution.
Table 3.1: Timestep sensitivity for step diffusivity correction methods for the well-mixed condition test (PR = particle reflection; TB = timestep balancing; MA = midstep adjustment; SS = step shortening. Reported $R$ values are the mass ratios at the end of the $10^4$ s simulation.)

<table>
<thead>
<tr>
<th>Correction method</th>
<th>$\Delta t$ (s)</th>
<th>Maintains well mixed condition?</th>
<th>Relative computational expense</th>
</tr>
</thead>
<tbody>
<tr>
<td>PR with MA</td>
<td>1</td>
<td>Yes ($R=1.002$)</td>
<td>1</td>
</tr>
<tr>
<td>PR with MA</td>
<td>1,000</td>
<td>Yes ($R=1.000$)</td>
<td>0.00370</td>
</tr>
<tr>
<td>PR with MA</td>
<td>10,000</td>
<td>Yes ($R=1.000$)</td>
<td>0.00258</td>
</tr>
<tr>
<td>PR with SS</td>
<td>1</td>
<td>Yes ($R=1.000$)</td>
<td>1.04</td>
</tr>
<tr>
<td>PR with SS</td>
<td>1,000</td>
<td>Yes ($R=0.999$)</td>
<td>0.0470</td>
</tr>
<tr>
<td>PR with SS</td>
<td>10,000</td>
<td>Yes ($R=0.999$)</td>
<td>0.0291</td>
</tr>
<tr>
<td>TB with SS</td>
<td>1 &amp; 0.01</td>
<td>Yes ($R=1.002$)</td>
<td>48.5</td>
</tr>
<tr>
<td>TB with SS</td>
<td>1,000 &amp; 10</td>
<td>No ($R=1.028$)</td>
<td>0.0745</td>
</tr>
<tr>
<td>TB with SS</td>
<td>10,000 &amp; 100</td>
<td>No ($R=1.211$)</td>
<td>0.0104</td>
</tr>
</tbody>
</table>

### 3.3.5 Validation and comparison of corrections

Accuracy and computational efficiency are used to determine which combination of correction methods is most useful. Table 3.1 contains the results of a timestep sensitivity analysis on the $10^4$ s test simulation used previously. The results for particle reflection combined with either placement method are similar; the well-mixed condition is maintained even if only a single timestep is used for the entire simulation (concentration profiles not shown but they are indistinguishable from Figure 3-14 in each case). For timestep balancing, the quality of the simulation decreases as the timestep increases because particles which have encountered the boundary must take a truncated step to reach the output time (see Section 3.3.4.2.2), which means that particle fluxes for this last step are not in balance. This flux error increases as the number of timesteps between outputs decreases (see Figure 3-14 for resulting concentration profiles and mass ratios). Thus, from the standpoint of maintaining the well-mixed condition, particle reflection with either placement approach is more useful. This conclusion is strongly supported by the computational efficiencies in Table 3.1: while particle reflection with either placement technique are comparably efficient, timestep balancing is much less efficient because a smaller timestep is used in the low diffusivity region.

The results in Table 3.1 suggest that the particle reflection corrections exhibit no timestep
sensitivity, but this is not true. Although Thomson [24] argues that satisfying the well-mixed condition is a sufficient test of a one-particle stochastic model such as the random walk, this conclusion does not necessarily apply here since the proposed correction techniques are specifically designed to maintain the uniform particle distribution (as pointed out by Wilson and Flesch [32], the well-mixed condition alone cannot fully determine the trajectory model). Nonetheless, it is reasonable to expect that the corrections will work even when concentrations are non-uniform. If particle density is higher on one side of the boundary, then the flux from that side will increase accordingly. The correction methodologies provide a way to subtract out the extra flux that is merely an artifact of the discontinuity; the remaining flux imbalance due to concentration gradients is left behind.

The correction methods were therefore also tested for a non-steady case. Thomson et al.
[26] found that their approach (equivalent to particle reflection with midstep adjustment) compared favorably when applied to non-uniform concentrations, using a small timestep simulation of a sharply varying but continuous diffusivity profile in a bounded domain as “truth”. Here an analytical solution is used as the basis for evaluation, in which the initial condition is

\[
c(z, t = 0) = \begin{cases} 
0 & z \leq z_b \\
c_0 & z > z_b 
\end{cases} \tag{3.58}
\]

The solution was derived by Crank [4] for an infinite domain:

\[
c(z', t) = \begin{cases} 
\frac{c_0}{1 + \sqrt{\frac{D_2}{D_1}}} \left[ 1 + \sqrt{\frac{D_2}{D_1}} \text{erf} \left( \frac{z'}{2D_1 t} \right) \right] & z \leq z_b \\
\frac{c_0}{1 + \sqrt{\frac{D_2}{D_1}}} \text{erfc} \left( \frac{z'}{2D_1 t} \right) & z > z_b 
\end{cases} \tag{3.59}
\]

where \( z' \) is the distance from the boundary, and \( D_1 \) and \( D_2 \) are the diffusivities for above and below the boundary at \( z = z_b \) (\( z' = 0 \)), respectively. Simulations were performed using a modified version of the test domain in which the impermeable boundaries were moved far beyond the discontinuity to emulate an unbounded domain (i.e., \( z = 0 \) and \( z = 200 \) cm are no longer boundaries and the simulation domain extends well beyond them). Figure 3-15 compares the analytical solution to simulation results using \( \Delta t = 1 \) s together with the correction methods developed herein and the Hoteit et al. scheme [14]. Each method is able to reproduce the analytical solution with this timestep.

The results begin to diverge from the analytical solution when the timestep is increased, as shown for \( \Delta t = 100 \) s in Figure 3-16. The timestep balancing with step shortening approach exhibits the greatest timestep sensitivity; there is a significant error in predicted concentration immediately below the \( z = 100 \) cm interface, and mass flux into the lower layer is overpredicted. The other three methods have approximately the same accuracy for this larger timestep (all underpredicting the flux into the low diffusivity region). Further increase in the timestep (not shown) reveals that the Hoteit et al. [14] approach diverges slightly faster than the particle reflection approach developed in 3.3.4.1, presumably due to the fact that the Hoteit et al. [14] approach reflects particles from the low diffusivity region as well (see Section 3.3.4.1.1). Coupled with the computational efficiency comparison in
Figure 3-15: Predicted concentration (top panels, averaged over last 20% of simulation and normalized by $c_0$) and ratio $R$ of mass below to mass above $z = 100$ cm (bottom panels) for the unsteady case for $\Delta t = 1$ s. The analytical solution is plotted (dashed line) along with simulation results using four different correction methods. Left panels: midstep adjustment with particle reflection (x-marks) and the approach of Hoteit et al. [14] (open circles). Right panels: step shortening with particle reflection (x-marks) and timestep balancing (open circles). Note that the domain used here extends well beyond the $0 \leq z \leq 200$ cm interval; the impermeable boundaries were moved because the analytical solution applies to an infinite domain.

Table 3.1, the particle reflection with midstep adjustment (first proposed by Thomson et al. [26]) exhibits the best performance.
Figure 3-16: Predicted concentration (top panels, averaged over last 20% of simulation and normalized by $c_0$) and ratio of mass below to mass above $z = 100$ cm (bottom panels) for the unsteady case for $\Delta t = 100$ s. The analytical solution is plotted (dashed line) along with simulation results using four different correction methods. Left panels: midstep adjustment with particle reflection (x-marks) and the approach of Hoteit et al. [14] (open circles). Right panels: step shortening with particle reflection (x-marks) and timestep balancing (open circles). Note that the domain used here extends well beyond the $0 \leq z \leq 200$ cm interval; the impermeable boundaries were moved because the analytical solution applies to an infinite domain.
3.4 Piecewise Linear Diffusivity

The case of a continuous, piecewise linear diffusivity profile is markedly different from the discontinuous diffusivity case in that the error introduced by the random walk model (3.1) can be made negligible if the timestep is decreased sufficiently. While the combination of small timesteps and smoothed profiles can address this problem (see Section 3.1), the goal of the present study is to develop a method by which the discontinuities can remain unsmoothed and reasonable accuracy can be retained without resorting to very small timesteps. The approach taken below closely mimics that of Section 3.3 in that it begins with an illustration of the error and the derivation of equations which describe various statistical distributions for particles crossing a boundary from a region of constant but non-zero \( \frac{dD}{dz} \). The same two sources of error identified in Section 3.3, particle flux and placement error, also apply here. The magnitude of the particle flux error across a discontinuity is quantified by means of a reflection probability, and a method for placing particles is proposed. The correction method is then tested in a number of situations.

3.4.1 Illustration of the problem

The test domain is shown in Figure 3-17. It differs from the domain in Figure 3-1 in that the diffusivity increases in the upper layer with a constant diffusivity gradient of 1 cm/s. Thus, the diffusivity is continuous but its gradient is not, changing abruptly from 0 to 1 cm/s at \( z = 100 \) cm.

The base case test simulation was initialized with \( 10^5 \) particles distributed uniformly over the domain. Simulations were conducted in the manner described in Section 3.3.3, and again the random walk model (3.1) is tested on its ability to maintain the uniform particle distribution. Figure 3-18 shows the average concentration profile (over the last 20% of the simulation) and mass ratio (as before, \( R = \frac{M_z<100cm}{M_z>100cm} \)) over the course of the \( 10^4 \) s simulations using \( \Delta t \) values of 1 s, 0.1 s, and 0.01 s. Although the well-mixed condition \( (C=1 \text{ and } R=1) \) is not maintained in any case, the solution improves with decreasing timestep and is fairly close to uniform for the \( \Delta t = 0.01 \) s case \( (R=1.03 \text{ at } t = 10^4 \) s). Reducing \( \Delta t \) to 0.001 s removes most of the remaining error (not shown). Thus, as noted previously, the discontinuity in
diffusivity gradient can be handled by decreasing the timestep enough, which could either be constant or part of a variable timestep scheme. However, it is difficult to a priori select a sufficiently small timestep or valid variable timestep scheme to make the discontinuity negligible, and computational expense can become considerable as the computation time is inversely proportional to the timestep. The correction methodology proposed below provides an approximate way to avoid this problem.

### 3.4.2 Correction methodology

In designing a correction method, a series of equations for particle statistics at the discontinuity was developed in much the same manner as in Section 3.3.2. The equations are considerably more complicated in this case and resulted in only a single useful correction technique. In the interest of brevity and readability, the correction approach is summarized below without derivation of the underlying equations; these derivations are deferred to Appendix A.

Following the approach taken in the step diffusivity case (see Section 3.3), the goal of the correction is to balance the flux at the discontinuity based on the error that arises when the particle density is uniform, and then to place particles on either side of the discontinuity in a manner which preserves the uniform density. Thus, the discontinuity should not be
Figure 3-18: Predicted time-averaged concentration (top panel) and mass ratio for the test domain with an initially uniform particle density using constant $\Delta t$ values of 1 s (solid line), 0.1 s (dashed line), and 0.01 s (dotted line). The steady state solution is $C=1$ and $R=1$.

“felt” because the particle flux should everywhere be the same as if no discontinuity existed. However, the deterministic “drift” term in (3.1) introduces an asymmetry in particle displacements about the initial position which has two important consequences in the present context. First, the simple rules for particle placement developed for the step diffusivity case (midstep adjustment and step shortening) do not work in this case, nor does the timestep balancing approach to flux correction. Second, the treatment of impermeable boundaries by particle bouncing is not accurate, i.e., it does not preserve a uniform particle density (see discussion in Section 3.2.3, [22], and [9]). Within the correction approach designed below, an impermeable boundary is merely a special case of a discontinuity in diffusivity gradient, and thus the correction addresses both the error caused by discontinuities as well as those caused by impermeable boundaries.

In the discussion that follows, a boundary coordinate system similar to (3.5) is employed to describe particle motion relative to a boundary at $z = z_b$, here adapted to include the diffusivity variation:

$$z' \equiv |z - z_b|$$

(3.60a)
\[ \xi' \equiv \begin{cases} -\xi & z < z_b \\ \xi & z > z_b \end{cases} \quad (3.60b) \]

\[ \frac{\partial D'}{\partial z} \equiv \begin{cases} -\frac{\partial D}{\partial z} & z < z_b \\ \frac{\partial D}{\partial z} & z > z_b \end{cases} \quad (3.60c) \]

Applying (3.60) to (3.1) yields

\[ \Delta z' = \frac{dD'}{dz} \Delta t + \xi' \sqrt{2D(z'_0)\Delta t}. \quad (3.61) \]

Here diffusivity is given by

\[ D(z') = D_b + \frac{dD'}{dz} z' \quad (3.62) \]

where \( D_b \equiv D(z_b) \) is the diffusivity at the boundary. Thus, a negative \( \frac{dD'}{dz} \) means that diffusivity decreases away from the boundary, and vice versa. Likewise, a negative \( \xi' \) or \( \frac{dD'}{dz} \) moves the particle toward the boundary (\( \Delta z' < 0 \)), and vice versa. As before, this coordinate system is not practical for simulation but is employed in describing the flux at the discontinuity because it eliminates the need to specify whether a particle crosses \( z_b \) from \( z < z_b \) or \( z > z_b \).

The discussion to follow makes frequent reference to a continuous domain, which must first be defined. For the step diffusivity case, equations were developed to describe the statistics of particles arriving at a boundary from an unbounded domain with constant diffusivity and uniform particle density, i.e., one which is infinite in extent and which contains no other boundaries or discontinuities. For the present case in which \( \frac{dD}{dz} \) is a non-zero constant, an unbounded domain cannot exist because for any \( \frac{dD}{dz} \) the diffusivity eventually goes to zero. Thus, when considering the flux at a boundary at \( z' = 0 \) from a region with \( \frac{dD'}{dz} < 0 \) (i.e., diffusivity decreasing away from the boundary), there is effectively an impermeable boundary that must exist at \( z' = \frac{D_b}{|\frac{dD}{dz}|} \). The term continuous domain will be used to refer to a domain with constant \( \frac{dD}{dz} \) everywhere and an impermeable boundary at the point where \( D = 0 \).

The correction method here also consists of two parts: a particle flux correction and a particle placement method. The former is handled by particle reflection in the same way as the step diffusivity case (see Section 3.3.4.1.1), i.e., a probability of reflection is
defined at the discontinuity based on the expected flux from either side if the domain were continuous (as defined above) with a uniform particle density $\rho$. The latter is addressed by a generalization of the midstep adjustment approach developed in Section 3.3.4.2.1, which is termed probability translation in the text that follows.

The implementation of the correction method requires the equation for $N_b(Z_f)$, the number of particles expected to cross the boundary at $z' = 0$ during a single timestep $\Delta t$ and reach a distance $Z_f$ beyond the boundary (see Appendix 3.A for its derivation):

$$
\frac{N_b(Z_f)}{\hat{\rho}} = H(-\alpha)H \left( 1 - \hat{D}_b - \hat{Z}_f \right) \left[ e^{\hat{D}_b + \hat{Z}_f - 1} - \hat{D}_b - \hat{Z}_f \right] + H(-\alpha) \left[ H(\hat{D}_b - \hat{z}^*) \left\{ \hat{D}_b + \hat{Z}_f \right\} + H(\hat{z}^* - \hat{D}_b)e^{(\hat{D}_b - \alpha \hat{Z}_f - 1)} \right] + \alpha H(\hat{z}^*)e^{(\hat{D}_b - \alpha \hat{Z}_f - 1)} - \frac{\hat{Z}_f}{2} \left[ \text{erf} \left( \frac{\xi'(0)}{\sqrt{2}} \right) + 1 \right] + \frac{\alpha}{\sqrt{2\pi}} e^{-\frac{\xi'(0)^2}{2}} \left[ \left\{ \hat{\lambda}(\xi'(0)) + \sqrt{\frac{\pi}{2}} \text{erf} \left( \frac{\hat{\lambda}(\xi'(0))}{\sqrt{2}} \right) \right\} \left\{ 1 - 2H(\hat{z}^*) \right\} - \xi'(0) \right] \tag{3.63}
$$

where $H(x)$ is the Heaviside function:

$$
H(x) = \begin{cases} 
0 & x < 0 \\
\frac{1}{2} & x = 0 \\
1 & x > 0 
\end{cases} \tag{3.64a}
$$

and

$$
\alpha \equiv \frac{dD'}{dz} \frac{d}{dz} = \text{sgn} \left( \frac{dD'}{dz} \right) \tag{3.64b}
$$

$$
\hat{Z}_f \equiv \frac{Z_f}{dD/dz} \Delta t \tag{3.64c}
$$

$$
\hat{\rho} \equiv \rho \left| \frac{dD}{dz} \right| \Delta t \tag{3.64d}
$$

$$
\hat{D}_b \equiv \frac{\hat{D}_b}{dD/dz} \Delta t \tag{3.64e}
$$

$$
\hat{z}^* \equiv \hat{Z}_f + \alpha (1 - 2\hat{D}_b) \tag{3.64f}
$$
\[ \xi'_c(z'_i) \equiv \frac{-z'_i - Z_f - \frac{dD'}{dz} \Delta t}{\sqrt{2(D_b + \frac{dD'}{dz} z'_i) \Delta t}} \quad \Rightarrow \quad \xi'_c(0) = \frac{-\dot{Z}_f - \alpha}{\sqrt{2D_b}} \tag{3.64g} \]

\[ \hat{\lambda}(\xi'_c) \equiv \sqrt{\xi'^2 - 2 + 2 \left( \dot{D}_b - \alpha \dot{Z}_f \right)} \tag{3.64h} \]

Also, \( \text{erfcx}(x) \) is the scaled complementary error function:

\[ \text{erfcx}(x) \equiv e^{x^2} \text{erfc}(x) = e^{x^2} (1 - \text{erf}(x)). \tag{3.65} \]

Equation (3.63) is used mainly in two ways. First, with \( Z_f = 0 \) it expresses the total number of particles arriving at a boundary, and thus can be used to characterize the flux imbalance caused by a discontinuity in \( \frac{dD}{dz} \). Second, it can be used to express \( P(0 \leq z_f \leq Z_f) \), the probability that a particle which crosses the \( z'_i = 0 \) boundary will have final position \( z'_f \) in the interval \( 0 \leq z'_f \leq Z_f \) beyond the boundary if the domain were continuous:

\[ P(0 \leq z_f \leq Z_f) \equiv 1 - \frac{N_b(Z_f)}{N_b(Z_f = 0)}. \tag{3.66} \]

In addition, the corresponding pdf \( p(Z_f) \) can be expressed:

\[ p(Z_f) \equiv \frac{\partial P(0 \leq z_f \leq Z_f)}{\partial Z_f}. \tag{3.67} \]

The recipe for the correction method is as follows:

1. Simulate particles with (3.1) using the same \( \Delta t \) everywhere in the domain.

2. At each discontinuity in the domain, calculate a probability of reflection \( P_r \) for particles arriving from either side of the discontinuity (referred to below as regions 1 and 2) based on the expected flux from a uniform particle density in a continuous domain. Following (3.48) with a constant \( \Delta t \), these are calculated as:

\[ P_{r,1} = \max \left\{ 0, 1 - \frac{N_{b,2}(Z_f = 0)}{N_{b,1}(Z_f = 0)} \right\} \tag{3.68a} \]

\[ P_{r,2} = \max \left\{ 0, 1 - \frac{N_{b,1}(Z_f = 0)}{N_{b,2}(Z_f = 0)} \right\} \tag{3.68b} \]
where the subscripts refer to regions 1 and 2, respectively. Thus, only particles coming from the region with the higher steady state flux are reflected, as was the case for step diffusivity. Impermeable boundaries have $P_r = 1$.

3. When a particle encounters a discontinuity or impermeable boundary, draw a random number $r$ from a $[0,1]$ uniform distribution. If $r \leq P_r$, the particle is reflected; if $r > P_r$, the particle is transmitted.

4. Calculate the distance $Z_{f,\text{orig}}$ beyond the boundary which the particle would travel if it were allowed to cross the boundary, i.e.,

$$Z_{f,\text{orig}} = |\Delta z' + z'_i|,$$

where $z'_i$ is always positive and $\Delta z'$ is always negative for crossing particles in boundary coordinates.

5. Calculate the probability $P(0 \leq z'_f \leq Z_{f,\text{orig}})$ by using (3.66), i.e., from the distribution of expected final positions on the other side of the boundary when the particle density is uniform in the originating region and the domain is continuous. Note that $P(0 \leq z'_f \leq Z_{f,\text{orig}})$ is defined even when the final position $z'_f = Z_{f,\text{orig}}$ is illegal (i.e., beyond the $D = 0$ boundary).

6. Find the value of $Z_{f,\text{new}}$ which has an equal probability in the region in which the particle is to be placed (i.e., the opposite region if transmitted or the originating region if reflected). This is accomplished by solving for $Z_{f,\text{new}}$ in

$$P_{\text{orig}}(0 \leq z'_f \leq Z_{f,\text{orig}}) = P_{\text{new}}(0 \leq z'_f \leq Z_{f,\text{new}}).$$

$P_{\text{new}}(0 \leq z'_f \leq Z_{f,\text{new}})$ is calculated using (3.66) with an appropriate value of $\frac{dD'}{dz'}$. If the particle is being transmitted ($R > P_r$), then use $-\frac{dD'}{dz'}$ of the new region. If the particle is being reflected ($R \leq P_r$), then use $-\frac{dD'}{dz'}$ of the originating region. The negative sign is needed because the distance $Z_{f,\text{new}}$ is taken from the distribution of particles coming from the other side of the boundary in a continuous domain (i.e., as
if there were no discontinuity). \( Z_{f,new} \) must be approximated iteratively, e.g., using the Newton-Raphson Method. Note that this step is followed even when the particle is reflected back into the same region because simple particle bouncing is not accurate when \( \frac{dD}{dz} \neq 0 \). This step is referred to as probability translation, since the probability of particle displacements is in effect translated across a discontinuity.

7. The final position of the particle is \( z'_f = Z_{f,new} \), which is converted to a final position in normal coordinates \( z_f \) from (3.60) based on whether the particle was transmitted or reflected.

Probability translation can be thought of as a generalization of the midstep adjustment approach introduced in the step diffusivity case, where the “velocity” of transmitted particles was adjusted to match the new region and the “velocity” of reflected particles was simply reversed (i.e., particle bouncing).

The probability translation approach proposed above is very similar in principle to the reflection approach originally proposed by Hurley and Physick [16] for impermeable boundaries (later developed in Thomson and Montgomery [25] and Thompson et al. [26] in the context of discontinuities), as well as the probability “mapping” approach of Weil [30, 32]. These studies were concerned with higher order stochastic models appropriate for simulating times shorter than the Lagrangian integral timescale, i.e., ones in which particles have some degree of “memory” and both particle velocity and position evolve stochastically (e.g., see [1] or introduction of [17]). While the details are different, the general approach is highly similar.

Before proceeding to the evaluation of the performance of the method, it should be noted that the derivation of (3.63) suggests that it may be prudent to impose a constraint on the timestep when using the correction method:

\[
\Delta t \leq \frac{D_b}{\left| \frac{dD}{dz} \right|}.
\]  

(3.71)

This follows from the fact that if the timestep violates this constraint, then the particle flux at the discontinuity quantified by (3.63) includes a contribution from an artificial “reflected” particle flux originating from the impermeable boundary that must exist at \( D = 0 \). As
explained in Appendix 3.A, this flux is deduced from the imbalance of non-reflected particles in a continuous domain and is required to meet the well-mixed condition. However, the test case simulations in the next section do not support the necessity of this constraint, and it is therefore not considered further.

3.4.3 Evaluation of correction method

The correction method (particle reflection with probability translation) is evaluated using the test diffusivity profile in Figure 3-17, as well as two other sample applications. The first is an approximation to the profile observed in laboratory experiments of a shear flow caused by a submerged aquatic canopy; the second example is a sample diffusivity profile generated by an OGCM. In addition, the reflection off of an impermeable boundary is examined as it has been the subject of past studies.

3.4.3.1 Test profile

The diffusivity profile in Figure 3-17 was simulated using the correction method proposed above. Figure 3-19 shows simulation results for for an initially uniform particle density with a \( \Delta t \) of 1 s, 10 s, and 100 s. The well-mixed condition is met in each case, i.e., the particle accumulations observed in Figure 3-19 are prevented (note the change in plotting scales between the figures). Using \( \Delta t = 1000 \) s (not shown) yields similar results, i.e., the ability of the correction methodology to preserve the well-mixed condition shows little timestep sensitivity. For comparison, the uncorrected \( \Delta t = 0.01 \) s simulation from Figure 3-18 has also been included in Figure 3-19; each corrected case is more successful at meeting the well-mixed condition. The corrected simulations are much more computationally efficient, needing only 1.1, 0.15, and 0.041 % of the uncorrected \( \Delta t = 0.01 \) s simulation time. Note that the scaling of simulation time with timestep for the corrected results is not linear as it is for the uncorrected results, since more probability translations per timestep are needed for larger timesteps; nonetheless, the savings are substantial.

However, as for the step diffusivity case, satisfying the well-mixed condition seems an incomplete test of the correction method, since it is specifically designed to meet this condi-
Figure 3-19: Predicted time-averaged concentration (top panel) and mass ratio for the test domain with an initially uniform particle density using the proposed correction methodology with constant $\Delta t$ values of 1 s (solid line), 10 s (dashed line), and 100 s (dotted line). The steady state solution is $C=1$ and $R=1$. For comparison, the uncorrected result with $\Delta t = 0.01$ s is also shown in the bottom panel (points).

Thus, non-steady conditions are also examined. Figure 3-23 shows simulations in which the test domain was initialized with a uniform particle density in the top half of the domain only. No analytical solution for this case was found, but an uncorrected simulation with a small timestep can be used in place of one. Convergence of the uncorrected result is apparent by the closeness of the $\Delta t = 0.01$ s and $\Delta t = 0.001$ s results (left panel). The corrected results are able to approximate the correct answer with much less computational expense. Although the corrected simulations begin to diverge from the true result as the timestep increases (underpredicting mass accumulation in the lower half of the domain), they retain fair accuracy even for large timesteps (based on intermediate time comparisons in Figure 3-23 coupled with the long time convergence implied by Figure 3-19 results). Similar quality of results were obtained when particles in the upper layer were initialized as a point source at $z = 150$ cm.

The reduction in computational expense achieved by the correction method can be judged on the basis of the results in Figure 3-23. Taking the corrected $\Delta t = 10$ s and uncorrected $\Delta t = 0.01$ s simulations as sufficiently accurate, then a factor of 600 reduction in simulation
Figure 3-20: Predicted time-averaged concentration (top panel) and mass ratio for the test domain with an initially uniform particle density in the top half of the domain. Left panels: uncorrected simulations using a $\Delta t$ of 1 s (solid line), 0.1 s (dashed line), 0.01 s (dotted line), and 0.001 s (points). Right panels: corrected simulations using a $\Delta t$ of 1 s (solid line), 10 s (dashed line), 100 s (dotted line), and 1000 s (dash-dot line). The uncorrected simulation using a $\Delta t$ of 0.001 s is also shown on the right (points) in place of an analytical solution.

Time can here be attributed to the correction method. Using the $\Delta t = 1$ s, the improvement is roughly a factor of 90.

3.4.3.2 Submerged canopy shear flow

The performance of the correction methodology was also studied using an idealization of the vertical diffusivity profile measured by Ghisalberti and Nepf [13], in which the shear induced turbulence due to a submerged aquatic vegetation canopy was examined (Figure 3-21 shows the measured profile; Figure 3-22 shows the adapted piecewise linear profile used in simulations). The profile has linear diffusivity variation immediately above and below the top of the canopy (where the maximum value is observed), and the diffusivity is constant.
Figure 3-21: Non-dimensional vertical diffusivity profile from Ghisalberti and Nepf [13] for flow over a submerged aquatic canopy.

for the top quarter of the experimental domain as well as for a thin bottom boundary layer.

To be consistent with the experiments carried out in [13], all $10^5$ particles were introduced at a single point, at $z = 15$ cm. Simulations were run for 200 s, which is long enough for the steady state solution (uniform concentration) to be reached. Results for a range of timesteps are presented in Figure 3-23 for both uncorrected (for $\Delta t = 1$, 0.1, 0.01, and 0.001 s) and corrected simulations (for $\Delta t = 1$, 5, and 10 s). Here the concentration profile is plotted at an intermediate time, $t = 50$ s, along with the time series of fractional mass in the top 8 cm ($M_{top}$) and bottom 3 ($M_{bot}$) cm of the domain (the regions of constant diffusivity) so that the steady state solution is $M_{top} = 0.84$ and $M_{bot} = 0.06$. The uncorrected simulations with $\Delta t = 0.01$ s and $\Delta t = 0.001$ s are very similar, indicating convergence by $\Delta t = 0.01$ s. The corrected $\Delta t = 1$ s simulation is comparable to the uncorrected $\Delta t = 0.1$ s result for intermediate times, but converges to the correct steady state solution. For larger timesteps, the corrected solution diverges from the true result for intermediate times but converges to the correct steady state solution (i.e., well-mixed). Adjusting for time spent on output, the simulation times for the corrected simulations with $\Delta t = 1$, 5, and 10 s were about 1.1, 0.2, and 0.17 %, respectively, of the time required for the uncorrected $\Delta t = 0.01$ s simulation. In addition, the corrected $\Delta t = 1$ s simulation required about 12 % of the uncorrected $\Delta t = 0.1$
s simulation time. Since the accuracy of the corrected simulations is comparable to that of the uncorrected simulations for intermediate times and better for longer times, the correction method can be judged to reduce simulation time by at least a factor of 8 for comparable accuracy.

It should be noted that in reality, this profile cannot be simulated as a pure diffusion problem since the turbulence is induced by vertical shear in the horizontal velocity. Thus, a horizontal advection term would be included in the random walk calculation, which because of the shear would lead to longitudinal dispersion (see [13]). This would place an additional constraint on the timestep, in that it would have to be small enough to adequately sample the velocity profile. However, even in this case the correction method is useful as it allows the timestep to be chosen, within reasonable limits as suggested by Figure 3-23, based on velocity shear considerations.

3.4.3.3 Diffusivity profiles from an OGCM

An additional illustration of the faster convergence of simulations with the proposed correction methodology is provided below using two severe examples of vertical diffusivity
Figure 3-23: Predicted time-averaged concentration (top panels) and mass ratio for the test domain with an initially uniform particle density in the top half of the domain (bottom panels). Left panels: uncorrected simulations using a $\Delta t$ of 1 s (solid line), 0.1 s (dashed line), 0.01 s (dotted line), and 0.001 s (points). Right panels: corrected simulations using a $\Delta t$ of 1 s (solid line), 10 s (dashed line), 100 s (dotted line), and 1000 s (dash-dot line). The uncorrected simulation using a $\Delta t$ of 0.001 s is also shown on the right (points) in place of an analytical solution.

profiles (Figure 3-24) observed in an ocean circulation simulation using the Lawrence Livermore National Laboratory version of the Modular Ocean Model (LLNL MOM, see [8] and [3]). The profiles were modified slightly for the present study in that the top and bottom diffusivities were set equal to their neighbors; within LLNL MOM, both of these values are actually zero such that no flux boundaries at the top and bottom of the ocean are imposed.\(^1\)

The left profile is generated by the “KPP” vertical mixing approach of Large et al. [20] and has a deep ocean boundary layer (OBL, $\sim$ 700 - 850 m) with convective instability which results in large diffusivities (to reduce the instabilities) with strong vertical gradients (needed

\(^1\)The modifications were made so that the analysis which follows could focus solely on the treatment of internal discontinuities in diffusivity gradient, and because the simulation of impermeable boundaries with non-zero diffusivity gradient is addressed specifically in Section 3.4.3.4
to satisfy the conditions that the diffusivity is zero at the ocean surface and continuous at the base of the OBL. The right profile has a shallow OBL (\(~\ 100\ m\)) but with large vertical gradients in diffusivity in the interior ocean due to either the subgrid scale mixing parameterization of Duffy et al. [8] or the shear instability parameterization of Large et al. [20] (the jagged nature of the profile suggests that the LLNL MOM vertical grid resolution may be insufficient). The magnitude of the diffusivities coupled with the sharp variability provide an unfriendly environment for a diffusing particle, i.e., one that would require a considerable degree of smoothing and/or a small timestep. However, if the vertical diffusivity is assumed to vary linearly between cell faces, simulation of these profiles can be aided by the proposed correction methodology.

Uncorrected and corrected test simulations using the OGCM diffusivity profiles are shown for a range of timesteps in Figures 3-25 to 3-28. Particles were binned using the vertical grid of the parent OGCM, and diffusivity was made constant in the top and bottom grid cells for simplicity. Here the convergence of the correction method is not tested by comparison to calculations with the parent OGCM because differences due to the numerical scheme and the assumed subgrid scale diffusivity variation influence the comparison (reproducing LLNL MOM calculations with a particle method is the subject of an ongoing companion study, see
Chapter 4). Rather, convergence is demonstrated by using smaller timesteps, as before.

Figure 3-25 shows that for an initially uniform distribution (using 5,000 particles), the first profile of Figure 3-24 can cause a significant mass accumulation to occur in the bottom grid cell (where diffusivity is 4-6 orders of magnitude lower than in the overlying water column). The top panels in Figure 3-25 shows the time-averaged concentration $C$ (normalized by the initial value, which is also the steady state value) while the bottom panel shows the ratio of the mass in the bottom OGCM grid cell (845 - 1039 m) to the initial and steady state value ($R$); both $C$ and $R$ should be one if there is no error. Dropping the timestep to 1 s reduces the observed accumulation to about 5% after 10 days in the uncorrected case, and further reduction to 0.1 s removes most of the remaining mass accumulation. By contrast, the corrected simulations are able to maintain the uniform particle density for a wide range of timesteps.

The performance of the corrected simulations for an initially non-uniform particle density was also tested (Figure 3-26). Here the domain was initialized with particles only in the depth range 229 - 309 m (which corresponds to a single OGCM grid cell). As expected, the uncorrected results are highly sensitive to the choice of timestep. The corrected results exhibit smaller timestep sensitivity, and are able to approximate the uncorrected $\Delta t = 0.1$ s simulation with much less computational effort. Even the $\Delta t = 1800$ s simulation yields reasonable results; by the end of the simulation, its predicted lower layer mass is only about 10% lower than the uncorrected $\Delta t = 0.1$ s result. The computation time for the corrected $\Delta t = 18, 180, and 1800$ s simulations required 0.80, 0.13, and 0.032 % of the uncorrected $\Delta t = 0.1$ s computation time. Thus, the correction method offers a large increase in computational efficiency for a modest cumulative flux error in this case.

The results for the second diffusivity profile are plotted in Figures 3-27 and 3-28; here the bottom panel shows the ratio of the mass in the upper 165 m of the domain to the steady state value (i.e., the region above the uppermost spike in diffusivity, which also has the largest value of diffusivity gradient). The thinner grid cells in the upper regions causes the results to be noisier, but nonetheless a mass accumulation is evident in the uncorrected results. Dropping the timestep to 1 s removes most of the mass accumulation in this case. The corrected results (right panels of Figure 3-27) show mass accumulation less than or
Figure 3-25: Predicted time-averaged concentration (top panels) and mass ratio (bottom panels) for the first LLNL MOM sample diffusivity profile with an initially uniform particle density. Left panels: uncorrected simulations using a $\Delta t$ of 180 s (solid line), 18 s (dashed line), 1 s (open circles), 0.1 s (points). Right panels: corrected simulations using a $\Delta t$ of 1800 s (solid line), 180 s (dashed line), 18 s (dotted line). The uncorrected simulation using a $\Delta t$ of 1 s and 0.1 s are also shown on the right for comparison (open circles and points).

equal to the uncorrected $\Delta t = 1$ s simulation using much larger timesteps. A similar trend is revealed in the results for a non-uniform initial condition (Figure 3-28); the corrected results approximate the uncorrected $\Delta t = 1$ s results using a much larger timestep. The computation time for the corrected $\Delta t = 18$, 180, and 1800 s simulations was 9.3, 1.5, and 0.37 % of the uncorrected $\Delta t = 1$ s computation time.

Considering the results of both diffusivity profiles, the correction method is able to reduce computational effort substantially. Taking the $\Delta t = 1800$ s as reasonably accurate in both cases, the correction method was observed to reduce computation time by factors of 270 and 2260 for comparable accuracy in the two cases tested.
3.4.3.4 Reflection off an impermeable boundary

The performance of the correction method in the case of reflection off an impermeable boundary is also presented because this issue has been the subject of a number of previous studies. After reviewing several reflection schemes, Wilson and Flesch [32] conclude that no reflection scheme satisfies the well-mixed condition for skew or inhomogeneous turbulence for higher order stochastic models (see Rodean [21], Wilson and Flesch [32], Weil [30], Thomson and Montgomery [25], Thomson et al. [26], and references therein for additional discussion in this context). As discussed in Section 3.2.3, this issue has been looked at by Ross and Sharples [22] and Ermak and Nasstrom [9] for particle models similar in form to (3.1). The former tried several approximate approaches to maintaining the well-mixed condition with
Figure 3-27: Predicted time-averaged concentration (top panels) and mass ratio (bottom panels) for the second LLNL MOM sample diffusivity profile with an initially uniform particle density. Left panels: uncorrected simulations using a $\Delta t$ of 180 s (solid line), 18 s (dashed line), 1 s (points). Right panels: corrected simulations using a $\Delta t$ of 1800 s (solid line), 180 s (dashed line), 18 s (dotted line). The uncorrected simulation using a $\Delta t$ of 1 s is also shown on the right for comparison (points). The steady state solution is $C=1$ and $R=1$.

varying degrees of success, while the latter combined (3.1) with an analytical solution for the case of a $D=0$ boundary with constant $\frac{dD}{dz}$.

The correction method proposed in Section 3.4.2 allows particles to be reflected at any boundary in a way that maintains the well-mixed condition. For the case of a domain with a $D = 0$ boundary at $z = 0$ and constant $\frac{dD}{dz}$ (i.e., a continuous domain as previously defined), the reflection probability is given by (3.66) and (3.67):

\[ p(Z_f) = \frac{\partial P(0 \leq z_f \leq Z_f)}{\partial Z_f} \]
\[ = -\frac{1}{N_b(Z_f = 0)} \frac{\partial N_b(Z_f)}{\partial Z_f}. \]  

(3.72)
Evaluating (3.63) and (3.91) with $-\frac{dD}{dz}$ and $D_b = 0$ yields:

$$N_b(Z_f = 0) = \frac{\hat{\rho}}{e} \quad (3.73a)$$

$$\frac{\partial N_b(Z_f)}{\partial Z_f} = \hat{\rho} H \left(1 - \hat{Z}_f\right) \left[e^{\hat{Z}_f} - 1\right] \quad (3.73b)$$

Thus, particles impinging on the $D = 0$ boundary will maintain the well-mixed condition if they are reflected according to the distribution:

$$p(\hat{z}) = H \left(1 - \hat{z}\right) \left[e^{\hat{z}} - e\right]. \quad (3.74)$$
where \( \hat{z} \equiv \frac{\hat{Z}_f}{\left| \frac{dD}{dz} \right| \Delta t} \) has been substituted for \( \hat{Z}_f \) since in this case the boundary is at \( z = 0 \) (and thus the reflection distance \( Z_f \) coincides with the normal \( z \) coordinate). As shown in Figure 3-29, all reflected particles are distributed in the interval \( 0 \leq \hat{z} \leq 1 \), which corresponds to \( 0 \leq z \leq \frac{dD}{dz} \Delta t \).

The ability of the correction method to maintain the well-mixed condition for this domain is demonstrated in Figure 3-30, which shows corrected and uncorrected (i.e., particle bouncing) results for a simulation of \( 3 \times 10^5 \) particles initially distributed uniformly in a 100 cm domain with impermeable boundaries at the top and bottom with \( D = 100 \) and 0 cm\(^2\)/s, respectively. The top panel shows the time-average concentration \( C \) over the simulation normalized by the initial concentration, while the bottom panel shows the ratio \( R \) of mass in the lower 5 cm of the domain to its well-mixed value. Thus, the well-mixed condition is met if \( C = 1 \) and \( R = 1 \). The uncorrected results exhibit an error near the bottom boundary which decreases as the timestep is decreased; in the range \( 0.1 s > \Delta t > 0.01 s \) this error becomes smaller than the grid resolution of 1 cm and is no longer discernible. The corrected results do not exhibit such a timestep sensitivity; the well-mixed condition is met even for \( \Delta t = 10 s \).
Figure 3-30: Predicted time-averaged concentration (top panels) and mass ratio (bottom panels) for a domain with impermeable boundaries at the top and bottom with $D = 100$ and $0 \text{ cm}^2/\text{s}$, respectively, and an initially uniform particle density. Left panels: uncorrected simulations using a $\Delta t$ of 10 s (solid line), 1 s (dashed line), 0.1 s (dotted line), and 0.01 s (points). Right panels: corrected simulations using a $\Delta t$ of 10 s (solid line), 1 s (dashed line), 0.1 s (dotted line).

The fact that the correction method reproduces the well-mixed condition is expected since it is derived specifically to meet this criterion. For a point source simulation, the corrected results also exhibit a timestep sensitivity before the steady state is reached, as shown in Figure 3-31 (here $C$ is shown after 20 s of the 150 s simulation). Although the corrected simulations are less timestep sensitive (i.e., closer to the actual solution for larger timesteps), for short times they do not converge to the correct answer (given by the uncorrected $\Delta t = 0.01$ s simulation) until $\Delta t = 0.1$ s. For this timestep, the uncorrected results are also accurate. Thus, even though the reflection scheme satisfies the well-mixed condition as the steady state is achieved, it does not converge faster to the true result for short times. Nonetheless, it is an improvement over the uncorrected result in that it retains greater accuracy in the short term for larger timesteps and achieves the correct steady state distribution regardless
Figure 3-31: Predicted concentration at $t = 20$ s (top panels) and mass ratio (bottom panels) for a domain with impermeable boundaries at the top and bottom with $D = 100$ and 0 cm²/s, respectively, with particles introduced as a point source at $z = 30$ cm. Left panels: uncorrected simulations using a $\Delta t$ of 10 s (solid line), 1 s (dashed line), 0.1 s (dotted line), and 0.01 s (points). Right panels: corrected simulations using a $\Delta t$ of 10 s (solid line), 1 s (dashed line), 0.1 s (dotted line). The uncorrected simulation using a $\Delta t$ of 0.01 s is also shown on the right (points) in place of an analytical solution.

Although it cannot match the accuracy that would be achieved for this case by the method proposed by Ermak and Nasstrom [9] (which combines (3.1) with the analytical solution for this particular case), the proposed correction method has the advantage of being able to handle the case where diffusivity is non-zero at the boundary. As an example, Figure 3-32 shows a simulation in which the diffusivity at the bottom boundary is increased to $D = 2$ cm²/s. Here the correction method converges to the correct result by about $\Delta t = 1$ s for short times (somewhat faster than the uncorrected results) and again achieves the well-mixed condition for large timesteps.
Figure 3-32: Predicted concentration at $t = 20$ s (top panels) and mass ratio (bottom panels) for a domain with impermeable boundaries at the top and bottom with $D = 102$ and 2 cm$^2$/s, respectively, with particles introduced as a point source at $z = 30$ cm. Left panels: uncorrected simulations using a $\Delta t$ of 10 s (solid line), 1 s (dashed line), 0.1 s (dotted line), and 0.01 s (points). Right panels: corrected simulations using a $\Delta t$ of 10 s (solid line), 1 s (dashed line), 0.1 s (dotted line). The uncorrected simulation using a $\Delta t$ of 0.01 s is also shown on the right (points) in place of an analytical solution.
3.5 Summary and Conclusions

A set of tools has been developed which allow for more efficient random walk simulations with step and piecewise linear diffusivity profiles. These tools provide an alternative to the traditional approach of smoothing the diffusivity profile and using a small and/or variable timestep to prevent artificial particle accumulations in regions of severe diffusivity inhomogeneity. Each correction method proposed was designed to satisfy the well-mixed condition for any step or piecewise linear diffusivity profile.

For the step diffusivity case, three correction methods were developed: particle reflection with midstep adjustment, particle reflection with step shortening, and timestep balancing with step shortening. Based on accuracy and computational efficiency, particle reflection with midstep adjustment was identified as the best performer. This approach is identical to that proposed in Thomson et al. [26], and is slightly preferred to the approach proposed by Hoteit et al. [14].

For the piecewise linear diffusivity case, a single correction method was developed: particle reflection with probability translation. Although the method requires an iterative approximation of particle displacement probabilities at each discontinuity, it has been demonstrated to offer a substantial reduction in computational effort in a wide range of applications. The approach affords a certain level of automation in that it can be applied to any diffusivity profile, generally offers faster convergence than (3.1) without correction, and is guaranteed to reach the correct steady state (i.e., the well-mixed state). These characteristics are highly useful when conducting particle simulations using diffusivity profiles that are computed externally by a parent circulation model, which may be highly variable in time and/or space.

Future work should be directed at comparing the techniques proposed herein with the diffusivity smoothing, variable timestep, and non-Gaussian forcing approaches that have been implemented in other studies (see discussion in Section 3.2), with the ultimate goal of determining how best to maximize computational efficiency and accuracy for different classes of diffusivity profiles. Hybrid approaches may be attractive in some cases. For example, even if the full particle reflection with probability translation technique is not implemented, the reflection probabilities defined in (3.68) can be used to quantify the flux error arising from
a discontinuity and thus provide a way to identify where smoothing is needed or to choose a sufficiently small timestep.
3.A Appendix: Derivation of the $N_b(Z_f)$ Equation

The following section outlines the derivation of the $N_b(Z_f)$ equation (3.63) used in the correction methodology described in Section 3.4.2. $N_b(Z_f)$ is the number of particles originating in a continuous domain (defined in Section 3.4.2) with uniform particle density $\rho$ expected to cross an arbitrary boundary during a single timestep $\Delta t$ and reach a distance $Z_f$ beyond that boundary.

3.A.1 Derivation approach

The derivation uses the same framework and similar logic as the step diffusivity case (see Section 3.3.2), but is made considerably more complicated by the presence of the diffusivity gradient term. In the interest of brevity, the $N_b(Z_f)$ equation is derived directly; this contrasts the approach in the step diffusivity case where (3.15) was first derived (corresponding to $Z_f = 0$, or the total number of particles reaching the boundary at $z = z_b$) and then extended to (3.27) (which allows for $Z_f > 0$, and can thus be used to quantify only those particles which also reach a distance $Z_f$ beyond the boundary).

The most important difference between the derivation of (3.27) in the step diffusivity case and the derivation that follows is the treatment of reflected particles. For the step diffusivity case, the correction methods used the expected distribution of particles from an unbounded (infinite) domain as the basis for placing particles after they encounter a discontinuity or impermeable boundary. As discussed in Section 3.3.4.2, this works because particle bouncing is known to be a perfect solution for regions of constant diffusivity such that any plane near the boundary does not “feel” the presence of the boundary because the reflected flux is the same as the flux that would come from an unbounded domain. However, as noted in Sections 3.2.3, 3.4.2 and 3.4.3.4, particle bouncing is not an accurate solution when $\frac{dD}{dz} \neq 0$ at the boundary, meaning that the application of (3.1) to a continuous domain (as defined in Section 3.4.2) leads to inaccuracies. As a result, when considering a region with diffusivity decreasing away from the boundary ($\frac{dD}{dz} < 0$, using the boundary coordinate system of (3.60)), the quantity $N_b(Z_f)$ must be deduced in a two step process. In particular, $N_b(Z_f)$
is defined as the sum of “reflected” and “non-reflected” particles:

\[ N_b(Z_f) = N_{b,r}(Z_f) + N_{b,nr}(Z_f). \]  \hspace{1cm} (3.75)

\( N_{b,nr}(Z_f) \) is the number of particles which originated in the region \( 0 < z' < z'_{\text{max}} \), undergo displacement toward the \( z' = 0 \) boundary (i.e., \( \Delta z' < 0 \)), reach the \( z' = 0 \) boundary, and also reach a distance \( Z_f \) beyond this boundary in a single timestep. As defined in Section 3.4.2, a continuous domain has a boundary at \( z' = z'_{\text{max}} = D \frac{dD'}{dz'} \) when \( \frac{dD'}{dz'} < 0 \) (because \( D(z'_{\text{max}}) = 0 \)). \( N_{b,r}(Z_f) \) is the number of particles which reach the \( z' = 0 \) boundary and a distance \( Z_f \) beyond this boundary after reflecting off of the impermeable boundary at \( z' = z'_{\text{max}} \). When considering the flux from a region with \( \frac{dD'}{dz'} > 0 \) (diffusivity increasing away from the boundary), \( z'_{\text{max}} = \infty \) because there are no other boundaries in a continuous domain and thus \( N_{b,r}(Z_f) = 0 \) and \( N_b(Z_f) = N_{b,nr}(Z_f) \).

The problem is that there is no way of solving for \( N_{b,r}(Z_f) \) directly, as was possible in the step diffusivity case. However, this quantity can be deduced in the following manner.

1. Consider the continuous domain defined in Section 3.4.2. Let the \( D = 0 \) impermeable boundary be located at \( z = 0 \), such that the diffusivity is given by \( D(z) = \frac{dD}{dz} z \).

2. Consider an arbitrary plane within this domain at \( z = z_b \). In the boundary coordinate system, the region on the \( z > z_b \) side of this “boundary” has \( \frac{dD'}{dz'} > 0 \), while the \( z < z_b \) side has \( \frac{dD'}{dz'} < 0 \). For the text that follows, recall that primed variables indicate local boundary coordinates, while unprimed variables indicate normal coordinates. Also, recall that \( Z_f \) is a distance beyond a boundary and not a coordinate; therefore it is the same in either coordinate system.

3. Derive an expression for \( N_{b,nr}(Z_f) \) for the arbitrary boundary at \( z = z_b \). This is equivalent to \( N_b(Z_f) \) for the region with \( \frac{dD'}{dz'} > 0 \), and equivalent to \( N_b(Z_f) - N_{b,r}(Z_f) \) for the region with \( \frac{dD'}{dz'} < 0 \).

4. Derive an expression for \( N_{b,r,z_b=0}(Z_{f,0}) \) for particles reflecting off of the impermeable \( D = 0 \) boundary at \( z = 0 \) (i.e., \( z_b = 0 \)), where \( Z_{f,0} \) is the distance traveled away from
the $D = 0$ boundary after reflection. This can be calculated as

$$N_{b,r,z_b=0}(Z_f,0) = N_{b,r,z_b=Z_f,0}^+(Z_f = 0) - N_{b,r,z_b=Z_f,0}^-(Z_f = 0). \quad (3.76)$$

Here $N_{b,r,z_b=Z_f,0}^+(Z_f = 0)$ is $N_{b,nr,z_b}(Z_f)$ evaluated at a “boundary” at $z = z_b = Z_f,0$ for the $z > z_b$ region (which has $\frac{dD'}{dz} > 0$) using $Z_f = 0$. Likewise, $N_{b,r,z_b=Z_f,0}^-(Z_f = 0)$ is the same quantity for the $z < z_b$ region (which has $\frac{dD'}{dz} < 0$).

5. The expression for $N_{b,r}(Z_f)$, the number of “reflected” particles which travel a distance $Z_f$ beyond boundary at $z = z_b$, can then be calculated by evaluating $N_{b,r,z_b=0}(Z_f,0)$ relative to the boundary at $z = z_b$, i.e., using $Z_f,0 = Z_f + \frac{D_b}{dD'/dz}$. This follows from the fact that in the continuous domain the $D = 0$ boundary is at $z = 0$, and thus a boundary $z = z_b$ will have $D_b = z_b \left| \frac{dD'}{dz} \right|$, which means that $z_b = \frac{D_b}{dD'/dz}$. The result is

$$N_{b,r}(Z_f) \equiv H(-\alpha)N_{b,r,z_b=0} \left( Z_f + \frac{D_b}{dD'/dz} \right). \quad (3.77)$$

where $\alpha \equiv sgn \left( \frac{dD'}{dz} \right)$ and $H$ is the Heaviside function (3.64a), which is used to enforce $N_{b,r}(Z_f) = 0$ when $\frac{dD'}{dz} > 0$, i.e., there is only a “reflected” particle flux from the side of the boundary which contains the impermeable $D = 0$ boundary.

This approach follows from the fact that the flux across any plane in one timestep must balance if the uniform particle density is to be preserved everywhere in the domain. Thus, for a boundary at $z = z_b$, the number of particles coming from $z > z_b$ (quantified by $N_{b,nr,z_b=Z_f,0}^+(Z_f = 0)$) must be balanced by the sum of three groups of particles:

a. particles initially in the region $0 < z < z_b$ which reach $z = z_b$ by positive displacement ($\Delta z > 0$).

b. particles initially in the region $0 < z < z_b$ which reach $z = z_b$ after negative displacement ($\Delta z < 0$) and reflection off the impermeable boundary at $z = 0$.

c. particles initially in the region $z > z_b$ which cross the $z = z_b$ boundary, reflect off of the impermeable boundary at $z = 0$, and again cross the $z = z_b$ boundary.

117
The first group above are the non-reflected particles and is quantified by \( N_{b,nr,z_b=Z\neq 0}(Z = 0) \), while groups b and c are the reflected particles and are jointly quantified by \( N_{b,r,z=0}(Z = 0) \). This balance of particles is expressed in (3.76) in Step 4 above. Step 5 then translates \( Z_f,0 \) (the distance from the reflecting boundary at \( z = 0 \)) into a distance relative to a boundary at \( z = z_b \) (i.e., it takes the total number of reflected particles and subtracts out those that do not reflect far enough to reach \( z = z_b \)).

The derivation of \( N_b(Z_f) \) will follow the steps outlined above. The equations for \( N_{b,nr}(Z_f) \) and \( N_{b,r}(Z_f) \) are derived in Section 3.A.2 and 3.A.3, respectively, and then combined and summarized in Section 3.A.4. The derived equations are verified in Section 3.A.5. Lastly, the derivative of \( N_b(Z_f) \) is provided in Section 3.A.6 because it is needed in the numerical implementation of the correction method in Section 3.4.2.

**3.A.2 Derivation of \( N_{b,nr}(Z_f) \)**

As explained above, \( N_{b,nr}(Z_f) \) quantifies the non-reflected particles crossing a boundary at \( z = z_b \) and reaching a distance \( Z_f \) beyond that boundary from a region in a continuous domain (as defined in Section 3.4.2) in a single timestep with uniform particle density \( \rho \). The approach taken to derive this quantity is similar to that used to derive (3.15) and (3.27) in Section 3.3.2.

We begin by considering the motion of a particle initially at some location \( z'_i \) within the region adjacent to a boundary at \( z' = 0 \) (primed variables indicate the boundary coordinate system of (3.60)). \( \frac{D}{dz} \) is constant everywhere in the region, such that \( D(z') \) is everywhere given by (3.62). If the region has \( \frac{D}{dz} > 0 \), then diffusivity increases away from the boundary and it extends to \( z' \to \infty \). If the region has \( \frac{D}{dz} < 0 \), then diffusivity decreases away from the boundary so that it must be bounded at \( z' = z'_{\text{max}} = \frac{D_b}{|D|} \) (see discussion in Sections 3.4.2 and 3.A.1); however, for the present derivation we treat this boundary to be permeable such that particles are not reflected at \( z'_{\text{max}} \) and we allow particles to travel beyond \( z'_{\text{max}} \) even though the diffusivity would by (3.62) be negative beyond this distance. Alternately, one can call this boundary “absorbing”; either way its important characteristic is that particles are not reflected back into the domain. Thus, when \( \frac{D}{dz} < 0 \) particles initially only occupy the region \( 0 < z' < z'_{\text{max}} \) meaning that there are a finite number of particles in the region
under consideration. By contrast, there are an infinite number of particles in the region if \( \frac{dD'}{dz} > 0 \) since the domain is in this case unbounded. Following the derivations of (3.15) and (3.27) in Section 3.3.2, the particle will reach the \( z' = 0 \) boundary and travel a distance \( Z_f \) beyond the boundary only if it draws a \( \xi' \) value that is less than or equal to a critical \( \xi'_c \):

\[
-z'_i - Z_f = \frac{dD'}{dz} \Delta t + \xi'_c \sqrt{2D \Delta t} \quad \Rightarrow \quad \xi'_c(z'_i) \equiv -\frac{z'_i - Z_f - \frac{dD'}{dz} \Delta t}{\sqrt{2D \Delta t}} \tag{3.78}
\]

Note that here \( D = D(z'_i) \). The probability of this occurring is given by (3.8), as in the step diffusivity case. However, unlike the step diffusivity case, \( \xi'_c(z'_i) \) is no longer guaranteed to be a negative number. If \( \frac{dD'}{dz} \) is negative and \( \left| \frac{dD'}{dz} \right| \Delta t \) is greater than \( Z_f + z'_i \), then \( \xi'_c(z'_i) \) is positive. In such cases, the drift term \( \left( \frac{dD'}{dz} \Delta t \right) \) causes the particle to sometimes reach a distance \( Z_f \) beyond the \( z' = 0 \) boundary even when the stochastic term pushes the particle away from the boundary (i.e., when \( \xi' > 0 \)).

An expression for \( N_{b,nr}(Z_f) \) is derived by the same logic used to derive (3.27) for the step diffusivity case and is thus given by (3.11) using the \( \xi'_c(z'_i) \) defined by (3.78). However, a closed-form solution for this equation could not be found.

In order to find a closed-form solution for \( N_{b,nr}(Z_f) \), the alternative conceptualization of the problem described for the step diffusivity case must be employed, i.e., the one embodied in (3.13) to (3.14) and (3.26). As before, \( N_{b,nr}(Z_f) \) is found by defining for all values of \( \xi' \) a region \( L(\xi') \) from which all particles will reach the boundary and a distance \( Z_f \) beyond it. The number of particles expected to reach from this region is given by (3.13), and thus

\[
N_{b,nr}(Z_f) = \rho \int_{-\infty}^{\infty} L(\xi') \ p(\xi') \ d\xi'. \tag{3.79}
\]

(3.79) is a generalization of (3.14) in which the upper bound of integration has been changed from 0 to \( +\infty \) since, as demonstrated below, \( L(\xi') \) is no longer necessarily equal to zero for all positive values of \( \xi' \) as in the step diffusivity case.

The challenge in this case lies in defining \( L(\xi') \), which is complicated considerably by the
diffusivity gradient. Region $L(\xi')$ is generally defined by

$$L(\xi') = \max\left[0, \min\left[\max\{z'_{c-}, z'_{c+}\}, z'_\text{max}\right] - \max\left[\min\{z'_{c-}, z'_{c+}\}, 0\right]\right] \quad (3.80)$$

where $z'_{\text{max}}$ is the physical extent of the domain (i.e., the region in which particles reside initially, see previous discussion regarding $z'_{\text{max}}$), and $z'_{c-}$ and $z'_{c+}$ are the negative and positive roots of

$$z'_c = \Delta t \left[ \frac{dD'}{dz'} \left( \xi'^2 - 1 \right) - \frac{Z_f}{\Delta t} \pm \xi' \sqrt{\left( \frac{dD'}{dz'} \right)^2 (\xi'^2 - 2) + \frac{2D_b}{\Delta t}} \right]. \quad (3.81)$$

(3.81) is derived by solving (3.78) for $z'_i$, which is then interpreted as the critical distance $z'_c$ at which particles can reach the boundary and $Z_f$ beyond it for a given value of $\xi'$. More intuitively, $N_{b,nr}$ is simply the area under the $\xi'(z'_c)$ curve, constrained by $z' = 0$ and $z'_{\text{max}}$.

As shown in Figure 3-33, there are 5 distinct cases for the shape of the $\xi'(z'_c)$ curve in a continuous domain (2 cases for $\frac{dD'}{dz'} > 0$ and 3 cases for $\frac{dD'}{dz'} < 0$). The resulting expressions for $L(\xi')$ are given in Table 3.2. In order to describe the shape of the $\xi'(z'_c)$ curve, a new variable has been introduced, $z'^r$, which is the inflection point of the curve:

$$\frac{d(\xi'^r)}{dz'_c} = 0 \quad \Rightarrow \quad z'^r = \frac{dD'}{dz'} \Delta t + Z_f - \frac{2D_b}{\frac{dD'}{dz'}}. \quad (3.82)$$

To aid in the understanding of the $L(\xi')$ expressions in Table 3.2, consider Case B (see Figure 3-33). For $\xi' < \xi'_c(0)$, the magnitude of $\xi'$ is large enough that only particles within a distance $z'_{c-}(\xi')$ can reach the boundary and $Z_f$ beyond it (i.e., all particles with $z'_i < z'_{c-}(\xi')$ can reach, and all others cannot). In the interval $\xi'_c(0) < \xi' < \xi'_c(z'^r)$, it is still the case that particles with $z'_i > z'_{c-}(\xi')$ are unable to reach a distance $Z_f$ beyond the boundary. In addition, particles with $z'_i < z'_{c+}(\xi')$ are unable to reach $Z_f$ beyond the boundary because their diffusivity (3.62) is too small to allow the stochastic term in (3.61) to overcome the “drift step” which (because $\frac{dD'}{dz'} > 0$) moves particles away from the boundary. Thus, only particles in the region $z'_{c+}(\xi') < z' < z'_{c-}(\xi')$ can reach, i.e., $L(\xi') = z'_{c-}(\xi') - z'_{c+}(\xi')$. At the inflection point $\xi' = \xi'_c(z'^r)$, $L(\xi')$ goes to zero because there are no particles for which the
stochastic term can overcome the drift term, and thus $L(\xi') = 0$ for $\xi' \geq \xi_c'(z'_{\text{max}})$. The $L(\xi')$ expressions were derived by applying this same thought experiment to all possible shapes of the $\xi_c'(z'_i)$ curve.

(3.79) was solved by first obtaining the following general solutions\(^2\)

$$
\int_b^a \frac{z'_{\text{max}}}{\sqrt{2\pi}} e^{-\xi' z^2} d\xi = \frac{z'_{\text{max}}}{2} \left[ \text{erf} \left( \frac{\xi'}{\sqrt{2}} \right) \right]_b^a \quad (3.83)
$$

$$
\int_b^a \frac{2}{\sqrt{2\pi}} e^{-\xi' z^2} d\xi = \frac{\Delta t}{2\pi} \left[ -\frac{dD'}{dz} \xi' e^{-\xi' \frac{z^2}{2}} - \sqrt{\frac{\pi}{2}} \frac{Z_f}{\Delta t} \text{erf} \left( \frac{\xi'}{\sqrt{2}} \right) \pm \left( -e^{-\xi' \frac{z^2}{2}} \sqrt{\left( \frac{dD'}{dz} \right)^2 (\xi'^2 - 2) + \frac{2D_b}{\Delta t}} \right) \right]_b^a \quad (3.84)
$$

and then constructing the solution of $N_{b,nr}$ for each case in Table 3.2. To get the solution

\(\)\(^2\)Using the MAPLE and Mathematica software packages, after some rearrangement of terms.
Figure 3-33: The 5 possible shapes of the $\xi'_c(z'_i)$ curve in the continuous domain. $\hat{z}^*$ is shown for cases B and D (dotted line); $\hat{z}_{\text{max}}$ is shown for cases C-E (dashed line).
in a more compact form, all cases were combined into one equation by means of the sign
function and the Heaviside function (3.64a). The solution to (3.79) can be expressed as:

\[
\frac{N_{b,r}(Z_f)}{\hat{\rho}} = H(-\alpha) \left[ H(\hat{D}_b - \hat{z}^*) \left\{ \hat{D}_b + \hat{Z}_f \right\} + H(z^* - \hat{D}_b) e^{(\hat{D}_b - \alpha \hat{z}_f - 1)} \right] \\
+ \alpha H(\hat{z}^*) e^{(\hat{D}_b - \alpha \hat{z}_f - 1)} - \frac{\hat{Z}_f}{2} \left[ \text{erf} \left( \frac{\xi_c(0)}{\sqrt{2}} \right) + 1 \right] \\
+ \frac{\alpha}{\sqrt{2\pi}} e^{-\frac{\xi_c(0)^2}{2}} \left\{ \left\{ \hat{\lambda}(\xi_c(0)) + \sqrt{\frac{\pi}{2}} \text{erfcx} \left( \frac{\hat{\lambda}(\xi_c(0))}{\sqrt{2}} \right) \right\} \{1 - 2H(\hat{z}^*)} \right) \\
- \xi_c(0) \right] 
\]

\hspace{1cm} (3.85)

where \( \text{erfcx}(x) \equiv e^{x^2} \text{erfc}(x) \) is the scaled error function, and where the non-dimensional
variables \( \alpha, \hat{\rho}, \hat{D}_b, \hat{z}^*, \) and \( \hat{\lambda} \) have previously been defined in Equation (3.64). The reader
should note that \( \hat{\rho} \) is a constant; it is kept on the left side of (3.85) merely to reduce clutter.

3.A.3 Derivation of \( N_{b,r}(Z_f) \)

As explained in Section 3.A.1, \( N_{b,r}(Z_f) \) quantifies the reflected particles crossing a bound-
ary at \( z = z_b \) in a single timestep and reaching a distance \( Z_f \) beyond that boundary from a
region with \( \frac{dD}{dz} < 0 \) in a continuous domain with uniform particle density \( \rho \). Inserting (3.85)
into (3.76) yields (after non-dimensionalizing variables)

\[
N_{b,r,z_b=0}(Z_{f,0}) = \hat{\rho} H \left( 1 - \hat{Z}_{f,0} \right) \left[ e^{\hat{Z}_{f,0} - 1} - \hat{Z}_{f,0} \right]. 
\]

\hspace{1cm} (3.86)

where \( \hat{\rho} \) and \( \hat{Z}_{f,0} \equiv \frac{Z_{f,0}}{|\frac{dD}{dz}| \Delta t} \) have been used instead of their dimensional counterparts to be
consistent with (3.85). The \( H \left( 1 - \hat{Z}_{f,0} \right) \) dicatates that there are no reflected particles for
\( \hat{Z}_{f,0} \geq 1 \) when the impermeable \( D = 0 \) boundary is at \( z = 0 \), which in dimensional coor-
dinates translates to no reflected particles for \( z \geq |\frac{dD}{dz}| \Delta t \). See Section 3.4.2 for additional
discussion of this result.

Using (3.86) with (3.77) yields

\[
N_{b,r}(Z_f) = \hat{\rho} H(-\alpha) H \left( 1 - \hat{D}_b - \hat{Z}_f \right) \left[ e^{\hat{D}_b + \hat{Z}_f - 1} - \hat{D}_b - \hat{Z}_f \right]. 
\]

\hspace{1cm} (3.87)
In deriving (3.87), the following identity was used:

\[ z_b = \frac{D_b}{\frac{dD}{dz}} \quad \rightarrow \quad \hat{z}_b \equiv \frac{z_b}{\frac{dD}{dz} \Delta t} = \frac{1}{\Delta t} \frac{D_b}{\frac{dD}{dz}} = \hat{D}_b \]  

(3.88)

Also, the equation is multiplied by \( H(-\alpha) \) because for any boundary, a reflected particle flux can only come from the side for which \( \frac{dD}{dz} < 0 \), i.e., diffusivity decreasing away from the boundary.

3.A.4 Summary of \( N_b(Z_f) \)

The overall equation for \( N_b(Z_f) \) can now be written by inserting (3.85) and (3.87) into (3.75), which yields:

\[
\frac{N_b(Z_f)}{\hat{\rho}} = H(-\alpha) H \left( 1 - \hat{D}_b - \hat{Z}_f \right) \left[ e^{\hat{D}_b + \hat{Z}_f - 1} - \hat{D}_b - \hat{Z}_f \right] \\
+ H(-\alpha) \left[ H(\hat{D}_b - \hat{Z}^*) \left\{ \hat{D}_b + \hat{Z}_f \right\} + H(\hat{Z}^* - \hat{D}_b) e^{(\hat{D}_b - \alpha \hat{Z}_f - 1)} \right] \\
+ \alpha H(\hat{Z}^*) e^{(\hat{D}_b - \alpha \hat{Z}_f - 1)} - \frac{\hat{Z}_f}{2} \left\{ \text{erf} \left( \frac{\xi_c(0)}{\sqrt{2}} \right) + 1 \right\} \\
+ \alpha e^{-\xi_c(0)^2 \frac{\xi_c(0)}{2 \sqrt{2} \pi}} \left\{ \left\{ \hat{\lambda}(\xi_c(0)) + \sqrt{\frac{\xi_c(0)}{2}} \text{erfcx} \left( \frac{\hat{\lambda}(\xi_c(0))}{\sqrt{2}} \right) \right\} \left\{ 1 - 2H(\hat{Z}^*) \right\} \\
- \xi_c(0) \right\}. 
\]

(3.89)

3.A.5 Verification of \( N_b(Z_f) \)

The verification of (3.89) is conducted independently for the non-reflected (3.85) and reflected (3.87) fluxes. This is necessary because the reflected particle flux is derived from consideration of the non-reflected flux, i.e., it is an artificial construct which is imposed to allow the random walk equation to satisfy the well-mixed condition in the continuous domain. Because direct application of (3.1) to the continuous domain does not satisfy the well-mixed condition, simulations cannot be conducted to directly quantify this reflected flux. Instead, the reflected flux equation (3.87) is verified indirectly by (1) demonstrating that it allows the well-mixed condition to be satisfied in the continuous domain (see Section 3.4.3.4), and
Table 3.3: $D_b$, $\frac{dD'}{dz'}$, and $\Delta t$ values used to verify (3.85) for each case.

(2) verification of (3.85), from which (3.87) was derived.

The validity of (3.85) to quantify the non-reflected particle flux is demonstrated in Figure 3-34 for each of the 5 examples used in Figure 3-33, i.e., one representative from each possible shape of the $\xi'_c(z'_i)$ curve (defined in Table 3.2). The boundary diffusivity, diffusivity gradient, and timestep values used to represent each case are listed in Table 3.3. $10^7$ particles were initialized in the domain for each case, distributed uniformly over a distance appropriate to emulate the continuous domain. The distribution of observed $Z_f$ for a single timestep is plotted against the distribution predicted by (3.85). In each case, the agreement is excellent, consistent with the results of similar tests using a wide variety of combinations of $D_b$, $\frac{dD'}{dz'}$, and $\Delta t$.

### 3.4.6 Derivative of $N_b(Z_f)$

Implementation of the correction method proposed in Section 3.4.2 requires solving Equation (3.70), which must be done iteratively. In the present study, this was accomplished using the Newton-Raphson method. This requires evaluating $\frac{\partial P(0 \leq z_f \leq Z_f)}{\partial Z_f}$ which from the definition of $P(0 \leq z_f \leq Z_f)$ in (3.66) can be written as

$$
\frac{\partial P(0 \leq z_f \leq Z_f)}{\partial Z_f} = - \frac{1}{N_b(Z_f = 0)} \frac{d[N_b(Z_f)]}{dZ_f}
$$

(3.90)
Figure 3-34: Comparison of predicted (line) and observed (points) $P_{nr}(0 < z'_f < Z_f)$ for “non-reflected” particles for each of the 5 cases presented in Table 3.2 and Figure 3-33.
since \( N_b(Z_f = 0) \), the total number of particles arriving at the boundary, is not a function of \( Z_f \). To evaluate, (3.90), the derivative of \( N_b(Z_f) \) is needed:

\[
\frac{1}{\rho} \frac{\partial N_b(Z_f)}{\partial Z_f} = H(-\alpha)H \left( 1 - \hat{D}_b - \hat{Z}_f \right) \left[ e^{\hat{D}_b + \hat{Z}_f - 1} - 1 \right] \\
+ H(-\alpha) \left[ H(\hat{D}_b - \hat{z}^*) + H(\hat{z}^* - \hat{D}_b)e^{(\hat{D}_b - \alpha \hat{Z}_f - 1)} \right] \\
- H(\hat{z}^*)e^{(\hat{D}_b - \alpha \hat{Z}_f - 1)} - \frac{1}{2} \left[ \text{erf} \left( \frac{\xi'_c(0)}{\sqrt{2}} \right) + 1 \right] \\
+ \frac{e^{-\xi'_c(0)^2}}{2\sqrt{\pi \hat{D}_b}} \left[ \alpha + \hat{Z}_f - \alpha \xi'_c(0)^2 \right] \\
+ \left\{ \alpha \xi'_c(0) \hat{\lambda}(\xi'_c(0)) - \sqrt{\pi \hat{D}_b} \text{erfcx} \left( \frac{\hat{\lambda}(\xi'_c(0))}{\sqrt{2}} \right) \right\} \left\{ 1 - 2H(\hat{z}^*) \right\}
\]

(3.91) (3.92)

where again \( \hat{\rho} \) is kept on the left side of the equation merely to reduce clutter. In any case, \( \hat{\rho} \) cancels when (3.90) is expanded, i.e., the distribution of crossing distances does not depend on \( \rho \) because the particle density is uniform.
Bibliography


Chapter 4

Emulating OGCM Tracer Transport with a Random Walk Particle Tracking Method for the Calculation of Residence Time Statistics

4.1 Introduction

Ocean tracer transport calculations are commonly performed using Ocean General Circulation Models (OGCMs), a class of models which compute ocean circulation using an Eulerian grid-based approach. For active tracers, i.e., temperature and salinity, the transport calculation is a required part of the ocean circulation calculation. For passive tracers, however, the transport calculation can be achieved by other means. In particular, Lagrangian methods are sometimes used as an alternative. These approaches make use of the circulation field computed by a parent hydrodynamic model (e.g., an OGCM) and can take many different forms depending on the application (see discussion in [44]). Two advantages offered by Lagrangian methods include a greater flexibility in representing turbulent diffusion and a natural way to visualize transport pathways of tracer mass. The present study explores how this latter advantage can be exploited to diagnose certain characteristics of the OGCM
computed tracer transport. The goal of the study is to develop a Lagrangian approach with general utility to oceanographic investigations that rely on OGCMs to understand the origin or eventual fate of tracer at a given location within the computational domain. An example is identifying two complementary quantities: the “age” of tracer mass in a given region (i.e., when did the tracer last contact the surface?) and its future residence time (i.e., when will the tracer contact the surface?).

This study is by no means the first to consider the use of a Lagrangian technique in this regard; many studies in the past have proposed and successfully applied a variety of Lagrangian approaches with this general goal in mind. What distinguishes the present study is the particular form of the Lagrangian method used, namely, a random walk particle tracking (RWPT) model that is specifically designed to mimic the internal calculations of an OGCM. The first part of the study is devoted to the derivation and implementation of such a model for use with the output of one particular OGCM, the Lawrence Livermore National Laboratory version of the Modular Ocean Model (LLNL MOM).

The second part of the study illustrates how an RWPT model can be exploited to diagnose OGCM residence time statistics by focusing on a practical application, namely, the characterization of injection sites for deep ocean carbon sequestration. As discussed more fully in IPPC [43] and Chapter 5 of this thesis, ocean sequestration is a proposed climate change mitigation strategy in which CO₂ from large point sources is discharged to the ocean at a suitable location such that it is “sequestered” from the atmosphere for long periods of time. In order to be an effective mitigation strategy, a significant fraction of the injected CO₂ should remain in the deep ocean for periods of centuries to millenia. Although effectiveness can be enhanced by modifications to the discharge method (e.g., [1, 72]), site selection is a central consideration to maximizing sequestration times. Specifically, the injection site should have a long residence time, or equivalently a high sequestration efficiency. The estimation of sequestration efficiencies is a specific example of the more general residence time investigations described above, and provides a good illustration of the benefits that can be realized from a Lagrangian framework. Indeed, it was the original motivation for the present study.

A logical approach to investigating sequestration efficiency is to use OGCMs to simulate
injections, since these workhorses of oceanographic investigation are calibrated to reproduce
the large scale circulation of the oceans. They are well suited to simulating the evolution of
a discrete source, and grid resolution can be refined on a regional scale as needed to improve
the quality of the circulation calculation (e.g., [23]). The drawback to these calculations is
the computational expense of longterm simulations, coupled with the fact that a separate
simulation is required for each site to be evaluated (or equivalently, multiple tracers within a
single simulation). This makes the task of characterizing many sites arduous. For example, in
a comparative analysis of sequestration efficiency involving eight OGCMs, the Global Ocean
Storage of Anthropogenic Carbon (GOSAC) study [42] considered only seven injection sites
at three different depths. The investigation was conducted by comparing simulations of
longterm point source releases of CO₂. The inefficiency of performing calculations in this
manner was addressed in a novel manner by Hill et al. [39], who applied an adjoint sensitivity
method to an OGCM to characterize, in a single simulation, the sequestration efficiencies
that would be predicted by the OGCM for the entire domain. Within the ocean sequestration
context, the present work investigates the extent to which an RWPT approach can be used
to achieve this same goal, by exploiting the economy-of-scale that can be achieved by using
each simulated particle to generate residence time statistics for many different sites.

In summary, the present study has two specific goals: (1) to develop the particle tracking
equivalent to the tracer calculations of a specific OGCM, and (2) to explore how particle
statistics can be exploited to compute sequestration efficiencies in an efficient manner. These
specific goals can be thought of as initial steps toward the more general goal of creating an
advanced diagnostic of OGCM tracer transport suitable for general source water and resi-
dence time calculations that would be difficult to conduct with an OGCM alone. Section
4.2 provides an overview of the tracer transport calculations of the LLNL MOM. Section
4.3 provides an overview of Lagrangian techniques in an oceanographic context, proposes
the RWPT equivalent of the LLNL MOM tracer calculations, and describes its present im-
plementation. Direct validation results for the RWPT scheme are provided in Section 4.4.
The ocean sequestration application is presented in Section 4.5, which serves as a valida-
tion of both the RWPT model predictions and the book-keeping method used to diagnose
sequestration efficiency. Concluding thoughts are provided in Section 4.6 along with goals
for future research efforts. While the present RWPT implementation yields generally good agreement with LLNL MOM calculations, some further development is ultimately recommended to resolve remaining inaccuracies. These recommendations are provided in Appendix 4.A, which provides a roadmap for ongoing and future improvement of the RWPT implementation. Overall, the methods presented herein are thought to hold promise for a range of oceanographic applications, and continued investigation is recommended.
4.2 Overview of LLNL MOM

The Lawrence Livermore National Laboratory (LLNL) has customized the Geophysical Fluid Dynamics Laboratory (GFDL) Modular Ocean Model (MOM version 1.1, see Pacanowski and Griffies [60]), a finite difference, z-level model which solves the hydrostatic primitive equations, and which is one of the most widely used ocean models [27, 73]. As described in Duffy and Caldeira [26], Duffy et al. [27], Caldeira and Duffy [11], and Wickett et al. [73], notable enhancements include the addition of a sea ice model [57], an explicit free surface approach [48], the Gent-McWilliams eddy-induced transport [33], and a revised vertical mixing scheme [27]. The configuration used for the present study was found to perform well relative to other OGCMs in the GOSAC study, based on its ability to reproduce observed radiocarbon and CFC-11 distributions in the ocean.

Since the goal of the present study is to create a particle model which can emulate the tracer transport of this OGCM, an overview is given below of some of the major characteristics of the LLNL MOM transport calculation, with particular emphasis on those aspects relevant to the particle model development.

4.2.1 LLNL MOM tracer equation

The transport of a passive tracer concentration is given by the advection-diffusion equation:

$$\frac{\partial c}{\partial t} + \nabla \cdot (uc) = -\nabla \cdot F$$

(4.1)

where \(u\) is the current velocity and \(F\) is the tracer flux vector:

$$F = -E \cdot \nabla c$$

(4.2)

Here \(E\) is the diffusivity tensor. The LLNL MOM uses planetary spherical coordinates, such that

$$\nabla \equiv \frac{\hat{\lambda}}{R \cos \phi} \frac{\partial}{\partial \lambda} + \frac{\hat{\phi}}{R} \frac{\partial}{\partial \phi} + \hat{z} \frac{\partial}{\partial z}$$

(4.3)
where $\lambda$ is longitude (positive eastward from the prime meridian), $\phi$ is the latitude\(^1\) (positive northward from the equator), $z$ is the distance above the bottom of the ocean\(^2\), and $R_e = 6,371$ km is the radius of the Earth.

### 4.2.1.1 Isopycnal diffusion

Appendix B of Pacanowski and Griffies [60] provides a thorough discussion of isopycnal diffusion and its adaptation to the MOM, as reviewed below. Following this work, the derivations that follow are in Cartesian coordinates, i.e., a fixed reference frame tangent to the local geopotential. The resulting diffusion tensor is compatible with (4.3).

Diffusion in the ocean has been observed to occur mainly along, rather than across, isopycnal surfaces (sometimes referred to as isoneutral surfaces) (e.g., [51]). Thus, rather than describe tracer spreading in the horizontal and vertical directions, it is more natural to describe spreading in the isopycnal and diapycnal directions. Within an orthonormal isoneutral coordinate system, the diffusivity tensor is symmetric:

$$
\mathbf{E} = \begin{pmatrix}
A_I & 0 & 0 \\
0 & A_I & 0 \\
0 & 0 & A_D
\end{pmatrix}
$$

where $A_I$ and $A_D$ are the isopyncal and diapycnal diffusion coefficients, respectively (note that oceanic transport is highly anisotropic, with $\frac{A_D}{A_I} \sim 10^{-7}$).

#### 4.2.1.1.1 Isoneutral reference frame

A Cartesian reference frame can be defined by its unit basis vectors $e = (\hat{x}, \hat{y}, \hat{z})$. The corresponding basis vectors for the orthonormal isopycnal (or isoneutral) frame are

$$
\tilde{e} = (\tilde{e}_1, \tilde{e}_2, \tilde{e}_3) = \left( \frac{\hat{z} \times \nabla \rho}{|\hat{z} \times \nabla \rho|}, \frac{\hat{e}_3 \times \hat{e}_1}{|\hat{e}_3 \times \hat{e}_1|}, \frac{\nabla \rho}{|\nabla \rho|} \right)
$$

\(^{1}\)Note that $\phi$ is latitude as opposed to colatitude.

\(^{2}\)The radial coordinate $r = R_e + z$ does not appear in (4.3) because $\frac{\partial}{\partial r} = \frac{\partial}{\partial z}$, and the maximum ocean depth ($z_{max}$) is much smaller than $R_e$, so that $R_e + z \approx R_e$. 

136
where $\nabla = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$ is the Cartesian gradient operator. The physical interpretation of (4.5) is that $\dot{e}_3$ is parallel to the density gradient making it the dianeutral direction, while $\dot{e}_1$ and $\dot{e}_2$ are orthogonal to $\dot{e}_3$ and each other, making them the isoneutral directions. Note that since isopycnal diffusion is considered isotropic, there are an infinite number of choices for the isoneutral directions; the only constraint is that they must be orthogonal to the dianeutral direction and to each other.

It is useful to define the isopycnal slope vector as:

$$S \equiv -\nabla_p \rho$$

(4.6)

where $\rho$ is the potential density (locally referenced) and $\nabla_p = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)$ is the horizontal gradient operator. The magnitude of the slope is $S = \sqrt{S_x^2 + S_y^2}$.

### 4.2.1.1.2 Diffusion tensor in a geopotential reference frame

As derived in Griffies and Pacanowski [60], rotating the diffusion tensor from an isoneutral reference frame to a fixed geopotential reference frame gives the following diffusion tensor:

$$E = \frac{A_I}{1 + S^2} \left( \begin{array}{ccc} 1 + S_y^2 + \frac{\Delta \rho}{\Delta_I} S_x^2 & \left( \frac{\Delta \rho}{\Delta_I} - 1 \right) S_y & \left( 1 - \frac{\Delta \rho}{\Delta_I} \right) S_x \\ \left( \frac{\Delta \rho}{\Delta_I} - 1 \right) S_x S_y & 1 + S_x^2 + \frac{\Delta \rho}{\Delta_I} S_y^2 & \left( 1 - \frac{\Delta \rho}{\Delta_I} \right) S_y \\ \left( 1 - \frac{\Delta \rho}{\Delta_I} \right) S_x & \left( 1 - \frac{\Delta \rho}{\Delta_I} \right) S_y & \frac{\Delta \rho}{\Delta_I} + S^2 \end{array} \right).$$

(4.7)

Equation (4.7) can be simplified considerably by making a small angle approximation (which is appropriate for the ocean which has a characteristic aspect ratio of $10^{-4}$). The slope $S \ll 1$ follows from the usual definition of $S = \tan(\theta) \approx \theta$ for small $\theta$. The approximated diffusion tensor is

$$E = \left( \begin{array}{ccc} A_I & 0 & (A_I - A_D) S_x \\ 0 & A_I & (A_I - A_D) S_y \\ (A_I - A_D) S_x & (A_I - A_D) S_y & A_D + A_I S^2 \end{array} \right).$$

(4.8)
It should be noted that the small angle approximation in (4.8) breaks down in regions of vertical convection, where the isopycnal slope may approach infinity. Isopycnal diffusion becomes unrealistic in these regions, and as a result the LLNL MOM “tapers” the isopycnal diffusion coefficient when the slope becomes too steep, effectively turning off isopycnal mixing in these regions (see Section 4.2.2.2.2).

Although (4.8) is the form used by the LLNL MOM in theory, in practice $\frac{A_D}{A_I} \ll 1$, and so the diffusion tensor is further simplified to

$$E = \begin{pmatrix} A_I & 0 & A_IS_x \\ 0 & A_I & A_IS_y \\ A_IS_x & A_IS_y & A_I S^2 + A_D \end{pmatrix}.$$  

(4.9)

Lastly, as demonstrated in Pacanowski and Griffies [60], the mild isopycnal slopes also imply that the difference between diapycnal and vertical diffusivity is negligible, i.e., the diffusivity tensor can be written as

$$E = \begin{pmatrix} A_I & 0 & A_IS_x \\ 0 & A_I & A_IS_y \\ A_IS_x & A_IS_y & A_I S^2 + K_v \end{pmatrix}.$$  

(4.10)

where $K_v$ is the vertical diffusivity. See Appendix B of Pacanowski and Griffies [60] for a complete derivation.

### 4.2.1.2 Gent-McWilliams eddy-induced transport

Gent and McWilliams (GM) [33] suggested an improvement to coarse resolution OGCMs to better mimic the physics observed in OGCM configurations that resolve mesoscale eddies. In particular, the stirring effect of mesoscale eddies is incorporated, which gives rise to a flux that is normal to the tracer gradient [36]. It applies to active as well as passive tracers, and has the tendency to reduce the available potential energy stored in sloping isopycnals, flattening them over time [36, 37, 66]. It has become standard in most OGCMs, e.g., all eight OGCMs in the GOSAC study implemented this process in some manner [42].
The GM parameterization can be implemented in one of two mathematically equivalent manners. The original formulation by GM [33] was to introduce an additional advective velocity such that the tracer transport equation becomes

\begin{equation}
\frac{\partial c}{\partial t} + \nabla \cdot [(u + u_{GM}) c] = -\nabla \cdot F \tag{4.11}
\end{equation}

where

\begin{equation}
u_{GM} = \left(-\frac{\partial (\kappa S_x)}{\partial z}, -\frac{\partial (\kappa S_y)}{\partial z}, \frac{\partial (\kappa S_x)}{\partial x} + \frac{\partial (\kappa S_y)}{\partial y}\right) \tag{4.12}
\end{equation}

is the GM velocity, which should be non-divergent and zero on all boundaries.

The other way to implement the GM transport is to modify the diffusion tensor to include a skew-diffusive flux [36], changing (4.10) into

\begin{equation}
E = \begin{pmatrix}
A_I & 0 & (A_I - \kappa) S_x \\
0 & A_I & (A_I - \kappa) S_y \\
(A_I + \kappa) S_x & (A_I + \kappa) S_y & A_I S^2 + K_v
\end{pmatrix} \tag{4.13}
\end{equation}

Here \( \kappa \) is the isopycnal layer “thickness diffusivity” of the GM parameterization. Note that, like all skew diffusive fluxes, the GM flux is strictly (4.13) is sometimes referred to as the Redi-GM tensor, combining Redi’s [64] isopycnal diffusion tensor with the GM flux. This tensor is used in LLNL MOM, i.e., the advective GM formulation is not used. As discussed in Griffies [36], the diffusive formulation has two main advantages over the advective formulation. First, it does not suffer from as much noise because spatial gradients in isopcnal slope/diffusivity do not need to be computed. Second, under the common assumption of \( \kappa = A_I \), the tensor simplifies and actually reduces OGCM simulation time.

### 4.2.2 LLNL MOM tracer calculation

The following section provides an outline of the LLNL MOM tracer calculation, with emphasis on how the quantities in (4.38) are defined internally, since such details are important when creating the Lagrangian framework. A brief discussion of OGCM numerics is also provided, but only to the extent that it is relevant to the present endeavor. The reader is
referred to the MOM manual for additional detail [60].

4.2.2.1 LLNL MOM grid

The OGCM uses an "Arakawa B-grid", in which velocities are calculated on U-cells and scalar variables (e.g., tracer concentrations) are calculated on T-cells. T-cells and U-cells are staggered horizontally, but not vertically. Since tracer calculations are performed on the T-cells, tracer fluxes are computed on T-cell faces. Figures 4-1, 4-2, and 4-3 show the grid structure from different angles. Because the Lagrangian model will attempt to mimic OGCM tracer calculations, the text that follows defines variables with respect to their location on the T-grid unless otherwise noted.

Each grid cell is referenced by an \( \{i,j,k\} \) index, where \( i \) is the zonal index (\( i = 1 \) is the easternmost grid cell, which for a global domain begins at the prime meridian), \( j \) is the meridional index (\( j = 1 \) is the southernmost latitude band), and \( k \) is the vertical index (\( k = 1 \) is the surface layer). In the text that follows, the location of a T-cell gridpoint (i.e., the center of a T-cell) will be referred to as \( \{\lambda_{i,j,k}, \phi_{i,j,k}, z_{i,j,k}\} \). In addition, the western and eastern cell faces will be referred to by indices \( i - \frac{1}{2} \) and \( i + \frac{1}{2} \); the southern and northern cell faces by \( j - \frac{1}{2} \) and \( j + \frac{1}{2} \); and the top and bottom cell faces by \( k - \frac{1}{2} \) and \( k + \frac{1}{2} \), respectively.

Grid spacing in LLNL MOM can be variable in all three spatial directions. However, in the coarse-gridded configuration used in the present study, the horizontal spacing is constant (in the angular sense) with \( \Delta \lambda = 4^\circ \) and \( \Delta \phi = 2^\circ \). The vertical dimension is discretized into 24 layers with thickness increasing exponentially with depth. Overall, the LLNL MOM configuration used herein has a T-grid with 194,400 gridcells (\( N_x = N_y = 90; N_z = 24 \)), of which roughly half are water (100,169 water cells).

Because the grid is defined by constant longitude and latitude intervals, the zonal width of grid cells is a function of latitude and the meridional width is constant.

\[ \Delta x_j = \Delta x_{equator} \cos \phi = R_e \Delta \lambda \cos \phi \]  

\[ (4.14a) \]

\(^3\)The following description does not use the same variable names and conventions as the MOM documentation, although they are similar. Naming conventions were chosen to be as convenient as possible for the present purpose.
Figure 4-1: LLNL MOM computational grid. Current velocities are calculated at grid points at the center of the *U-grid* (open circles and dashed lines) and tracer concentrations are calculated at grid points at the center of the *T-grid* (filled circles and heavy gridlines). The grids are aligned vertically but staggered horizontally such that the *U-grid* points coincide with the *T-grid* vertices, and vice versa.

\[ \Delta y = R_e \Delta \phi \]  

(4.14b)

where \( \Delta \lambda \) and \( \Delta \phi \) have units of radians and \( R_e = 6,370 \) km is the radius of the Earth. Thus, the area of the southern and northern faces of a grid cell are not equal, having widths \( \Delta x_{j-1/2} = R_e \Delta \lambda \cos \phi_j \) and \( \Delta x_{j+1/2} = R_e \Delta \lambda \cos \phi_j \), respectively.

The vertical extents of the 24 layers are given in Table 4.1, with a maximum depth\(^4\) of 5 km. For convenience, two variables to describe vertical discretization are introduced (Figure 4-3): \( \Delta z_k^t \) is the height of the T-grid cell; \( \Delta z_k^c \) is the distance between the centers of layer \( k-1 \) and layer \( k \).

---

\(^4\)The LLNL MOM configuration uses a smoothed topography, which means abyssal regions with depth > 5 km are not represented.
Figure 4-2: LLNL MOM computational grid (plan view). The *U-grid* (open circles and dashed lines) and *T-grid* (filled circles and heavy gridlines) are staggered horizontally. Zonal and meridional grid spacing are constant when expressed in degrees.

Figure 4-3: LLNL MOM computational grid (profile view). The *U-grid* (open circles and dashed lines) and *T-grid* (filled circles and heavy gridlines) are aligned vertically. Zonal grid spacing is constant when expressed in degrees. Vertical grid spacing is variable.
Table 4.1: LLNL MOM vertical discretization used in present study (rounded to the nearest m).

<table>
<thead>
<tr>
<th>k index</th>
<th>Grid cell bottom (m)</th>
<th>k index</th>
<th>Grid cell bottom (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25</td>
<td>13</td>
<td>1,258</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>14</td>
<td>1,501</td>
</tr>
<tr>
<td>3</td>
<td>79</td>
<td>15</td>
<td>1,769</td>
</tr>
<tr>
<td>4</td>
<td>117</td>
<td>16</td>
<td>2,060</td>
</tr>
<tr>
<td>5</td>
<td>166</td>
<td>17</td>
<td>2,374</td>
</tr>
<tr>
<td>6</td>
<td>229</td>
<td>18</td>
<td>2,708</td>
</tr>
<tr>
<td>7</td>
<td>310</td>
<td>19</td>
<td>3,060</td>
</tr>
<tr>
<td>8</td>
<td>410</td>
<td>20</td>
<td>3,429</td>
</tr>
<tr>
<td>9</td>
<td>532</td>
<td>21</td>
<td>3,811</td>
</tr>
<tr>
<td>10</td>
<td>677</td>
<td>22</td>
<td>4,202</td>
</tr>
<tr>
<td>11</td>
<td>846</td>
<td>23</td>
<td>4,600</td>
</tr>
<tr>
<td>12</td>
<td>1,039</td>
<td>24</td>
<td>5,000</td>
</tr>
</tbody>
</table>

4.2.2.2 Variable definitions

Only the variables necessary to describe tracer transport on the LLNL MOM grid are included below; see Pacanowski and Griffies [60] for further detail.

4.2.2.2.1 Advevtive flux Horizontal current velocities \{u,v\} are computed at the center of "U-cells" and then averaged onto T-cell faces (see Figure 4-2):

\[
\bar{u}_{i,j,k} = \frac{1}{2} (u_{i,j,k} + u_{i,j-1,k}) \tag{4.15a}
\]

\[
\bar{v}_{i,j,k} = \frac{1}{2} (v_{i,j,k} + v_{i-1,j,k}) \tag{4.15b}
\]

for the LLNL MOM configuration used in the present study where the horizontal resolution is constant. The LLNL MOM assumes an incompressible fluid, and thus the vertical velocity is calculated at the bottom of T-cell faces from the divergence of the cell-face average horizontal velocities by integrating the continuity equation.

\[
\nabla \cdot \mathbf{u} = 0 \quad \Rightarrow \quad w_{i,j,k} = \int_0^{h_{i,j,k}} \frac{1}{a \cos \phi} \left( \frac{\partial u}{\partial \lambda} + \frac{\partial (vcos\phi)}{\partial \phi} \right) dz
\tag{4.16}
\]

143
The integration is performed from the bottom to the top of the domain with a no flux boundary condition \( w = 0 \) at the ocean bottom. Note that the LLNL MOM configuration does not use the rigid lid approximation, and thus there is a free surface \( h_{i,j,k} \) that varies in time and vertical velocity at the surface may be non-zero.

4.2.2.2 Diffusive flux  In order to compute the diffusive fluxes, each term in the small angle approximation of the diffusion tensor (4.13) must be calculated. All terms except \( K_v \) are due to isopycnal diffusion, so we begin with these.

Isopycnal diffusion coefficient  The isopycnal diffusivity is calculated as a base value \( (A_{I,0}) \) multiplied by a “taper” coefficient which is based on the local isopycnal slope. The purpose of tapering the isopycnal diffusion coefficient is to avoid unrealistic vertical fluxes in regions where the slope is too steep and isopycnal diffusion is no longer thought to be realistic, e.g., in regions of convective instability. The particular tapering scheme implemented in MOM is:

\[
A_I(S) = \frac{A_{I,0}}{2} \left[ 1 - \tanh \left( \frac{S - \delta_{dm}}{S_{dm}} \right) \right] H \left( \frac{\partial \rho}{\partial z} \right)
\]  

(4.17)

where \( S = \sqrt{S_x^2 + S_y^2} \), \( \delta_{dm} \) and \( S_{dm} \) are constants, and \( H \) is the Heaviside function:

\[
H(\alpha) \equiv \begin{cases} 
0 & \alpha < 0 \\
\frac{1}{2} & \alpha = 0 \\
1 & \alpha > 0 
\end{cases}
\]  

(4.18)

In the current LLNL MOM configuration, \( A_{I,0} = 10^7 \text{ cm}^2/\text{s} \), \( \delta_{dm}=0.004 \), and \( S_{dm}=0.001 \). Figure 4-4 shows \( A_I \) as a function of isopycnal slope using these values. Because the function quickly transitions from \( A_{I,0} \) to zero when the slopes become problematic (dropping from 0.99\(A_{I,0}\) to 0.01\(A_{I,0}\) in the interval \( 0.0017 \lesssim S \lesssim 0.0063 \)), (4.17) is always used to calculate \( A_I \), even when isopycnal slopes are not steep. The value of \( A_I \) is defined on each of the cell faces after the isopycnal slopes have been calculated. Thus, \( A_I \) is spatially variable in regions of the ocean where isopycnal slopes are steep.

The value of the Gent-McWilliams ”thickness diffusivity” \( \kappa_{GM} \) is treated in the same manner as \( A_I \), and is traditionally assumed to be equal to \( A_I \) (see Section 4.2.1.2). Al-
though recent research suggests that this is not physically realistic ([30], S. Griffies personal communication), this assumption is made in the current configuration of LLNL MOM.

**Isopycnal slope**  The values of the isopycnal slope in $x$ ($S_x$) and $y$ ($S_y$) are computed at each cell face from (4.6) using the local density gradients, which are calculated from potential density values at neighboring grid points, as described below. Once the slope has been calculated, $A_I$ is calculated using (4.17).

**Zonal direction ($i + 1/2$ cell faces)**

For computing the first row of the diffusion tensor, i.e., for zonal tracer fluxes ($x$-direction), density gradients are calculated at the center of the eastern cell faces of grid cell ${i, j, k}$ in the following manner:

\[
\left( \frac{\delta \rho}{\delta x} \right)_x = \mathcal{L}_{i,j,k} \mathcal{L}_{i+1,j,k} \frac{\rho_{i+1,k,j} - \rho_{i,j,k}}{\Delta x_{j+1/2}}
\]  

\[
\left( \frac{\delta \rho}{\delta y} \right)_x = \mathcal{L}_{i,j-1,k} \mathcal{L}_{i,j+1,k} \mathcal{L}_{i+1,j-1,k} \mathcal{L}_{i+1,j+1,k} \frac{\rho_{i+1,j+1,k} - \rho_{i,j-1,k} + \rho_{i+1,j+1,k} - \rho_{i+1,j-1,k}}{4\Delta y}
\]  

Figure 4-4: Isopycnal diffusivity ($A_I$) as a function of isopycnal slope in LLNL MOM ($S = \sqrt{S_x^2 + S_y^2}$).
\[
\left( \frac{\delta \rho}{\delta z} \right)_x = \frac{\rho_{i,j,k-1} - \rho_{i,j,k+1} + \rho_{i+1,j,k-1} - \rho_{i+1,j,k+1}}{4 \Delta z^t_k}, \tag{4.19c}
\]

where \( L(i, j, k) \) is a land mask (0 for land, 1 for water), and \( \Delta x_{j+1/2} \) is the x-width of the northern face of cells in row \( j \). (4.19c) is actually only valid for interior ocean points; density values are extrapolated to calculate density gradients for the top and bottom cells.

**Meridional direction (\( j + 1/2 \) cell faces)**

For computing the second row of the diffusion tensor, i.e., for meridional tracer fluxes \((y\)-direction), density gradients are calculated at the center of the northern cell faces of grid cell \( \{i, j, k\} \) in the following manner:

\[
\left( \frac{\delta \rho}{\delta x} \right)_y = \mathcal{L}_{i+1,j+1,k} \mathcal{L}_{i-1,j+1,k} \mathcal{L}_{i+1,j,k} \mathcal{L}_{i-1,j,k} \frac{\rho_{i+1,j+1,k} - \rho_{i-1,j+1,k} + \rho_{i+1,j,k} - \rho_{i-1,j,k}}{4 \Delta x_j} \tag{4.20a}
\]

\[
\left( \frac{\delta \rho}{\delta y} \right)_y = \mathcal{L}_{i,j,k} \mathcal{L}_{i+1,j,k} \frac{\rho_{i,j+1,k} - \rho_{i,j,k}}{\Delta y} \tag{4.20b}
\]

\[
\left( \frac{\delta \rho}{\delta z} \right)_y = \frac{\rho_{i,j,k-1} - \rho_{i,j,k+1} + \rho_{i+1,j,k-1} - \rho_{i+1,j,k+1}}{4 \Delta z^t_k}. \tag{4.20c}
\]

Note that as in the x-direction, (4.20c) is only valid for interior ocean points, and density gradients for the top and bottom cells are calculated by extrapolation.

**Vertical direction (\( k - 1/2 \) cell faces)**

For computing the third row of the diffusion tensor, i.e., for vertical tracer fluxes \((z\)-direction), density gradients are calculated at the center of the top cell face of grid cell \( \{i, j, k\} \) in the following manner:

\[
\left( \frac{\delta \rho}{\delta x} \right)_z = \frac{1}{4 \Delta x_j} \left( \mathcal{L}_{i-1,j,k-1} \mathcal{L}_{i,j,k-1} \left[ \rho_{i,j,k-1} - \rho_{i-1,j,k-1} \right] \right.
\]

\[
+ \mathcal{L}_{i,j,k-1} \mathcal{L}_{i+1,j,k-1} \left[ \rho_{i+1,j,k-1} - \rho_{i,j,k-1} \right]
\]

\[
+ \mathcal{L}_{i-1,j,k} \mathcal{L}_{i,j,k} \left[ \rho_{i,j,k} - \rho_{i-1,j,k} \right]
\]

\[
+ \mathcal{L}_{i,j,k} \mathcal{L}_{i+1,j,k} \left[ \rho_{i+1,j,k} - \rho_{i,j,k} \right] \tag{4.21a}
\]

146
\[
\left(\frac{\delta \rho}{\delta y}\right)_z = \frac{1}{4\Delta y} \left(\mathcal{L}_{i,j-1,k-1} \mathcal{L}_{i,j,k-1} \left[\rho_{i,j,k-1} - \rho_{i,j-1,k-1}\right]
+ \mathcal{L}_{i,j,k-1} \mathcal{L}_{i,j+1,k-1} \left[\rho_{i,j+1,k-1} - \rho_{i,j,k-1}\right]
+ \mathcal{L}_{i,j-1,k} \mathcal{L}_{i,j,k} \left[\rho_{i,j,k} - \rho_{i,j-1,k}\right]
+ \mathcal{L}_{i,j,k} \mathcal{L}_{i+1,j,k} \left[\rho_{i+1,j,k} - \rho_{i,j,k}\right]\right) (4.21b)
\]

\[
\left(\frac{\delta \rho}{\delta z}\right)_z = \mathcal{L}_{i,j,k-1} \mathcal{L}_{i,j,k} \frac{\rho_{i,j,k-1} - \rho_{i,j,k}}{\Delta z_{k-1}^c}. (4.21c)
\]

**Vertical diffusivity**  The vertical diffusivity, \(K_v\), is computed at the bottom of each cell face from local density gradients and surface fluxes. This term combines mixing due to various processes, as described below. Looking ahead to Section 4.3, the most important characteristics of \(K_v\) are that it can be large and have a high degree of spatial variability.

The vertical diffusivity in the present configuration of LLNL MOM consists of three parts: sub-grid scale mixing in the ocean interior, resolved shear instability in the ocean interior, and mixing due to surface fluxes and convective instability within the Oceanic Boundary Layer (OBL). A brief description of each is provided below.

- **Sub-grid scale vertical diffusivity** in the interior of the ocean is calculated using the scheme described in Duffy *et al.* [26], where the diffusivity is inversely proportional to the local Brunt-Väisälä frequency \((N)\):

\[
K_{v_{sg}} = \frac{a}{N} = \frac{a}{\sqrt{-g \frac{\partial \rho}{\partial z}}}. (4.22)
\]

Here \(a\) is a constant, \(\rho\) is the local density, and \(\rho_0\) is a reference density. In the present configuration of LLNL MOM, \(a = 10^{-3}\) cm\(^2\)/s and \(K_{v_{sg}}\) is constrained to have a minimum background value of 0.2 cm\(^2\)/s and a maximum value of \(10^3\) cm\(^2\)/s.

- **Mixing due to resolved shear instability** is calculated from a local gradient Richardson number following the calculation described in Large *et al.* [50]:

\[
Ri_g = \frac{N^2}{\left(\frac{\partial u}{\partial z} + \frac{\partial v}{\partial z}\right)^2} (4.23a)
\]
\[ K_{v_{si}} = \begin{cases} 
K_{v_{si},0} & Ri < 0 \\
K_{v_{si},0} \left(1 - \left(\frac{Ri}{Ri_0}\right)^3\right) & 0 < Ri_g < Ri_0 \\
0 & Ri_g > Ri_0 
\end{cases} \]  

(4.23b)

where \( Ri_0 \) is a critical Richardson number for shear instability and \( K_{v_{si},0} \) is the maximum vertical mixing due to shear instability. LLNL MOM uses \( Ri_0 = 0.7 \) and \( K_{v_{si},0} = 50 \text{ cm}^2/\text{s} \), as recommended by Large et al. [50]. The total interior mixing, \( K_v = K_{v_{si}} + K_{v_{sg}} \), is constrained to have a maximum value of 1,000 cm\(^2\)/s.

- The mixing in the OBL follows the nonlocal K profile parameterization (KPP) of Large et al. [50], in which the vertical flux of tracer is given by:

\[ \overline{wx}(d) = -K_{v_{OBL}} \left(\frac{\partial X}{\partial z} - \gamma_x\right) \]  

(4.24)

where \( X \) is the mean tracer concentration, \( \overline{wx}(d) \) is the time-averaged kinematic turbulent tracer flux, \( d = -z \) is the depth, and \( K_{v_{OBL}} \) and \( \gamma_x \) are the vertically variable diffusivity and non-local transport term for tracer within the OBL. The method was adapted from an approach commonly used to model the atmospheric boundary layer (ABL), and has two characteristics which are noteworthy within the present discussion:

1. The diffusivity \( K_{v_{OBL}} \) is depth dependent and calculated as:

\[ K_{v_{OBL}}(\sigma) = hw_x(\sigma)G(\sigma) \]  

(4.25)

where \( h \) is the OBL thickness, \( \sigma \equiv \frac{d}{h} \) is a fractional depth within the OBL, \( w_x(\sigma) \) is a depth dependent turbulent velocity scale for tracers and \( G(\sigma) \) is a shape function. The linear dependence of \( K_{v_{OBL}} \) on \( h \) is imposed to allow deeper OBLs (with larger and more efficient eddies) to induce more mixing. The thickness of the OBL is calculated iteratively such that a bulk Richardson number equals a critical value at the base of the OBL, effectively ending the influence of OBL eddies at the depth at which they become stable relative to the local buoyancy and velocity. \( w_x(\sigma) \) is scaled off of the surface friction velocity \( u^* \), and increases,
decreases, or remains constant with depth depending on whether unstable, stable, or neutrally stable conditions prevail. $G(\sigma)$ is a cubic polynomial which is fitted to satisfy $G(\sigma = 0) = 0$ (i.e., turbulent eddies cannot cross the surface) and to force the value and gradient of $K_{vOBL}$ at the base of the OBL ($\sigma = 1$) to match the value and gradient of $K_v$ at this depth, i.e., the interior ocean vertical diffusivity computed using (4.22) and (4.23b). This ensures that the vertical tracer flux is continuous across the interface between the OBL and interior ocean. There is no hard-wired upper bound on the OBL diffusivity as in the interior ocean, and values routinely exceed 1,000 cm$^2$/s, reaching as high as $10^6$ cm$^2$/s in regions of severe convective instability, with sharp spatial variations in the overall vertical diffusivity $K_v$.

2. There is a non-local aspect to the tracer transport during unstable or convective surface tracer forcing, i.e., only when a destabilizing density perturbation occurs at the ocean surface. The process is deemed non-local because it is unrelated to the tracer concentration gradient within the OBL. It is imposed via the parameter $\gamma$, which scales off of the surface tracer flux.

### 4.2.2.2.3 LLNL MOM output variables

The goal of the present study is to create a Lagrangian approach that can be implemented with readily available output data from the LLNL MOM. Figures 4-5 and 4-6 summarize the basic output variables that appear in the random walk equations (4.42).

The advective output variables are $u$, $v$, and $w$ velocities, defined at grid cell corners as described previously. From this output the cell-face average velocities, $\overline{u}$, $\overline{v}$, and $\overline{w}$ are calculated in the manner described in Section 4.2.2.2.1.

The diffusive output variables include $K_v$, the vertical diffusivity and $S_x$ and $S_y$, the isopycnal slopes in the $x$ and $y$ directions, respectively. Values of $S_x$ and $S_y$ are output at $x$ and $y$ cell faces, respectively, and $S_x$ and $S_y$ are also both output at the vertical cell faces. The horizontal cell-face values are used by the OGCM to compute the horizontal component of the isopycnal tracer flux, whereas the vertical cell-face values are used to compute the vertical component of isopycnal tracer flux (i.e., the bottom row of the tensor in (4.13)).
Figure 4-5: Advection variables available from standard LLNL MOM output.

Figure 4-6: Diffusion variables available from standard LLNL MOM output.
However, these basic diffusion variables are insufficient to conduct the random walk calculations, and some significant output modifications have been implemented to accommodate the particle simulations:

- The isopycnal slopes which the model outputs are tapered, i.e., they have been multiplied by the taper coefficient given in (4.17). This is inconsistent with the internal calculations of the OGCM, in which $A_I$ is instead the tapered quantity. Thus, the OGCM has been modified to output untapered slope values.

- The OGCM does not by default output the taper coefficients (or isopycnal diffusivity), so the code has been modified to do so.

- The vertical diffusivity $K_v$ output by the model is actually the quantity $K_v + A_I S^2$. While these parameters can be lumped together in the Eulerian OGCM, they are treated very differently in the random walk equivalent. As shown in (4.42), the $K_v$ term is stochastic, where as the $A_I S^2$ term is deterministic. Thus, the OGCM has been modified to also output the internally computed $A_I S^2$ at the vertical cell faces.

- The OGCM has also been modified to output density gradients in all three directions at each of the three cell faces. As discussed in the next section, these quantities are potentially useful because they allow the isopycnal slopes $S_x$ and $S_y$ to be computed externally to the OGCM.

4.2.2.3 Numerical scheme and additional configuration

The discussion that follows is restricted to some major features of the LLNL MOM numerics that are deemed most relevant to the present study. The finite difference version of the transport equation is solved using the ”leapfrog” centered time method and a combination of techniques in space. For advection terms, a standard centered difference approach (2nd order) is used in the horizontal and the flux-corrected transport (FCT) method is used in the vertical. One consequence of the centered difference approach is numerical dispersion, which causes dispersive ripples in the concentration field for sharp gradients in tracer concentration, i.e., it is best suited for smoothly varying scalars. This effect is to some extent remedied by
the FCT approach, which is a combination of an upwind and centered difference approach (see [35]) and is better suited to handle sharp concentration fronts. The convergence of the meridians at the poles can give rise to Courant number violations for the advective terms, which requires some special treatment. Specifically, a Fourier filtering technique is applied to the concentration fields along latitude bands near the poles, redistributing mass in the zonal direction to remove instability ripples introduced by Courant violations. For diffusion, all terms except for the (3,3) component of the diffusivity tensor (4.10) are treated in a fully explicit manner; the (3,3) term is treated in a fully implicit manner using a matrix inversion technique. Thus, of the isopycnal fluxes (i.e., the horizontal and vertical components of along-isopycnal diffusion), only the $A_I S^2$ (somewhat like an apparent vertical diffusivity) is treated implicitly. This term can be fairly large relative to the vertical diffusivity, routinely exceeding the lower end of the diffusivity spectrum computed by the Large et al. [50] scheme (Section 4.2.2.2.2). It is noted that the diffusion calculation takes advantage of the computational benefit offered by the Griffies [36] skew diffusive flux formulation of the GM terms (Section 4.2.1.2), and also uses the common assumption $\kappa = A_I$ in the diffusivity tensor (4.13).

The parameterization of the model is as follows for the grid resolution described in Section 4.2.2.1. The isopycnal diffusivity is set to a base value $A_I,0$ of $10^7$ cm$^2$/s, and is tapered following the approach in Section 4.2.2.2.2. In the vertical, a background diffusivity of 0.2 cm$^2$/s is specified, and scaled upward following the Large et al. [50] scheme. Timesteps of 0.5 hours and 1 day are used for velocity and tracer transport calculations, respectively. A free surface is allowed to exist, i.e., a rigid lid approximation is not imposed. Surface forcings (heat, freshwater, and wind) are specified using interpolated monthly mean climatology, supplemented by a sea ice model under ice conditions and a restoring to mean surface salinity values under non-ice conditions (see [12] for more details). The state of the ocean was allowed to evolve over a 6,200 year spinup calculation, i.e., any LLNL MOM simulation presented herein starts from this spinup condition so that circulation patterns are in a quasi steady state.
4.2.2.4 Spatial distributions of key terms

In order to provide the reader with a sense for the spatial variability of some key quantities relevant to LLNL MOM tracer transport, Figures 4-7 to 4-9 illustrate the annual average values of isopycnal slope, isopycnal diffusivity, and vertical diffusivity. For isopycnal slope and vertical diffusivity, annual averages were calculated and then averaged spatially in each direction. For isopycnal diffusivity, which is controlled by the isopycnal slope, Figure 4-8 shows spatial averages of the fraction of time over the year that any tapering occurred (defined as $A_I < 0.95A_{I,0}$). Overall, the three figures show similar spatial distributions. The largest slopes and highest vertical diffusivities generally occur in the upper ocean (i.e., within the OBL), generating severe vertical gradients. The spatial distribution suggests that the most challenging regions to model are likely to be the Southern Ocean and the North Atlantic, and to a lesser degree the North Pacific (near the northern subtropical gyre). In addition, the need for tapering of the isopycnal diffusivity is clear, since average isopycnal slopes can reach as high as $10^3$ (tapering occurs for slopes greater than about 0.0017). Lastly, it should be noted that average vertical diffusivity computed by the Large et al. [50] scheme varies by several orders of magnitude over short vertical scales. For specific examples of computed diffusivity profiles, the reader is referred to Figure 3-24.
Figure 4-7: Annual mean LLNL MOM isopycnal slope ($S = \sqrt{S_x^2 + S_y^2}$) for the year of output used to drive the RWPT model in the present study. Top panel: depth averaged slope; Right panel: zonally averaged slope; Bottom panel: meridionally averaged slope. All values are are plotted on a logarithmic scale (base 10).
Figure 4-8: Spatial averages of the fraction of time that isopycnal diffusivity is tapered over the year of output used to drive the RWPT model in the present study. Top panel: depth averaged values; Right panel: zonally averaged values; Bottom panel: meridionally averaged values. In this analysis, a diffusivity was considered to be tapered if the taper coefficient was less than 0.95. Note that only cells with a non-zero average fraction are plotted; cells shown as white space had no tapering at any time during the year.
Figure 4-9: Annual mean LLNL MOM vertical diffusivity for the year of output used to drive the RWPT model in the present study. Top panel: depth averaged diffusivity; Right panel: zonally averaged diffusivity; Bottom panel: meridionally averaged diffusivity. All values are in cm$^2$/s and are plotted on a logarithmic scale (base 10).
4.3 Particle model

The present work differs from most past investigations in that it attempts to build a Lagrangian model which can mimic the tracer calculations of an OGCM. Although the motivation of this effort is to take advantage of the Lagrangian framework in the estimation of residence time statistics and related quantities, the first step in this process is to select the appropriate model form and develop equations accordingly. This section describes the development and implementation of the model; validation and application of the model are deferred to Sections 4.4 and 4.5.

The choice of which Lagrangian model to use for a particular application depends to a large extent upon how turbulent mixing is to be represented. Consequently, Lagrangian modeling endeavors for oceanographic purposes differ mainly in the complexity of their representation of turbulent diffusion. On the one end of the complexity spectrum are the streamline (or streakline) tracing techniques that are commonly used, in which diffusion is ignored and particles are advected by the mean flow field (e.g., Boning et al. [8], Blanke and Raynaud [6], Blanke et al. [7], Drifjhout et al. [24], de Vries and Döös [18], Drifjhout et al. [25], Marsh and Megann [54]). These calculations have been successively refined to achieve high computationally efficiency, to allow forward or backward time trajectories in stationary or time-varying [18] flow fields, and to model convection events by assigning a random particle depth within the convective column [22, 24]. An eddy-induced transport or bolus velocity (to represent stirring by unresolved mesoscale eddies) has been added in both a z-level [25] and isopycnic [54] coordinate system. However, none of these methods explicitly resolve particle diffusion. Rather, the goal of these studies was to efficiently (and elegantly) map out the particle trajectories implied by the mean circulation field, allowing for special parameterization of sub-grid scale stirring and convection. On the other end of the spectrum are the higher order stochastic particle models which represent turbulent diffusion using a correlated random walk. Here particle velocities and in some cases accelerations evolve stochastically and have some degree of “memory”, suitable for simulating timescales shorter than the Lagrangian integral timescale (e.g., Berloff and McWilliams [5], Dewey et al. [20], Brickman et al. [9], van Dop et al. [71]). This group of models seeks to achieve more
physically realistic tracer spreading by going beyond the eddy diffusivity representation of turbulence, which has been shown to be inconsistent with observed ocean tracer spreading for short times (e.g., Sundenmeyer and Price [69], Ledwell et al. [51, 52], Davis [15, 16]) and even for long times (e.g., [5]). Going beyond this spectrum entirely are the Lagrangian methods which predict the flow field as well (e.g., Haertel et al. [38], Dukowicz and Smith [28]).

In the present study, the uncorrelated random walk (or random displacement) method is used (random walk particle tracking, RWPT, in the text that follows). RWPT falls in the middle of the complexity spectrum described above. Unlike the other approaches, the RWPT method is the Lagrangian equivalent of the advection-diffusion equation, consistent with the notion of absolute diffusion (see Chapter 2), and is therefore the only method which can be used, at least in theory, to truly reproduce the tracer calculations of an OGCM. It is important to note that the choice of RWPT is completely dictated by this goal. Studies employing the simpler streamline tracing techniques use Lagrangian methods to map out transport pathways to an extent not possible with a standard Eulerian approach, ignoring the effect of sub-grid scale diffusion. Studies employing a higher order (correlated) random walk may also take advantage of this benefit but choose the Lagrangian method such that turbulent diffusion can be represented in a more complex manner that the eddy diffusivity parameterization. Although RWPT is the crudest (lowest order) of the Lagrangian stochastic models (see review by Berloff and McWilliams [5]), it is selected herein with a specific purpose in mind. As such, it is perhaps best thought of as an advanced internal diagnostic to an OGCM, ideally being capable of extracting information from the OGCM which is otherwise difficult to extract (e.g., residence times for many sites and source water identification). Therefore, the remainder of this study focuses on the application of RWPT in this regard, and alternative Lagrangian model formulations are not considered further.

4.3.1 Random walk equations

An overview and derivation of the basic random walk approach is provided in Chapters 2 and 3 of this thesis. Summarizing these derivations, Lagrangian particles can approximate
the advection-diffusion equation by moving in accordance with

$$\Delta x = A(x(t), t) \Delta t + B(x(t), t) \xi \sqrt{\Delta t}, \quad (4.26)$$

where $A$ and $B$ are determined by rearranging the advection-diffusion equation to be equivalent in form to the following Fokker Planck equation:

$$\frac{\partial p(x, t|x_0, t_0)}{\partial t} + \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \left[ A_i(x, t)p(x, t|x_0, t_0) \right] = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} \left[ B_{ik}(x, t)B_{jk}(x, t)p(x, t|x_0, t_0) \right]. \quad (4.27)$$

Here $\Delta t$ is the timestep and $\xi = (\xi_1, \xi_2, \xi_3)$ is a vector of random numbers drawn from a standard Gaussian distribution (zero mean and unit standard deviation). Although strictly speaking each particle in an RWPT simulation represents an individual realization of the trajectory of the same “packet” of scalar mass or fluid, individual particles can more intuitively be considered a fraction of the total initial mass such that there is a correspondence between particle density and scalar concentration (see [44]).

In order to emulate the tracer transport calculations of LLNL MOM, two key features must be incorporated into the general RWPT framework above:

1. Tracer diffusion is oriented along isopycnal surfaces, as well as in the vertical direction.

2. Tracers are subject to the GM eddy-induced transport [33, 34, 36], a parameterizated stirring process induced by unresolved mesoscale eddies.

To the author’s knowledge, previous applications of RWPT to ocean or coastal models (e.g., [75, 9]) have not incorporated these nuances. Spivakovskaya [67] developed a set of random walk equations for simulating iso- and diapycnal diffusion, and tested them for simple cases with analytical solutions, but not within the context of an OGCM.

The following sections develop the random walk equations that will be used to emulate the OGCM. The derivation is performed in Cartesian coordinates (i.e., a local geopotential reference frame), and then adapted to be consistent with the spherical coordinates used in LLNL MOM.
4.3.1.1 Random walk in an isoneutral reference frame

Before deriving the random walk equations on the fixed geopotential frame, it is instructive to consider the random walk on an isoneutral frame. Looking ahead, this exercise will yield results which will later be useful in determining the tensor $\mathbf{B}$ in (4.26) for a fixed reference frame.

Consider a domain with a spatially constant isoneutral slope, and ignore any GM terms. The simplest way to describe transport within the domain is to define coordinate axes consisting of two isoneutral directions ($\tilde{x}$, $\tilde{y}$) and a dianeutral direction ($\tilde{z}$). Assuming that isoneutral diffusion is isotropic, the choice of $\tilde{x}$ and $\tilde{y}$ is arbitrary, i.e., any set is valid provided that they are on the isoneutral plane and perpendicular to each other as well as $\tilde{z}$. Tracer transport is then given by

$$\frac{\partial c}{\partial t} + \nabla \cdot (\tilde{u}c) = \nabla \cdot \left( \tilde{E} \cdot \nabla c \right) \quad (4.28)$$

where $\tilde{E}$ is given by (4.4), $\tilde{u}$ is a vector of velocities ($\tilde{u}, \tilde{v}, \tilde{w}$), and $\tilde{\nabla} = \left( \frac{\partial}{\partial \tilde{x}}, \frac{\partial}{\partial \tilde{y}}, \frac{\partial}{\partial \tilde{z}} \right)$ is the gradient operator within the isoneutral frame.

Equation (4.28) can be rearranged to be similar in form to (4.27) (assuming a non-divergent flow field, $\tilde{\nabla} \cdot \tilde{u} = 0$):

$$\frac{\partial c}{\partial t} + \frac{\partial}{\partial \tilde{x}} \left[ c \left( \frac{\partial A_I}{\partial \tilde{x}} + \tilde{u} \right) \right] + \frac{\partial}{\partial \tilde{y}} \left[ c \left( \frac{\partial A_I}{\partial \tilde{y}} + \tilde{v} \right) \right] + \frac{\partial}{\partial \tilde{z}} \left[ c \left( \frac{\partial A_I}{\partial \tilde{z}} + \tilde{w} \right) \right] = \frac{\partial^2 A_{Ic}}{\partial \tilde{x}^2} + \frac{\partial^2 A_{Ic}}{\partial \tilde{y}^2} + \frac{\partial^2 A_{Dc}}{\partial \tilde{z}^2} \quad (4.29)$$

Thus, (4.26) is the appropriate random walk equation provided that values of $p(\mathbf{x}, t | \mathbf{x}_0, t_0)$, $\mathbf{A}$, and $\mathbf{B}$ can be chosen such that (4.27) and (4.29) are equivalent. From inspection of the two equations, it is clear that $p(\mathbf{x}, t | \mathbf{x}_0, t_0) = c$ and

$$\mathbf{A} = \begin{pmatrix} \frac{\partial A_I}{\partial \tilde{x}} + \tilde{u} \\ \frac{\partial A_I}{\partial \tilde{y}} + \tilde{v} \\ \frac{\partial A_I}{\partial \tilde{z}} + \tilde{w} \end{pmatrix} \quad (4.30)$$
To determine the components of tensor $B$, it is useful to write out the right hand term of (4.27) explicitly:

$$\frac{\partial^2}{\partial x_i \partial x_j} \left( \frac{1}{2} B_{ik} B_{jk} f \right) = \frac{1}{2} \left\{ \frac{\partial^2}{\partial x^2} \left[ f \left( B_{xx}^2 + B_{xy}^2 + B_{xz}^2 \right) \right] + \frac{\partial^2}{\partial x \partial y} \left[ f \left( B_{xx} B_{yx} + B_{xy} B_{yy} + B_{xz} B_{yz} \right) \right] + \frac{\partial^2}{\partial x \partial z} \left[ f \left( B_{xx} B_{zx} + B_{xy} B_{zy} + B_{xz} B_{zz} \right) \right] + \frac{\partial^2}{\partial y^2} \left[ f \left( B_{yx}^2 + B_{yy}^2 + B_{yz}^2 \right) \right] + \frac{\partial^2}{\partial y \partial z} \left[ f \left( B_{yx} B_{zy} + B_{yy} B_{zy} + B_{yz} B_{zz} \right) \right] + \frac{\partial^2}{\partial z^2} \left[ f \left( B_{zx}^2 + B_{zy}^2 + B_{zz}^2 \right) \right] \right\}$$

(4.31)

where $f = p(x, t|x_0, t_0)$ has been introduced for notational simplicity.

Comparing the right side of (4.29) to (4.31), and letting $f = c$, we get the following set of equalities which must be satisfied in order for (4.27) and (4.29) to be equivalent:

$$B_{xx}^2 + B_{xy}^2 + B_{xz}^2 = 2A_I \quad (4.32a)$$

$$B_{yx}^2 + B_{yy}^2 + B_{yz}^2 = 2A_I \quad (4.32b)$$

$$B_{zx}^2 + B_{zy}^2 + B_{zz}^2 = 2A_D \quad (4.32c)$$

$$B_{xx} B_{xx} + B_{xy} B_{xy} + B_{xz} B_{xz} = 0 \quad (4.32d)$$

$$B_{yx} B_{yx} + B_{yy} B_{yy} + B_{yz} B_{yz} = 0 \quad (4.32e)$$

$$B_{yz} B_{xz} + B_{yy} B_{xy} + B_{yz} B_{xz} = 0 \quad (4.32f)$$

It is important to note that we now have six equations with nine unknowns. This means
that the choice of $B$ is unconstrained. Any value that satisfies (4.32) should be valid (although we probably want to impose the constraint that all elements of $B$ must be real). Thus, there exists an infinite number of random walk equations that can be used to simulate (4.29). This point is also mentioned in Gardiner [32]. Looking at the general form of the random walk equation (4.26) reveals why an infinite number of possible equations exist. The random variates $\xi$ represent the diffusion of the particle due to random fluctuations. In so doing, they define a Cartesian coordinate system that is unique to the random walk. This random walk coordinate system may or may not be aligned with the isoneutral coordinate system on which $\tilde{E}$ is defined. Rotation of the random walk coordinate system relative to the isoneutral coordinate system coordinate system by an arbitrary angle will change the value of the tensor $B$, but (4.26) remains valid as long as $B$ satisfies (4.32).

The most straightforward choice for $B$ is

$$B = \begin{pmatrix} \sqrt{2A_I} & 0 & 0 \\ 0 & \sqrt{2A_I} & 0 \\ 0 & 0 & \sqrt{2A_D} \end{pmatrix}.$$  (4.33)

Thus, the random walk equations for the isopycnal reference frame are:

$$\Delta \tilde{x} = \left[ \frac{\partial A_I}{\partial \tilde{x}} + \tilde{u} \right] \Delta t + \xi_1 \sqrt{2A_I} \Delta t$$  (4.34a)

$$\Delta \tilde{y} = \left[ \frac{\partial A_I}{\partial \tilde{y}} + \tilde{v} \right] \Delta t + \xi_2 \sqrt{2A_I} \Delta t$$  (4.34b)

$$\Delta \tilde{z} = \left[ \frac{\partial A_I}{\partial \tilde{z}} + \tilde{w} \right] \Delta t + \xi_3 \sqrt{2A_D} \Delta t$$  (4.34c)

Equation (4.34) predicts particle motion along and across a constant isopycnal slope using an isoneutral reference frame. However, because isopycnal slopes in the ocean are not spatially constant, (4.34) is not sufficient to describe particle motion in a real ocean. Complete equations will therefore be derived in the next section using a fixed reference frame. Before proceeding, it is instructive to rotate the particle displacements predicted by (4.34) onto a fixed reference frame $(x, y, z)$, i.e., if we take (4.34) as a valid local approximation.
Using small angle approximations, similar to those used in converting (4.7) to (4.8), to crudely rotate each term in (4.34) yields

\[ \Delta x = \left[ \frac{\partial A_I}{\partial x} + u \right] \Delta t + \xi_1 \sqrt{2A_I} \Delta t \] (4.35a)

\[ \Delta y = \left[ \frac{\partial A_I}{\partial y} + v \right] \Delta t + \xi_2 \sqrt{2A_I} \Delta t \] (4.35b)

\[ \Delta z = \left[ S_x \frac{\partial A_I}{\partial x} + S_y \frac{\partial A_I}{\partial y} + \frac{\partial A_D}{\partial z} + w \right] \Delta t + S_z \xi_1 \sqrt{2A_I} \Delta t + S_y \xi_2 \sqrt{2A_I} \Delta t + \xi_3 \sqrt{2A_D} \Delta t \] (4.35c)

If these were taken as random walk equations for the local geopotential reference frame, they imply that vector \( \mathbf{A} \) and tensor \( \mathbf{B} \) in (4.27) have the following values on the local geopotential reference frame:

\[ \mathbf{A} = \begin{pmatrix}
\frac{\partial A_I}{\partial x} + u \\
\frac{\partial A_I}{\partial y} + v \\
S_x \frac{\partial A_I}{\partial x} + S_y \frac{\partial A_I}{\partial y} + \frac{\partial A_D}{\partial z} + w
\end{pmatrix}. \] (4.36)

\[ \mathbf{B} = \begin{pmatrix}
\sqrt{2A_I} & 0 & 0 \\
0 & \sqrt{2A_I} & 0 \\
S_x \sqrt{2A_I} & S_y \sqrt{2A_I} & \sqrt{2A_D}
\end{pmatrix}. \] (4.37)

In the next section, formal derivation of the random walk equations for the local geopotential reference frame (i.e., where isopycnal slopes are not spatially constant) will demonstrate that (4.37) is valid for this case as well, but that (4.36) must be augmented with some additional \( \mathbf{A} \) terms.

### 4.3.1.2 Random walk in a fixed reference frame

The derivation in the preceding section was incomplete because it did not consider the spatial variability of isopycnal slopes, nor did it include any GM parameterization. In this section, general random walk equations are derived for a fixed reference frame corresponding to the local geopotential.

Combing the tracer transport equation (4.1) with the diffusion tensor (4.13) yields the
tracer equation in this coordinate system:

\[
\frac{\partial c}{\partial t} + \mathbf{u} \nabla \cdot c = \frac{\partial}{\partial x} \left[ A_I \frac{\partial c}{\partial x} + (A_I - \kappa) S_x \frac{\partial c}{\partial x} \right] + \frac{\partial}{\partial y} \left[ A_I \frac{\partial c}{\partial y} + (A_I - \kappa) S_y \frac{\partial c}{\partial y} \right] + \frac{\partial}{\partial z} \left[ (A_I + \kappa) S_z \frac{\partial c}{\partial z} + (A_I + \kappa) S_y \frac{\partial c}{\partial y} + (K_v + A_I S^2) \frac{\partial c}{\partial z} \right]
\]  

where a non-divergent flow field ($\nabla \cdot \mathbf{u} = 0$) has been assumed. Equation (4.38) can be rearranged to be similar in form to (4.27):

\[
\frac{\partial c}{\partial t} + \frac{\partial}{\partial x} \left[ c \left( \frac{\partial A_I}{\partial x} + \frac{\partial}{\partial z} [(A_I - \kappa) S_x c] + u \right) \right] + \frac{\partial}{\partial y} \left[ c \left( \frac{\partial A_I}{\partial y} + \frac{\partial}{\partial z} [(A_I - \kappa) S_y c] + v \right) \right] + \frac{\partial}{\partial z} \left[ c \left( \frac{\partial}{\partial x} [(A_I + \kappa) S_x] + \frac{\partial}{\partial y} [(A_I + \kappa) S_y] + \frac{\partial}{\partial z} [K_v + A_I S^2] + w \right) \right] = \frac{\partial^2 A_I c}{\partial x^2} + \frac{\partial^2 A_I c}{\partial y^2} + \frac{\partial^2}{\partial x \partial y} [2A_I S_x c] + \frac{\partial^2}{\partial y \partial z} [2A_I S_y c] + \frac{\partial^2}{\partial z^2} \left( K_v + A_I S^2 \right) c
\]

Thus, (4.26) is the appropriate random walk equation provided that values of $p(\mathbf{x}, t | \mathbf{x}_0, t_0)$, $A$, and $B$ can be chosen such that (4.27) and (4.39) are equivalent. Upon inspection of (4.39), it is clear that $p(\mathbf{x}, t | \mathbf{x}_0, t_0) = c$ and

\[
A = \begin{pmatrix}
\frac{\partial A_I}{\partial x} + \frac{\partial}{\partial z} [(A_I - \kappa) S_x] + u \\
\frac{\partial A_I}{\partial y} + \frac{\partial}{\partial z} [(A_I - \kappa) S_y] + v \\
\frac{\partial}{\partial x} [(A_I + \kappa) S_x] + \frac{\partial}{\partial y} [(A_I + \kappa) S_y] + \frac{\partial}{\partial z} [K_v + A_I S^2] + w
\end{pmatrix}.
\]

Comparing the right side of (4.39) to (4.31) (and again letting $f = p(\mathbf{x}, t | \mathbf{x}_0, t_0) = c$) yields the following set of equalities which must be satisfied in order for (4.27) and (4.39) to be equivalent:

\[
B_{xx}^2 + B_{xy}^2 + B_{xz}^2 = 2A_I
\]

\[
B_{yx}^2 + B_{yy}^2 + B_{yz}^2 = 2A_I
\]

\[
B_{zx}^2 + B_{zy}^2 + B_{zz}^2 = 2 \left( K_v + A_I S^2 \right)
\]

\[
B_{xx} B_{xy} + B_{xy} B_{xz} + B_{xz} B_{zz} = 2A_I S_x
\]
\[ B_{yx} B_{xx} + B_{yy} B_{xy} + B_{yz} B_{xz} = 2 A_I S_y \]  
\[ B_{yx} B_{xx} + B_{yy} B_{xy} + B_{yz} B_{xz} = 0 \]

As in (4.32), we are left with six equations and nine unknowns. Again, this simply means that there are an infinite number of valid choices for \( \mathbf{B} \), and thus an infinite number of random walk equations which can be used to simulate (4.39).

However, all versions of the tensor are not equally convenient. One simple form that satisfies (4.41) is (4.37):

\[
\mathbf{B} = \begin{pmatrix}
\sqrt{2A_I} & 0 & 0 \\
0 & \sqrt{2A_I} & 0 \\
\sqrt{2A_I S_x} & \sqrt{2A_I S_y} & \sqrt{2K_v}
\end{pmatrix}
\]

where \( A_D \) has been replaced by \( K_v \) to be consistent with the diffusion tensor used by LLNL MOM (see Section 4.2.1.1.2).

Inserting (4.40) and (4.37) into (4.26), the random walk equations for the isopycnal diffusion scheme implemented in MOM are:

\[
\Delta x = \left[ \frac{\partial A_I}{\partial x} + \frac{\partial}{\partial z} [(A_I - \kappa) S_x] + u \right] \Delta t + \xi_1 \sqrt{2A_I \Delta t} \] (4.42a)

\[
\Delta y = \left[ \frac{\partial A_I}{\partial y} + \frac{\partial}{\partial z} [(A_I - \kappa) S_y] + v \right] \Delta t + \xi_2 \sqrt{2A_I \Delta t} \] (4.42b)

\[
\Delta z = \left[ \frac{\partial}{\partial x} [(A_I + \kappa) S_x] + \frac{\partial}{\partial y} [(A_I + \kappa) S_y] + \frac{\partial}{\partial z} [K_v + A_I S^2] + w \right] \Delta t \\
+ \xi_1 S_x \sqrt{2A_I \Delta t} + \xi_2 S_y \sqrt{2A_I \Delta t} + \xi_3 \sqrt{2K_v \Delta t} \] (4.42c)

It is interesting to note the similarity of this equation to (4.35), save for a few terms in the definition of \( \mathbf{A} \) which address the spatial variability of the isopycnal slope and incorporate the Gent-McWilliams sub-grid scale mixing parameterization.

It is also noteworthy that the advective form of the GM parameterization is recovered in (4.42) even though we started with the skew diffusive flux form of the diffusion tensor.
Thus, the derivation above confirms the equivalence of the two different forms of the GM parameterization demonstrated in Griffies [36].

Lastly, it should be noted that the equations simplify further with the common assumption $\kappa = A_I$, as is the case in the LLNL MOM configuration used in the present study. As in the Eulerian case, inclusion of the GM terms under this assumption actually removes terms from the calculation. For generality, however, the $\kappa$ terms are left in the equations above, since $\kappa = A_I$ is not a universal assumption; recent research suggests that tapering $\kappa$ in the oceanic boundary layer is a physically unrealistic approach ([30], S. Griffies personal communication).

4.3.1.3 Adaptation to spherical coordinates

As discussed in Section 4.2.1, the LLNL MOM uses planetary spherical coordinate in its tracer simulations. Thus, it is necessary to adapt the random walk equations (4.42) to this coordinate system for simulation on the LLNL MOM grid. At first glance, this seems easily accomplished by simply converting the horizontal displacements $\Delta x$ and $\Delta y$ into displacements in longitude ($\Delta \lambda$) and latitude ($\Delta \phi$) using (4.14a) and (4.14b), i.e., using the same approach used internally in the OGCM to compute gradients. Experience with this approach yielded good results for short times but for long times a net poleward drift at high latitudes was observed, identifying the need for a correction term.

Ideally the random walk equations should be rederived using a spherical Fokker-Planck equation. However, because the form of the equation required was not found in the literature, a somewhat simpler approach was used to correct the poleward drift. Based on the work of Raible and Engel [63] and Brillinger [10], a drift correction term of the form $-\frac{\Delta x}{R_e} \tan(\phi) \Delta t$ must be added to the meridional displacement (where $\phi$ is latitude and $R_e$ is the radius of the Earth). In addition, to be consistent with the notion of isopycnal diffusion, the vertical projection of this term must be added to the vertical displacement, $-S_y \frac{\Delta \lambda}{R_e} \tan(\phi) \Delta t$. Thus, overall the RWPT equations implemented in the present study are:

$$\Delta x \equiv R_e \Delta \lambda \cos \phi = \left[ \frac{\partial A_I}{\partial x} + \frac{\partial}{\partial z} [(A_I - \kappa) S_x] + u \right] \Delta t + \xi_1 \sqrt{2A_I} \Delta t$$  (4.43a)
Here the gradients are left in terms of $\partial_x$ and $\partial_y$ for notational simplicity; these quantities are calculated from LLNL gridded output using (4.14a) and (4.14b). The procedure of incrementing particle position in longitude $\lambda$ and latitude $\phi$ is straightforward using the above equations. Also, for consistency with the present LLNL MOM configuration, the Gent-McWilliams thickness diffusivity $\kappa$ is set equal to the isopycnal diffusivity $A_I$ (which removes terms from (4.43a) and (4.43b)).

The need for the meridional correction introduced above can be understood intuitively. Consider an ocean with no land and spatially uniform horizontal diffusivity in which a tracer is introduced at time $t = 0$ along a latitude band in the northern hemisphere (the tracer concentration is zero everywhere else). For $t > 0$, there will be a meridional tracer flux in the north and south directions due to the resulting concentration gradient. The mass flux per unit meridional area will be the same since the diffusivity and concentration gradients are initially the same in the north and south directions. However, the total mass flux will be greater to the south than the north because the meridional area is greater to the south due to the converging meridians. For an Eulerian scheme, this is handled naturally since the area of grid cell faces varies as a function of latitude. For Lagrangian particles undergoing horizontal diffusion, however, a correction is needed. For the first timestep after $t = 0$, the number of particles that go north and south are the same, since all particles are initially on a single latitude band and since the diffusivity is spatially constant. Thus, the mass flux to the north and the south are the same, which is incorrect. The correction term derived in [63] and [10] adjusts particle displacements to have a net drift toward the equator. Raible and Engel [63] explain the term in a slightly different manner: “the Brownian motion of the particle...has two independent components, one goes along the meridian...and the other one goes along the great circle that intersects the meridian at a right angle”. Using either interpretation,
the net effect of the term is to prevent the mass accumulation that would otherwise occur at high latitudes due to meridional convergence. Since the LLNL MOM uses an isopycnal rather than horizontal diffusivity, there is also a projection of this term in the vertical. From the equations developed in Sections 4.3.1.1 and 4.3.1.2 and the discussion which follows in Section 4.3.1.4, it is clear that the small angle approximation used by LLNL MOM causes all horizontal displacements due to isopycnal diffusion to be the same as if the diffusivity were strictly horizontal, and that the vertical projection of isopycnal diffusion has magnitude \(S_x\) or \(S_y\) times the horizontal projection. The meridional and vertical projections of this correction term included in (4.43b) and (4.43c) were inferred in this manner.

Unfortunately neither [63] nor [10] derived the corrections that result when the horizontal diffusivity is spatially variable as is the case in the LLNL MOM (where isopycnal diffusivity is “tapered” in regions of steep isopycnal slope). Thus, it is likely that there are additional correction terms that may exist which have to do with meridional gradients in the isopycnal diffusivity. Since, as discussed later in this section, isopycnal diffusivity is only spatially variable for a rather narrow range of isopycnal slopes, it is reasonable to assume that any additional correction terms would have only a minor impact on the longterm performance of the particle model. For example, in point source releases of tracer particles the poleward drift caused by the absence of the meridional correction term described above was only discernible after a fairly long time (about 10-30 years, depending on release location). Given that any terms related to a gradient in isopycnal diffusivity would only come into play in regions of steep isopycnal slopes, its effect is expected to be subtle since large parts of the ocean do not experience tapering (see Figure 4-8). Furthermore, the impact of additional terms will in all likelihood be dwarfed by inaccuracies introduced by the sub-grid scale assumptions that must be made in these regions of the oceans (see discussion in Section 4.3.2.1). Thus, further derivation of additional correction terms is deferred to future investigations, and (4.43) will be taken as the appropriate RWPT equations in the present study.

4.3.1.4 Discussion of the random walk equation

It is useful to consider the meaning of the individual terms in (4.43). These equations differ from those employed in most applications of the random walk to environmental trans-
port simulation in that the principle axes are not aligned with the axes of diffusion. As a result, there are offdiagonal terms in the diffusion tensor, which give rise to extra terms in the random walk equations (4.43). The meaning of these extra terms can be understood by comparing (4.43) to a simpler set of random walk equations derived for the case where the principle axes are aligned with the axes of diffusion. Although many such examples can be found in the literature (e.g., Dimou and Adams [21], Kinzelbach [49], Brickman and Smith [9], Israelsson et al. [44]), in the present case it is most instructive to consider (4.34), the random walk equations developed for a fixed isoneutral reference frame in Section 4.3.1.1, restated below for convenience (with advection ignored)

\[
\Delta \tilde{x} = \frac{\partial A_I}{\partial x} \Delta t + \xi_1 \sqrt{2 A_I \Delta t} \\
\Delta \tilde{y} = \frac{\partial A_I}{\partial y} \Delta t + \xi_2 \sqrt{2 A_I \Delta t} \\
\Delta \tilde{z} = \frac{\partial A_I}{\partial z} \Delta t + \xi_3 \sqrt{2 A_D \Delta t}.
\]

For each direction, there is a stochastic term to represent the diffusion process and a deterministic “drift correction” or “pseudovelocity” term. As discussed in [44] and Chapter 3 and references therein, the purpose of the deterministic term is to prevent particle accumulation in regions of low diffusivity by adjusting the center of mass of a diffusing particle cloud toward higher diffusivity. Thus, the stochastic and deterministic terms form a pair which allow accurate simulation of diffusion in the limit of a small timestep (depending on the domain and the diffusivity profile, see Chapter 3). Examining (4.43) reveals that these terms are also present here: \( \frac{\partial A_I}{\partial x} \Delta t + \xi_1 \sqrt{2 A_I \Delta t} \) for \( x \); \( \frac{\partial A_I}{\partial y} \Delta t + \xi_2 \sqrt{2 A_I \Delta t} \) for \( y \); and \( \frac{\partial K_v}{\partial z} \Delta t + \xi_3 \sqrt{2 K_v \Delta t} \) for \( z \). From this starting point, we are in a position to understand the meaning of the other terms in (4.43). To this end, (4.43) is expanded and rearranged below without advection terms, the GM terms or the meridional correction term previously described (i.e., using \( u = 0, \kappa = 0 \) and \( \phi = 0 \)):

\[
\Delta x = \left\{ \frac{\partial A_I}{\partial x} \Delta t + \xi_1 \sqrt{2 A_I \Delta t} \right\} + \left\{ \frac{\partial}{\partial z} \left[ A_I, S_x \right] \Delta t \right\} \\
(4.44a)
\]
\[ \Delta y = \left\{ \frac{\partial A_I}{\partial y} \Delta t + \xi_2 \sqrt{2 A_I \Delta t} \right\} + \left\{ \frac{\partial}{\partial z} [A_I S_y] \Delta t \right\} \]  
\[ (4.44b) \]

\[ \Delta z = \left\{ \frac{\partial K_v}{\partial z} \Delta t + \xi_3 \sqrt{2 K_v \Delta t} \right\} \]

\[ + \left\{ S_x \left[ \frac{\partial A_I}{\partial x} \Delta t + \xi_1 \sqrt{2 A_I \Delta t} \right] \right\} + \left\{ S_y \left[ \frac{\partial A_I}{\partial y} \Delta t + \xi_2 \sqrt{2 A_I \Delta t} \right] \right\} \]

\[ + \left\{ A_I \left[ \frac{\partial S_x}{\partial x} \Delta t + \frac{\partial S_y}{\partial y} \Delta t \right] \right\} + \left\{ \frac{\partial A_I S^2}{\partial z} \Delta t \right\}. \]  
\[ (4.44c) \]

The first bracketed group of terms in each equation in (4.44) is the stochastic and deterministic pair noted above for each direction. The vertical projection of these paired terms in (4.44a) and (4.44b) also shows up in the vertical particle displacement (second and third bracketed terms in (4.44c)), i.e., this is the vertical component of the along-isopycnal motion. Note that because of the small angle approximation (see Section 4.2.1.1.2), there is no slope adjustment to the along-isopycnal displacements in \( x \) and \( y \), i.e., the displacements are the same as if \( A_I \) was a horizontal diffusivity. The fourth term in (4.44c) causes an additional particle displacement in the vertical due to horizontal gradients in the isopycnal slope. Somewhat analogous to the way the deterministic term in the first bracketed term moves particles toward higher diffusivity, this term moves particle toward steeper isopycnal slopes which prevents particle accumulation in regions where the slope is small. Likewise the last term in (4.44c) moves particles toward areas of higher diffusivity or isopycnal slope which prevents accumulation in areas of low diffusivity and/or slope. The last two terms in (4.44a) and (4.44b) serve a similar purpose, moving particles horizontally away from regions of lower diffusivity and/or isopycnal slope based on the vertical gradient. Thus, all the terms work to adjust the particle trajectories to account for changes in the diffusivity and isopycnal slope along the particle paths. If implemented correctly, the RWPT should (at least in theory) be able to meet the well-mixed condition, i.e., maintain an initially uniform particle distribution.

It is also useful to consider how the random walk equation changes when different groups of terms are ignored, as outlined below in no particular order.

- Setting \( K_v = 0, \kappa = 0 \) and \( u = 0 \) yields the pure diffusion problem of particles
undergoing isopycnal diffusion in the presence of spatially variable isopycnal slopes.

- Setting all terms except $K_v$ to zero yields a one dimensional diffusion problem which is prone to a high degree of spatial variability and as such is subject to the considerations of Chapter 3 on piecewise linear diffusivity profiles.

- Setting all terms except $u$ equal to zero yields the streakline tracing work of numerous studies mentioned previously.

- Setting $A_f = 0$, $K_v = 0$, and $u = 0$ leaves only the GM advective fluxes, which should be non-divergent (e.g., [36])

If the random walk implementation is correct, each of these subsets should in theory be able to satisfy the well-mixed condition.

The above considerations suggest two key classes of tests that can be performed to test the success of the random walk implementation. First, and most importantly, are predicted tracer distributions similar to those of the OGCM? Second, can the well-mixed condition be maintained with all the terms and with the individual subset of terms listed above? Note that the latter can be true even if the former is not true, meaning that the two tests combined will shed light on whether observed differences in tracer distributions are due to the implementation of the random walk (i.e., not preserving the well-mixed condition) or if the differences are due to a failure to correctly estimate variables in a manner consistent with the internal calculations of LLNL MOM.

It may, however, be unreasonable to expect the well-mixed condition to be met perfectly by a random walk implementation in a real domain, as suggested by Brickman and Smith [9]. For example, sub-grid scale variable interpolation and the influence of complex geometry is likely to introduce artificial particle accumulations to some extent.

4.3.2 RWPT implementation

The Devil, they say, is in the details. Because of the fundamental differences between the Eulerian and Lagrangian modeling approaches, it is far from obvious how to best implement the random walk in a manner that will accurately reproduce the tracer transport predicted
by an Eulerian OGCM. The choices to be made go well beyond choosing a numerical scheme with sufficient accuracy. Since the parent OGCM is the source of all hydrodynamic input variables for the Lagrangian model, and because OGCM variables are defined on a relatively coarse grid, implementation of the Lagrangian model inevitably involves making assumptions about the sub-grid scale variation of each hydrodynamic variable, and consequently about the sub-grid scale motion of Lagrangian particles.

Implementing the random walk equations (4.42) in a manner that can accurately reproduce the OGCM tracer calculations consists of two main parts. First, variables should ideally be defined such that each term in (4.43) can be calculated in a manner consistent with all other terms in the equation, and without unreasonable computational expense. Second, the numerical scheme should move particles within the defined variable fields without artificial biases. Although the two are closely linked since variable definitions dictate the spatial variability of each term in the numerical scheme, the discussion that follows first addresses variable definition/interpolation and then the numerical approach used to move particles.

4.3.2.1 Variable definition and interpolation

The challenge lies in defining the quantities that appear in (4.43) in a manner that mimics the LLNL MOM tracer calculation. The problem that arises is the fundamental difference between the variables needed by each scheme. Griffies [36] points out this problem in the context of the GM parameterization. As noted previously, the random walk approach must rely on the advective form of the GM transport, which is notoriously noisy and prompted the skew flux formulation of Griffies [36].

4.3.2.1.1 Advective velocity  The current velocity for a particle at \( \{x_p, y_p, z_p\} \) is estimated in the following way:

\[
\begin{align}
    u_p &= \bar{u}_{i-1,j,k} + \frac{x_p - x_{i-1/2}}{\Delta x_j} \left( \bar{u}_{i+1/2,j,k} - \bar{u}_{i-1/2,j,k} \right) \\
    v_p &= \bar{v}_{i,j-1,k} + \frac{y_p - y_{j-1/2}}{\Delta y} \left( \bar{v}_{i,j+1/2,k} - \bar{v}_{i,j-1/2,k} \right)
\end{align}
\]  

(4.45)
Thus, within each cell, the velocity in a particular direction is assumed to only vary in that direction. This approach is computationally inexpensive and assures that the OGCM’s non-divergent velocity field is preserved, as discussed in Section 4.2.2.2.1. This is the same interpolation approach is used in Döös [22] and subsequent papers (including the development of the TRACMASS model).

Other options for velocity interpolation have been used in previous studies. Berloff and McWilliams [5] use a bicubic interpolation, Boning and Cox [8] use trilinear interpolation, and Taylor [70] simply uses nearest neighbor velocity to reduce computation time. Experience with the random walk model developed herein has shown that particle simulations are not very sensitive to the interpolation scheme used. A trilinear interpolation scheme has also been coded in the present model, but is not used because it did not show improved performance yet increased computation time significantly.

### 4.3.2.1.2 Vertical diffusivity

As described previously, the OGCM’s vertical diffusivity \( K_v \) is a lumped parameter into which all vertical transport processes except resolved advection and isopycnal diffusion have been incorporated. It is calculated by the OGCM at each vertical cell face by considering the vertical density structure of each column of cells. In stably stratified, quiescent regions of the ocean interior, the vertical diffusivity approaches a background value of 0.2 cm\(^2\)/s. In highly turbulent regions with weak or unstable stratification, the vertical diffusivity can become up to 7 orders of magnitude larger, with values in the 1000 to 10000 cm\(^2\)/s range being common in the oceanic boundary layer. The large diffusivities destroy unstable density gradients by causing a large flux between overlying grid cells.

In the Lagrangian model, vertical diffusivity is treated as varying linearly in the vertical direction (between the values defined at the cell faces), i.e.,

\[
K_{v,p} = K_{v,i,j,k-1/2} + \frac{z_p - z_{k-1/2}}{\Delta z_k} (K_{v,i,j,k+1/2} - K_{v,i,j,k-1/2})
\]  

(4.46)

where \( K_{v,p} \) is the diffusivity at the particle location \( z_p \). This simple definition is chosen
because $K_v$ represents mixing at vertical cell faces based on the density structure, and thus should be horizontally constant across each vertical cell face. The linear vertical variation between the faces is arbitrary, but is chosen because of its simplicity and transparency. The drawback to this approach is that the diffusivity gradient is discontinuous, which is a difficulty for the random walk model as it can cause artificial particle accumulations. Although past random walk studies have addressed this problem by smoothing the diffusivity profile and/or using a variable timestep to avoid such problems, these approaches are not pursued here; instead, the numerical scheme described in Chapter 3 is used. This methodology was specifically designed to handle sharp discontinuities in diffusivity gradient; in fact, it was motivated by the diffusivity field computed by LLNL MOM.

4.3.2.1.3 Isopycnal diffusivity and slope All remaining terms in (4.43) are associated with diffusion along isopycnal surfaces. At this time, the definition of variables related to isopycnal diffusion for Lagrangian particles is somewhat arbitrary, and the most accurate approach remains an unresolved issue in the present research. The details of how variables are defined, which dictates how they are interpolated, is one of the key challenges in implementing the random walk equations in a manner that can mimic the calculations of a parent OGCM. The isopycnal diffusivity and slopes are here treated together because as previously noted $A_I$ is a function of $S_x$ and $S_y$ in the LLNL MOM.

For convenience, it is useful to examine the random walk equations with the vertical diffusivity and advection terms removed:

\[
\Delta x = \left[ \frac{\partial A_I}{\partial x} + \frac{\partial}{\partial z} [(A_I - \kappa) S_x] \right] \Delta t + \xi_1 \sqrt{2A_I} \Delta t 
\]

\[
\Delta y = \left[ \frac{\partial A_I}{\partial y} + \frac{\partial}{\partial z} [(A_I - \kappa) S_y] \right] \Delta t + \xi_2 \sqrt{2A_I} \Delta t - \frac{A_I}{Re} \tan \phi 
\]

\[
\Delta z = \left[ \frac{\partial}{\partial x} [(A_I + \kappa) S_x] + \frac{\partial}{\partial y} [(A_I + \kappa) S_y] + \frac{\partial}{\partial z} [A_I S^2] \right] \Delta t
+ \xi_1 S_x \sqrt{2A_I} \Delta t + \xi_2 S_y \sqrt{2A_I} \Delta t - S_y \frac{A_I}{Re} \tan \phi 
\]
All of these terms are dependent on the isopycnal slope vector (even the GM thickness diffusivity $\kappa$ since it is usually assumed equal to $A_I$). Thus, the definition of isopycnal slopes is a crucial assumption in the random walk model. This is not, however, a straightforward task as there are a number of constraints that we would ideally like to place on the variable definitions from the perspective of a Lagrangian particle:

1. In order to define all terms in (4.47), non-zero spatial gradients should be able to exist as needed in each direction, i.e., $S_x$ and $S_y$ cannot be assumed constant in any direction within a grid cell.

2. As in the OGCM, the flux through any cell-face should be the same everywhere along the face, with the same flux magnitude as in the parent OGCM. This is the case for the velocity and vertical diffusivity definitions described previously.

3. The GM thickness diffusivity $\kappa$ should be equal to $A_I$ as it is in the OGCM, i.e., $\kappa$ should also be a function of the isopycnal slope.

4. The GM velocity defined in (4.12) should be non-divergent, i.e., $\nabla \cdot \mathbf{u}_{GM} = 0$.

5. To minimize simulation time, interpolations should be computationally efficient, and the number of interpolations should be minimized.

The problem is that some of these constraints are mutually exclusive. The first constraint requires isopycnal slopes and diffusivity to vary in each spatial direction while the second constraint requires them to be constant along cell faces. The fourth constraint implies the first (so that each GM terms can be defined), but it seems unreasonable to expect the fourth constraint to ever be satisfied since the third constraint means that two of the GM terms disappear. Non-divergence could be accomplished by calculating $w_{GM}$ from the divergence of the $u_{GM}$ and $v_{GM}$ as in [14], but this would mean that some gradient terms in (4.47) would be calculated differently depending on whether they are multiplied by $A_I$ or $\kappa$, even though these two coefficients are equal according to the third constraint.

It should be noted how much easier it is to define these terms on an Eulerian grid. As described in Section 4.2.2, the OGCM computes cell-face average isopycnal slopes and uses
these to calculate an isopycnal diffusivity that is tapered as needed. Calculating fluxes across
the cell faces does not require that the gradient of isopycnal slope or isopycnal diffusivity be
defined anywhere; only the terms in the diffusion tensor (4.13) must be calculated.

It does not seem possible to define isopycnal slopes and diffusivities in a way that is
entirely consistent with the parent OGCM. Instead it seems that we must choose from the
following set of general approaches:

1. Use one set of OGCM $S_x$ and $S_y$ output values (see Figure 4-6, either those defined
on the horizontal faces or those defined on the vertical faces), and interpolate the
isopycnal slope only in one direction. This means that within each grid cell, some of
the spatial gradient terms would equal zero. These terms could either be left as zero,
or an average value could be used by interpolating across neighboring grid cells.

2. Use one set of OGCM $S_x$ and $S_y$ output values (see Figure 4-6, either those defined on
the horizontal faces or those defined on the vertical faces), and define all other variables
from these by interpolation as needed. This means interpolating in three dimensions,
ignoring the boundaries between T-grid cells.

3. Perform either of the above interpolations using raw OGCM density gradients at the
cell faces instead, from which the isopycnal slope and diffusivity can be calculated
locally.

4. Perform the interpolation using OGCM density values directly. Through multiple
interpolation calculations, it would be possible to calculate the density gradient in
each spatial direction at the particle’s position, and also the spatial gradient of the
density gradient in each direction, which is necessary to calculate the gradient of the
isopycnal slopes.

5. Some combination of the above in which allows some of the variables are defined as step
functions and the associated gradient terms are treated in an approximate manner.

However, defining isopycnal slopes is only one part of the problem; the isopycnal diffusivity $A_I$ must also be defined. As discussed in Section 4.2.2.2.2, within LLNL MOM $A_I$ is a
constant except in regions of steep isopycnal slopes where it is brought to zero rapidly. Thus, for most of the ocean, \( \frac{\partial A_I}{\partial x} = 0 \), \( \frac{\partial}{\partial x} [A_I S_x] = A_I \frac{\partial S_x}{\partial x} \), and \( \frac{\partial}{\partial y} [A_I S_y] = A_I \frac{\partial S_y}{\partial y} \). As demonstrated in Section 4.2.2.2.2, \( A_I \) goes from 99% to 1% of its base value \( A_{I,0} \) in a fairly narrow slope interval \( 0.0017 \lesssim S \lesssim 0.0063 \).

When implementing the RWPT with the LLNL MOM output, the \( A_I \) to use for a Lagrangian particle can be calculated in one of two ways: (1) from a local isopycnal slope value at the particle position using (4.17) or (2) by interpolation of output values of \( A_I \) to the particle position.

The first of these approaches is tempting since it follows the manner in which isopycnal slopes are computed internally in the LLNL MOM. However, it gives rise to problems because the rapid variation of the taper function (4.17) in the transition region can lead to very large gradients in diffusivity. Furthermore, implementing this technique places unreasonable demands on the isopycnal slope interpolation. Specifically, the taper function makes \( A_I \) zero if the slope is too large in magnitude or if it is illegal (i.e., the vertical density gradient is negative), and thus when interpolating isopycnal slope values to the particle position, the interpolation scheme would not only need to know the magnitude of the slope but also whether or not it is illegal. An illegal slope in \( x \) or \( y \) cannot be identified merely by its sign, since the horizontal density gradients can be either negative or positive (see the definition of isopycnal slope in (4.6)). As a result, using the taper function to estimate \( A_I \) at a particle position would require estimating the density gradient in each direction at the particle position. This could in turn be accomplished by interpolating gridded density values in much the same manner the OGCM does when estimating slopes at flux faces, or it could be done by interpolating the density gradients computed by the OGCM on the grid cell faces, upon which the slopes used by LLNL MOM are based (see Section 4.2.2.2.2).

However, these density gradient values are weighted averages of density points in neighboring grid cells, and thus interpolating them to a point would not be entirely accurate anyway. Using the gridded density data would be slightly more accurate from a physical standpoint, but the computational expense would be large. More importantly, both techniques would map out an isopycnal slope field that would be different than the one experienced by LLNL MOM, i.e., one with much greater variation in isopycnal slopes and consequently pockets of
sharply varying isopycnal diffusivity that are not in the LLNL MOM simulations. Thus, any approach using the taper function to estimate $A_I$ at a particle position from output values of isopycnal slope, density gradient, or density data leads to increased computational expense, a more noisy diffusivity field, and spurious gradient terms in the random walk model, with likely no benefit to the ultimate goal of allowing the particle model to emulate the tracer transport of the OGCM. These conclusions are supported by several attempts to implement such schemes in the present work.

The other approach to calculating $A_I$ at the particle position is to interpolate it from gridded $A_I$ values output by the LLNL MOM. This approach is attractive because the $A_I$ values are computed internally by the LLNL MOM from the density gradients and are therefore already “tapered” to prevent unrealistic vertical fluxes due to isopycnal diffusion. Although this approach is used in the present study, it is not without problems. Most importantly, by interpolating $A_I$ and isopycnal slope independently, illegally large values of $A_I$ and $S_x$ or $S_y$ can be paired, causing a violation of the constraint placed on $A_I$ by the taper function (4.17). Thus, this approach is feasible only if the interpolation of isopycnal slope is handled in a manner which respects the constraint introduced by the taper function.

Based on these considerations, the scheme currently used to estimate isopycnal slopes $S_x$ and $S_y$ and diffusivity $A_I$ is as follows:

- Isopycnal diffusivity in the x-direction (zonal direction) is calculated by interpolation in x between two neighboring values on the x-faces of grid cells, and is assumed constant in y and z across the grid cell. Likewise, isopycnal diffusivity in the y-direction (meridional direction) is calculated by interpolation in y between neighboring values on the y-faces of grid cells, and is assumed constant in x and z across the grid cell. Thus different values of $A_I$ are used in the x and y directions, $A_{I,x}$ and $A_{I,y}$.

- $S_x$ and $S_y$ are defined in a similar manner, i.e., by interpolation in only a single direction and by assuming that they are constant in the other two spatial directions within a grid cell. Interpolation is performed using *untapered* slope values (note that LLNL MOM by default outputs tapered slopes, even though this is inconsistent with the model’s internal calculations). To prevent pairing a non-zero diffusivity with a steep slope, the
untapered slopes are adjusted in the following manner prior to interpolation:

\[
S_x = \begin{cases} 
S^*_x & A_{I,x} \geq 0.01A_{I,0} \\
0.0063 & A_{I,x} < 0.01A_{I,0}
\end{cases} \quad (4.48a)
\]

\[
S_y = \begin{cases} 
S^*_y & A_{I,y} \geq 0.01A_{I,0} \\
0.0063 & A_{I,y} < 0.01A_{I,0}
\end{cases} \quad (4.48b)
\]

where \(S^*_x\) and \(S^*_y\) are the untapered slopes. This slope adjustment is based on the shape of the taper function, effectively clipping it at 0.01\(A_{I,0}\) to prevent non-zero diffusivity from being paired with a steep slope.

Although the above isopyncal variable definition/interpolation scheme takes some liberties such as defining an \(A_{I,x}\) and \(A_{I,y}\) separately, it has the advantage of allowing most (but not all) of the isopyncal diffusion terms in the random walk equations (4.43) to be defined in a manner that seems consistent with LLNL MOM’s tracer calculations, and which makes interpolation relatively efficient because variables are only interpolated in a single direction. The scheme allows all terms with gradients in \(x\) and \(y\) to be defined naturally, i.e., \(\frac{\partial A_I}{\partial x}\), \(\frac{\partial A_I}{\partial y}\), \(\frac{\partial}{\partial x}[(A_I + \kappa) S_x]\), and \(\frac{\partial}{\partial y}[(A_I + \kappa) S_y]\). Horizontal interpolation to a particle’s position is performed between cell faces in the same manner as for advective velocities in (4.15).

The main drawback to the scheme is that isopyncal diffusivity and slope varies as a step function in the vertical, so that the following terms are left undefined: \(\frac{\partial}{\partial z}[(A_I - \kappa) S_x]\), \(\frac{\partial}{\partial z}[(A_I - \kappa) S_y]\), and \(\frac{\partial}{\partial z}[A_I S^2]\). Of these terms, the first two are of lesser concern since the Gent-McWilliams thickness diffusivity \(\kappa\) is set equal to \(A_I\) in the LLNL MOM, which makes these terms identically zero under normal simulation conditions regardless of the interpolation scheme selected. Thus, the problem is primarily the \(\frac{\partial}{\partial z}[A_I S^2]\) term, and secondarily the two other vertical gradient terms for the special case of \(\kappa = 0\) (used during sensitivity analysis in Appendix 4.A). As a crude approximation, the present RWPT implementation estimates these vertical gradient terms by interpolating \(A_{I,x}\), \(S_x\), \(A_{I,y}\), and \(S_y\) to the particle’s position in \(x\) and \(y\) at the two nearest non-land cell vertical mid-points, and then defining a constant vertical gradient between these two cell mid-points (i.e., the vertical gradient is defined between the two nearest non-land \(z_k^c\) levels). Also, because an \(A_{I,x}\) and
\( A_{I,y} \) have been defined within the present scheme, \( \frac{\partial A_{I,y} S^2}{\partial z} \equiv \frac{\partial A_{I}(S_x^2 + S_y^2)}{\partial z} \) is instead calculated as \( \frac{\partial (A_{I,x} S_x^2 + A_{I,y} S_y^2)}{\partial z} \).

Another drawback to this approach is that the isopycnal diffusion terms which cause vertical particle displacement are based on the \( S_x \) and \( S_y \) values computed at horizontal cell faces. This is different from the internal OGCM calculation, in which all vertical flux terms are computed from the \( S_x \) and \( S_y \) values defined on the vertical grid cell faces. Looking ahead to Appendix 4.A, ongoing investigations into alternative schemes focus on using these vertical face values instead. This approach was not used in the present RWPT implementation for the reason cited above, (i.e., that most RWPT isopycnal terms involve horizontal rather than vertical gradients), but is now being revisited because simulation results with the present RWPT implementation suggest that resolving the vertical terms \( \frac{\partial}{\partial z} [A_{I} S_x], \frac{\partial}{\partial z} [A_{I} S_y], \) and \( \frac{\partial}{\partial z} [A_{I} S^2] \) may significantly improve the RWPT model’s vertical tracer distribution.

To summarize, defining isopycnal slope and diffusivity at a particle location such that LLNL MOM tracer calculations can be emulated is a major difficulty in implementing the RWPT. The present scheme uses a fairly simple approach which keeps the number of interpolations to a minimum and which allows most terms in the RWPT equation to be defined in a natural way. The drawback is that isopycnal slope and diffusivities are treated as step functions in the vertical, which prevents terms involving vertical gradients from being defined in anything more than an approximate manner, most notably the vertical displacement \( \frac{\partial}{\partial z} [A_{I} S^2] \Delta t \) in (4.43c).

4.3.2.2 Numerical scheme

There are several choices which must be made with regard to the numerical scheme beyond the variable interpolation considerations described above. These are mainly related to the order of evaluation of terms, the treatment of variable diffusivity, and the treatment of land boundaries.

In evaluating the displacement of a particle, terms are treated in the following order: (1) isopycnal diffusion, (2) vertical diffusion, and (3) advection. Operator splitting in this fashion is not uncommon in Lagrangian methods, e.g., the random walk scheme of Leone et al. [53] or the Eulerian-Lagrangian scheme of Baptista [4]. Thus, during a particular step,
the particle’s isopycnal slope and diffusivity are first calculated and the particle position is updated to an intermediate position. Then, the vertical diffusivity is estimated at the intermediate position and the particle position is updated to a new intermediate position. Lastly, the advective velocity is calculated at the new intermediate position and the particle position is updated to its final position at \( t + \Delta t \). In addition, there are a number of nuances in the numerical technique for each group of terms, as described below.

As discussed in Sections 4.2.2.2.2 and 4.3.2.1, vertical diffusivity as parameterized in the LLNL MOM (as well as other OGCMs) can be highly variable, at times changing by as much as 6 orders of magnitude over a single grid cell. This creates an unfriendly environment for a Lagrangian particle and requires special treatment. While studies in the past have smoothed the diffusivity profiles and used a small timestep to address such problems (or if convection is simulated by a convective index instead of a large diffusivity, assigned a random depth within the column), the present study uses a novel technique which is described in detail in Chapter 3. The technique is an approximate method which balances fluxes across discontinuities in diffusivity gradient by particle reflection and which places particles in a manner such that the well-mixed condition is preserved without needing to smooth profiles or impose an impractically small timestep. It is for this reason that the vertical diffusivity is treated independently from isopycnal diffusion, i.e., so that the first bracketed term in (4.44c) can be evaluated alone as a 1D diffusion problem.

Since isopycnal diffusivity is linearly interpolated between horizontal grid cell faces (see Section 4.3.2.1), discontinuities in the horizontal gradient of isopycnal diffusivity also exist. To be consistent, the horizontal displacements of the first bracketed terms in (4.44a) and (4.44a) are also treated by the method developed in Chapter 3. These terms are treated first, and then the remaining drift correction terms are added to the particle displacement (which are all vertical displacements since the last bracketed terms in (4.44a) and (4.44a) disappear when \( \kappa = A_I \) is imposed). Experience has shown that the RWPT results are not very sensitive to this modification, but it is implemented nonetheless for consistency.

For the horizontal diffusion of the particles, land boundaries are treated in the following manner. The isopycnal slope and diffusivity at a horizontal land boundary are assigned a value of zero and \( A_{I,0} \) (i.e., the untapered value), respectively. Although the OGCM inter-
nally imposes a zero diffusivity at impermeable boundaries to eliminate diffusive flux, this is not done for particles because this would create low diffusivity zones near boundaries where particles would preferentially collect until a uniform particle density was reached (i.e., the steady state). Since the OGCM does not resolve sub-grid scale concentration distributions, and since the RWPT implementation in the present study attempts to mimic the OGCM calculation, it is more appropriate to make the diffusivity non-zero at the boundary. Unless the neighboring diffusivity is tapered due to a severe isopycnal slope, this approach also makes the diffusivity constant (homogeneous) across the boundary grid cell, which is consistent with the recommendations of [9]. A zero isopycnal slope is imposed at land boundaries for three reasons: (a) because it yielded the best results, (b) because it is consistent with the tapering function, and (c) because it is theoretically consistent with the notion of vanishing $u_{GM}$ and $v_{GM}$ velocities at solid boundaries. The no flux boundary condition is imposed by reflecting particles, either by perfect reflection in the case of homogeneous diffusivity across the boundary grid cell, or by the methodology outlined in Chapter 3 if a gradient exists across the boundary cell. Also, in order for a particle to be able to negotiate a horizontal corner, the $\Delta x$ and $\Delta y$ displacements are converted to velocities (by dividing by $\Delta t$) such that a time to reach adjacent boundaries in each direction can be calculated. It can then be determined whether the particle will round a (convex) corner or be reflected off one of the walls. This is similar to the method implemented by Zhang [75] to handle corners. Finally, when a particle is reflected horizontally off of a wall, the vertical projection of this along-isopycnal motion is also reflected, i.e., the particle is reflected back along the isopycnal slope (using the value at the particle’s initial position).

For the vertical diffusion of particles, the diffusivity at the top and bottom of the ocean were set equal to the value at the neighboring cell faces. As described above, this is done to prevent the creation of low diffusivity zones which would be inconsistent with the OGCM tracer calculation. The no flux boundary condition was satisfied by particle reflection using the methodology described in Chapter 3.

Obviously, the boundary treatment schemes described above are arbitrary; they were selected over the course of many iterations because they tended to minimize artificial particle accumulations, and yield the best agreement between the RWPT and OGCM calculations.
There is no correction method implemented to handle vertical discontinuities in isopycnal diffusivity or slope, even though these variables are treated as step functions in the vertical (see Section 4.3.2.1). Because these terms arise from diffusion along isopycnals, it is not possible to develop a 1D correction method of the type developed in Chapter 3 (which is applied here to handle vertical discontinuities in vertical diffusivity gradient and horizontal discontinuities in horizontal isopycnal diffusivity gradients, as described above). Instead, a timestep is selected that is sufficiently small to prevent particles from crossing more than one grid cell in the vertical from the combined isopycnal vertical displacement terms. Based on an analysis of the taper function used to reduce isopycnal diffusivity when isopycnal slope becomes too large, a timestep of 1,800 seconds was selected. This timestep is small enough that the vertical projection of the along isopycnal stochastic step (the second and third bracketed terms in (4.44c)) with the largest possible combination of $S_x = S_y$ and $A_I$ is less than the thickness of the thinnest LLNL MOM gridcell (25 m) for $|\xi_1| = |\xi_2| = 3$, where $P(|\xi| > 3) \approx 0.0027$. It should be noted that choosing a timestep this large is only possible because the extreme discontinuities which exist in the vertical diffusivity field are addressed by the technique in Chapter 3, without which it would be necessary to smooth vertical diffusivity profiles and use a much smaller timestep in severe regions. Also, to prevent any sort of bias due to variable time-stepping [9, 74], a constant timestep of 1,800 is used throughout. Large-scale simulations of test sites have indicated that results are not sensitive to further decreases in timestep.

The treatment of particle advection warrants some discussion as well. In the present scheme, particle velocity is not updated during the advective step, i.e., Euler’s method is used. Lagrangian studies often use a higher order method to handle advection in variable velocity fields, e.g., a Runge Kutta method. This is especially important for studies which do not consider sub-grid scale diffusion; here particle trajectories should follow the streamlines as precisely as possible since this is the only transport process. For example, Döös [22] developed a way to solve exactly for the particle trajectories using gridded velocity variables, but this method cannot be applied here since the particle must take finite regular steps for the diffusive terms. Also, the computational expense incurred by a higher order method such as a Runge Kutta method is not justified in the present application because less accuracy is
needed due to the “smearing” effect of horizontal diffusion. More importantly, the timestep used in the present study (1,800 s) is small enough that the error introduced should be minimal (e.g., a 3,600 s timestep was applied by Boning and Cox [8] using Euler’s method with acceptable accuracy). Model testing of a 2nd order Runge Kutta scheme yielded no discernible improvement in the RWPT results. Particles reaching impermeable boundaries during the advective step are reflected in the same manner as during the diffusive step (this is necessary because Euler’s method is used; this would never happen if a perfect integration of particle advective trajectories was achieved, since the flow field is non-divergent).

Lastly, it should be noted that higher order schemes such as Runge Kutta are not applicable to treating the diffusive terms in (4.43). As discussed in Chapter 3, the derivation of the random walk equations makes use of Ito calculus, in which the particle displacement is a non-anticipating function (which is consistent with a Markov process and yields the diffusion equation [65]). Although the correction methodology developed in Chapter 3 to handle discontinuities is a way by which a particle’s diffusive characteristics are updated during a step, traditional approaches used to numerically integrate ordinary differential equations such as Runge Kutta do not apply to the diffusive terms in the random walk equation.
4.4 Direct model evaluation results

The present study judges the quality of the RWPT results by their ability to:

- reproduce 100-year simulations of OGCM tracer transport calculations of point releases at multiple sites.

- satisfy the well-mixed condition, i.e., preserve a uniform particle density without severe accumulations.

- reproduce OGCM estimated CO$_2$ sequestration efficiencies, which was tested by conducting a series of simulations in which tracer is released from a point source and allowed to decay at the surface (representing volatilization).

The present section describes only the results of the first two validation approaches as they are direct evaluations of the RWPT calculation. The third validation approach is indirect in that it tests not only the quality of the particle transport calculations but also the bookkeeping methodology employed to conduct the sequestration efficiency calculations. As such, these validation results are discussed in Section 4.5 in conjunction with the application of the RWPT model. Also, the results presented below are further discussed in Appendix 4.A, in the context of proposing future refinements to improve model accuracy.

4.4.1 Tracer release tests

The ability of the particle model to reproduce the transport calculations of the LLNL MOM was mainly judged by conducting a series of tracer release experiments with each model at 5 different depth intervals at each of 8 horizontal regions throughout the ocean (see Tables 4.2 and 4.3, and Figures 4.2 and 4.3). The horizontal and vertical extent of the source regions was defined by the size of the OGCM grid cell at that location. For the LLNL MOM, a tracer concentration of 1 was specified as the initial condition for the grid cell containing the point source. For the particle model, about 50,000 particles were distributed uniformly across the same grid cell.

For all simulations, the RWPT model used a one-year hydrodynamic record from the OGCM, which was repeated 100 times, i.e., for each year of the simulation. In order to
Figure 4-10: Tracer test release simulation sites.

Figure 4-11: Initial condition for the 5 depth intervals for the tracer test release simulations (each depth is simulated separately, or as individual tracers within the same simulation).
Table 4.2: Horizontal extent of tracer test source regions.

<table>
<thead>
<tr>
<th>Region ID</th>
<th>Longitude interval</th>
<th>Latitude interval</th>
<th>LLNL MOM index (i,j)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>20°E - 24°E</td>
<td>60°S - 58°S</td>
<td>(6,16)</td>
</tr>
<tr>
<td>b</td>
<td>180°W - 176°W</td>
<td>10°N - 12°N</td>
<td>(46,51)</td>
</tr>
<tr>
<td>c</td>
<td>172°W - 168°W</td>
<td>34°N - 36°N</td>
<td>(48,63)</td>
</tr>
<tr>
<td>d</td>
<td>32°W - 28°W</td>
<td>28°N - 30°N</td>
<td>(83,60)</td>
</tr>
<tr>
<td>e</td>
<td>76°E - 80°E</td>
<td>12°S - 10°S</td>
<td>(20,40)</td>
</tr>
<tr>
<td>f*</td>
<td>52°W - 48°W</td>
<td>54°N - 56°N</td>
<td>(78,73)</td>
</tr>
<tr>
<td>g</td>
<td>92°W - 88°W</td>
<td>60°S - 58°S</td>
<td>(68,16)</td>
</tr>
<tr>
<td>h*</td>
<td>132°W - 136°W</td>
<td>34°N - 36°N</td>
<td>(34,65)</td>
</tr>
</tbody>
</table>

* Shallower site; only 3 depths simulated here.

Table 4.3: Vertical extent of tracer test source regions.

<table>
<thead>
<tr>
<th>Depth ID</th>
<th>Depth interval (m)</th>
<th>LLNL MOM k index</th>
<th>Depth label*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0 - 25</td>
<td>1</td>
<td>“12.5m”</td>
</tr>
<tr>
<td>2</td>
<td>165 - 229</td>
<td>6</td>
<td>“200m”</td>
</tr>
<tr>
<td>3</td>
<td>846 - 1,039</td>
<td>12</td>
<td>“1000m”</td>
</tr>
<tr>
<td>4</td>
<td>2,708 - 3,060</td>
<td>19</td>
<td>“3000m”</td>
</tr>
<tr>
<td>5</td>
<td>4,600 - 5,000</td>
<td>24</td>
<td>“5000m”</td>
</tr>
</tbody>
</table>

* Descriptive depth used for plot labeling

make the OGCM simulations as directly comparable to the RWPT simulation as possible, the OGCM was run in a special manner. Rather than restarting the OGCM from the last year of the 6,200 year spinup simulation and running it forward for 100 years with a new tracer initial condition, the OGCM was run as a series of 100 one-year simulations. The tracer distribution at the end of each year was used as the initial condition for the following year, but the hydrodynamic initial condition was the same for each year, i.e., using the state of the ocean at the end of the 6,200 year simulation. Thus, the ocean circulation for year 6,201 was looped 100 times in both the RWPT and OGCM simulation. This was done to ensure that both the RWPT and OGCM tracer calculations used the same circulation fields, which would not necessarily have been the case if the OGCM were run continuously (e.g., circulation patterns could change gradually as the physical state of the ocean continues to evolve, or interannual patterns could be captured by the OGCM). A one-year repeat interval was chosen for a number of practical computational reasons, and is consistent with the fact
that the OGCM is forced by average monthly mean climatology.

Comparisons of RWPT and OGCM predicted tracer distributions for each of the 36 locations are given in Figures 4-12 to 4-47 after 100 years of simulation. Each simulation is referenced by the location of the point source used as the initial condition. For example, the simulation result for “d4” corresponds to a source spanning horizontal region d (see Table 4.2 and Figure 4-10) and depth interval 4 (see Table 4.3 and Figure 4-11). In each figure, the models are compared in the following manner. The top three maps on the left side of the page show the concentration after averaging in one spatial dimension (clockwise from top left, averaging was performed in the vertical, zonal, and meridional dimensions). The bottom three maps on the left side show the same quantities as estimated by the RWPT model, after binning particle onto the LLNL MOM grid. In order to facilitate the comparison of the two models, the average concentrations have been normalized by the maximum average concentration computed by the OGCM and particle model in that dimension (i.e., different normalization values are used for each averaging dimension, but the same normalization value is used for the particle and OGCM results in each averaging dimension). The three right panels compare the OGCM and RWPT results after averaging in two spatial dimensions (the top panel is averaged zonally and vertically, the middle panel is averaged meridionally and vertically, and the bottom panel is averaged zonally and meridionally). Again, the concentrations are normalized by the maximum value between the OGCM and RWPT models in each plot. For reference, the initial tracer release location is shown on the right panels (dotted line).

This method of plotting is chosen to facilitate model to model comparisons and to give a comprehensive view of the quality of agreement of the tracer distributions in each direction. Spatial averaging of the particle results in this manner is a useful way to reduce the impact of noise in the predicted concentration field on the evaluation metrics; grid cell by grid cell comparisons are strongly affected by noise. For some locations (e.g., c1 in Figure 4-13), the top left map can be difficult to read. This is an artifact of the plotting method; if the particle results are noisy, then the maximum value used for normalization can be substantially higher than the average. This is most likely when tracer mass is concentrated in small grid cells, i.e., near surface grid cells at low or high latitudes. Nonetheless, this mode of comparison
was on the whole found to be the most informative because it shows the raw particle output in a variety of dimensions, all normalized in a similar manner. Over the course of model development, this was also an effective way to diagnose pathological particle accumulations.

The reader may note that in some cases the OGCM results have slightly negative average concentrations in part of the domain (as signified by a negative value on the colormap legend). This is due to the advection scheme used by the LLNL MOM tracer transport algorithm, which gives rise to numerical oscillations for sharp fronts such as the initial condition used herein. The impact of these oscillations are reduced over time, and although they may be a source of divergence between the RWPT and OGCM results, their impact is expected to be transient. Some consideration of the impact of the advection schemes used by LLNL MOM (standard centered differences with polar filtering in the horizontal and flux-corrected transport in the vertical) is included in Section 4.A.2.3.

The quality of agreement between the RWPT and OGCM results varies by location. Qualitatively, the comparisons are favorable for locations b, c, d, e, f, and h; and less favorable for location a and g. In terms of depth range, releases at the surface (depth range 1, 0-25 m) generally compare less favorably than releases at depth (depth ranges 2-5). As a metric, the 100-year simulations have the benefit of being sufficiently long to allow significant tracer spreading yet short enough such that the final tracer distributions are markedly different depending on the horizontal and vertical location of the initial condition.

The tracer test results demonstrate that in general the RWPT model does a good to excellent job of capturing the zonal and meridional spreading of the tracer. Even for the most challenging locations, a and g, the major features of the horizontal tracer spreading are captured, with maximum errors in average meridional concentrations on the order of 30%, and somewhat closer agreement in the zonal direction than the meridional direction.

The largest differences between the RWPT and OGCM models occur in the vertical dimension, where the RWPT model overpredicts average surface concentrations for many of the shallow releases and some of the deep releases. Particle concentrations are naturally more noisy in the surface layers, and there is evidence that the present scheme tends to empty the top layer somewhat and cause a slight accumulation immediately below the surface layer. However, even if this effect is averaged across the top few layers, the mass accumulation near
the surface is clear for many of the sites. In the milder cases, the error is on the order of 5-10% (e.g., b1-b3, c1-c3, e1-e3, g1-g2, and h1-h2). For the more severe cases the surface concentration error is on the order of 25-150% (e.g., a1-a4, g3-g5). Beneath the surface mixed layer, the agreement between the OGCM and RWPT results is generally better, with lower relative errors even for the worst locations (i.e., a and g). For some locations, the vertical tracer distribution agreement is very good (e.g., a5, b3-b5, c4-c5, d2-d4, e3-e5, f1-f3, h3).

The stochastic nature of RWPT simulations make them difficult to debug unless the error is an obvious one (e.g., a dead zone where particles cease moving), making it difficult to identify the cause of observed differences between the RWPT and OGCM results. Nonetheless, looking ahead to Appendix 4.A, the main culprit for the most severe errors appears to be regions of steep isopycnal slopes, e.g., in the Southern Ocean near locations a and g. For deeper releases, it is believed that shortcomings in the RWPT model’s ability to accurately simulate along isopycnal motion in these regions give rise to an initial error in the vertical tracer distribution, which in turn gives rise to errors in the lateral distribution of the tracer. For shallower releases such as a1, it is unclear whether this same progression is true, i.e., the initial error could instead be caused by errors in horizontal tracer spreading.

Overall, it is emphasized that, despite the observed errors, the RWPT model successfully reproduces the large scale features of the OGCM tracer spreading in the majority of cases. Some noteworthy examples include:

- the complex vertical plume structure in b4, d2, and f3.

- the strong dependence of tracer transport into the Arctic Ocean on release depth for locations b and c (via the Bering Straits) and d (via the Greenland and Labrador Seas).

- the North Atlantic Deep Water (NADW) formation region in and around the Labrador Sea (e.g., f1) from which tracer travels southward at depths of around 1-3 km, consistent with classical thermohaline circulation.

- the strong dependence of tracer transport into the Southern Ocean on release depth from location d: for d1-d3 tracer mass is transported northward while tracer from d4 joins the underlying southward current carrying NADW.
- the extensive horizontal/vertical spreading of tracers originating from g1 and g2, which are carried along the Antarctic Circumpolar Current and then northward along the surface in the Atlantic before sinking in the NADW formation region.

- the marked difference in the distribution of tracers originating from Sites a1 and a2, which differ initially only by about 150-200 m in depth. Tracer from a2 is mainly transported downward (presumably as part of the Antarctic Bottom Water (AABW) formation region in the Weddell Sea) and then northward along the ocean bottom into the Atlantic basin (beneath the southward NADW transport noted above, e.g., for f1). Although tracer from a1 partly shows this same transport, a much larger fraction travels northward via surface currents. Tracer from a1 in effect enters the Atlantic via two branches: one traveling along the bottom (like tracer from a2) and the other traveling along the surface, reaching the NADW formation region within 100 years (like g1 and g2). This split is captured by the RWPT simulation despite the proximity of the two release points and the aforementioned challenging environment of the upper Southern Ocean.

- the large variability in Pacific meridional tracer spreading with depth for locations b and c, with the transport from shallower releases (1-2) spread extensively by the North Pacific subtropical gyre.

- the difference in zonal transport between locations b3 and b5.

- the slow upwelling across the Pacific for deeper releases from locations b and c.

- the transport of mass into the Indian Ocean via the Indonesian Throughflow from location b2 and c2 and to a lesser extent b3 and c3, but not from b4 or c4.

- the dramatic difference in the tracer spreading between releases at e3 and e4.

- the extensive ventilation of location g in the Southern Ocean, and the ventilation of tracer from deeper releases in the Indian Ocean (e.g., e4) as they enter the Southern Ocean.
• the general ability of particles to negotiate complex land boundaries such as in the Sea of Japan (e.g., h3), in the Gulf of Mexico and Caribbean Sea (e.g., f3), around the Indian Peninsula (e.g., e3), and the Indonesian Throughflow (e.g., b2 and c2).
Figure 4-12: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location a1 (dotted line) after 100 years of simulation.
Figure 4-13: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location a2 (dotted line) after 100 years of simulation.
Figure 4-14: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location a3 (dotted line) after 100 years of simulation.
Figure 4-15: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location a4 (dotted line) after 100 years of simulation.
Figure 4-16: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location a5 (dotted line) after 100 years of simulation.
Figure 4-17: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location b1 (dotted line) after 100 years of simulation.
Figure 4-18: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location b2 (dotted line) after 100 years of simulation.
Figure 4-19: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location b3 (dotted line) after 100 years of simulation.
Figure 4-20: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location b4 (dotted line) after 100 years of simulation.
Figure 4-21: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps
and solid line) predicted tracer concentrations for release location b5 (dotted line) after 100
years of simulation.
Figure 4-22: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location c1 (dotted line) after 100 years of simulation.
Figure 4-23: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location c2 (dotted line) after 100 years of simulation.
Figure 4-24: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location c3 (dotted line) after 100 years of simulation.
Figure 4-25: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location c4 (dotted line) after 100 years of simulation.
Figure 4-26: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location c5 (dotted line) after 100 years of simulation.
Figure 4-27: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location d1 (dotted line) after 100 years of simulation.
Figure 4-28: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location d2 (dotted line) after 100 years of simulation.
Figure 4-29: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location d3 (dotted line) after 100 years of simulation.
Figure 4-30: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location d4 (dotted line) after 100 years of simulation.
Figure 4-31: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location d5 (dotted line) after 100 years of simulation.
Figure 4-32: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location e1 (dotted line) after 100 years of simulation.
Figure 4-33: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location e2 (dotted line) after 100 years of simulation.
Figure 4-34: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location e3 (dotted line) after 100 years of simulation.
Figure 4-35: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location e4 (dotted line) after 100 years of simulation.
Figure 4-36: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location e5 (dotted line) after 100 years of simulation.
Figure 4-37: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location f1 (dotted line) after 100 years of simulation.
Figure 4-38: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location f2 (dotted line) after 100 years of simulation.
Figure 4-39: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location f3 (dotted line) after 100 years of simulation.
Figure 4-40: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location g1 (dotted line) after 100 years of simulation.
Figure 4-41: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location g2 (dotted line) after 100 years of simulation.
Figure 4-42: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location g3 (dotted line) after 100 years of simulation.
Figure 4-43: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location g4 (dotted line) after 100 years of simulation.
Figure 4-44: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location g5 (dotted line) after 100 years of simulation.
Figure 4-45: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location h1 (dotted line) after 100 years of simulation.
Figure 4-46: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location h2 (dotted line) after 100 years of simulation.
Figure 4-47: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location h3 (dotted line) after 100 years of simulation.
4.4.2 Uniformity tests

The tracer release tests described above attempt to validate the RWPT results by comparing them to OGCM tracer calculation. While this is perhaps the most relevant comparison since the goal of the model development is to design an RWPT model capable of reproducing the OGCM calculation, it is complicated by sources of error in the OGCM numerics (e.g., the negative concentration oscillations caused by sharp fronts in concentration). It is therefore useful to test the RWPT model on its own merits. Although comparison to an analytical solution is not possible given the complex geometry and time varying flow and diffusivity fields, the RWPT can be tested on its ability to meet the well-mixed criterion, i.e., to preserve the uniformity of an initially uniform concentration field. As discussed in Section 4.3.1.4, the RWPT model should be able to satisfy the well-mixed condition as a whole or by using different subsets of variables.

In order to achieve a uniform particle density for the initial condition, the following approach was taken to initialize the uniformity tests.

- The ocean was divided up into $N_\phi$ latitude bands, with central latitude

$$\phi_c(j) = \frac{180}{N_\phi} \left( j - \frac{1}{2} \right) - 90$$

(4.49)

- The number of zonal particle positions for the southernmost latitude band ($N_\lambda(1)$) was specified, and the number of particles for all other latitude bands was scaled off of this number:

$$N_\lambda(j) = N_\lambda(1) * \frac{\cos \phi_c(j)}{\cos \phi_c(1)}$$

(4.50)

The central longitude is then assigned (ignoring land boundaries) as

$$\lambda_c(j, i) = \frac{360}{N_\lambda(j)} \left( i - \frac{1}{2} \right)$$

(4.51)

- At each zonal position, $N_d$ particles are given a random depth in the interval 0-5 km, which is the maximum depth of the ocean as resolved by LLNL MOM. Thus, the total
The number of particles distributed across the domain (ignoring land boundaries) is

\[ N_p^* = N_d \sum_{j=1}^{N_\phi} N_{\lambda}(j) \]  

(4.52)

- The actual number of particles, \( N_p \), is arrived at by removing those particle initial positions which are in land using the LLNL MOM grid and topography.

Figures 4-48 to 4-50 show the results of a 200 year simulation using an initial condition created in this manner at \( t = 0 \), \( t = 100 \) years, and \( t = 200 \) years, respectively. In this case \( N_\phi = 90 \), \( N_{\lambda}(1) = 10 \), and \( N_d = 5 \) were used such that \( N_p^* = 164,360 \), which after land filtering yields \( N_p = 84,499 \). The initial condition (Figure 4-48) is included here to illustrate that the smaller grid cell sizes in the high and low latitudes give rise to considerable noise. This noise would be reduced by using more particles, but it should be noted that this same effect contributes to some of the noise observed in the results from Section 4.4.1. To achieve smoother concentration fields, multiple years of output are averaged together when analyzing the results of these uniformity tests.

Averaging concentrations over the last 50 years of the end of the simulation (\( t = 200 \) years, Figure 4-50) reveals a fundamental bias in the present implementation of the RWPT model; particles preferentially migrate toward the surface, as if acted upon by an artificial upwelling velocity. The average surface concentration exceeds the average concentration at depth by about 30%. Thus, the RWPT model does not meet the well-mixed condition. Based upon intermediate results in Figure 4-49, the effect is progressive over the course of the simulation (e.g., surface concentrations are only about 20% higher after 100 years, see Figure 4-49). The observed bias is consistent with the tracer release tests of Section 4.4.1, where surface concentrations were observed to routinely exceed those computed by the OGCM.

On a smaller scale, there is also a clear tendency for the particle model to preferentially accumulate particles in the second layer, leading to a sharp gradient between the first two thin layers (25 m thick). Experience has shown that this is likely caused by the troublesome \( \frac{\partial A_i S^2}{\partial z} \) term, which as discussed previously is poorly defined in the present isopycnal variable definition/interpolation scheme (see Appendix 4.A for additional discussion). The impact
Figure 4-48: RWPT uniformity test initial condition. Maps show average concentration normalized by the maximum concentration in each direction. Line plots show average concentration normalized by the mean concentration in each direction.
Figure 4-49: RWPT uniformity test results averaged in time over interval 50-100 years. Maps show average concentration normalized by the maximum concentration in each direction. Line plots show average concentration normalized by the mean concentration in each direction.
Figure 4-50: RWPT uniformity test results averaged in time over interval 150-200 years. Maps show average concentration normalized by the maximum concentration in each direction. Line plots show average concentration normalized by the mean concentration in each direction.
of this bias is, however, secondary to the overall vertical bias noted above in that it mainly concerns the tracer balance between the top two layers.

In addition to the vertical bias, there is also some evidence of particle accumulation at high and/or low latitudes. However, the pattern is by no means as clear as in the vertical. The average from the interval 150 to 200 years indicates particle accumulation near the North Pole and particle deficit near the South Pole, while the average concentrations from 50 to 100 years suggest minimal accumulation near the North Pole and significant accumulation near the South Pole. These effects can therefore at least partially be attributed to the noisiness of the particle concentration field, which is most pronounced in low and high latitudes where the grid cells are smaller and generally shallower. For the 100-year tracer release simulations, such a consistent bias was not observed in the meridional direction. None of the release points suffered from significant over-prediction (relative to the OGCM) of concentrations at high latitudes. Release point a shows some accumulation near Antarctica but results at other locations suggest that this over-prediction is more likely to be due to the complex land geometry of Drake Passage than a consistent bias. For example, whether the particle model under- or over-predicts low latitude concentrations at locations e and g depends on the release depth, which suggests that low latitude results are subject to noise in the particle results. Overall, it is concluded that the RWPT implementation does reasonably well at satisfying uniformity in the zonal and meridional directions\(^5\), although the issue warrants revisiting once the aforementioned vertical bias is resolved.

Although it is clear that the well-mixed condition is not met by the particle model, this result is not altogether unexpected given the choices that were necessary with regard to variable definition and interpolation. Brickman [9] suggests that the well-mixed condition may be too stringent a test as the process of spatially interpolating output variables invariably leads to accumulations, and that a particle model may still be sufficiently accurate in spite of this. Indeed, the most important question for the present study is not whether the RWPT model can exactly satisfy the well-mixed condition, but rather whether the errors inherent in the RWPT model (manifested by violations of the well-mixed condition) compromise sim-

\(^5\)Without the meridional correction term described in Section 4.3.1.3, an observable poleward bias consistently occurs in both the uniformity and tracer release tests with average concentrations near the poles routinely reaching twice the expected value within 50 years of simulation.
ulation accuracy to an extent that the model cannot be used for its intended purpose, i.e., to reproduce the tracer transport calculations of the OGCM. The present RWPT implementation was selected primarily on the basis of the 100-year tracer test release simulations (see Section 4.4.1). Of all schemes tested, the present one achieves the best overall agreement between the RWPT and OGCM results, while still respecting the basic random walk equations and the tracer transport formulation of the OGCM. As noted previously, the present scheme’s ability to reproduce the OGCM tracer transport calculations is impressive in many respects.

The fact that the uniformity test yields a clear bias, and that this artificial upwelling is consistent with the observed discrepancies between RWPT and OGCM predictions for point sources, suggests that improvements to the RWPT scheme are possible. Appendix 4.A addresses this issue by attempting to diagnose which terms in the model are the cause of the existing bias, and proposes a path for ongoing and future improvements. The reader should take note that additional evaluation of the quality of agreement between the RWPT and OGCM predictions are provided in the next section (4.5), in the context of a practical application of the model to estimate carbon sequestration efficiencies. The analysis and discussion provided in Appendix 4.A, while central to understanding sources of error in the current RWPT implementation and how it may be improved, is not a necessary prerequisite for understanding the application in Section 4.5. Nonetheless, some readers may find it useful to review Appendix 4.A before proceeding.
4.5 Application: estimation of CO₂ sequestration efficiencies

As noted in Section 4.1, a central consideration in implementing ocean storage is the sequestration efficiency of the injection site, i.e., the fraction of injected CO₂ remaining in the ocean over time. Even if the viability of a site might ultimately be controlled by biological impact concerns (see Chapter 5), the expected efficiency would dictate whether the site is even worth considering. As such, numerous past investigations have applied OGCMs to characterize the sequestration characteristics of particular sites (e.g., [2, 3, 42, 13, 12, 20, 19, 23, 46, 47, 55, 56, 58, 59, 68, 73]). In an effort to achieve some consensus, the GOSAC study [42] applied eight coarse-gridded OGCMs with standardized simulation protocols to seven sequestration sites near major population centers, evaluating three different injection depths at each site. Results have been mixed; although some common trends have emerged (see Section 4.5.2.3), OGCMs still disagree about the relative efficiency of particular locations [42, 43], which can be attributed to differing OGCM resolutions, structure and parameterizations [55].

One inherent limitation to these types of investigations is that they are computationally expensive, and as a result it has only been practical to evaluate a fairly small number of sites. It can be argued that this is not a severe constraint, since proximity to a large CO₂ source presumably places a limit on how much of the ocean might realistically be viable for sequestration. Still, it is possible that OGCMs can disagree about the efficiency of a specific site, even if their larger scale circulation patterns are consistent. Testing this hypothesis requires evaluating more sites, which can quickly become an arduous task with traditional OGCM point source simulations, since each additional site requires a substantial computational effort. The adjoint sensitivity approach of Hill *et al.* [39] is vastly more efficient, capable of calculating the sequestration efficiencies of an entire domain with a computational effort equivalent to simulating a handful of sites in the traditional manner. The resulting predictions were unique in that spatial and temporal trends within the OGCM could be visualized to an extent not previously possible.

Here we illustrate how the RWPT model can be employed to achieve this same goal.
The approach involves conducting simulations which ensure that each grid cell of interest (e.g., all cells at a particular depth) is sampled by a sufficiently large number of particles so as to allow the true distribution of transport pathways from that cell to be adequately represented by the ensemble of ensuing particle trajectories. The sequestration efficiency for individual cells is calculated via a book-keeping method that considers the time particles spend near the ocean surface after leaving a particular cell. Although residence times are not computed explicitly in the application presented herein, the sequestration efficiency metric is closely related to the notion of a mean residence time (see below), and thus estimating sequestration efficiency is a specific example of the more general application of estimating OGCM residence times.

Like past sequestration investigations, the simulations considered herein will make a number of simplifications regarding the modeled CO$_2$. The complex biogeochemical interactions that govern carbon transport in the oceans will not be modeled and an abiotic ocean will be assumed, as in the GOSAC study [42]. The only sink of carbon will be loss to the atmosphere in the upper layer of the ocean, and once lost, it cannot reenter the ocean. Following Hill et al. [39], this loss can be treated as a first order decay in the absence of a return flux from the atmosphere, by assuming that the injected CO$_2$ is a small perturbation to surface layer pCO$_2$ and that the atmospheric pCO$_2$ is constant. Hill et al. [39] estimates a characteristic timescale of decay of 1 year as reasonable, and applies a decay rate of 1 yr$^{-1}$ to describe tracer loss in a 50 m thick surface layer. This same approach will be adopted below, i.e., outgassing to the atmosphere will be parameterized as a first order decay with rate $k$.

Because the sequestration efficiency metric used herein is based on a calculation that ignores “pushback” from the atmosphere or any changes in buffering capacity resulting from the perturbation to the surface carbonate chemistry, it is best suited to characterize the relative efficiency of sites. Previous studies have varied in their complexity of treating the air-water interface. For example, the GOSAC study used a prescribed time-variable atmospheric concentration trajectory; thus, even though no atmospheric “pushback” of outgassed CO$_2$ was included, the GOSAC simulations did consider future build-up of atmospheric CO$_2$ and the corresponding reduction in outgassing as well as the increase in surface dissolved inorganic carbon (and consequently a reduction in buffering capacity). The present study
could be modified to take such considerations into effect by using a surface decay rate \( k \) that varies in time and/or space, but only in an approximate manner since the true flux depends on the local pCO\(_2\) gradient. Thus, the sequestration efficiencies calculated herein are best suited as measures of relative efficiency between sites.

The reader should also note that the sequestration efficiency metric used herein, like those of most past studies, is somewhat artificial in that it does not allow CO\(_2\) to re-enter the ocean; eventually the efficiency of all sites will reach zero as all injected CO\(_2\) interacts with the atmosphere. Under present conditions, the atmosphere and oceans are expected to equilibrate over time such that 80% of all CO\(_2\) will reside in the oceans and 20% in the atmosphere [43] \textit{regardless} of whether it is initially stored in the ocean or the atmosphere. Thus, one could argue that in reality the long-term (millenia) sequestration efficiency of all ocean CO\(_2\) injection sites is about 80%. However, such a definition is hardly useful in site selection; the point of ocean sequestration is to reduce peak atmospheric CO\(_2\) concentrations while we develop cleaner energy sources and thus the metric of interest in selecting sites is the efficiency of one site relative to another. Evaluation of the overall benefit of the injection would require accounting for a multitude of other factors such as the response of the atmosphere, the avoided climate impacts, the economic and energy cost of sequestration, etc. Rather, the goal of the present study is to use the sequestration efficiency calculation as an illustration of the value of the RWPT model developed in earlier chapters. As such, we take advantage of simplifications made by past studies of ocean carbon sequestration.

### 4.5.1 Sequestration efficiency and mean residence time

The concept of a mean residence time is commonly used in tracer release experiments, as discussed in Hilton \textit{et al.} [40] and references therein. For an instantaneous tracer release, a time-varying mass function \( m(t) = \int c(\mathbf{x}, t) \, dV \) can be defined for the domain of interest, and the mean residence time of tracer mass is the zeroth moment of this function:

\[
\tau = \frac{1}{m_0} \int_0^{\infty} m(t) \, dt \tag{4.53}
\]
where \( m_0 = m(t = 0) \). Alternately, the mean residence time can be computed as the first moment of the rate of mass loss from the system \( r(t) \),

\[
\tau = -\frac{1}{m_0} \int_0^\infty \frac{dm}{dt} t \, dt = \int_0^\infty r(t) t \, dt.
\]  

(4.54)

These two definitions are interchangeable since by definition \( r(t) = -\frac{1}{m_0} \frac{dm}{dt} \), and demonstrate that the residence time can be calculated from observations of either \( m(t) \) or \( r(t) \). These basic definitions can also be applied, with some extension, to the case of a continuous tracer release in either steady or unsteady conditions (see [40]).

In the context of carbon sequestration, it is convenient to introduce the notion of a sequestration efficiency for an instantaneous source, the fraction of injected mass remaining as a function of time since injection:

\[
q(t) = \frac{m(t)}{m_0}.
\]  

(4.55)

Although the tracer release in realistic sequestration scenarios is continuous or periodic depending on the method of injection (e.g., [1]), it is treated here using the framework for instantaneous sources, since a continuous source can be treated as a series of instantaneous sources. The analysis of residence times for true continuous sources typically considers tracers which exist in a steady state, where the mean residence time is simply the steady state mass of tracer in the system divided by the loading rate. For carbon sequestration calculations, however, the focus is not on the steady state concentration of injected CO\(_2\) that would eventually be achieved but rather how long CO\(_2\) injected today would remain sequestered from the atmosphere. Thus, it is best conceptualized as an instantaneous source.

The reader should keep in mind that throughout the present study, the term \textit{sequestration efficiency} is used to describe an instantaneous source unless otherwise noted. This choice is made for generality, since the sequestration calculation is but one example of a larger class of residence time type calculations, and as discussed in Hill et al. [39], the notion of a mean residence time generally applies to an instantaneous source. As such, the definition of
sequestration efficiency used here is not the same as in most other studies of ocean carbon sequestration, where it is usually defined as the retained fraction of total mass injected over time (e.g., [39, 55]). To keep these quantities distinct, the term *injection efficiency* \( q_i(t) \) is borrowed from the GOSAC study [42] to describe the efficiency of a continuous, temporary or intermittent source:

\[
q_i(t) = \frac{m(t)}{\int_0^t m_i(t) dt}
\]

where \( m_i(t) \) is the mass injection rate and \( t \) is the time since the injection first started. For example, in the OCMIP protocols used for the GOSAC study [42] and numerous other studies (e.g., [55]), \( m_i(t) = 0.1 \) PgC/yr for the first 100 years of simulation, and zero thereafter, i.e., the source is temporary.

The power of the Lagrangian method to estimate residence time statistics is derived from the discrete nature of the calculation; the travel time of each particle from its initial position to its final position is always known. Two methods by which this characteristic can be used to estimate mean residence times and sequestration efficiencies are outlined below. The first and simplest form of the calculation is the traditional *first passage* problem, where the goal is to estimate the mean travel time between a source and exit region over a large number of particle trajectories. This corresponds to a residence time estimate for a tracer which is conservative everywhere except for the exit region, where it instantly disappears. For the purpose of the present discussion, this can be thought of as a tracer which undergoes first order decay with a rate constant \( k \) that is zero outside of the exit region and infinite within the exit region. While such first passage calculations are certainly relevant to the evaluation of carbon sequestration, they are rather different than the CO\(_2\) residence time or sequestration efficiency that a typical OGCM would estimate. As noted previously, OGCM calculations would generally compute a flux based on a local gradient of CO\(_2\) across the air-sea interface, but in the present calculations the simplifying assumption of Hill *et al.* [39] is used, namely, it is assumed that CO\(_2\) loss at the air-sea interface can be modeled as a first order decay with an appropriately selected decay rate. Thus, the second method outlined below allows a finite \( k \) to be used in the calculation of sequestration efficiency and mean residence time.
4.5.1.1 First passage calculations

Given a set of \( N \) particles initially distributed over a source region \( S = \{ x_{S,1} \leq x \leq x_{S,2} \} \) (\( x_{S,1} = x_{S,2} \) for a point source), the travel time required to reach the exit region \( E = \{ x_{E,1} \leq x \leq x_{E,2} \} \) can be calculated for each particle, \( t(x_n \in E) \). The distribution of arrival times provides an estimate of the mean travel time \( T \) from \( S \) at \( t = 0 \) to \( E \) over \( N \) realizations:

\[
T = \frac{1}{N} \sum_{n=1}^{N} t(x_n \in E),
\]

Equation (4.57) is a discrete form of (4.54) with \( m_0 = N \) and

\[
r(t) = \delta(t - t(x_1 \in E)) + \delta(t - t(x_2 \in E)) + \ldots + \delta(t - t(x_N \in E)),
\]

and thus \( T \) is the numerical estimate of the mean residence time \( \tau \). Computationally, \( T \) is calculated by placing \( N \) particles in \( S \) at \( t = 0 \) and simulating particle motion until all particles have encountered region \( E \). This is presumably the method of calculation used by Spivakovskaya [67], in which the residence time prediction of a random walk model for settling particles was shown to agree with the analytical solution for a one-dimensional ocean domain.

In the context of carbon sequestration, the time to reach the ocean surface from a CO\(_2\) injection site is the quantity of interest (i.e., \( E = \{ z < z_s \} \), where \( z_s \) is the depth of the surface layer, and \( S \) is the extent of the injection source after near-field mixing has occurred). Also, since ocean circulation can be assumed quasi steady state with cyclical variability, the quantity of interest is the average time for particles to reach \( E \) over many different starting times rather than at a specific time as in (4.57). Thus the more relevant quantity in this context is

\[
\bar{T} = \frac{1}{N} \sum_{n=1}^{N} \sum_{l=1}^{L_n} \left[ t(x_{n,l} \in E) - t(x_{n,l} \in S) \right],
\]

where \( L_n \) is the number of times \( t(x_n \in S) \) that particle \( n \) was known to have been in \( S \) for which the first subsequent time \( t(x_n \in E) \) that the particle reached \( E \) is also known. Computationally, \( \bar{T} \) is calculated by sampling particle positions at a regular interval and also
Figure 4-51: Illustration of first passage travel time estimation in Equation (4.59). When a particle enters the surface layer (solid circles; Region E in (4.59)), a travel time from the injection layer (Region S in (4.59)) can be calculated for all previous times at which the particle was known to reside in that layer (hatched circles). In the case shown, a single particle trajectory generates 7 estimates of travel time from S to E.

recording all times at which each particle enters region E. Thus, each regular output provides a new ‘initial’ position for that particle. The calculation in (4.59) is more computationally efficient than the one in (4.57) because a single particle may generate multiple estimates of travel time for the same source region, as illustrated in Figure 4-51. Furthermore, a single simulation can be used to estimate \( \bar{T} \) for multiple source regions, since each arrival time in Region E yields an estimate of travel time for all known prior locations of that particle. Equation (4.59) can be regarded as an extension of (4.57) that takes full advantage of the nature of the Lagrangian simulation; (4.57) is recovered from (4.59) if only the positions of particles at the start of the simulation are considered (i.e., \( L_n = 1 \) and \( t(x_{n,t} \in S) = 0 \) for all particles). There are two special cases in which \( \bar{T} \) and \( T \) are estimates of the same quantity. The first is a steady flow field, where the travel time is not a function of \( t(x_{n,t} \in S) \). The second is a cyclic (repeated) flow field in which the particle positions are always sampled at a particular time in the flow field cycle (i.e., with a sampling frequency equal to the inverse of the flow field period).

It is important to note that the mean travel time \( \bar{T} \) in (4.59) is only meaningful if the
flow field is in steady state or quasi-steady state. If not, then the mean time $T$ in (4.57) must be used, i.e., travel time must be computed only for a specific starting time. For most ocean applications, and certainly for the carbon sequestration calculation, a quasi steady state with cyclical variability can be assumed and $\bar{T}$ is a valid quantity. Computationally, consideration should be given to the initial particle distribution used and the sampling frequency of intermediate particle positions so that the starting times for the individual travel time estimates ($t(x_n \in S)$) are somewhat evenly distributed over the cycle of variability in the flow field. One way that this may be helped is to distribute the introduction of the $N$ particles in time over the first cycle of variability, rather than just introducing them all at the start of the simulation.

The simulation length used in estimating $\bar{T}$ must be considered carefully. Simply ending the simulation at a predetermined time and using all starting times $t(x_n \in S)$ for which a subsequent arrival time $t(x_n \in E)$ exists would bias $\bar{T}$ to be low, since the longest travel times would be truncated by the end of the simulation and not be included in the calculation. To prevent this bias, a cutoff time $t_{\text{max}}$ should be chosen beyond which starting times are no longer considered, and the simulation should be continued until all valid starting times ($t(x_n \in S) < t_{\text{max}}$) have a corresponding arrival time in region $E$. Only in this case are the longest trajectories captured such that $\bar{T}$ can be considered a true estimate of the mean residence time $\tau$ for tracer mass in $S$. The value of $t_{\text{max}}$ should be chosen so that an adequate number of starting times $t(x_n \in S)$ have been sampled. This same ambiguity is not encountered in simulations to estimate $T$ in (4.57), which must only be continued until all $N$ particles reach region $E$.

4.5.1.2 Sequestration efficiency calculations

As discussed previously, the calculation of sequestration efficiency in the present framework requires a modification of the first passage calculation to include a first order tracer decay in the exit region. The goal is to calculate the fractional mass that remains in the ocean as a function of time after release, i.e., $m(t)$ in (4.53). Thus, rather than tracking the times at which each particle enters the exit region $E$, the total time spent by each particle in $E$ is tracked and recorded at regular intervals over the entire simulation. For a particle
in the source region $S$ at $t=0$, the mass of that particle evolves by

$$M_n = M_{n,0} \, e^{-k\tilde{t}_n(t)},$$

(4.60)

where $\tilde{t}_n(t)$ is the total time spent in exit region $E$ and $M_{n,0}$ is the initial mass. Given $N$ particles initially in the source region, the mass remaining in the system is thus

$$M(t) = \sum_{n=1}^{N} M_{0,n} \, e^{-k\tilde{t}_n(t)},$$

(4.61)

where $M(t)$ is the numerical estimate of $m(t)$ in (4.53). The initial mass of each particle is equal in these calculations, and thus the estimated sequestration efficiency for region $S$ is

$$Q(t) = \frac{M(t)}{M_0} = \frac{1}{N} \sum_{n=1}^{N} e^{-k\tilde{t}_n(t)}$$

(4.62)

Computationally, it is noteworthy that a single simulation of particle trajectories can be used to estimate $Q$ for any value of $k$. Equation (4.62) is analogous to (4.57) in that it only considers particles in the source region at the start of the simulation. Indeed, the estimate of mean residence time in (4.57) can be duplicated using the approach in (4.62) by letting $k \to \infty$ and taking the zeroth moment of $Q(t)$, i.e., by using the definition of $\tau$ in (4.53).

As in the preceding section, the quantity of greater interest in a steady or cyclic flow field is the average sequestration efficiency over many different starting times ($\bar{Q}(t)$), which is given by

$$\bar{Q}(t) = \frac{1}{N} \sum_{n=1}^{N} \sum_{l=1}^{L_n} e^{-k[\tilde{t}_n(t')-\tilde{t}_n(t'(x_{n,l} \in S))]},$$

(4.63)

where $t$ is actual time and $t'$ is the simulation time. In other words, the sequestration efficiency for region $S$ at time $t$ since injection is calculated by considering all particles that passed through $S$ at some point during the simulation, where particles with $L_n > 1$ contribute multiple data points. For each particle starting time in $S$ during the simulation ($t'(x_{n,l} \in S)$), the fractional mass is initially 1 and decays by $e^{-k[\tilde{t}_n(t')-\tilde{t}_n(t'(x_{n,l} \in S))]},$ i.e., based on the time spent by that particle in the exit region after the starting time $t'(x_{n,l} \in S)$. The
approach in (4.63) may be considered a generalization of the first passage calculation in (4.59) in that the same estimate of residence time would be arrived at by letting $k \to \infty$ and taking the zeroth moment of $\bar{Q}(t)$.

The particle initialization and sampling issues discussed in the implementation of (4.59) for cyclic flow fields also applies to (4.63), i.e., that particle initialization and sampling should be spread out across the cycle of variability being studied. Also, if (4.63) is to be used to estimate a mean residence time, then the considerations of simulation length and cutoff time also apply here such that bias in $\tau$ is avoided. However, such considerations only affect the calculation of $\tau$; simulations which truncate long trajectories do not impart bias in $Q(t)$ or $\bar{Q}(t)$. Rather, all particles in region $S$ at the start of the simulation will yield values of $Q(t)$ or $\bar{Q}(t)$ up until the end of the simulation (even if they never reach the exit region $E$).

The end of the simulation truncates the $Q(t)$ or $\bar{Q}(t)$ function but does not bias it. It also does not affect the accuracy of the $Q(t)$ calculation, since the number of data points is in this case constant, i.e., it is the number of particles in $S$ at the start of the simulation. It does, however, affect the quality of the calculation of $Q(t)$, since the number of data points used to define $\bar{Q}(t)$ is not constant in $t$. The number of data points available decreases with increasing $t$, since at any time $t'$ in the simulation the number of particle starting times for which $t' - t' (x_{n,t} \in S) \geq t$ decreases as $t$ increases.

### 4.5.1.3 A note on evaluating continuous sources

All preceding equations for sequestration efficiency are only applicable to instantaneous sources because this is the quantity most relevant to the general utility of the RWPT approach in residence time estimation. Nonetheless, since past sequestration studies have generally focused on what in the present study is deemed an injection efficiency (see (4.56)) for a continuous or temporary source, some thoughts on how this fits into the above discussion is provided. It is assumed that the source is either continuous with a constant injection rate $m_i^*$ (as in [39]) or temporary with a constant injection rate $m_i^*$ that ends after some initial period $t_i^*$ (as in [42]). Both can be represented as

$$m_i(t) = m_i^* H(t_i^* - t)$$

(4.64)
where \( t_i^* \) is a finite constant for a temporary source and \( t_i^* = t \) for a continuous source. Here \( H \) is a Heaviside function:

\[
H(x) = \begin{cases} 
0 & x < 0 \\
1 & x \geq 0 
\end{cases}
\]  

(4.65)

Even if particles are injected at different times during a simulation (e.g., over the course of one cycle of a cyclic flow field), they still provide an estimate of the instantaneous sequestration efficiency if used in conjunction with (4.63). Estimating injection efficiency for a source of the form (4.64) could be achieved by releasing particles at regular intervals until \( t_i^* \) and applying

\[
Q_i(t) = \frac{\sum_{n=1}^{N} H(t - t_n^0) e^{-k t_n (t - t_n^0)}}{\sum_{n=1}^{N} H(t - t_n^0)}
\]  

(4.66)

where \( Q_i(t) \) is the numerical estimate of \( q_i(t) \) and \( t_n^0 \) is the release time for particle \( n \) of \( N \) total particles to be released over the simulation. However, this would not be efficient because, like (4.62), it does not allow a single particle to generate multiple datapoints as in (4.63). Rather, the easiest way to do the calculation is to first estimate \( \bar{Q}(t) \), the average observed sequestration efficiency for an instantaneous source from (4.63), and then use this function to evaluate the fraction of injected mass remaining via a convolution integral:

\[
Q_i(t) = \frac{\int_{t_i^0}^{t_i^*} \bar{Q}(t - \tau') m_i^* d\tau'}{\int_{t_i^0}^{t_i^*} m_i^* d\tau'} = \frac{\int_{t_i^0}^{t_i^*} \bar{Q}(t - \tau') d\tau'}{t}
\]  

(4.67)

where the latter equality exists because \( m_i^* \) is a constant. This function must be estimated numerically since the function \( Q(t) \) is itself a numerical estimate. This of course assumes that the instantaneous sequestration time for any release time is well described by the average value \( \bar{Q}(t) \).

### 4.5.1.4 Benefits of the Lagrangian approach

It is important to note the computational benefits offered by the Lagrangian method in both the first passage and sequestration efficiency calculations. Most importantly, a single
simulation can in theory be used to characterize the mean residence time or sequestration efficiency for every region of the domain and any value of the exit region decay coefficient $k$. Furthermore, this process can be made much more computationally efficient by allowing particles to generate multiple starting times throughout the simulation (i.e., as in (4.59) and (4.63)). While computer resources will dictate the length of the simulation and the number of particles that can be used, the benefits offered by the particle tracking approach are considerable if employed creatively.

In addition, relative to OGCMs the Lagrangian simulation has the computational advantage of not being tied to a grid. While the accuracy of both OGCM and Lagrangian calculations can be improved by increasing the resolution of the OGCM’s hydrodynamic calculations, there is a vast difference in the computational expense of higher grid resolution for each case. For the OGCM, the transport simulation time increases dramatically with increasing grid resolution because there are more grid cells to compute and smaller timesteps must be used to ensure stability. For the Lagrangian model, the main penalty is simply more hydrodynamic data to search through in finding interpolation variables at each particle position. Thus, an RWPT simulation could be used to diagnose the sequestration efficiencies of a high resolution OGCM by looping through a short hydrodynamic record (e.g., 1 year as in the present calculations) even if the grid is too fine for even a single OGCM long-term calculation to be realistically possible.

### 4.5.2 Sequestration simulations

The RWPT was employed to estimate sequestration efficiencies using the approach outlined in the Section 4.5.1. In order to test the RWPT calculations, a series of OGCM simulations were conducted to provide validation, the results of which are summarized in Section 4.5.2.2. The utility of the overall approach is demonstrated in Section 4.5.2.3, where spatial distributions of sequestration efficiencies are presented. Numerical considerations are discussed briefly in Section 4.5.2.4. Finally, related applications and the relative advantages of the adjoint method are also considered.
4.5.2.1 Simulation approach

Two sets of RWPT sequestration simulations were conducted. The goal of the first set was to compare the RWPT and OGCM computed sequestration efficiency at the test sites, and to illustrate the sequestration calculation approach. Instantaneous releases of particles were simulated at the deeper of the tracer release sites described in Section 4.4.1: horizontal locations a-g and depth ranges 3-5 (a total of 20 release locations). The two shallowest depth ranges (1-2) were not tested in this manner because they are either at or near the surface, for the most part yielding very short sequestration times. The RWPT model was initialized and run for 2,000 years in the same fashion as described in Section 4.4.1, using 1,000 particles per location. For convenience, all 20 locations were run simultaneously in the same simulation. In order to allow the sequestration calculations described in Section 4.5.1, the RWPT code tracks the total time spent by each particle in the surface layer (0-25 m) over the course of the simulation. Output was generated annually (January 1) for both the instantaneous particle positions and the aggregate time spent in the surface layer.

The second set of RWPT simulations sought to characterize the global distribution of sequestration efficiencies at multiple depth levels in a manner that takes advantage of the computational benefits of the Lagrangian approach. Three depth intervals were evaluated, each corresponding to an LLNL MOM grid layer: 677 - 846 m (layer 11), 1,258 - 1,501 m (layer 14), and 2,708 - 3,060 m (layer 19). These were chosen to be roughly equivalent to the 800 m, 1,500 m, and 3,000 m depths evaluated in the GOSAC study [42]. These simulations were also run for 2,000 years with an annual output interval. For computational convenience, 4 separate simulations were conducted for each depth interval, each with 3 particles per horizontal grid cell initially (placed randomly within each cell). The overall set thus consisted of 12 simulations, but they were all analyzed together, which allowed particles originating at a particular depth level to also be used in the sequestration calculation for the other two depths, if possible.

In both sets of simulations, the sequestration efficiency function $Q(t)$ was calculated using the approach embodied in (4.63). In computing sequestration efficiency, the exit region was defined as the upper 25 m of the ocean (which corresponds to the top layer of the LLNL
MOM grid), and two mass decay rates were tested in this exit region: 73 yr$^{-1}$ and 2 yr$^{-1}$ (the justification for these values is presented below in connection to the corresponding OGCM simulations). Because the output interval of the instantaneous particle positions is the same as the loop interval for the hydrodynamic forcing (1 year), the calculation actually estimates the sequestration efficiency of particles starting at a specific time (January 1 of a 'typical' year), and thus $\bar{Q}(t) = Q(t)$ in this case. This was done so that the results could be compared to an OGCM calculation of an instantaneous point source for validation (discussed below). The alternative would have been to run a continuous release with the OGCM and sample RWPT particle initial positions at shorter intervals throughout the year, but this approach was not taken here since the main goal of this application is to demonstrate and validate the method, and the latter approach would introduce an unwanted source of divergence between OGCM and particle results. Namely, particle sampling over the course of the year would have to be sufficiently resolved to properly capture the influence of multiple starting times, which would be difficult to guarantee without incurring a substantial increase in computational burden. Thus, for all applications below it is implicitly assumed that, for a given location, the sequestration efficiency for any time is well described by the sequestration efficiency for an instantaneous January 1 releases. Future applications, in which fundamental method validation is no longer a primary concern, could relax this assumption by sampling particle positions with greater frequency, or perhaps better yet, at a random time during each year of simulation.

It should be noted that the manner in which each simulation set is initialized is dictated by the goal of the computation. In the first set, a relatively large number of particles was used at each site, since there were only 20 sites and the focus was on validating RWPT results and illustrating the sequestration calculation. In the second set, particles were initially spread out over a given depth level across all wet cells, since the goal was to illustrate the sequestration efficiency calculation across the global domain. While the first set can generate reasonable sequestration efficiency estimates using the initial particle positions only (i.e., using (4.62)), the second set has only a few particles per grid cell at the beginning of the simulation and is therefore much more dependent on using subsequent particle “initial positions” (i.e., those generated later in the simulation) via the sequestration calculation in (4.63).
To validate the first set of RWPT simulations described above, a series of corresponding LLNL MOM simulations was conducted. For each injection location, an instantaneous point source was simulated for 1,500 years using a special LLNL MOM code in which tracer mass in the surface grid cell undergoes a first order decay with rate \( k \). This decay was specified as a surface mass flux \( F = -khC \) where \( h \) is the thickness of the surface layer (\( \sim 25 \) m) and \( C \) is the tracer concentration in that layer. Simulations were conducted for two different \( k \) values for each of the 20 injection locations. In particular, characteristic decay timescales of 5 days and 1/2 year were tested, which corresponds to \( k \) values of \( 73 \text{ yr}^{-1} \) and \( 2 \text{ yr}^{-1} \). The larger \( k \) value causes tracer mass to disappear quickly from the surface layer, and was selected to roughly correspond to an Eulerian approximation of the first passage calculations described in Section 4.5.1, i.e., one in which mass is immediately removed at the exit region (an infinite \( k \) cannot be used here because it would cause instability in the LLNL MOM). The smaller decay value is approximately equivalent to the one used in Hill et al. [39] to simulate volatilization of CO\(_2\) at the ocean surface and is thus more realistic for CO\(_2\) sequestration efficiency\(^6\). The RWPT results were post-processed using each of these two \( k \) values for direct comparison to the OGCM predictions.

In order to provide a larger validation dataset for the second set of RWPT sequestration calculations, 1,000-year OGCM simulations were run for 30 additional horizontal locations, using only the \( k \) appropriate for CO\(_2\) sequestration (\( 2 \text{ yr}^{-1} \)). These horizontal locations can be grouped into distinct categories:

- **GOS1 - GOS7** were chosen to correspond roughly to the GOSAC study locations (Bay of Biscay, New York, Rio de Janeiro, San Francisco, Tokyo, Jakarta, and Bombay). Simulations were performed for depth intervals 677 - 846 m, 1,258 - 1,501 m, and 2,708 - 3,060 m; these are approximately the 800 m, 1500 m, and 3000 m GOSAC Study depths [42]. See Figure 4-52.

- **RAN1 - RAN20** were selected with a random number generator. Simulations were performed here only for depth interval 1,258 - 1,501 m. See Figure 4-53.

\(^6\)Hill *et al.* used a decay rate \( k \) of \( 1 \text{ yr}^{-1} \) for a surface layer thickness \( h \) of about 50 m. In the present study the surface layer is about 25 m thick, and thus a \( k \) of \( 2 \text{ yr}^{-1} \) is used such that a similar mass transfer rate \( kh = 50 \text{ m/yr} \) is achieved.
Figure 4-52: GOS sites for validating the RWPT sequestration efficiency calculations, which are approximately colocated with the GOSAC study locations (Bay of Biscay, New York, Rio de Janeiro, San Francisco, Tokyo, Jakarta, and Bombay).

- ADD1 - ADD3 are additional test areas near features of potential interest. Each of the three GOSAC depths were tested here. See Figure 4-53.

Taken as a whole, this second set of OGCM simulation results introduces 50 release locations at which the RWPT and OGCM can be compared, beyond the 20 release sites included in the first set.

One issue of potential concern in computing sequestration efficiencies using the RWPT model is the impact of the nonlocal transport parameterization of the Large et al. [50] KPP vertical mixing scheme used by the OGCM. As discussed in Section 4.2, the nonlocal term dictates the vertical distribution of surface fluxes during periods of convective instability. The present RWPT implementation does not incorporate this process\(^7\). This could

\(^7\)Due to its nonlocal nature, it is not clear how this transport term could best be mimicked by the RWPT. In the sequestration efficiency calculation, it would require deducting mass from particles beneath the surface layer.
potentially hinder the ability of the RWPT to reproduce OGCM computed sequestration efficiencies since the CO$_2$ mass loss in the OGCM simulations is parameterized as a flux through the ocean surface, i.e., it would be subject to the nonlocal transport term. To evaluate the importance of this term, 500-year OGCM sequestration simulations with and without the term were performed for tracer release locations a3-h3, g4, and g5, using a decay rate $k = 2$ yr$^{-1}$. The maximum observed difference in mass remaining over the simulation was 0.028%. The observed small impact of the nonlocal transport term is likely due to the fact that it only applies to regions experiencing convective instability, i.e., it is often zero. It is therefore concluded that the nonlocal transport term will have a negligible impact on comparisons between RWPT and OGCM predicted sequestration efficiencies.
4.5.2.2 Comparison of RWPT and OGCM predictions

4.5.2.2.1 Base tracer release test sites  Figures 4-54 through 4-63 show a comparison of RWPT and LLNL MOM results for the first set of sequestration simulations. In each case, the top and middle panels show the fractional mass remaining in the ocean as a function of time (i.e., sequestration efficiency) for \( k = 73 \text{ yr}^{-1} \) and \( k = 2 \text{ yr}^{-1} \), respectively. In these panels, three lines are plotted: (1) the OGCM estimate, (2) the RWPT estimate considering only those particles in the injection grid cell at \( t = 0 \), and (3) the RWPT estimate considering only those particles in the injection grid cell after \( t = 0 \). The number of particles used in each RWPT result are shown in the bottom panel. For the OGCM results, the mass remaining is plotted at 10 year intervals. For the RWPT results, plotting intervals of 50, 100, and 200 years were used for times 0-200 years, 200-1,000 years, and 1,000-1,500 years, respectively. These variable plotting intervals were used to shorten the processing time of the particle results.

The results are qualitatively consistent with those observed in the 100-year tracer release tests (Section 4.4.1) in that the quality of agreement between the RWPT and OGCM results varies by location, with generally favorable agreement for Sites b, c, d, e, f, and h and less favorable agreement for locations a and g. In terms of depth range, there is no clear trend; in some cases the bottom release was less favorable than mid-depth releases (e.g., d5 and g5), while the reverse was true elsewhere (e.g., a5). The results are also qualitatively consistent with the 100-year tracer release tests in that sequestration efficiency is generally underpredicted (i.e., mass is lost too quickly). This is expected given the vertical bias or artificial upwelling observed in the RWPT results, which generally causes an overprediction of near-surface tracer concentration. The absolute and relative error of the RWPT predictions are characterized later in this section, but for now it is noted that the RWPT model successfully captures the shape of the sequestration function in most cases.

Although some issues clearly remain in correcting the particle transport scheme so that all locations can be simulated with similar accuracy, the sequestration efficiency calculation approach is validated by the observed results. One interesting example is Site d3 (Figure 4-58), where the shape of the mass function is most different for the two values of \( k \) tested.
For the larger $k$ value (73 yr$^{-1}$), the sequestration efficiency drops below 0.2 after about 350 years; for the smaller $k$ value (2 yr$^{-1}$), it takes about 1,500 years to reach the same mass loss. Although the RWPT computes a greater mass loss than the OGCM (mass loss is overpredicted by about 3-6% relative error at 1,500 years, resulting in about a 30-50% relative error in sequestration efficiency at this time), it does reproduce the shapes of the function for each $k$ value. Computationally, the RWPT has the benefit of generating each curve by post-processing the same simulation results with a different $k$, whereas a new OGCM simulation is required for each new $k$ value. Thus, unlike the OGCM, a trivial amount of additional computation time is needed to calculate this curve for any $k$ value in the RWPT case.

The results in Figures 4-54 - 4-63 also validate the general approach of using each particle output as a new set of “starting positions”, as evidenced by the generally close agreement between the RWPT estimate using $t = 0$ particles and the estimate using $t > 0$ particles. As expected, the quality of agreement varies with the number of particles, and the quality of agreement for a given number of particles varies by location. It is known that the noise of RWPT concentration predictions generally follows an $N^{-1/2}$ scaling, where $N$ is the number of particles [44]; this same scaling is demonstrated to be roughly true on average for these sequestration calculations but large variability was also observed (see Section 4.5.2.4). The fact that more particles are required for some sites is reasonable since sites have different spatial extents and are located in areas of differing spatial variability in flow and diffusivity (e.g., the average travel time for particles originating from two sites may be equal but the standard deviation may be very different). Overall, reasonable agreement seems to be achieved here using about 1,000 particle per site, but in any real application sensitivity analysis should be performed to verify that the metrics of interest are not appreciably influenced by the number of particles used initially. Most importantly, the results demonstrate that individual particles can indeed be used to generate residence time statistics for multiple sites within the domain. This characteristic is one main advantage of the Lagrangian approach to transport modeling.
Figure 4-54: LLNL MOM and RWPT sequestration efficiency for Sites a3 (left) and a4 (right) using surface decay rates $k = 73 \text{ yr}^{-1}$ (top) and $k = 2 \text{ yr}^{-1}$ (middle). Dashed line: OGCM; solid line: RWPT using $t = 0$ particle positions; dotted line: RWPT results using $t > 0$ positions. Bottom panel: number of particles in RWPT result.
Figure 4-55: LLNL MOM and RWPT sequestration efficiency for Sites a5 (left) and b3 (right) using surface decay rates $k = 73 \text{ yr}^{-1}$ (top) and $k = 2 \text{ yr}^{-1}$ (middle). Dashed line: OGCM; solid line: RWPT using $t = 0$ particle positions; dotted line: RWPT results using $t > 0$ positions. Bottom panel: number of particles in RWPT result.
Figure 4-56: LLNL MOM and RWPT sequestration efficiency for Sites b4 (left) and b5 (right) using surface decay rates $k = 73 \, \text{yr}^{-1}$ (top) and $k = 2 \, \text{yr}^{-1}$ (middle). Dashed line: OGCM; solid line: RWPT using $t = 0$ particle positions; dotted line: RWPT results using $t > 0$ positions. Bottom panel: number of particles in RWPT result.
Figure 4-57: LLNL MOM and RWPT sequestration efficiency for Sites c3 (left) and c4 (right) using surface decay rates $k = 73 \text{ yr}^{-1}$ (top) and $k = 2 \text{ yr}^{-1}$ (middle). Dashed line: OGCM; solid line: RWPT using $t = 0$ particle positions; dotted line: RWPT results using $t > 0$ positions. Bottom panel: number of particles in RWPT result.
Figure 4-58: LLNL MOM and RWPT sequestration efficiency for Sites c5 (left) and d3 (right) using surface decay rates $k = 73$ yr$^{-1}$ (top) and $k = 2$ yr$^{-1}$ (middle). Dashed line: OGCM; solid line: RWPT using $t = 0$ particle positions; dotted line: RWPT results using $t > 0$ positions. Bottom panel: number of particles in RWPT result.
Figure 4-59: LLNL MOM and RWPT sequestration efficiency for Sites d4 (left) and d5 (right) using surface decay rates $k = 73$ yr$^{-1}$ (top) and $k = 2$ yr$^{-1}$ (middle). Dashed line: OGCM; solid line: RWPT using $t = 0$ particle positions; dotted line: RWPT results using $t > 0$ positions. Bottom panel: number of particles in RWPT result.
Figure 4-60: LLNL MOM and RWPT sequestration efficiency for Sites e3 (left) and e4 (right) using surface decay rates $k = 73 \text{ yr}^{-1}$ (top) and $k = 2 \text{ yr}^{-1}$ (middle). Dashed line: OGCM; solid line: RWPT using $t = 0$ particle positions; dotted line: RWPT results using $t > 0$ positions. Bottom panel: number of particles in RWPT result.
Figure 4-61: LLNL MOM and RWPT sequestration efficiency for Sites e5 (left) and f3 (right) using surface decay rates $k = 73\;\text{yr}^{-1}$ (top) and $k = 2\;\text{yr}^{-1}$ (middle). Dashed line: OGCM; solid line: RWPT using $t = 0$ particle positions; dotted line: RWPT results using $t > 0$ positions. Bottom panel: number of particles in RWPT result.
Figure 4-62: LLNL MOM and RWPT sequestration efficiency for Sites g3 (left) and g4 (right) using surface decay rates $k = 73 \text{ yr}^{-1}$ (top) and $k = 2 \text{ yr}^{-1}$ (middle). Dashed line: OGCM; solid line: RWPT using $t = 0$ particle positions; dotted line: RWPT results using $t > 0$ positions. Bottom panel: number of particles in RWPT result.
Figure 4-63: LLNL MOM and RWPT sequestration efficiency for Sites g5 (left) and h3 (right) using surface decay rates $k = 73 \text{ yr}^{-1}$ (top) and $k = 2 \text{ yr}^{-1}$ (middle). Dashed line: OGCM; solid line: RWPT using $t = 0$ particle positions; dotted line: RWPT results using $t > 0$ positions. Bottom panel: number of particles in RWPT result.
4.5.2.2 GOSAC study sites  As described in Section 4.5.2.1, an additional validation dataset of 50 release locations is provided for the second set of RWPT simulations, i.e., those aimed at characterizing sequestration efficiencies for all grid cells at a given depth level. Presented below are results for the seven horizontal locations (GOS1 - GOS7) that are approximately colocated with the GOSAC study locations (Bay of Biscay, New York, Rio de Janeiro, San Francisco, Tokyo, Jakarta, and Bombay). Figures 4-64 to 4-67 show comparisons of RWPT and OGCM predictions for each of the three depths evaluated (∼800 m, ∼1,500 m, and ∼3,000 m) for all seven locations.

The level of agreement between the OGCM and RWPT predicted sequestration efficiency (plotted lines) is qualitatively similar to that of the first set of sequestration results (Figures 4-54 to 4-63). The RWPT model underpredicts sequestration efficiency relative to the OGCM after 1,000 years at all sites to varying degrees. For the best locations (GOS1, GOS4, GOS5, and GOS7, corresponding to Bay of Biscay, San Francisco, Tokyo, and Bombay), the sequestration efficiency is generally underpredicted by less than 5%. For the worst locations (GOS2, GOS3, and GOS6, corresponding to New York, Rio de Janeiro and Jakarta), the underprediction is on the order of 10-15% (in an absolute sense). Such errors can be large in a relative sense for low sequestration efficiencies, e.g., if after 1,000 years the RWPT underpredicts efficiency by 10% and the OGCM predicts a 30% efficiency, then the relative error can be taken as 33% (e.g., GOS3 for a 1,500 m release depth). Characterization of the relative and absolute errors is provided at the end of this section.

It is interesting to view the RWPT and OGCM results in relation to the range of predictions reported in the GOSAC study, which has been included on Figures 4-64 to 4-67 after 500 years. The GOSAC simulations are for a 100-year source, and thus the results shown in the figures are 400 years after the termination of the loading. In an absolute sense, the GOSAC results are not directly comparable with the simulations in the present study. The primary reason is that the metrics are different: the GOSAC results are injection efficiencies of a temporary source (i.e., the fraction of total injected mass remaining in the ocean), whereas the OGCM and RWPT plotted lines are sequestration efficiency (i.e., the fraction remaining for an instantaneous source). Secondary reasons are that (a) the air-sea interface was treated in a more rigorous manner in the GOSAC study by considering local
pCO$_2$ gradients with projected atmospheric CO$_2$ concentrations, (b) the grid cells selected for the present study are only roughly collocated with the GOSAC locations, and (c) the GOSAC simulations were continuous releases whereas the present study evaluates instantaneous January 1 releases, and thus will not capture any variability in sequestration efficiency that could be experienced over the looped year-long hydrodynamic forcing.

Still, comparison of the RWPT/OGCM results to the GOSAC results are interesting in both a relative and absolute sense. In a relative sense, the difference between the RWPT and OGCM predictions is in all cases smaller than the range in the GOSAC results (at least half, and usually much less). Since the GOSAC study is an intercomparison between eight different OGCMs, this implies that the observed differences between the predictions of LLNL MOM and the RWPT model are likely smaller (perhaps considerably so) than the differences one might expect to find between two different OGCMs, or even between different configurations of the same model (e.g., see Mignone et al. [55]). This provides some context in which to interpret the importance of the shortcomings of the present RWPT implementation (detailed in Section 4.4 and Appendix 4.A). To allow a comparison in an absolute sense, Figures 4-64 - 4-67 also show the RWPT equivalent injection efficiency for a 100-year GOSAC-style source using the approach in (4.67). The difference between sequestration and injection efficiency is generally small and decreases over time, and thus this distinction will be dropped for the remainder of the analysis (i.e., only the sequestration efficiency metric will be used). However, the fact that the LLNL MOM prediction falls outside of the GOSAC range for two releases (GOS2 at 800 and 1,500 m) even though it is one of the eight GOSAC OGCMs, indicates that the other aforementioned inconsistencies between the RWPT/OGCM and GOSAC simulations limit the extent to which they can be compared in an absolute sense. Nonetheless, the RWPT/OGCM predicted efficiencies fall within the range of the GOSAC predictions for all other release points (see Section 4.5.2.3 for additional comparisons).
Figure 4-64: LLNL MOM (dashed line) and RWPT (solid line) computed sequestration efficiency (for CO$_2$, $k = 2$ yr$^{-1}$) for Sites GOS1 (Bay of Biscay) and GOS2 (New York) for an instantaneous source at three injection depths. The approximate range of predicted 500-year injection efficiencies from the GOSAC study at a similar location is also shown (vertical line), for a temporary source which ends after 100 years of injection. The RWPT predicted injection efficiency for the GOSAC temporary source is also shown (points).
Figure 4-65: LLNL MOM (dashed line) and RWPT (solid line) computed sequestration efficiency (for CO$_2$, $k = 2$ yr$^{-1}$) for Sites GOS3 (Rio de Janeiro) and GOS4 (San Francisco) for an instantaneous source at three injection depths. The approximate range of predicted 500-year injection efficiencies from the GOSAC study at a similar location is also shown (vertical line), for a temporary source which ends after 100 years of injection. The RWPT predicted injection efficiency for the GOSAC temporary source is also shown (points).
Figure 4-66: LLNL MOM (dashed line) and RWPT (solid line) computed sequestration efficiency (for CO$_2$, $k = 2$ yr$^{-1}$) for Sites GOS5 (Tokyo) and GOS6 (Jakarta) for an instantaneous source at three injection depths. The approximate range of predicted 500-year injection efficiencies from the GOSAC study at a similar location is also shown (vertical line), for a temporary source which ends after 100 years of injection. The RWPT predicted injection efficiency for the GOSAC temporary source is also shown (points).
Figure 4-67: LLNL MOM (dashed line) and RWPT (solid line) computed sequestration efficiency (for CO₂, k = 2 yr⁻¹) for Site GOS7 (Bombay) for an instantaneous source at three injection depths. The approximate range of predicted 500-year injection efficiencies from the GOSAC study at a similar location is also shown (vertical line), for a temporary source which ends after 100 years of injection. The RWPT predicted injection efficiency for the GOSAC temporary source is also shown (points).
4.5.2.2.3 Remaining sequestration sites The remaining validation dataset for sequestration simulations consists of the “RAN” and “ADD” locations described in Section 4.5.2.1, which comprise a total of 29 release locations. The individual results for these sites can be found in Appendix 4.B. The agreement between the RWPT and the LLNL MOM is qualitatively similar for these stations as in the datasets compared previously. Sequestration efficiency is generally underpredicted, although the RWPT correctly reproduces the shape of the sequestration curves at most sites and achieves near perfect agreement for some sites (e.g., RAN2 and RAN20).

4.5.2.2.4 Overall error between RWPT and LLNL MOM predictions The magnitude of the observed error between the RWPT and LLNL MOM predicted sequestration efficiency after 500 and 1,000 years is summarized in Figure 4-68 and Table 4.4. In the top panel of Figure 4-68, the RWPT prediction is plotted against the LLNL MOM result, demonstrating that the RWPT consistently underpredicts sequestration efficiency, which is consistent with the previous observation of an observed artificial upwelling in the present RWPT implementation. On average, the relative error in sequestration efficiency is 9 and 18 % after 500 and 1,000 years, respectively. The relative error tends to be larger for lower sequestration efficiencies (middle panels); as the denominator in the calculation decreases, the relative error increases even if the absolute error is the same. This trend is reversed if the relative error of the mass loss (= 1 - sequestration efficiency) is instead considered (bottom panels). There is also a temporal trend in the observed errors; as time increases the relative error in sequestration efficiency increases, whereas the opposite is true for mass loss. Despite the RWPT’s tendency to underpredict sequestration efficiencies, there is overall a decent agreement between the RWPT and OGCM predictions, with many of the datapoints falling on or close to the 1:1 line after 500 years of simulation.

In addition to the temporal trend, there is also a spatial trend. The spatial relative error distribution after 1,000 years is shown in Figure 4-69, averaged over depth where multiple depths were simulated. There is a clear gradient in relative error in the meridional direction, with the highest errors observed in the low latitudes near the Southern Ocean. This is consistent with findings in Section 4.4, where the worst performance of the tracer release
Figure 4-68: Error summary for RWPT 500-year and 1,000-year sequestration efficiency calculations (for CO$_2$, $k = 2$ yr$^{-1}$), shown as RWPT vs OGCM sequestration efficiency (top panel), as well as RWPT relative error on a sequestration efficiency (middle panel) and mass lost (bottom panel) basis.
Table 4.4: Summary of relative error in RWPT sequestration calculations for all 70 comparison sites.

<table>
<thead>
<tr>
<th>Time (years)</th>
<th>Mean Error in Sequestration Efficiency (%)</th>
<th>Median Error in Sequestration Efficiency (%)</th>
<th>Mean Error in Mass Loss (%)</th>
<th>Median Error in Mass Loss (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>2.2</td>
<td>1.1</td>
<td>51</td>
<td>17</td>
</tr>
<tr>
<td>500</td>
<td>9.0</td>
<td>6.7</td>
<td>16</td>
<td>8.2</td>
</tr>
<tr>
<td>1,000</td>
<td>18</td>
<td>17</td>
<td>12</td>
<td>8.5</td>
</tr>
</tbody>
</table>

Tests was observed in this region. As discussed in previous sections, the relatively poorer performance of the RWPT model in these regions is likely attributed to regions of steeply sloping isopycnals with large spatial gradients (leading to a more significant impact of the Gent-McWilliams eddy-induced transport, see Figure 4-7). Since ongoing refinements of the RWPT implementation (Appendix 4.A) are specifically aimed at improving the treatment of isopycnal terms, it is expected that future improvements will reduce error in sequestration calculations in these regions.

4.5.2.3 Global sequestration efficiencies

Global CO$_2$ sequestration efficiencies over 1,000 years were computed using the second set of RWPT simulations described in Section 4.5.2.1 (with CO$_2$ surface decay rate $k = 2$ yr$^{-1}$). Figure 4-70 shows the number of particle states available at each depth to compute the sequestration efficiency at various times. The nature of the particle book-keeping approach causes more particle states to be available for earlier times. At 1,000 years, the maximum time considered, most sites have 1,000 states or more available, i.e., accuracy should be on par with earlier validation. The fewest particle states are available for shallow sites at low or high latitudes, since these are the smallest grid cells in the domain. The domain-wide minimum number of available states for 1,000 years was 399 for a depth of 800 m. Even with 399 particle states, the results are expected to be reasonable given the validation results for Sites a3, g3, and g4 (where a roughly equivalent number of $t > 0$ states were encountered, see Figures 4-54 and 4-62).

Following the GOSAC study’s lead, we begin by evaluating the average depth dependence
Figure 4-69: Spatial distribution of relative errors of RWPT predicted sequestration efficiency (for CO$_2$, $k = 2$ yr$^{-1}$) for all simulated locations after 1,000 years. The error was averaged over depth where multiple depths were simulated, resulting in a total of 38 data-points. The error is in almost all cases an underprediction of sequestration efficiency.

of sequestration efficiency. Figure 4-71 shows the mean sequestration efficiency as a function of time for each of the three GOSAC depths, and confirm the GOSAC finding of a strong depth dependence, which has also been noted by most past investigations (e.g., [2, 3, 13, 23, 39, 43, 55, 68, 73]). The minimum, mean, and maximum efficiencies at each depth and time are summarized in Table 4.5. The mean efficiency at 3,000 m is 0.97, 0.72, and 0.44 after 100, 500, and 1,000 years, respectively, compared to 0.87, 0.46, and 0.26 at 1,500 m and 0.67, 0.28, and 0.16 at 800 m. As an approximate rule of thumb, for the interval 500 - 1,000 years, the mean efficiency remains about 60% higher for 1,500 m than for 800 m, and about 60% higher for 3,000 m than for 1,500 m (i.e., efficiencies at 3,000 m are about 160% greater than at 800 m). This is roughly comparable to the depth dependence of the 500-year injection efficiency found by Caldeira et al. [13] for LLNL MOM simulations of the 7 GOSAC sites. Figure 4-71 also shows the range of efficiencies over time for the GOSAC sites and for the ocean overall.
Figure 4-70: Number of particle states used to calculated the sequestration efficiencies at approximate depths of 800 m (top), 1,500 m (middle), and 3,000 m (bottom) after 100 (left), 500 (middle), and 1,000 (right) years. The minimum number of states for each panel is shown in parentheses.
Table 4.5: Global statistics for RWPT predicted sequestration efficiencies.

<table>
<thead>
<tr>
<th>Time (yrs)</th>
<th>800 m Range</th>
<th>800 m Mean</th>
<th>1,500 m Range</th>
<th>1,500 m Mean</th>
<th>3,000 m Range</th>
<th>3,000 m Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.25 - 0.96</td>
<td>0.67</td>
<td>0.53 - 0.99</td>
<td>0.88</td>
<td>0.78 - 1.00</td>
<td>0.97</td>
</tr>
<tr>
<td>500</td>
<td>0.07 - 0.72</td>
<td>0.28</td>
<td>0.27 - 0.75</td>
<td>0.46</td>
<td>0.53 - 0.90</td>
<td>0.72</td>
</tr>
<tr>
<td>1,000</td>
<td>0.03 - 0.50</td>
<td>0.16</td>
<td>0.15 - 0.52</td>
<td>0.27</td>
<td>0.33 - 0.68</td>
<td>0.44</td>
</tr>
</tbody>
</table>

The mean efficiency for the GOSAC sites is surprisingly close to the ocean-wide mean at each depth, considering only seven sites were used. The range in the efficiency at the GOSAC sites is about 1/3 to 3/4 of the ocean-wide range, with the largest differences at intermediate times for 800 m and for longer times for the deeper depths. The range in reported GOSAC injection efficiencies compares reasonably well with the range predicted herein. The fact that the GOSAC sites fail to capture the full range of sequestration efficiencies is further demonstrated in Figure 4-72, where the RWPT predicted GOSAC sequestration efficiencies are plotted together with the ocean-wide probability density function for 100 and 500 years (the GOSAC simulation length). At 100 years, the GOSAC site efficiencies are generally higher than the mean, i.e., they are slightly more favorable than average. At 500 years, they are more evenly distributed about the mean. They come close to capturing the upper limit of efficiency at 100 years, but less so at 500 years, where sites with about a 1/4 to a 1/3 higher efficiency exist for 800 m and 1,500 m releases. The figure also demonstrates that the range of sequestration efficiency decreases with depth and increases with time. The temporal growth of the distribution is slowest for 3,000 m, which exhibits much less variability after 100 years than the other depths.

The spatial distribution of sequestration efficiencies are plotted in Figure 4-73 for the three depths at times 100, 500, and 1,000 years. The value of the present calculation becomes clear, revealing intricate spatial patterns. The patterns from the previous figures are reinforced, i.e., there is a dramatic improvement in efficiency with depth, and injections at 3,000 m exhibit considerably less variability. In addition, it becomes clear that spatial gradients in sequestration efficiency are smaller at 3,000 m than at shallower depths, where efficiency can vary rapidly over fairly small geographic scales.

In order to understand which regions of the ocean have the best sequestration character-
Figure 4-71: Temporal trends in mean RWPT predicted sequestration efficiency for release depths of approximately 800 m (top), 1,500 m (middle), and 3,000 m (bottom). The ocean-wide volume-weighted mean (solid line) and range (circles) are plotted along with the unweighted mean (dashed line) and range (vertical dashed lines) for the GOSAC sites (using RWPT predictions). The reported GOSAC injection efficiencies at 100 and 500 years are also plotted (diamonds).
Figure 4-72: Probability density function (unweighted) for predicted 100-year (left) and 500-year (right) sequestration efficiencies from release depths of approximately 800 m (top), 1,500 m (middle), and 3,000 m (bottom). The RWPT predicted efficiencies at the 7 GOSAC sites are shown as dashed vertical lines. The volume-weighted mean sequestration efficiency for each depth is indicated by the solid semicircles.
Figure 4-73: Global sequestration efficiency for an instantaneous source at approximate depths of 800 m (top), 1,500 m (middle), and 3,000 m (bottom) after 100 (left), 500 (middle), and 1,000 (right) years. The volume-weighted mean efficiency for each panel is shown in parentheses.
istics, it is useful to plot the efficiencies on a relative basis (Figure 4-74). Here the efficiencies at each depth/time have been normalized by the maximum value, thus comparing all sites to the most efficient one. The conclusions from before are again apparent; for all three times the spatial variability decreases with depth, as do the spatial gradients. There is, however, some distinct difference in spatial patterns between the depth levels.

For the shallowest depth (800 m), the North Atlantic Deep Water (NADW) formation region in and around the Labrador Sea and the Antarctic Bottom Water (AABW) formation region in the Weddell Sea exhibit the highest sequestration efficiencies over time. The high sequestration efficiency for the North Atlantic region was also found by Hill et al. [39] for the Labrador Sea and by Drange et al. [23] for the Norwegian Sea. Indeed, the results show a substantial level of agreement with the 935 m injection efficiency maps of Hill et al. [39] (Figure 4-75), generated for a continuous source with a coarser resolution OGCM (MITgcm, 4x4 degrees, 15 depth levels, no Arctic Ocean) than the present LLNL MOM configuration (4x2 degrees, 24 depth levels). Hill et al. noted a reversal between the relative efficiency of the North Atlantic and the North Pacific over time, the latter being better initially (100 years) and the former being better later (500 years and beyond). They hypothesize that the North Atlantic initially performs worse because of the fast upwelling of waters at the base of the thermocline along the western boundary, which are transported along the subtropical gyre and ventilated on a decadal timescale before reaching the NADW formation region. In contrast, the Pacific has less overturning and larger gyres with longer timescales. At longer times, however, the global thermohaline circulation dominates such that the deep water formation region in the North Atlantic causes better sequestration efficiency than the slow but steady upwelling in the Pacific at the other end of the thermohaline circulation. This interpretation is nearly identical to that of Dewey et al. [20] for 1,000 m injections near Tokyo and New York. The 800 m RWPT results are somewhat consistent with this, but to a lesser degree. The region of poorer 100-year efficiency in the western North Atlantic does not extend as far north, and the very high sequestration efficiency region extends south of the Labrador Sea into the subpolar North Atlantic and the Greenland Sea (it also extends north into the Arctic, which was not included in the Hill et al. grid). Also, the RWPT predictions show a much stronger zonal gradient in the Pacific, with low sequestration efficiencies (even
Figure 4-74: Global relative sequestration efficiency for an instantaneous source at approximate depths of 800 m (top), 1,500 m (middle), and 3,000 m (bottom) after 100 (left), 500 (middle), and 1,000 (right) years. In each panel, the sequestration efficiencies are normalized by the maximum efficiency in the panel (shown in parentheses), thus showing the efficiency of sites relative to the most efficient site at that depth/time.
lower than the western North Atlantic) occurring in the North Pacific subtropical gyre along the western boundary. At 500 years, the agreement is considerably better; both show the Pacific gradient and lower efficiencies in the Pacific relative to the North Atlantic. At 1,000 years, the zonal gradient again disappears from the Hill et al. result, but this could simply be an artifact of the coarse plotting scale. In the southern hemisphere, there is good agreement for all three times. Both models show the high sequestration efficiency in the Weddell Sea and the areas “downstream” of Drake’s Passage and the Antarctic Peninsula. They also show the same characteristic band of poor sequestration efficiency in the northern regions of the Southern Ocean where it meets the other ocean basins. Moving north of that band, both models predict similar meridional gradients in sequestration efficiency in the Indian, Pacific, and Atlantic basins. In general, the RWPT results, perhaps because they are derived from a better resolved parent OGCM, predict a stronger zonal gradient, with lower efficiencies generally occurring along the western boundaries, presumably due to western intensification [20, 42]. For example, low sequestration efficiencies at this shallow depth can be found along the Gulf Stream (North Atlantic), Kuroshio (North Pacific), Agulhas Current (South Indian), Brazil Current (South Atlantic), and East Australia Current (South Pacific), although in the latter three cases the low efficiencies may be attributable to proximity to the turbulent Southern Ocean.

As injection depth increases, the reversal between the North Pacific and North Atlantic does not occur; at 1,500 m they are comparable and at 3,000 m the Pacific has higher efficiencies (contradicting the 2,000 m residence time estimates of Hill et al. [39], where the Atlantic remains more efficient). In addition, the zonal gradient in predicted efficiency decreases as depth increases, and as does the advantage of the deep water formation regions. At 3,000 m, the Pacific offers some of the highest long-term efficiencies in the domain. Also, the band of low efficiencies in the Southern Ocean is less pronounced at 3,000 m, with high efficiencies computed beneath that band. For all depths, the Arctic Ocean has among the highest efficiencies in the domain.

Although marked differences in spatial patterns between OGCMs have been reported [58, 42, 43, 55], the spatial trends noted here are consistent with some of the main findings of past studies which evaluated individual point sources. Using LLNL MOM, Caldeira et
Figure 4-75: Hill et al. [39] injection efficiency for a continuous source at 935m.

al. [13] found the Pacific GOSAC sites to be more efficient than the Atlantic ones, and also noted reduced spatial variability with increasing depth. The GOSAC study [42] found this same trend for the deeper releases, and also found the southern hemisphere sites to be less efficient due with proximity to the Southern Ocean. For 3,000 m releases, Kheshgi et al. [46, 47] also found the Pacific to be more efficient than the Atlantic. As an example of a contradictory finding, Stegen et al. [68] found the Pacific to be better even for shallower releases, and found the greatest efficiency in the Equatorial Pacific.

Although the present implementation has some remaining issues which cause an artificial upwelling, it is unlikely that the major conclusions stated above are compromised by this error. The spatial trends in Figure 4-69 illustrate that the largest errors are encountered in
the Southern Ocean, and that error generally decreases moving north. There is no clear zonal
gradient in error, which suggests that the model can diagnose relative differences between
the North Atlantic and North Pacific basins. The favorable sequestration efficiencies of the
deep water formation regions in the North Atlantic and Southern Ocean would if anything be
understated by the RWPT results, since their error is in almost all cases an underprediction
relative to the OGCM. It is likely that the band of low sequestration efficiencies in the
Southern Ocean is overstated since this is also the region of the greatest underpredictions,
but its existence is supported by a similar band in the results of Hill et al. [39]. Indeed, even
though different OGCMs with different resolutions are used, there is a good level of agreement
between the RWPT 800 m results and those of Hill et al. at 935 m, which suggests that the
remaining errors in the RWPT calculation do not substantially compromise the predicted
spatial/temporal patterns in relative sequestration efficiency.

The spatial patterns shown in Figure 4-74 are somewhat impacted when the surface decay
$k$ value is increased or decreased by factor of 4. Figure 4-76 shows the relative sequestration
efficiency with the $k$ value reduced to 0.5 yr$^{-1}$ (i.e., slower outgassing at the surface). The 800
m advantage of the Atlantic ocean is more pronounced with higher sequestration efficiencies
spreading south; this is consistent with the hypotheses of Hill et al. [39] and Dewey et al. [20]
in that less of the carbon that is ventilated along the western boundary is lost before it can
enter the abyss via the NADW formation region. Likewise, at 1,500 m the Atlantic shows
higher efficiency than the Pacific. At 3,000 m, the Pacific still shows the higher longterm
efficiencies at 1,000 years, but the advantage over the Atlantic is slight. Unlike the North
Atlantic, the spatial pattern in the North Pacific is only marginally impacted by a reduced
$k$; as ventilated carbon exits the Kuroshio into the North Pacific Current, there is no deep
water formation region which allows it to evade the surface, and outgassing is only marginally
slowed. Since most of the Pacific is characterized by slow steady upwelling, changes to the
outgassing rate does not strongly impact the relative efficiency of sites there. Conversely, if
the $k$ is instead increased by a factor of 4 (not shown), the efficiency of the Atlantic is reduced
in each case, with the Pacific showing higher efficiency at 1,500 m and a more pronounced
longterm advantage at 3,000 m. The band of low efficiencies in the Southern Ocean at 800
and 1,500 m persists for each $k$ value, although its prominence at a given time decreases as
Table 4.6: Sensitivity of 1,000-year mean RWPT predicted sequestration efficiency to surface decay rate ($k$).

<table>
<thead>
<tr>
<th>Depth (m)</th>
<th>$k = 0.5 \text{ yr}^{-1}$ Mean</th>
<th>$k = 2 \text{ yr}^{-1}$ Mean</th>
<th>$k = 8 \text{ yr}^{-1}$ Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>800</td>
<td>0.25</td>
<td>0.16</td>
<td>0.11</td>
</tr>
<tr>
<td>1,500</td>
<td>0.36</td>
<td>0.27</td>
<td>0.22</td>
</tr>
<tr>
<td>3,000</td>
<td>0.54</td>
<td>0.44</td>
<td>0.40</td>
</tr>
</tbody>
</table>

$k$ decreases. The sensitivity of the mean sequestration efficiency to $k$ is shown in Table 4.6.

Overall, the analysis of sequestration efficiencies highlights the importance of achieving a sufficient sequestration depth, which has the dual benefit of increasing average sequestration time and reducing spatial variability, which in turn reduces the possibility of encountering a “hotspot” of poor sequestration efficiency. Looking ahead to the CO$_2$ discharge scenarios considered in Chapter 5, attaining a sufficient depth would be important for both stationary and moving (i.e., a “towed pipe”) discharges, and reinforces the potential value of CO$_2$ hydrates as a means to enhance sequestration depth/efficiency while minimizing ecological disruption. Moreover, sequestration strategies should be tailored to take advantage of prevailing regional circulation patterns to the extent possible.

Beyond the sequestration issue, the above application demonstrates the inherent value of the statistics that can be derived from the RWPT model. The results allow a deeper understanding of the OGCM predicted transport than would otherwise be realistically attainable with standard OGCM tracer simulations.

### 4.5.2.4 Numerical considerations

One drawback to RWPT simulations is that they can be computationally expensive. In particular, predicted concentration fields of the type presented in Section 4.4.1 can be too noisy to be of use unless a sufficient number of particles is used. Thus, for that mode of simulation, they are often more expensive than their Eulerian counterparts. It is in residence time applications such as the one presented above that the computational benefits of RWPT can be realized. In the application above, grid cells were initialized with only a handful of particles, and calculations take advantage of the fact that a single particle generates data
Figure 4-76: Global relative sequestration efficiency with reduced surface decay $k = 0.5 \ yr^{-1}$ for an instantaneous source at approximate depths of 800 m (top), 1,500 m (middle), and 3,000 m (bottom) after 100 (left), 500 (middle), and 1,000 (right) years. In each panel, the sequestration efficiencies are normalized by the maximum efficiency in the panel (shown in parentheses), thus showing the efficiency of sites relative to the most efficient site at that depth/time.
points for multiple grid cells throughout the simulation.

With present computing technology, each of the twelve 2,000-year simulations to characterize sequestration efficiencies required about 2-3 weeks of computer time, with an additional 1-2 days of post-processing time. A significant fraction of this time (~30%-50%, depending on the system) was, however, consumed by disk input/output. The main reason for this is that the 365-day hydrodynamic coupling record was too large to keep in memory in the base code used here. One recoding effort successfully allowed the record to be kept in memory, but run-time was the same or worse due to heavy reliance on virtual memory (“paging”). Future editions of the code can likely be streamlined and optimized to run faster, and even without streamlining run times should become considerably less as larger amounts of physical memory becomes standard on model production computers.

From both a simulation time and accuracy standpoint, selecting the appropriate number of particles to use in an RWPT simulation is an important step, and is a case specific decision. Whether or not an appropriate number has been used can be determined by sensitivity analysis; increasing the number of particles should not impact conclusions drawn from the modeling study. As discussed and verified in [44], the relative error (noise) of an RWPT simulation on a concentration basis scales as $N^{-1/2}$, where $N$ is the number of particles. To verify whether this same scaling applies to the sequestration calculations presented above, the results of each of the 70 validation test sites (see above) were reanalyzed using subsets of the data. Specifically, for each site, subsets of increasing size were selected from the first half of the particle state data and used to calculate the 500-year sequestration efficiency. This result was then compared to the value computed using the entire dataset, and a relative error was calculated for each site as a function of the number of particle states. Averaged over all sites, the relative error was found to scale with $N^{-0.59}$ (i.e., the average exponent was -0.59), with an average $R^2$ correlation value of 0.62. There was, however, considerable variability, with exponents ranging from -0.1 to -1.33, and $R^2$ values ranging from 0.02 to 0.9. Nonetheless, the analysis demonstrates that on the whole a similar scaling applies to these types of calculations.

In addition, simulation times can likely be improved by optimizing the timestep, perhaps to a significant degree. In all RWPT simulations presented herein, an 1,800 s timestep
was used. The main controlling factor in choosing such a short timestep was the vertical
diffusion calculation; 1D numerical experiments with extreme diffusivity profiles indicated
that a reasonable approximation of much smaller $\Delta t$ results could be achieved with a 1,800
s timestep, when applying the correction methodology described in Section 4.3.2.1.2. In the
course of developing the RWPT model and testing various permutations, this small $\Delta t$ was
used throughout to remove timestep impacts from being a significant source of divergence in
the RWPT vs OGCM comparisons; sensitivity analysis with the full RWPT model showed
that further decrease of the timestep did not cause an appreciable difference in these com-
parisons. However, it is likely that the timestep can be increased with only a small increase
in error, especially since the correction methodology used for vertical diffusion preserves
the “well-mixed condition” for any timestep, albeit with increasing error for spatially vari-
able concentration. For the present simulations, the practical upper bound of $\Delta t$ is 1 day,
since this is the temporal resolution of the OGCM hydrodynamic coupling data. Given that
RWPT run times scale almost linearly\textsuperscript{8} with $\Delta t$, future studies involving this model would
benefit from a rigorous timestep sensitivity analysis, as the RWPT model transitions from a
development/testing phase to an application phase.

Lastly, it is reemphasized that a great numerical advantage of the RWPT approach over
an OGCM approach is that it could be employed with a high resolution hydrodynamic forcing
with only a modest increase in computational effort. It is likely that the errors observed in the
present RWPT simulations (Section 4.4 and Appendix 4.A) would be reduced with improved
resolution because it would lessen or even altogether remove the Gent McWilliams terms from
the calculation. Moreover, as the velocity field becomes better resolved, the importance of
all isopycnal diffusion terms is reduced and thus the calculations are less influenced by the
arbitrary choices that must be made in defining the sub-grid scale variation of isopycnal
slopes and diffusivities. There would be some computational penalty for searching through a
larger grid for interpolation variables at each particle position, but presumably the timestep
would not need to be reduced further. Most importantly, the same number of particles
could be used, provided particle results were still computed on a coarser grid such as the

\textsuperscript{8}the scaling is not linear because of the correction methodology applied, and because of significant in-
put/output times
one used herein; in this case the spatial resolution of the particle results would not improve
but the accuracy of the predictions would improve because they would be based on better
hydrodynamic forcing. Thus, a sequestration efficiency calculation, or more generally any
tracer transport calculation, could be conducted using hydrodynamic forcing from a grid
that is too resolved to practically conduct long-term tracer transport calculations with the
OGCM alone (due to the nonlinear increase of OGCM computation time with increased grid
resolution).

4.5.3 Related applications and beyond

As noted in Section 4.1, the sequestration efficiency calculation is but one example of a
larger class of residence time calculations. In the application above, mean residence times
(4.53) were not explicitly calculated because that would require assumptions about the tail
of the mass remaining function, i.e., the sequestration efficiency curve. This is not a technical
barrier, e.g., Hill et al. [39] used an exponential fit to extrapolate beyond the end of their
simulation. It is, however, somewhat arbitrary and would therefore introduce a complicating
factor in interpreting the results, and would not aid in the present evaluation of the RWPT
model’s predictive capability. Still, the reader should note that the sequestration efficiency
curve estimated above is the basis for a mean residence time calculation, and unless the
simulations were run for a very long time, some extrapolation would be needed (ignoring the
tail would underpredict the residence time). The same is true for a traditional OGCM-based
calculation.

Another important class of applications to which the RWPT can be applied is the identi-
fication of source waters, i.e., identifying the source of tracer mass in a particular location.
In principle, it should be possible to post-process the results of a uniformity test to map out
the distribution of particle positions reaching a cell backwards in time, thus mapping the
source of tracer mass to that cell as a function of backward time. However, if the RWPT
model is not able to maintain a statistically uniform particle density over time, then the
results could be biased, i.e., regions with higher particle density could be overrepresented as
source regions, or underrepresented if particles get “stuck”. Future research into how such
studies might be conducted would be useful.
The identification of source waters is related to the concept of particle age, in fact it is a generalization since age requires a well-defined boundary. As discussed by Engqvist et al. [29], age and residence time can be considered the forward time and backward time components of the total transit time (or hydraulic residence time) of a particle in a domain. The particle trajectory methods which ignore diffusion (e.g., [22, 6, 24, 18, 7, 25, 54, 29], some of which handle convection and eddy-induced transport) are equally well-suited to handle the backwards and forward leg of a particle’s journey; indeed, often both the complete forward and backward trajectories are calculated and compared as an error diagnostic.

In theory, it should be possible to conduct backward time simulations with the RWPT approach presented herein as well, although this is not without controversy. For example, Engqvist et al. [29] concludes time-reversed simulation for any trajectory method is only possible if diffusion is completely disregarded. Likewise, it can be argued that backward time simulation of a diffusive process amounts to anti-diffusion, which is unrealistic since diffusion is an irreversible process. But this constraint does not seem true in the present case because we simulate the absolute diffusion of particles, as opposed to their relative diffusion (see discussion in Chapter 2 or [44]). Absolute diffusion is what the OGCM simulates, as dictated by the Eulerian nature of the calculation. Consequently, if RWPT particles are all released from the same point in space and time, they do not technically constitute a point source simulation. Rather, each particle is an individual realization of a turbulent flow field, and together many particles map out the ensemble average plume. This diffusive process is modeled by a stochastic movement of the RWPT particle, the magnitude of which is dictated by the time and space at which the particle exists, but not by the position of other particles in the simulation. As such a backward time trajectory could be valid, even though a diffusive process is modeled. This notion is supported by Flesch et al. [31] in the context of backward time Lagrangian stochastic models to estimate gaseous emissions. It is also analogous to the backward integration of the advection-diffusion equation with a “time-reversed flow” [17] used by the adjoint sensitivity approach (e.g., Hill et al. [39], Delhez et al. [17]) as well as the closely related Green’s function “boundary propagator” approach of Holzer and Hall [41], Primeau and Holzer [62], and references therein. Future research should be directed at resolving this issue, since backward integration would make RWPT a considerably more
Beyond residence time and source water investigations, it might be possible to use the RWPT approach to reduce simulation time for some tracer transport calculations. For example, ocean biogeochemistry simulations track the transport of many different chemical compounds simultaneously, and thus require many concurrent tracer simulations within an OGCM framework. If the evolution of each compound could instead be diagnosed from a single particle simulation, then overall simulation time could be greatly reduced. For example, in the sequestration efficiency calculation, particles were used to evaluate sequestration efficiency for any value of surface decay $k$ without recalculation of the particle trajectories. The problem in employing RWPT in this manner is, however, that the approach is not well-suited to simulate any chemical dynamics which are not first order. A first order decay such as the one used to approximate outgassing in the sequestration calculations is convenient in the RWPT approach because it does not require knowledge of chemical concentration; rather, the mass of each particle can simply be reduced. It was used in place of a traditional diffusive loss calculation at the air-water interface because this would require knowing the concentration of CO$_2$ near the surface, such that the gradient of pCO$_2$ across the interface could be defined. This would require binning particle densities to compute concentration at regular intervals throughout the simulation, performing the diffusive flux calculation, and then redistributing the mass across the particles in each surface grid cell. It would invalidate or at least greatly complicate the approach described in Section 4.5.1 in which one particle can be used to generate multiple estimates of sequestration efficiency. However, there may be ways of adapting RWPT simulations to be useful in handling chemical kinetics that are not first order. For example, perhaps the reactive part of the transport could be evaluated far less frequently than the particle transport and thereby reduce computational effort. Future research should explore how RWPT might be used in this regard.

4.5.4 Adjoint vs RWPT

Further investigation is needed to understand the relative benefits of the adjoint and RWPT methods. The adjoint sensitivity method used by Hill et al. [39], which is applied in a general residence time context for semi-enclosed water bodies by Delhez et al. [17], is clearly
a powerful and elegant technique. Closely related is the “boundary propagator” approach of Holzer and Hall [41] to estimating tracer age and transit times, discussed in the context of OGCMs by Khatiwala [45], Primeau [61] and by Primeau and Holzer [62]. Relative to the RWPT approach, these techniques have the advantage of staying within the Eulerian framework of an OGCM grid. This makes it likely that no arbitrary choices regarding the sub-grid scale variation of variables are required. Given the difficulties described in RWPT implementation in the present study, this could be a significant advantage.

The grid-based framework of the adjoint methods is, however, also a disadvantage in that the computational expense of adjoint methods increases nonlinearly with increased grid resolution, similar to traditional OGCM calculations. As discussed in Sections 4.5.1.4 and 4.5.2.4, the RWPT calculations may be conducted with higher resolution hydrodynamic forcing with only a modest increase in simulation time. This follows from the fact that the same number of particles can be used, so that the main additional computational burden is a larger data record from which to look up particle interpolation variables. Thus, while the adjoint method is only useful for OGCM configurations with which long-term simulations are also feasible, the RWPT approach can be used to conduct long-term residence time calculations with more finely resolved hydrodynamics. Moreover, as the grid resolution increases, the influence of the sub-grid scale assumptions within the RWPT model decreases, thereby likely improving the quality of the RWPT predictions.

In addition, there may be situations in which RWPT enjoys other advantages. For the carbon sequestration simulations, for example, the RWPT results could be immediately reinterpreted for various surface decay rates $k$, whereas such evaluations might require new adjoint simulations. In this particular example, the RWPT approach offered a benefit similar to the Green’s function approach of Khesgi et al. [47], whereby a set of OGCM results for point source releases could be reevaluated with multiple atmospheric CO$_2$ scenarios. Another advantage to the Lagrangian framework is that the impact of individual terms in the transport equation can readily be studied (e.g., diffusion can be turned off without compromising numerical stability). Future research aimed at characterizing the relative benefits and disadvantages of RWPT and adjoint methods would be useful.
4.6 Conclusions and recommendations for future research

The main goals of the present study were largely achieved. First, an RWPT model was successfully derived and implemented within the LLNL MOM framework. Although the model would benefit from additional refinement (Appendix 4.A), it is largely successful at reproducing OGCM computed transport, capturing the major features of tracer spreading from a variety of test release sites. This supports both the validity of the derived RWPT equations as well as many aspects of the model implementation, e.g., the novel vertical diffusivity handling technique developed in Chapter 3. Second, the utility of the RWPT approach to efficiently generate OGCM residence time statistics was demonstrated in the computation of global CO$_2$ sequestration efficiencies.

The recommended model refinements, some of which are already underway, focus on improving the treatment of the isopycnal terms such that the noted artificial upwelling problem can be reduced or removed. In particular, a revised treatment of the Gent McWilliams eddy-induced transport velocity is recommended, along with a vertically oriented isopycnal variable definition/interpolation scheme. Results from an earlier RWPT implementation suggest that the latter modification has good potential.

The fact that additional refinement is recommended should not be interpreted to mean that the ultimate goal of mimicking OGCM transport is unattainable; a high level of agreement has already been achieved. Rather, it highlights the iterative nature of RWPT model development, attributable to the large number of arbitrary choices that must be made in implementing the RWPT equations. The quality of the RWPT results relative to the OGCM has improved substantially over the iterative development process, and the recommended refinements seek to continue this improvement. The main remaining challenge is specifying the sub-grid scale variation of variables in a way that is consistent with the OGCM grid-based calculation. This process is complicated by a number of “fixes” that have been added to standard OGCM frameworks over the years, including diffusivity tapering, vertical mixing by means of large diffusivities with great spatial variability, and the addition of a parameterized eddy-induced transport. While these have improved OGCM accuracy, they tend to
create an unfriendly and awkward environment for a diffusing Lagrangian particle, in a sense forcing it to “play by someone else’s rules”. As parent OGCM resolution improves, the importance of these issues should decrease (the OGCM configuration used in the present study was relatively coarse by current standards, 4° x 2° resolution). Since RWPT simulation time would likely only be modestly impacted by an increase in grid resolution, using a more finely resolved hydrodynamic forcing would have the double benefit of reducing the importance of remaining errors in the RWPT implementation as well as making predictions more accurate because of the better resolved physics of the circulation field.

A comprehensive view of the LLNL MOM predicted distribution of CO₂ sequestration efficiencies was achieved in the application section, yielding results generally consistent with past OGCM studies. The simplified sequestration efficiency metric from Hill et al. [39] was used, i.e., one that assumes a small perturbation to surface pCO₂ and ignores “pushback” from the atmosphere. The major findings are that sequestration at 3,000 m is substantially more effective than shallower depths, with a mean efficiency of 44% after 1,000 years, and with much less spatial variability. For an 800 m injection, the deep water formation regions in the North Atlantic and Southern Oceans offer the highest efficiencies, and in general the Atlantic outperforms the Pacific. At 3,000 m, the Pacific outperforms the Atlantic, with efficiencies on par with the deep water formation regions. In addition, the Arctic Ocean generally has high sequestration efficiency. From an implementation standpoint, the analysis suggests that discharge technologies should seek to maximize depth in regions of high sequestration efficiency and low spatial variability. In addition, a spatially distributed source might be the best alternative because it reduces the influence of local “hotspots” of poor sequestration efficiency (the present study used a coarse grid resolution; finer eddy-resolving models would reveal smaller scale heterogeneity).

Overall, it is concluded that the RWPT model outlined herein could be a valuable tool for future oceanographic investigations aimed at diagnosing OGCM transport pathways (e.g, residence time calculation, source water identification) or understanding the relative importance of transport processes (i.e., transport terms in the particle model can be “turned off” to an extent not possible within the OGCM). While most Lagrangian particle models can to some extent be used in these manners, the present model has the advantage of being specif-
ically designed to mimic the parent OGCM. As such, it should be particularly well-suited as an OGCM diagnostic tool. While other Lagrangian approaches may have advantages such as better computational efficiency or more physically realistic diffusion, they do not attempt to mimic the OGCM transport calculation to the same extent.

A main alternative to RWPT, which appears to offer many of the same advantages, is adjoint sensitivity and related approaches (e.g., [17, 39, 41, 45, 61, 62]). Relative to RWPT, the adjoint approach has the advantage of staying within the OGCM framework, and thus may not require any arbitrary choices regarding the sub-grid scale variation of variables. On the other hand, RWPT incurs a much smaller computational penalty as the grid resolution of the parent OGCM is increased, and thus can allow long-term residence time evaluations for grid resolutions that would be impractical for the adjoint method. Also, adjoint approaches are likely not able to “turn off” transport terms to the extent possible with an RWPT model. Future research on characterizing the relative benefits of RWPT and adjoint methods is recommended.

Further research into how the RWPT model is best used in oceanographic investigations is also recommended. Issues such as particle seeding and book-keeping require careful consideration, optimizing computational efficiency while avoiding bias in the statistics extracted from RWPT results. In this regard, the possibility of backward-time RWPT simulation should be investigated, since it has the potential to substantially enhance the utility of the RWPT approach in source water investigation. Beyond residence time and source water investigations, RWPT may be an attractive option to improve the simulation time of traditional OGCM tracer transport calculations where many different chemical species must be tracked, although additional research would need to be directed at developing efficient approaches to simulating non-first-order chemical transformations within RWPT.

Lastly, implementing the RWPT approach within additional OGCMs would be useful. This would provide insight into whether the implementation issues encountered in the present study are germane to all OGCMs or specific to the LLNL MOM configuration used herein. It would also illustrate the extent to which the RWPT approach can serve as an intercomparison tool between different OGCMs.
4.A Appendix: improving the RWPT implementation

Because a large number of fairly subjective choices are required in defining and interpolating the OGCM variables for the particle model, the development of the RWPT model has been a highly iterative process in which various treatments of each term have been explored. Throughout this process it was common that previously unidentified shortcomings in the treatment of specific terms were revealed as improvements were made to other terms. The present RWPT implementation has resulted from this process of successive refinement, and for the purpose of the present study, it is the end result. It is, however, important to note that several recent improvements in the implementation have allowed a clear vertical transport bias to be revealed, i.e., the artificial particle upwelling identified in Section 4.4.2. Previous formulations generally suffered from particle deficits near the surface with particle accumulations at mid depth and near the bottom, but did not demonstrate a consistent bias. The fact that the recent improvements (e.g., to the isopycnal slopes and diffusivities at boundaries, and the treatment of $\frac{\partial A IS^2}{\partial z}$) have significantly improved the agreement between RWPT and OGCM predictions while revealing the vertical bias is an important development. It clearly identifies the goal of future development, namely, to find a variable definition/interpolation scheme which removes vertical bias without degrading the present level of agreement between the RWPT and OGCM predictions. Furthermore, the fact that, aside from the bias, the RWPT results compare favorably with the OGCM results suggests that the model to model agreement can be further refined by addressing the bias. The following section first describes the present understanding of the vertical bias, and then interprets these findings in the context of guiding future model refinements.

Before proceeding, however, it is noted that the importance of the remaining errors in the RWPT implementation would likely be reduced by improving the grid resolution of the parent OGCM. In addition, using better resolved hydrodynamics would likely improve the quality of RWPT predictions because the physics of the circulation field would be better resolved. Thus, the improvements outlined below may not be necessary if the RWPT model developed in Section 4.3 is employed with a more finely resolved parent OGCM.
4.A.1 Analysis of bias in RWPT implementation

The path of future model development is made considerably easier by the fact that an observable vertical bias has emerged in the RWPT results. Without a clear bias, the source of disagreement between the RWPT and OGCM predictions must be attributed to pathologies within either the RWPT or OGCM tracer calculations, or both. In the present case, however, it is known that a bias within the RWPT model exists, and that removing this bias is likely to cause the RWPT to OGCM comparisons to improve given the tracer test results in Section 4.4 and the sequestration efficiency calculations of Section 4.5. Extensive sensitivity analysis was conducted to understand which terms in the RWPT implementation cause the observed bias. In particular, specialized simulations were conducted where groups of terms in (4.43) were turned off. The major findings of this analysis are described below.

At present, the vertical bias is mainly caused by the isopycnal terms. Figure 4-77 shows a uniformity test (conducted in the same manner as for the base case) in which isopycnal slopes are set to zero everywhere, i.e., the isopycnal diffusivity becomes a horizontal diffusivity that is untapered everywhere. The vertical bias in Figure 4-50 is removed. This finding is supported by 100-year tracer tests using the same assumption, as shown in Figure 4-78 and Figure 4-79 for Sites a3 and g4, respectively. The OGCM results in these figures were generated with a modified code in which passive tracers experienced zero isopycnal slope, but the simulation of active tracers was left unaltered to prevent any changes in the flow or vertical diffusivity field. These locations were selected because they exhibited some of the largest errors in vertical tracer distribution in the base case. Comparing these simulations to the base case (Figures 4-14 and 4-43), it is noted that: (1) the impact of the isopycnal terms is significant; (2) the RWPT and OGCM agreement improves when the isopycnal terms are left out; and (3) the overprediction of surface concentrations is not present with no isopycnal terms (in fact, a ∼10% underprediction occurs at a3). Thus, results indicate that the RWPT bias is mainly due to the isopycnal terms. Without the isopycnal terms, the model agreement is generally good despite the sharp gradients and discontinuities in vertical diffusivity which tend to occur in these regions (Figure 4-9), i.e., the treatment of vertical diffusivity in the RWPT model seems effective.
Figure 4-77: RWPT uniformity test results averaged in time over interval 150-200 years, with zero isopycnal slope. Maps show average concentration normalized by the maximum concentration in each direction. Line plots show average concentration normalized by the mean concentration in each direction.
Figure 4-78: Comparison of LLNL MOM (top 3 maps and dashed line) and particle model (bottom 3 maps and solid line) results for tracer release location a3 after 100 years of simulation. Isopycnal slopes are zero everywhere in these simulations.
Figure 4-79: Comparison of LLNL MOM (top 3 maps and dashed line) and particle model (bottom 3 maps and solid line) results for tracer release location g4 after 100 years of simulation. Isopycnal slopes are zero everywhere in these simulations.
Of the isopycnal terms, the Gent-McWilliams (GM) eddy-induced transport terms are mainly responsible for the observed vertical bias. Figure 4-80 shows a uniformity test in which only the GM terms were turned off, i.e., $\kappa = 0$ in (4.43). The remaining terms in (4.43) should satisfy the well-mixed condition. The vertical bias in the base case (Figure 4-50) is for the most part removed. There is a deficit of particles near the ocean bottom ($\sim 10\%$), but this bias does not extend throughout the water column. Likewise, there is a deficit in the surface layer relative to the adjacent layer, but this bias also does not extend throughout the water column. Both of these deficits are absent from the simulations with zero isopycnal slope, i.e., they can also attributed to the isopycnal terms.

To confirm the impact of the GM terms, select 100-year tracer release simulations were rerun without these terms for both the RWPT model and the OGCM. For the most part, the model to model agreement improved. For example, Figures 4-81 and 4-82 show the results for tracer locations a3 and g4. At these locations there is a marked improvement in overall agreement compared with the base case (Figures 4-14 and 4-43), and the tendency to over-predict surface concentrations is greatly reduced. A similar improvement was noted for many other locations tested (e.g., a1-a4, g3-g5), with the main exception being some near bottom releases (e.g., a5, b5, c5). Although removing the GM terms does not yield perfect model agreement, it demonstrates that the artificial upwelling is mainly caused by these terms. This provides a clear target for future RWPT refinement.

Experience has also demonstrated that the $\partial A_p S^2 / \partial z$ term is problematic within the RWPT implementation. It is likely the main cause of the remaining biases beyond those introduced by the GM terms, i.e., the aforementioned underprediction of tracer in the bottom/top layers. As discussed in Section 4.3.2.1, this term is poorly defined in the present variable definition/interpolation scheme, and is also a target for future refinements.

Despite the vertical bias, the ability of the RWPT model to capture the influence of the various groups of terms is noteworthy. This is best illustrated by reviewing the 100-year tracer release results for Sites a3 and g4 in the following order: (1) zero isopycnal slopes (Figures 4-78 and 4-79); (2) no GM terms (Figures 4-81 and 4-82), and (3) base case (Figures 4-14 and 4-43). For example, consider Site a3, where tracer is released at about 1 km depth. With no isopycnal slopes (Figure 4-78), peak tracer concentration remains in the upper 1
Figure 4-80: RWPT uniformity test results averaged in time over interval 150-200 years, excluding the Gent-Mcwilliams eddy-induced transport terms. Maps show average concentration normalized by the maximum concentration in each direction. Line plots show average concentration normalized by the mean concentration in each direction.
Figure 4-81: Comparison of LLNL MOM (top 3 maps and dashed line) and particle model (bottom 3 maps and solid line) results for tracer release location a3 after 100 years of simulation. The Gent-McWilliams eddy induced transport terms are removed from both models.
Figure 4-82: Comparison of LLNL MOM (top 3 maps and dashed line) and particle model (bottom 3 maps and solid line) results for tracer release location g4 after 100 years of simulation. The Gent-McWilliams eddy induced transport terms are removed from both models.
km of the ocean with a maximum at the surface and significant zonal and meridional spreading. Introducing isopycnal slopes without the GM terms (Figure 4-81) causes an enhanced downward transport such that the peak occurs at the ocean bottom, with a secondary peak at the surface. Horizontal spreading is notably reduced as a larger fraction of tracer mass becomes trapped by the Antarctic Peninsula. Also, the characteristic near-bottom deficit of concentration previously noted emerges with the addition of the isopycnal terms. Finally, the addition of the GM transport (Figure 4-14) greatly enhances the downward transport, removing the secondary surface peak and further reducing horizontal spreading. The characteristic overprediction of surface concentrations emerges, and the concentration deficit near the ocean bottom increases. Here the RWPT model underpredicts the downwelling effect of the GM terms, i.e., as if particles experience an artificial upwelling. Nonetheless, the RWPT model generally reproduces the effect of each set of terms.

4.A.2 Paths for future refinement

The preceding analysis suggests that the RWPT implementation could be improved by refining the treatment of isopycnal terms. The goal is to remove the vertical bias without degrading agreement between the models.Outlined below are two proposed paths to achieving this, both of which are already under investigation, along with some secondary issues that could warrant investigation in the future.

4.A.2.1 Enforcing the non-divergence of the Gent-McWilliams terms

As discussed in Section 4.2.1.2, the GM transport can be implemented in two ways in OGCMs:

1. an advective flux of the form,

   \[ \mathbf{u}_{\text{GM}} = \left( -\frac{\partial (\kappa S_x)}{\partial z}, -\frac{\partial (\kappa S_y)}{\partial z}, \frac{\partial (\kappa S_x)}{\partial x} + \frac{\partial (\kappa S_y)}{\partial y} \right), \]  
   \hspace{1cm} (4.68)

   where \( \nabla \cdot \mathbf{u}_{\text{GM}} = 0 \) and the normal velocity is zero at all boundaries.
2. a skew diffusive flux in which the small-angle isopycnal diffusion tensor becomes

\[
E = \begin{pmatrix}
A_I & 0 & (A_I - \kappa) S_x \\
0 & A_I & (A_I - \kappa) S_y \\
(A_I + \kappa) S_x & (A_I + \kappa) S_y & A_I S^2 + K_v
\end{pmatrix}
\] (4.69)

where \( \kappa \) is the GM thickness diffusivity.

LLNL MOM employs the latter formulation, widely adopted by OGCMs due to its superior computational characteristics. Specifically, the diffusive form does not require the calculation of gradients in isopycnal diffusivity or slope, and actually reduces computational time when \( \kappa = A_I \).

This same choice between formulations is, however, not enjoyed by the RWPT model. As discussed in Section 4.3.1.2, both GM formulations yield the same RWPT equations, i.e., (4.43). Thus, the GM terms must be specified in the advective form in the RWPT model. Although a cancellation of terms occurs in the horizontal directions between the GM terms and other isopycnal terms, the vertical GM term must be calculated. The noisiness of such interpolations is one reason why the skew diffusive flux formulation is easier to implement [36]. Indeed, much of the difficulty of implementing RWPT stems from the fact that all isopycnal variables (GM and non-GM) within the RWPT model require spatially interpolating isopycnal diffusivity and/or slope.

The noisiness of the advective form of the GM terms caused them to be implemented in a special manner within OGCMs, which differs from (4.68). As per Danabasoglu and McWilliams [14], \( w_{GM} \) was usually calculated as the divergence of the horizontal velocities, i.e., analogous to the calculation of vertical advection within 3D OGCMs. Furthermore, the vertical GM velocity was set to zero at the surface and bottom of the ocean, thus satisfying the constraint that the GM velocity be zero at all boundaries. However, imposing the double boundary condition over-constrains the vertical integration, meaning that the non-divergence constraint was not satisfied in the top or bottom layer, depending on whether the integration started from the ocean surface or bottom (surface was apparently more common).
In the present RWPT implementation, only the vertical Gent-McWilliams velocity is computed \( w_{GM} = \frac{\partial (\kappa S_x)}{\partial x} + \frac{\partial (\kappa S_y)}{\partial y} \); the horizontal components \( u_{GM} \) and \( v_{GM} \) are canceled by their isopycnal diffusion counterparts because \( \kappa = A_I \). In principle, the present variable definition/interpolation scheme (Section 4.3.2.1) should be reasonably well-suited to estimate \( w_{GM} \) because the diffusivity and slope are continuous in the horizontal directions (unlike the vertical directions where they are step functions). In reality, however, \( w_{GM} \) seems to be causing a vertical bias in the RWPT results. One reason for this may be that non-divergence is not strictly enforced. The horizontal GM velocities are not well-defined in the present variable definition/interpolation scheme because they are based on the vertical gradient of isopycnal slope and diffusivity, but this was considered of secondary importance since these terms are canceled by their isopycnal diffusion counterparts when \( \kappa = A_I \).

Thus, one approach to refining the RWPT implementation is to change the treatment of the GM terms such that non-divergence is enforced. In this approach, the vertical random walk equation (4.43c) is changed to

\[
\Delta z = \left[ \frac{\partial (A_I S_x)}{\partial x} + \frac{\partial (A_I S_y)}{\partial y} + \frac{\partial}{\partial z} \left[ K_v + A_I S^2 \right] + w + w_{GM} \right] \Delta t \\
+ \xi_1 S_x \sqrt{2A_I \Delta t} + \xi_2 S_y \sqrt{2A_I \Delta t} + \xi_3 \sqrt{2K_v \Delta t} - S_y \frac{A_I}{Re} \tan \phi \tag{4.70}
\]

and \( w_{GM} \) is instead calculated from the divergence of \( u_{GM} \) and \( v_{GM} \) as defined in (4.68). Following the approach in [14] (i.e., the approach used by OGCMs before the diffusive was developed), the vertical integration would be performed from the surface to the bottom with \( w_{GM} = 0 \) enforced at the bottom as well. Although this leaves the non-divergence criterion unsatisfied in the bottom layer, it will guarantee non-divergence throughout most of the water column.

Testing of this refinement is ongoing, but initial attempts at implementation have proven unsuccessful, yielding pathological particle accumulations in certain regions. Nonetheless, there is evidence that this revised treatment could improve the RWPT implementation with further development. An example of both the pathological accumulation and the improvement is provided by Figure 4-83, which shows the results of a 100-year tracer release for Site g3. Compared to the base case (Figure 4-42) the vertical tracer distribution is much
improved in that surface tracer concentration is not overpredicted. However, a severe accumulation of particles occurs at approximately 60°N, perhaps because non-divergence is not enforced in the bottom layer. Additional investigation is required, and it is noted that some of the challenges of implementing the advective form of the GM transport within LLNL MOM are mentioned in Duffy et al. [27].

One hindrance to the efficacy of this refinement could be the present variable definition/interpolation scheme for isopycnal variables, in which isopycnal variables are treated as step functions in the vertical. Thus, the vertical gradient terms $u_{GM}$, $v_{GM}$, and $\frac{\partial A_I S^2}{\partial z}$ can only be treated in an approximate manner. The other proposed refinement to the RWPT implementation is therefore to improve this scheme to better resolve vertical gradients.

4.A.2.2 Modifying the isopycnal variable definition/interpolation scheme

One of the main challenges in designing an RWPT model to mimic the calculations of an OGCM is to prescribe the sub-grid scale variation of each variable. The OGCM’s grid-based tracer calculation only requires that variables be calculated on the grid cell faces. The value on each cell face represents a spatial average of the model’s state variables, carefully averaged in a way that is consistent with the governing finite difference equations. In contrast, the RWPT implementation must prescribe the sub-grid scale variation of each variable, which inevitably involves interpolating between gridded variables. Although the unconstrained nature of the sub-grid scale variation allows creativity, it also requires a choice to be made among a seemingly infinite number of options, none of which is guaranteed to achieve the ultimate goal of closely mimicking the OGCM. The present variable definition/interpolation scheme (Section 4.3.2.1) was one of many approaches tried during model development, and was selected based on certain favorable characteristics. Although it produces results that are impressive in many ways, it imparts a vertical bias through the treatment of isopycnal terms. This bias seems mainly caused by the GM terms, and to a lesser degree by the $\frac{\partial A_I S^2}{\partial z}$ term. While some improvement may be possible by tweaking the treatment of these terms within the present variable definition/interpolation scheme (as in 4.A.2.1), it is more likely that the scheme itself must be changed.

As explained in Section 4.3.2.1, a choice was made to interpolate gridded isopycnal slope
Figure 4-83: Comparison of LLNL MOM (top 3 maps and dashed line) and particle model (bottom 3 maps and solid line) results for tracer release location g3 after 100 years of simulation. Here the Gent-McWilliams terms are treated in the alternate manner described in Section 4.A.2.1.
values rather than estimating local isopycnal slopes by interpolating OGCM density values. The decision was based partly on computational expense but also on the inevitable noisiness of the slopes resulting from density interpolations, prone to sharp spatial gradients not present in the OGCM calculation. Despite the lingering problems in treating the isopycnal terms, this conclusion has not changed; using sub-grid scale interpolation of density values to calculate slopes is not a recommended approach. Rather, it is appropriate to interpolate between the gridded values because they are used directly by the OGCM.

It is not, however, clear which gridded slopes will yield the best results. As a brief review of Section 4.3.2.1, the present approach uses

- the $S_x$ and $S_y$ values on the zonal and meridional faces, respectively, and treats these as varying linearly across the cell face in one direction only, i.e., $S_x$ is a step function in $y$ and $z$ and $S_y$ is a step function in $x$ and $z$.

- independent isopycnal diffusivities are imposed in $x$ and $y$ ($A_{I,x}$ and $A_{I,y}$), based on the slope values on the cell faces.

The scheme was selected to yield a close approximation of the OGCM flux through horizontal cell faces. Of the isopycnal terms in (4.43c), all gradient terms except $\frac{\partial A_{I}S^2}{\partial z}$ are resolved in this scheme when the GM terms are included with $\kappa = A_{I}$. Results in Section 4.4.1 demonstrates that in general the model successfully reproduces OGCM horizontal tracer spreading. However, the preceding analysis suggests that RWPT results would benefit from better resolving vertical isopycnal terms. For example, experience has shown that the RWPT results are sensitive to the treatment of the $\frac{\partial A_{I}S^2}{\partial z}$ term, and that at present this term is responsible for at least some of the near-bottom and near-surface vertical bias that emerges in the uniformity and tracer release tests (e.g., Figures 4-80 and 4-81). In addition, when the GM terms are removed $\kappa = 0$, problems arise with near-bottom horizontal tracer spreading, presumably due to the introduction of $\frac{\partial (A_{I}S_x)}{\partial x}$ and $\frac{\partial (A_{I}S_y)}{\partial y}$ into the simulation. An example is tracer release at Site c5 (Figure 4-84), where horizontal tracer spreading without the GM terms is notably worse than the base case (Figure 4-26). Such errors are thought to be the cause of the pathological particle accumulations noted for the initial implementation of the alternate GM term treatment (Figure 4-83).
Figure 4-84: Comparison of LLNL MOM (top 3 maps and dashed line) and particle model (bottom 3 maps and solid line) results for tracer release location c5 after 100 years of simulation. The Gent-McWilliams eddy induced transport terms are removed from both models.
Perhaps RWPT accuracy would improve if a different set of isopycnal interpolation variables is used. Future research should explore using the isopycnal slopes defined at the center of the vertical cell faces instead. Although early RWPT development moved away from this option, much experience has been gained and several key unrelated improvements have been implemented subsequently (e.g., land boundary treatment and the meridional transport correction term). As such, revisiting the issue is warranted. Results suggest that vertical tracer transport may be more sensitive to terms involving vertical gradients ($\frac{\partial \tilde{A} I S^2}{\partial z}$, $u_{GM}$, and $v_{GM}$) than those involving horizontal gradients ($\frac{\partial (\tilde{A} I S_x)}{\partial x}$ and $\frac{\partial (\tilde{A} I S_y)}{\partial y}$), probably because the resolved spatial scale is much smaller in the vertical, i.e., more severe gradients can be encountered.

One approach is to use separate isopycnal slope and diffusivity values for the calculation of gradients in the horizontal and vertical directions. Some previous efforts implemented this in the treatment of the $\frac{\partial \tilde{A} I S^2}{\partial z}$ term, using the OGCM values on the vertical cell faces for this term only. This earlier version had some benefits which are noteworthy, including a less obvious vertical bias (see Figure 4-85) and better overall sequestration results (Figure 4-86). The predicted sequestration efficiencies are not uniformly underpredicted and at 1,000 years the mean and median sequestration efficiency were 11 and 7.1 %, compared to 18 and 17 % for the present code (based on 40 test sites). However, the 100-year tracer release tests clearly indicate that the present RWPT implementation better predicts tracer spreading for most of the test sites, suggesting that even though the older version performs better in terms of sequestration efficiency at the test sites, it does so for the “wrong reasons”. As an example, consider the results for location d4 in terms of sequestration efficiency (Figure 4-87; notably better than the current base case in Figure 4-59) and 100-year tracer distribution (Figure 4-88; notably worse than the current base case in Figure 4-30). This older approach was abandoned partly on this basis, and partly on a theoretical basis. As discussed in Section 4.3.1.4, each gradient term in (4.43c) exists to prevent spatial gradients in slope and/or diffusivity that are unresolved by the stochastic terms from causing unrealistic particle accumulations. The gradient terms cannot serve this purpose if they are based on slope/diffusivity fields different from those used to estimate these terms at the particle location, as was the case in the previous code. Nonetheless, these older results support
the notion that switching to a vertically oriented isopycnal variable definition/interpolation scheme might help to reduce the vertical bias in the present RWPT implementation.

Ongoing efforts are exploring a scheme in which isopycnal variables are exclusively defined by the OGCM values at the vertical cell faces, and treated as step functions in the horizontal. This resolves $\frac{\partial A_I S^2}{\partial z}$ and is expected to recapture some of the benefits noted for the previous code version. It would also be more compatible with the alternate GM term treatment proposed in Section 4.A.2.1 (since $u_{GM}$ and $v_{GM}$ would be better resolved). The challenge is the treatment of horizontal gradients, which would have to be handled in an approximate manner. Adjustments will have to be made to handle land boundaries and perhaps to enforce zero-diffusion barriers in the presence of large isopycnal slopes. Nonetheless, if the vertical bias can be removed with only a marginal degradation in horizontal tracer spreading, overall accuracy should improve.

4.A.2.3 Secondary issues

The previously proposed refinements, used either separately or in combination, are believed to hold the greatest potential for removing the present vertical bias and thereby improve the RWPT model’s ability to mimic the LLNL MOM. If this can be achieved, a number of secondary issues may come into play in evaluating remaining errors, as noted below.

The scheme used to handle sharp, discontinuous gradients in vertical diffusivity, which was specifically developed for the present application (see Section 4.3.2.2 and Chapter 3), is an approximate solution that balances fluxes such that the steady state well-mixed condition is eventually reached. However, as demonstrated in Chapter 3, the approximate nature of this scheme can impart inaccuracy over short timescales in unsteady conditions, i.e., when sharp gradients in particle distributions exist. The magnitude of this error is timestep dependent. In addition, there is an error between RWPT and OGCM calculations which exists even in the limit of an infinitesimal timestep, caused by fundamental differences between the implicit finite difference calculation on the OGCM grid and a particle simulation that interpolates diffusivity between cell faces. The present study uses a timestep (1,800 s) which is selected to keep the first type of error small even for very challenging diffusivity
Figure 4-85: RWPT uniformity test results averaged in time over interval 150-200 years, using an older code version. Maps show average concentration normalized by the maximum concentration in each direction. Line plots show average concentration normalized by the mean concentration in each direction.
Figure 4-86: Error summary for RWPT 500-year and 1,000-year sequestration efficiency calculations using an older code (for CO₂, k = 2 yr⁻¹), shown as RWPT vs OGCM sequestration efficiency (top panel), as well as RWPT relative error on a sequestration efficiency (middle panel) and mass lost (bottom panel) basis.
Figure 4-87: LLNL MOM and RWPT sequestration efficiency for Site d4 using an older code with surface decay rates $k = \frac{182}{5}$ yr$^{-1}$ (top) and $k = 1$ yr$^{-1}$ (middle). Dashed line: OGCM; solid line: RWPT using $t = 0$ particle positions; dotted line: RWPT results using $t > 0$ positions. Bottom panel: number of particles in RWPT result.
Figure 4-88: LLNL MOM (top 3 maps and dashed line) and RWPT model (bottom 3 maps and solid line) predicted tracer concentrations for release location d4 (dotted line) after 100 years of simulation using an older code.
profiles. The agreement observed for tracer release simulations with zero isopycnal slope was deemed acceptable (Section 4.A.1), and thus the performance of the vertical diffusivity scheme is taken as sufficient in the present RWPT model, i.e., concerns over the treatment of isopycnal terms far outweigh concerns over residual shortcomings in the vertical diffusivity scheme. Nonetheless, some of the differences between OGCM and RWPT results in the zero isopycnal slope case may be caused by the treatment of vertical diffusivity, and as the treatment of isopycnal terms is improved, such differences may warrant closer examination. For example, treating vertical diffusivity as a step function in the vertical (while still applying the particle balancing scheme developed in Chapter 3) may reduce both types of error noted above. This is clearly of secondary importance given the good agreement already achieved by the RWPT model, but is nonetheless noted in case it gains importance in future RWPT implementations.

Another issue that may be relevant in reconciling differences between future RWPT and OGCM results is the accuracy of the advection scheme within the latter. The present LLNL MOM configuration uses a standard centered difference scheme in the horizontal and the flux-corrected transport (FCT) scheme in the vertical (e.g., see Gerdes et al. [35]). To illustrate the influence of the OGCM numerical scheme, LLNL MOM simulations were rerun for location g after changing vertical advection to use (a) standard centered differences (2nd order) and (b) upwind advection (1st order). As advection schemes go, these represent extreme endpoints, with (a) being the most dispersive and (b) being the most diffusive. Only a marginal difference was observed between the base FCT case and the standard centered difference case. There was, however, a marked difference for the upwind results for all release depths, with the quality of agreement degrading for g1-g3 degrading and improving for g4-g5, most notably for g5 (Figure 4-89). It is tempting to conclude that the remaining error between RWPT and OGCM results is merely a result of OGCM numerics. This is, however, not true; the uniformity test (Figure 4-50) clearly demonstrates a bias in the RWPT results which is unaffected by the choice of OGCM advection scheme. Thus, caution must be used in attributing differences between RWPT and OGCM results to the OGCM numerics. Still, as the RWPT implementation is improved to remove the vertical bias, OGCM numerics should be considered before further refinement is attempted (i.e., at some point, minor variations
within the RWPT scheme cannot be distinguished by comparison to the OGCM due to variability in the choice of OGCM advection scheme).

A related OGCM numerical issue which can give rise to differences between the OGCM and RWPT predictions is the polar filtering applied in the zonal direction to correct errors caused by Courant number violations near the poles. This should be included in any future evaluation of OGCM advection schemes. However, it is not expected to be a significant factor; OGCM tracer release tests at Sites a and g showed negligible differences when polar filtering was disabled, despite their proximity to the South Pole.

In addition, an issue which may eventually warrant consideration is whether there are missing terms in the random walk equations (4.43) related meridional gradients in isopyc-nal diffusivity. As discussed in Section 4.3.1.3, the random walk equations were adapted to spherical coordinates by incorporating a meridional correction term \(-\frac{A}{R_e}\tan\phi\) in (4.43b) derived by Raible and Engel [63] and Brillinger [10], which corrects particle motion to account for the convergence of the meridians. Their derivations, however, did not consider a gradient in meridional diffusivity, which should also give rise to some adjustment to the meridional transport, and thereby vertical transport since meridional along-isopycnal displacements have a vertical component. Future efforts might be directed at rederiving the random walk equations directly in spherical coordinates, although this is not a trivial task since an appropriate Fokker-Planck equation was not found in the literature. It is, however, the author’s opinion that the impact of these corrections will be minor. They can certainly be ruled out as major contributors to the observed vertical bias, on the basis of several lines of evidence:

- the RWPT model currently does a good job of simulating horizontal tracer spreading,
- the vertical bias largely disappears when the GM terms are removed (despite the existence of meridional diffusivity gradients),
- any correction terms related to the gradient will have less impact than the meridional correction term already incorporated, which has only a marginal impact on the vertical bias if removed, and
Figure 4-89: Comparison of LLNL MOM (top 3 maps and dashed line) and particle model (bottom 3 maps and solid line) results for tracer release location g5 after 100 years of simulation. In this case, the OGCM result was calculated using upwind advection in the vertical.
the correspondence between RWPT and OGCM sequestration efficiencies at high latitudes is excellent (e.g., Sites RAN2 and RAN20 in Figures 4-90 and 4-92), and any missing correction terms would be largest at low and high latitudes like the existing meridional correction term.

More generally, it would be prudent to revisit the issue of a possible meridional bias in the uniformity tests once the vertical bias is addressed. The uniformity results for the base case (Figures 4-49 and 4-50) suggest a possible accumulation near the North Pole, although at least some of this can be attributed to noise in the RWPT results. The uniformity test for the zero isopycnal slope case (Figure 4-77) and the no GM case (Figure 4-80) do not show the same accumulation, from which one might infer that the base case accumulations are real and induced by the GM terms. The question then becomes whether a meridional bias is caused by the vertical bias, or vice versa (i.e., the chicken, or the egg?). The 100-year tracer release tests, which showed good agreement in meridonal tracer spreading, strongly suggest the vertical bias is the root cause of any meridional bias, if indeed it exists. Thus, the uniformity tests should be revisited once the GM-induced vertical bias has been addressed.
4.B Appendix: additional sequestration figures\textsuperscript{9}

\textsuperscript{9}Presented below are the sequestration simulations for the "RAN" and "ADD" injection locations, see Sections 4.5.2.1 and 4.5.2.2.
Figure 4-90: LLNL MOM (dashed line) and RWPT (solid line) computed sequestration efficiency (for CO$_2$, $k = 2$ yr$^{-1}$) for releases at 1,500 m at sites RAN1 - RAN6.
Figure 4-91: LLNL MOM (dashed line) and RWPT (solid line) computed sequestration efficiency (for CO$_2$, $k = 2$ yr$^{-1}$) for releases at 1,500 m at sites RAN6 - RAN12.
Figure 4-92: LLNL MOM (dashed line) and RWPT (solid line) computed sequestration efficiency (for CO$_2$, $k = 2$ yr$^{-1}$) for releases at 1,500 m at sites RAN13 - RAN18
Figure 4-93: LLNL MOM (dashed line) and RWPT (solid line) computed sequestration efficiency (for CO$_2$, $k = 2$ yr$^{-1}$) for releases at 1,500 m at sites RAN19 and RAN20.
Figure 4-94: LLNL MOM (dashed line) and RWPT (solid line) computed sequestration efficiency (for CO₂, $k = 2 \text{ yr}^{-1}$) for releases at site ADD1 and ADD2 at three different release depths.
Figure 4-95: LLNL MOM (dashed line) and RWPT (solid line) computed sequestration efficiency (for CO$_2$, $k = 2$ yr$^{-1}$) for releases at site ADD3 at three different release depths.
Bibliography


Chapter 5

Evaluation of the Environmental Viability of Direct Injection Schemes for Ocean Carbon Sequestration

5.1 Introduction

Carbon capture and storage (CCS) is increasingly being mentioned as one of the major options available to help reduce the build-up of greenhouse gases in the atmosphere [96, 56]. The ocean is potentially the largest carbon sink, and has the distinction that it is already being “used” under business as usual (BAU) operation: the net carbon flux from the atmosphere to the ocean is about one-third of the anthropogenic emission to the atmosphere, and over time more than two-thirds of the carbon we emit to the atmosphere will eventually wind up in the ocean. As a result of this practice, the surface ocean has already experienced a depression of about 0.1 pH units, which carries concern over impacts to coral and other near surface biota [56].

The logic behind direct ocean storage is that some of the CO$_2$ that we now put in the atmosphere could be input directly to the ocean, thus eliminating its deleterious effects on

---

climate. One general strategy calls for isolating the CO$_2$, e.g., in a deep lake where its negative buoyancy would reduce the exchange of CO$_2$ with the overlying water column. The CO$_2$ could also react with seawater to form solid CO$_2$ hydrates, which would further inhibit mass exchange. However, even with a hydrate covering, diffusion from the lake surface to the overlying water column would render such storage temporary [37, 49]. The other strategic endpoint, and the one considered herein, is to dilute the CO$_2$ by dispersing if over the largest possible volume such that the excess pCO$_2$ concentrations, and the changes in pH, will be as small as possible. In order that the sequestration be effective (i.e., that the CO$_2$ be retained in the ocean for as long as possible before being exchanged with the atmosphere), injection would need to be into intermediate and deep ocean depths, and away from areas of strong upwelling. Numerical simulations with ocean general circulation models suggest that most CO$_2$ injected at a depth of 3,000 m over a range of representative sites would be retained for more than 500 years [94, 56].

Clearly direct ocean injection is only worth considering if the impacts to organisms residing at intermediate and deep depths are substantially less than the avoided climatic impacts, plus impacts to near surface marine biota, that would accrue under BAU. Such environmental impacts constitute the general objective of this study. There are several other attributes of direct ocean injection. It has already been mentioned that the storage is temporary. However, while several centuries is certainly short by geological time scales, it should be long enough to help us find alternatives to fossil fuels. Second, while the ocean is not infinite, it is large. Over 500 GTC could theoretically be dispersed uniformly over the entire ocean, without depressing the average pH by more than 0.1 units [56] presumably causing a small effect that is comparable to that currently experienced by near surface waters [8]. As a point of reference, Pacala and Socolow [96] identify 175 GTC as the emissions needed to be avoided over the next 50 years in order to stabilize atmospheric concentrations below 500 ppm. Of course, the engineering task of dispersing the CO$_2$ uniformly is far from trivial. Finally, injection would be transparent. Unlike geological storage, where the CO$_2$ is hidden underground, the concentration of CO$_2$ injected into the ocean would gradually increase in ways that could be easily monitored. If an environmentally acceptable endpoint can be established, relatively simple monitoring can be used to determine when this endpoint is
From the above discussion it seems clear that any policy that includes direct ocean injection should be linked to a policy of substantially and quickly reducing CO\textsubscript{2} emissions. Policy implications of ocean injection are discussed further in the concluding chapter.

Marine impacts associated with direct injection can be categorized in several ways. First is the type of organism, which include fish, plankton, benthic organisms and micro-organisms. Our interest is mainly on plankton, because they can generally not avoid a plume, and because they generally reside within the water column where the exposure is greatest. Data also suggest that they are the most sensitive. Marine impacts can also be categorized in terms of degree and extent of impact, including acute impacts (mortality), sub-lethal impacts (e.g., reproductive effects), and ecosystem effects. Our interest, here, is mainly on acute effects, partly because these are the most dramatic, and also because they are easiest to measure. Some discussion of sub-lethal and ecosystems impacts is provided in Sections 5.2 and 5.5.

Auerbach \textit{et al.} (1997, [5]) examined available data on the acute impacts to planktonic marine organisms exposed to low pH. As part of this analysis they developed a procedure to integrate constant-concentration laboratory assay data with variable-concentration field exposures. In a companion study, Caulfield \textit{et al.} (1997, [19]) simulated the time-variable concentrations seen by passive organisms that were transported through plumes resulting from several scenarios for direct ocean injection. Using the approach of Auerbach \textit{et al.} [5], they identified regions of expected mortality for each scenario.

The bioassay data used by Auerbach \textit{et al.} [5] were collected largely to explore the impacts of acidic wastes in the coastal zone, and not CO\textsubscript{2} discharge in the deep ocean. However, within the past decade a large number of studies have been conducted to directly examine the biological impacts associated with ocean sequestration. Additional work has also been conducted to optimize injection scenarios to reduce CO\textsubscript{2} concentrations. The specific objective of the current research is to review these more recent studies and to re-evaluate impacts using the approach of Auerbach \textit{et al.} [5] and Caulfield \textit{et al.} [19]. Accordingly, Section 5.2 summarizes available data concerning acute impacts to marine organisms, and Section 5.3 describes the modeling approach used herein, which is adapted from the work of Auerbach \textit{et al.} [5], Caulfield \textit{et al.} [19], and several other studies. Section 5.4 presents
and interprets the predicted impact for the modeled discharge scenarios. Lastly, Section 5.5 summarizes the conclusions, provides recommendations for future investigation, and discusses policy implications.
5.2 \textbf{CO}_2 \textbf{Toxicity Studies}

Auerbach \textit{et al.} [5] and Caulfield \textit{et al.} [19] provided quantitative estimates of the acute biological impact of a \textit{CO}_2 injection into a marine environment. These studies were based on compiled toxicity data on the mortality of marine organisms, mainly zooplankton, due to a decrease in pH. Because none of these toxicity studies used \textit{CO}_2 as the acidifying agent, the relevance of these data to carbon sequestration depends on the extent to which the mortality due to a \textit{CO}_2 release is caused by the accompanying decrease in pH. Numerous studies over the past decade suggest that mortality for a given level of pH reduction is significantly enhanced when \textit{CO}_2 is the acidifying agent, thus implying that hypercapnia is a more important effect than the accompanying acidosis.

Another shortcoming of the data available to Auerbach \textit{et al.} [5] and Caulfield \textit{et al.} [19] is that they were limited to organisms that inhabit the upper ocean. Recent studies suggest that the \textit{CO}_2 tolerance of organisms in the upper ocean may be less than that of organisms in the deep ocean. Given that most ocean sequestration scenarios involve discharging \textit{CO}_2 into the deep ocean, this is an important distinction.

Before proceeding it is important to briefly describe the analytical framework that will be used so that it is clear to the reader what type of data are needed. The \textit{isomortality} approach developed by Auerbach is a way of adding up the exposure history of an organism (e.g., a zooplankter) and translate it into an acute impact. This is accomplished by constructing isomortality curves, i.e., lines of equal mortality, such as the ones shown in Figure 5-1, and using the exposure history to map out a cumulative harm trajectory in exposure time-stress level space (see Section 5.3 for details). Such curves can be mapped out using mortality statistics such as LC values, i.e., a concentration that is lethal to a certain percentage of the test population within a fixed exposure time (e.g., LC$_{50}$ for 50\% mortality), or equivalently LT values, i.e., an exposure time at which a fixed concentration will be lethal to a certain percentage of the test population (see [38, 118] for an overview). Thus, the primary goal of the following literature review is to identify sources of acute toxicity data for as many species as possible, which can then be used to construct revised isomortality functions. In addition, the review will attempt to synthesize the current state of knowledge regarding the
acute response of various organisms such that the isomortality analysis can be adjusted and evaluated accordingly.

5.2.1 pH toxicity studies

The existing dataset on the mortality of marine organisms in the presence of lowered pH due to an acidifying agent other than CO$_2$ is reviewed first. This consists mainly of the dataset compiled by Auerbach et al. [5] and a more recent study by Yamada and Ikeda [145].

5.2.1.1 Auerbach et al. (1997)

The following studies were included in Auerbach et al. [5]:

- Bamber (1987, [6]) studied the pH mortality of young carpet-shell clams.
- Bamber (1990, [7]) studied the pH mortality of three commercial bivalve mollusc species.
- Brownell [12] studied the pH mortality of the larvae of 8 marine fish species.
- Calabrese and Davis [14] looked at the pH tolerance of embryos and larvae of two commercial bivalve mollusc species.
- Grice et al. [47] looked at the mortality and reproductive effects of three zooplankton when exposed to acid waste of varying pH levels.
- Rose et al. [106] looked at the pH mortality of a marine copepod.
- Portmann [100] studied the pH mortality of brown shrimp.

Auerbach [4] provides little explanation as to how he extracted and condensed the data to arrive at a “representative” dataset. It is however clear that the dataset presented does not consist solely of the raw data but is rather a subset that has been manipulated to some extent. For example, some mortality data were adjusted based on observed decreases in reproductive rates. For the present purpose, we will simply take his compiled data as given. Looking ahead, these data will not be used in the present analysis since more appropriate
Figure 5-1: pH mortality data for the combined dataset of Auerbach et al. [5] for a variety of marine zooplankton and benthic species. Also shown are the isomortality lines developed by [5] from these data.

(CO$_2$-induced mortality) data now exist. Figure 5-1 shows the data extracted by Auerbach along with the isomortality curves developed from these data.

5.2.1.2 Yamada and Ikeda (1999)

Given the fact that the mortality dataset that existed at the time of Auerbach was restricted to shallow-water organisms, Yamada and Ikeda [145] looked at the pH sensitivity of 10 marine zooplankton species. Plotted in Figure 5-2 is a comparison of the Auerbach et al. [5] dataset and the data collected by Yamada and Ikeda [145]. The combined dataset is compared to CO$_2$-induced toxicity data later in this chapter, which shows that marine organism are more sensitive to CO$_2$ than pH depression alone. This unfortunately makes the dataset plotted in Figure 5-2 of limited value to the present analysis.
Figure 5-2: pH mortality data from the Auerbach [4] combined dataset (filled symbols) and the Yamada and Ikeda [145] dataset (open symbols): $\text{LC}_0$ (top) and $\text{LC}_{50}$ (bottom).
5.2.2 CO₂ toxicity studies of fish

The fish studies are divided into two groups: adult and developmental. The developmental group includes eggs, larvae, and juvenile/young fish. Although the set of fish articles reviewed below are restricted to those with original data, there exist a number of in-depth review articles on the biological effects of CO₂ on fish. In particular, Ishimatsu et al. [59, 61] synthesized much (but not all) of the data presented below on fish (developmental and adult stages) and offered insight into the state of knowledge on the effects of hypercapnia on fish.

5.2.2.1 CO₂ toxicity studies of adult fish

Four main studies of lethal hypercapnia on adult fish were found: Grøttum and Sigholt [48]; Lee et al. [79]; Hayashi et al. (2004a, [52]); and Hayashi et al. (2004b, [51]). In addition, there are some data in Takeda and Itazawa [124]. These studies were not limited to overall mortality of marine fish species; in each case various indicators of blood composition were measured in an effort to understand the physiological response of the fish to lethal and sub-lethal exposures. Of the four main studies, all except Grøttum and Sigholt [48] were motivated by carbon sequestration. The CO₂ exposures tested were similar in each case, with a maximum pCO₂ of nearly 8 kPa.

There have been many studies of sub-lethal effects of CO₂ on fish; Grøttum and Sigholt [48] refer to this body of work as comprehensive and recommend [53] for an overview. For example, sub-lethal exposures have been studied in freshwater species such as carp [27, 147, 123, 146], tench [66], channel catfish [13], brown bullhead [45], European eel [31, 85], white sturgeon [28, 29], and multiple species of rainbow trout [115, 16, 65, 97, 99, 64, 10, 98]. Examples of marine species for which sub-lethal hypercapnic effects have been studied include Pacific spiny dogfish [30, 26, 84], Conger conger (a marine teleost) [131], the seawater salmon [97], Atlantic salmon [39, 40], spotted wolffish [42], larger spotted dogfish [103], spotted skate [46] and cod [78]. Additional references may be found in the literature review provided by Kita and Ohsumi [72].

The studies of lethal effects are described briefly below. In each case, crude estimates of LC₅₀ values have been generated whenever possible by simple linear interpolation between
the reported mortalities. This approach was used for the adult fish dataset in favor of a more rigorous method (e.g., [38, 118]) for two reasons. First, the data are sparse with generally few partial kills making curve fitting problematic, and second, these values are not used further in the present study in any isomortality-type analysis. The resulting data for all adult fish species are combined and discussed at the end of this section.

5.2.2.1.1 Grøttum and Sigholt (1996) Motivated by the potential buildup of CO$_2$ that may occur during commercial farming of European seabass (Dicentrarchus labrax, a pelagic teleost [107]), Grøttum and Sigholt [48] studied the acute toxicity of carbon dioxide on this species, measuring mortality as well as plasma ion levels. 14 adult fish were placed in each of 7 flow-through tanks and were exposed to different steady concentrations of CO$_2$ for a period of 120 hours. The maximum pCO$_2$ studied was 62.3 mmHg (8.31 kPa), and pH values were also reported for each pCO$_2$ studied. LC$_{50}$ values were reported for 48, 72, 96, and 120 hour exposures, and the mortality curve for the 120 hour exposure was given. The study also reports plasma Cl$^-$, Na$^+$, and lactate concentrations.

5.2.2.1.2 Lee et al. (2003) Lee et al. [79] studied the lethal and sub-lethal effects of CO$_2$ on yellowtail (Seriola quinqueradiata, a pelagic teleost [51]). 5 and 6 yellowtail were exposed to a steady pCO$_2$ of 7 mmHg (0.93 kPa) and 38 mmHg (5.07 kPa), respectively, for a period of 72 hours. Most of the discussion in the study is devoted to the cardiorespiratory and blood-gas responses of the fish, based on data such as heart rate, cardiac output, blood pressure and pH, and blood levels of O$_2$, CO$_2$, lactate, hematocrit, and bicarbonate. The authors tentatively conclude that cardiac failure is the primary cause of mortality induced by high CO$_2$ concentrations.

5.2.2.1.3 Hayashi et al. (2004a) Hayashi et al. (2004a, [52]) compared the responses of Japanese flounder (Paralichthys olivaceous, a benthic teleost [51]) to high CO$_2$ exposure with the response to acidification using sulfuric acid. The goal of the study was to determine whether the toxicity of aquatic hypercapnia is due to the CO$_2$ exposure or to the accompanying decrease in ambient pH. Of the 11 flounder included in the study, 6 were exposed to seawater equilibrated with an air mixture containing 5% CO$_2$, and the remaining 5 were
exposed to seawater to which 1 N H$_2$SO$_4$ had been added. The resulting pH in both cases was 6.18 (2 pH units lower than the unaltered seawater). Although the authors did not report the resulting aqueous pCO$_2$ value, it is reported as 4.95 kPa in Hayashi et al. (2004b, [51]) using the same experimental setup. The differences in mortality for the two groups is dramatic: all of the fish in the CO$_2$ group died between 3 and 48 hours of exposure, while none of the fish in the H$_2$SO$_4$ group died after 72 hours of exposure. A similar difference was noted in the response of arterial pH, hematocrit, pCO$_2$, and plasma ion concentrations (HCO$_3^-$, Cl$^-$, Na$^+$, and K$^+$).

5.2.2.1.4 Hayashi et al. (2004b) The same authors as in the preceding article publish additional data in Hayashi et al. (2004b, [51]), focusing only on toxicity due to hypercapnia. Three fish species were studied: Japanese flounder (*Paralichthys olivaceus*), yellowtail (*Seriola quinqueradiata*), and starspotted dogfish (*Mustelus manazo*, a demersal elasmobranch). All three species were subjected to seawater equilibrated with an air mixture containing 1, 3, and 5% CO$_2$, and the dogfish were also exposed to an air mixture containing 7% CO$_2$ (the exposures correspond to pCO$_2$ values of 0.99, 2.97, 4.95, and 6.96 kPa, respectively). The pH of the unaltered seawater was 8.18, and the 1, 3, 5, and 7% exposures resulted in pHs of 7.01, 6.41, 6.18, and 6.02, respectively. The mortality results were:

- For the 18 Japanese flounder tested, there was no mortality for the 1 and 3% CO$_2$ cases after 48 and 72 hours of exposure, respectively. For the 5% CO$_2$ case, mortality was 17, 33, and 100% at 8, 24, and 48 hours of exposure, respectively.

- For the 15 yellowtail tested, there was no mortality for the 1 and 3% CO$_2$ cases after 72 hours of exposure. For the 5% CO$_2$ case, mortality was 20 and 100% after 3 and 8 hours of exposure, respectively.

- For the 20 dogfish tested, there was no mortality for the 1, 3, and 5% CO$_2$ cases after 72 hours of exposure. For the 7% CO$_2$ case, mortality was 20% at 72 hours.

Thus, the data indicate that the CO$_2$ toxicity is in this case species specific; the yellowtail were more susceptible than the Japanese flounder, and the starspotted dogfish had considerably higher resistance than the other two species.
Physiological parameters measured during the study included arterial pH, hematocrit, pCO\textsubscript{2}, and plasma ion concentrations (HCO\textsubscript{3}\textsuperscript{-}, Cl\textsuperscript{-}, Na\textsuperscript{+}, and K\textsuperscript{+}). The authors found a difference between the acid-base regulatory mechanism in the two teleosts and the elasmobranch, and also concluded that the observed acid-base regulation mechanisms differ from the generally accepted model for marine fish. The physiological response factors are discussed at length in [51].

It should be noted that these data are also presented in review articles by Ishimatsu et al. ([61, 59]), which contain some additional discussion of sub-lethal effects.

5.2.2.1.5 Takeda and Itazawa (1983) Takeda and Itazawa [124] studied methods for sedating fish for the purpose of live transport. One of these methods involved bubbling CO\textsubscript{2} and O\textsubscript{2} through water throughout the transport. Two species were tested, carp and porgie (scientific names not given in English; the article is in Japanese). This sedation approach proved feasible for carp but not for porgies as the latter group exhibited mortality at the pCO\textsubscript{2} required to induce sedation. Mortality statistics are reported for exposure periods of 4, 8, and 22 hours, with pCO\textsubscript{2} values ranging from 4 to 104 mmHg (0.53 - 13.9 kPa) and pH values ranging from 5.72 to 7.48. Exposure pCO\textsubscript{2} and pH were reported as observed ranges as opposed to specific values, making the data less precise than the other studies on adult fish. According to Hayashi et al. (2004b, [51]) the porgie species is actually red sea bream (Pagrus major), which is assumed to be correct in the present study.

5.2.2.1.6 Combined adult fish dataset Figure 5-3 shows all of the adult fish LC\textsubscript{50} data together as a function of pCO\textsubscript{2} and pH. Based on the data from [52], hypercapnia appears to be a stronger stressor than the accompanying decrease in pH. However, the other studies did not consider mortality due to pH depression by another acid, and thus only one datapoint on the hypercapnia vs. acidosis issue is available from the adult fish dataset. The issue will instead be addressed in the developmental stage fish and copepod datasets. The adult fish data are compared to the other data later in this chapter.

The adult fish dataset is augmented with two studies relevant to evaluating the expected acute effects of CO\textsubscript{2} discharges on fish, Tamburri et al. [127] and Vetter and Smith [135].
Figure 5-3: Mortality data for adult fish as a function of pCO$_2$ (top) and pH (bottom). Dataset includes European seabass (*Dicentrarchus labrax*, squares [48]); Japanese flounder (*Paralichthys olivaceus*, circles [52, 51]); yellowtail (*Seriola quinqueradiata*, triangles [79, 51]); and red sea bream (*Pagrus major*, diamonds [124]). Reported values are shown as filled symbols; estimated values are shown as open symbols.
Both studies attempted to observe the effect of high in situ pCO$_2$ on nektonic animals (i.e., species that can swim enough to overcome the ambient current), and in particular to determine whether or not they avoided regions of high pCO$_2$. Although the species studied were not all fish, their findings are included here because fish are the only nektonic organisms considered in the literature review. Tamburri et al. [127] attracted benthic macrofauna to beakers on the seafloor by releasing an odor solution and, once enough animals had gathered, the odor solution was passed through CO$_2$ hydrate to enrich the seawater. Avoidance of the plume was not observed for fish or invertebrates; fish swam through the plume in a zigzag fashion following the odor regardless of whether or not the elevated pCO$_2$ had been introduced. One species of fish, a hagfish, was observed to repeatedly enter the beaker and become anaesthetized, and then eventually recover and swim away. Acute toxicity was not observed even though the pH was lowered by about 2 pH units; respiratory distress was, however, noted in some species. Overall the authors conclude that fish and invertebrates do not avoid high CO$_2$ regions, and speculate that the induced mortality by CO$_2$ discharges could be significantly enhanced if nektonic scavengers seek out dead animals in high CO$_2$ regions and perish themselves. These findings were contradicted by the observations of Vetter and Smith [135] on the behavior of amphipods and synaphobranchid eels (the latter being a member of the fish family) near a natural hydrothermal vent discharging seawater with high pCO$_2$. They observed avoidance of baited traps near the vent, relative to a control site with additional baited traps, indicating avoidance for all species in the region. They did, however, note significant narcotic effects to amphipods placed in the vent which they attributed to high pCO$_2$. Once removed from the vent, the amphipods recovered. Overall, their observations did not support the creation of a mortality sink in the vicinity of a CO$_2$ discharge. Thus, the data are inconclusive regarding whether or not fish and other nektonic creatures would avoid or seek out CO$_2$ discharges.

5.2.2.2 CO$_2$ toxicity studies of developmental stage fish

Four articles with original data on the tolerance of developmental stage fish to high levels of CO$_2$ were found, all by the same lead author: Kikkawa et al. (2003, [69]), Kikkawa et al. (2004, [70]), Kikkawa et al. (2006a, [71]), and Kikkawa et al. (2006b, [68]).
5.2.2.2.1 Kikkawa et al. (2003) Kikkawa et al. (2003, [69]) compared the mortality of red sea bream (Pagrus major) eggs and larvae due to hypercapnia with the mortality due to an equivalent level of acidification using another acid. In the acidification portion of the study, pH levels of 6.2 and 5.9 were achieved by adding varying amounts of 1 N HCl to seawater. In the hypercapnia portion of the study, these same pH levels were achieved by equilibrating seawater with air mixtures containing 5 and 10% CO\(_2\), corresponding to pCO\(_2\) values of 4.95 and 9.90 kPa, respectively. The eggs (embryos) used in the study were at the stage where auditory vesicles form (21 hours after fertilization) and were subjected to a 6-hour exposure time; the larvae were at the preflexion stage (10-12 days after hatching) and were given a 24-hour exposure time.

The differences in mortality between the strong acid and CO\(_2\) exposures were dramatic for both eggs and larvae. For eggs, the HCl and CO\(_2\) exposure cases resulted in mortalities of 3.6 and 85.8% for a pH of 6.2, and 0.9 and 97.4% for a pH of 5.9, respectively. For larvae, the HCl and CO\(_2\) exposure cases resulted in mortalities of 1.6 and 61.2% for a pH of 6.2, and 5.0 and 100% for a pH of 5.9, respectively. These findings are consistent with the findings of Hayashi et al. [52] for adult Japanese flounder, for which hypercapnia and not the accompanying acidification was the cause of toxicity.

5.2.2.2.2 Kikkawa et al. (2004) By far the largest dataset on the CO\(_2\) toxicity of developmental fish comes from Kikkawa et al. (2004, [70]). In total four species were studied: red sea bream (Pagrus major), Japanese silla (also known as Japanese whiting, Sillago japonica), Japanese flounder (also known as bastard halibut, Paralichthys olivaceus), and eastern tuna (Euthynnus affinis). For red sea bream and Japanese silla, developmental stages tested included egg (cleavage and embryo), larva (preflexion, flexion, and postflexion), and juvenile. For Japanese flounder, egg (cleavage) and young were tested; for eastern tuna, only the egg (cleavage) stage was tested. Exposure tests for three of the species generally lasted 24 hours; Japanese flounder tests went as long as 72 hours. In most cases mortality data were collected after 15 minutes, 90 minutes, 6 hours, and 24 hours. Exposure concentrations were for the most part in the 1-10 kPa range, except for the more CO\(_2\) tolerant eastern tuna where they reached nearly 15 kPa. The authors reported LC\(_{50}\) concentrations
Figure 5-4: Reported CO₂ LC₅₀ data from Kikkawa et al. (2004, [70]) for developmental stages of red sea bream (*Pagrus major*), Japanese sillago (*Sillago japonica*), Japanese flounder (*Paralichthys olivaceus*), and eastern tuna (*Euthynnus affinis*). All developmental stages that were tested are combined in the figure.

in each case, as well as the mortality curves for the red sea bream and Japanese sillago experiments.

Taken as a whole, the combined dataset from this study (see Figure 5-4) does not reveal a clear trend. The data are more revealing when plotted by species and by developmental stage, as shown in Figure 5-5. Note that in the case of red sea bream and Japanese sillago, the eggs at cleavage stage and the juvenile fish are most sensitive to hypercapnia. The Japanese flounder egg data are similar to these, and the young data suggest a slightly higher tolerance, but this could also be due to slightly older fish (young flounder vs juvenile bream/sillago). The eastern tuna egg data show a higher CO₂ tolerance, as noted by [70].
Figure 5-5: Kikkawa et al. (2004, [70]) CO₂ LC₅₀ data by development stage for red sea bream (*Pagrus major*, top left); Japanese sillage (*Sillago japonica*, top right); Japanese flounder (*Paralichthys olivaceus*, bottom left); and eastern tuna (*Euthymnus affinis*, bottom right - note the different vertical scale). Symbols refer to different stages: filled triangles = egg (cleavage); filled squares = egg (embryo); filled diamonds = preflexion larva; filled circle = flexion larva; open squares = postflexion larva; x = juvenile (juv) or young (yng). All values were reported by the authors.
5.2.2.2.3 Kikkawa et al. (2006a) Building on previous results, Kikkawa et al. (2006a, [71]) investigated how the CO₂ toxicity of juvenile Japanese sillago (Sillago japonica) differs when the exposure concentration is time variable. As a baseline, Kikkawa et al. conducted a series of 18-hour experiments in which fish were immediately placed in seawater equilibrated with air mixtures containing 1, 3, 5, 7, and 9% fCO₂ (approximately 1, 3, 5, 7, and 9 kPa). This *one-step exposure* approach is the same as in all CO₂ mortality studies previously described. Kikkawa et al. then conducted a set of four *step-wise exposures* in which fish were exposed to CO₂ levels that were gradually increased in a series of discrete steps from 1% and 7-10% fCO₂, and then suddenly reintroduced into natural (normocapnic) seawater. The resulting CO₂ toxicity was markedly different, and two general observations were made. First, in all cases the mortality at peak CO₂ concentration was reduced when the exposure was gradual rather than sudden. For example, while mortality was near 100% after 15 minutes of the one-step 7% and 9% fCO₂ exposures, mortality in the first step-wise experiment was not observed until the maximum 9% fCO₂ was reached (6 hours into the experiment) and at 18 hours the mortality was only 67%. Second, the fish were very sensitive to a sudden drop of CO₂ concentration; of the fish that survived the peak CO₂ concentration, 100% died within 15 minutes of exposure to normocapnic seawater in three of the four experiments and 76.7% died in the last experiment.

The step-wise exposure data from [71] are not plotted as they are not compatible with one-step exposure data. However, the step-wise exposure findings cast doubt over the validity of the Auerbach et al. [5] isomortality approach as it pertains to fish and are considered later in this analysis.

5.2.2.2.4 Kikkawa et al. (2006b) The general findings of [71] were confirmed in Kikkawa et al. (2006b, [68]), where tomato clownfish embryos (Amphipron frenatus) were subjected to one-step and two-step exposures. One-step exposures of 14.3, 9.6, 6.8, 4.8, and 2.9 kPa were administered for experiment durations of 6, 24, 48, 72, and 96 hours. For these experiments, an LC₅₀ of 14.3 kPa, 10.3 kPa, and 7.0 kPa was reported for 48, 72, and 96 hour exposures. During the two-step exposure experiments, where embryos were first placed for 48 hours in one concentration and then moved suddenly to a second concentration for
another 48 hours, a substantial difference in response was observed. When going from 2.9 kPa to normocapnic water, the mortality after 96 hours was about 40%, compared to 0% for a constant one-step exposure. Likewise, mortality is about 75% after 96 hours when going from 4.8 kPa to normocapnia, compared to 5–35% for a one-step exposure of 4.8 kPa for 96 hours. Thus, embryos of tomato clownfish exhibit a sensitivity to sudden changes in pCO$_2$ akin to that observed for juvenile Japanese sillago in Kikkawa et al. (2006(a), [71]), albeit to a lesser degree. The adaptive ability observed in [71], whereby fish survived longer when gradually exposed to higher concentrations in multiple steps, was not observed here.

5.2.2.2.5 Combined developmental fish dataset The combined CO$_2$ mortality dataset for developmental fish is plotted in Figure 5-6, where the red sea bream data of [69], the Japanese sillago data of [71], and the tomato clownfish data of [68] have been added to the data from [70]. The large variability across developmental stages is apparent in the red sea bream and Japanese sillago data, and it is interesting to note that in these case the cleavage stage eggs and the juvenile fish show the highest sensitivity.
Figure 5-6: Combined CO$_2$ LC$_{50}$ dataset for developmental fish for red sea bream (*Pagrus major*, top left); Japanese sillago (*Sillago japonica*, top right); Japanese flounder (*Paralichthys olieaceus*, bottom left); and eastern tuna cleavage-stage eggs (*Euthymnus affinis*, bottom right) with tomato clownfish embryos (*Amphipron frenatus*, bottom right). Note the different vertical scale used for the bottom right panel. Symbols refer to different stages: filled triangles = egg (cleavage); filled squares = egg (embryo); filled diamonds = preflexion larva; filled circle = flexion larva; open squares = postflexion larva; x = juvenile (juv) or young (yng). Includes reported and estimated data from [68], [69], [70], and [71].
5.2.2.3 Combined fish CO$_2$ toxicity dataset

Data for developmental as well as adult fish are available for two species, red sea bream and Japanese flounder. The combined LC$_{50}$ dataset for these species is shown in Figure 5-7. The adult dataset only adds a small number of datapoints to each plot, and in each case exhibits more tolerance than most of the developmental dataset. With regard to constructing a revised isomortality curve of the type in [5] (see Section 5.3 for details), the following conclusions are drawn from the combined developmental and adult fish dataset:

- Fish are much more sensitive to hypercapnia than to an equivalent level of pH depression caused by another acid, as confirmed by [69] for developmental fish and by [52] for adult fish. Thus, acute impact estimates of CO$_2$ injection should be based on CO$_2$ mortality data rather than pH mortality data as in [5].

- Fish seem to be most sensitive to hypercapnia during the early and late developmental stages, as demonstrated for red sea bream and Japanese sillago by [70]. Thus, a conservative isomortality curve should be based on these life stages rather than the combined dataset.

- CO$_2$ tolerance is variable between different fish species. For example, adult starspotted dogfish were notably more tolerant than adult Japanese flounder and yellowtail [51] and eastern tuna eggs were more tolerant than red sea bream, Japanese sillago or Japanese flounder eggs [70]. Thus, a conservative isomortality curve should be based on the most sensitive species rather than the combined dataset.

- Hypercapnic mortality is strongly influenced by the time variability of the ambient CO$_2$ concentrations, as demonstrated by [71] and [68]. Mortality due to peak CO$_2$ concentrations may be reduced when the concentration is increased gradually, and a sudden return to normocapnic conditions can induce mortality. This casts doubt on the relevance of the integrated mortality concept, upon which the isomortality method of [5] is based, to fish.
Figure 5-7: Combined CO₂ LC₅₀ dataset for adult and developmental fish for red sea bream (*Pagrus major*, top) and Japanese flounder (*Paralichthys olivaceus*, bottom). Symbols refer to different stages: filled triangles = egg (cleavage); filled squares = egg (embryo); filled diamonds = preflexion larva; filled circle = flexion larva; open squares = postflexion larva; x = juvenile (juv) or young (yng); + = adult. Developmental stages include reported and estimated data from [69], [70], and [71]; adult data are estimated from [51], [52], and [124].
5.2.3 CO₂ toxicity studies of pelagic copepods

Three main studies have been identified which provide insight into the acute toxicity of pelagic copepods (a type of zooplankton): Kurihara et al. (2004(a), [74]), Watanabe [139] and Watanabe et al. [140]. The studies provide data on both developing and adult copepods, but are here treated as a group because the distinction is not as clear as in the fish studies.

5.2.3.1 Kurihara et al. (2004a)

Kurihara et al. [74] looked at the impacts of elevated CO₂ concentrations on the survival, egg production rates, and early development of Acartia steueri and Acartia erythraea, both shallow-water marine copepods. In the first part of the study, the survival rate of adult females and their egg production rate were measured during an 8-day exposure to seawater with pCO₂ values of 0.236, 0.536, and 1.036 kPa (control pCO₂ was 0.036 kPa). Although observed mortality increased with time in each case, this trend was also observed in the control samples, where mortality exceeded 60 and 40% for A. steueri and A. erythraea, respectively. For A. steueri, the differences in mortality between all elevated CO₂ cases and the control were not statistically significant. For A. erythraea, the 1.036 kPa exposure had slightly higher mortality rates but the difference from the control was only statistically significant in the early portion of the experiment. In contrast, the egg production rates were clearly impacted by elevated CO₂ exposure, decreasing with both time and CO₂ concentration. At the highest exposure tested (1.036 kPa), A. steueri and A. erythraea egg production rates were 40% and 5% of their control values, respectively, indicating a significant but different response by the two species.

The second part of the study looked at the hatch rate and survival of nauplii (larvae) of A. erythraea only. Eggs produced in each of the CO₂ conditions above were placed in vials of the same concentration, and the number of hatched and dead nauplii were counted after 24 hours. The hatch rate declined and the nauplii mortality increased when exposed to high CO₂, with differences from the control on the order of 30 and 20%, respectively (hatch rate was significantly different from the control at 1.036 kPa; nauplii survival at 0.536 and 1.036 kPa).
In terms of mortality data of the type collected for adult and developmental fish, this study does not yield much data. An egg mortality rate could be inferred by combining the hatch rate and nauplii mortality observations (e.g., a 30% decline in hatch rate plus a 20% decline in nauplii survival, could be interpreted as a 44% mortality even though adult female mortality was statistically similar to the control). Such adjustments are not made here since it would be inconsistent with the previous interpretations of adult and developing fish mortality. This issue is, however, revisited in Sections 5.3.3.2.2 and 5.5 since the observed reproduction impacts imply intergenerational consequences of CO$_2$ exposure.

5.2.3.2 Watanabe (2001)

The recuperation of several different species of copepods was studied after high pCO$_2$ exposure (200,000 $\mu$atm for 3-30 minutes) by Watanabe [139]. Each experiment used about 5 copepods of the same species, and their level of activity was observed and assigned a ranking based on a qualitative assessment. The scale went from 1 (not moving or reacting, appearing dead) to 7 (sensitive to touch stimulus). This activity or behavioral index was logged at regular intervals during and after the exposure. Because the metric is qualitative and the data are published only in Japanese with a short overview explanation provided by the author, certain assumptions regarding the interpretation of the data were necessary here. The most obvious result is that in most cases, the copepods generally returned to normal activity (7) within a relatively short time after being returned to normocapnic seawater (10-40 minutes) after reaching the lowest activity rank of 1 in the high pCO$_2$ environment. In some cases the longest exposed group activity only returned to about 4 after 40 minutes (when the experiment apparently ended). Taking the plotted behavioral index as a mean response, this suggests that mortality was minimal for short exposures to this pCO$_2$, and on the order of 0-50% for exposures up to 30 minutes over a fairly large range of copepod species. The study also looked at several reproductive factors but these are not interpreted here due to the language barrier.

The data provided in this study are highly relevant to understanding the fate of a copepod encountering the plume, but do not lend themselves to inclusion in an isomortality analysis. A constant toxicity test would have identified an LT$_{50}$ as being very short if the non-moving
animals were assumed to be dead, but this would be incorrect. As described in Section 5.3.1.3, Chen et al. [23] used this activity data to construct a model of copepod response to elevated pCO$_2$ which allowed recovery of organisms but not mortality. The implications of the observed recovery for the validity of an isomortality approach is open to debate and should be the focus of more study. Still, the data are not used quantitatively in the present study for the reasons noted above.

5.2.3.3 Watanabe et al. (2006)

The largest source of copepod (zooplankton) mortality data comes from Watanabe et al. [140]. The purpose of the study was to remedy the lack of existing CO$_2$ toxicity data on deep-sea zooplankton species, i.e., for species that dwell in the depth ranges considered by some CO$_2$ injection schemes. Samples were collected from five stations in the Western North Pacific: three stations were in the subtropical region ($\{25^\circ 26'\text{N}, 144^\circ 50'\text{E}\}; \{24^\circ \text{N}, 155^\circ \text{E}\}; \{14^\circ \text{N}, 155^\circ \text{E}\}$), one was in the subarctic region ($43^\circ \text{N}, 155^\circ \text{E}$), and the remaining station was in the transitional region ($36^\circ \text{N}, 155^\circ \text{E}$). At each station, deep and shallow samples were collected using a vertical tow method (i.e., vertically integrating the sample over a depth range). Depth ranges for the shallow and deep samples were 0-150 and 0-1000 m for the subarctic station; 0-500 and 500-1500 m for the other four stations. CO$_2$ exposure experiments were conducted on the ship soon after collection. Sample specimens were placed in seawater of varying CO$_2$ concentrations and incubated in the dark at temperatures similar to their native environments (at atmospheric pressure). Animal behavior and mortality was assessed after 6 and 12 hours, and at 12-hour intervals thereafter. A total of 16 exposure experiments were conducted (2-5 experiments per station, with at least one shallow sample and one deep sample), where each experiment consisted of a control ($530 \leq \text{pCO}_2 \leq 1,600 \mu\text{atm}, 8.02 \geq \text{pH} \geq 7.64$) and multiple high CO$_2$ exposures ($1,100 \leq \text{pCO}_2 \leq 98,000 \mu\text{atm}; 7.65 \geq \text{pH} \geq 6.02$). Zooplankton species were classified taxonomically and assigned to one of three groups: epipelagic (inhabiting shallow water), meso/bathypelagic (inhabiting deep water), and eurybathic (inhabiting a large depth range). The authors reported LT$_{50}$ values for each exposure level for all experiments and provided complete mortality curves for experiments 1, 3, 15, and 16. The taxonomical breakdowns of each experiment are also
Watanabe et al. [140] make the following general conclusions from the dataset:

- Copepods do exhibit increased mortality when exposed to high CO$_2$ concentrations, and the mortality increases with increasing exposure time. The sensitivity to exposure time seems much greater for zooplankton than for fish, which are thought to adjust to external hypercapnic stress through the action of gill cells.

- Mortality to CO$_2$ exposure is significantly higher than mortality due to acification by another agent, based on a comparison of the pH mortality data of [5] and [145] with the data collected in this study.

- Deep-living copepods appear to have better tolerance of high CO$_2$ concentrations than shallow-water copepods, which is contrary to the generally held belief that deep-sea organisms would be more sensitive. While observed differences can perhaps in part be attributed to the fact that some deep-living copepods go through a dormant stage or to the fact that toxicity of most chemicals is generally lower at lower temperatures such as the ones found in the deep ocean, the authors argue that these factors alone do not explain the observed difference in tolerance between deep- and shallow-water copepods. It is suggested that deep-sea organisms are adapted to better tolerate CO$_2$ because they are naturally exposed to higher CO$_2$ concentrations than surface organisms (pCO$_2$ peaks at about 1,000 m depth due to remineralization of sinking organic matter from primary production in the upper ocean). The higher tolerance of deep-sea organisms is most pronounced in the subarctic and transitional regions; in the subtropical region, there is a smaller but still significant difference between the CO$_2$ tolerance of deep- and shallow-water copepods (see Figure 5 in [140]).

The combined Watanabe dataset is plotted in Figure 5-8, where the data for the various regions can be distinguished. The lead author of [140] has kindly provided the remaining raw mortality data not included in the original paper. The treatment of this additional data and their application to the development of a new isomortality function is described in Section 5.3.3.2.2.
Figure 5-8: Reported LT$_{50}$ values for copepods in the Western North Pacific from Watanabe et al. [140]. Filled squares: shallow-living copepods from the subarctic and transitional regions; hatched squares: shallow-living copepods from the subtropical region; open squares: deep-living copepods.

5.2.4 CO$_2$ toxicity studies of benthic organisms

The isomortality curves developed by Auerbach et al. [5] were partially based on studies of benthic organisms. The response of such organisms to CO$_2$ exposure is relevant not only to sequestration schemes in which CO$_2$ is introduced on the sea floor, but also to mid-depth releases since some benthic organisms are planktonic in their developmental stages. Only one of the benthic studies, Sato et al. (2005, [112]), yields data directly relevant to isomortality modeling. For the other studies reviewed below, there is little in the way of controlled toxicity data that would be useful in the evaluation of discharge scenarios in the next chapter. Nonetheless, the studies are reviewed below because they have general relevance to acute impact modeling across a range of species.

5.2.4.0.1 Sato et al. (2005) The Sato et al. [112] study is highly relevant to the present discussion for two reasons. First, the authors report CO$_2$ mortality data on the harpacti-
coid (benthic) copepod *Metamphiascopsis hirsutus* for steady and time-variable exposures. Second, a more general form of the isomortality method of [5], the *extended probit mortality model* of Sato and Sato [111] and Sato et al. (2004, [109]) is tested using the experimental mortality data. The discussion below is restricted to the mortality data yielded by the study and its use in validating their model; theoretical considerations of how the probit model differs from the Auerbach *et al.* [5] isomortality calculations are discussed in Section 5.3.3.

The experimental portion of [112] consisted of 5 experiments of elevated pCO$_2$ and one control. In each case, 19-20 copepods were placed in individual vials containing seawater equilibrated with varying pressures of CO$_2$. Three of the experiments were one-step exposures to 2, 4, and 9% fCO$_2$ that lasted 96 hours. The other two experiments were two-step exposures; a fCO$_2$ of 4% was used for the first 48 hours of the experiments, followed by another 48 hours at 2 or 8% fCO$_2$. The raw mortality data are provided by the authors, from which mortality statistics have been estimated for the three one-step exposures.

The purpose of the two two-step exposure experiments was to test whether probit model mortality calculations based on one-step exposure data could accurately predict two-step exposure mortality. Sato *et al.* [112] found that the model performed “moderately well” in this respect and conclude that the time integration of CO$_2$ mortality data is valid for zooplankton (*Metamphiascopsis hirsutus* was chosen as a model organism for pelagic copepods based on unpublished observations, and pelagic copepods were taken to represent zooplankton in general). Although the two-step exposure data are not directly useful in constructing new isomortality curves, they provide an interesting comparison to the time variable exposure of juvenile Japanese sillon in [71]. While the fish showed greater tolerance to high CO$_2$ levels during step-wise exposures than one-step exposures, the mortality of *M. hirsutus* during a 4 to 8% two-step exposure seemed more additive, i.e., more consistent with the assumptions behind the time integration of the probit calculations in this study and of the calculations in [5]. This conclusion is qualitative at best since the corresponding one-step experiment accidentally used 9% fCO$_2$ instead of 8%, and there is only one datapoint. Nonetheless, the authors suggest that the ability of fish to compensate for gradual exposure to otherwise lethal pCO$_2$ (by accumulating bicarbonate ions in the blood) is not as well-developed in zooplankton, and thus the isomortality method (probit method in their parlance) may be
valid for zooplankton but not fish.

5.2.4.1 Barry *et al.* (2004), Carman *et al.* (2004), and Thistle *et al.* (2005)

The *in situ* response of sediment-dwelling meiofaunal communities to small scale CO$_2$ releases has been studied by Barry *et al.* (2004, [9]), Carman *et al.* (2004, [17]), and Thistle *et al.* (2005, [128]). In each case, small corrals (48 cm in diameter) were filled with liquid CO$_2$ and meiofaunal abundance in sediment cores after roughly month-long exposures was compared to control samples (taken far away from the CO$_2$ sources). Barry *et al.* [9] observed a significant mortality of meiofaunal organisms in the CO$_2$ exposed cores, most notably for flagellates, amoebae, and nematodes (ciliates and allogromid foraminifera did not decrease in biovolume, but since these taxa have much lower abundances it was thought that the experimental design was inadequate to detect small changes in abundance/biovolume of these organisms). Carman *et al.* [17] did not detect significant declines in the abundances of the major groups (harpacticoid copepods, nematodes, nauplii, kinorhynchs, polychaetes) or total meiofauna, but this was attributed to slow decomposition of meiofaunal carcasses which rendered the experimental procedure incapable of detecting mortality due to CO$_2$ exposure. Using the same samples, Thistle *et al.* [128] showed significant differences in mortality between control and CO$_2$ exposed harpacticoid copepods, using a new technique in which copepods were classified as alive or dead at the time of collection based on the appearance of striated muscles. Based on these results, Thistle *et al.* [128] and Carman *et al.* [17] concluded that meiofaunal communities were likely strongly impacted by CO$_2$ exposure, even though traditional abundance indicators failed to detect the signal due to slowly decaying carcasses. Thus, [9], [17], and [128] are consistent in their overall conclusion, namely that hypercapnia can induce high mortality on meiofaunal communities.

5.2.4.2 Takeuchi *et al.* (1997)

Takeuchi *et al.* [125] looked at the impact of high CO$_2$ concentrations on three species of nematodes and eleven species of marine bacteria. These were selected as representative of marine organisms because nematodes are the most abundant taxa and have the highest species diversity in the benthic ecosystem, and because bacteria are a major decomposer and
an important genetic resource. For nematodes, 7-day exposure experiments were conducted under starvation conditions at varying concentrations of CO₂ and compared to a control. Mortality curves are provided by the authors, but as a function of pH and not CO₂. The CO₂ concentrations used to achieve the given pH levels are not reported, and insufficient supplemental information is given to estimate corresponding CO₂ concentrations after the fact. For the bacteria, 12-hour incubations under eutrophic conditions at various pH levels were performed and compared to control incubations, but again the CO₂ concentrations used were not reported. All eleven bacteria species were tested at atmospheric pressure, and two were also tested at 350 atm which is more typical of pressures in a deep ocean environment. The following results were obtained:

- No significant nematode mortality in excess of the control (pH 8.0) was observed during exposures of pH 7.0 and 6.2. At pH 5.4, all three species showed significantly higher mortality, and the effect was more pronounced in two of the three species (Mesacanthion sp. and Symplocostoma sp. were more sensitive than Metachromadora sp.).

- Bacterial growth was strongly affected by CO₂ exposure, dropping to less than 50% of the control (pH 7.7) for pH < 6.0 for all bacteria, and for one species this decrease was achieved at pH ≤ 6.5. The authors suggest that the in situ sensitivity of bacteria may be higher because of the oligotrophic conditions (leading to higher stress) that prevail in the ocean.

- Bacterial sensitivity to CO₂ exposure was similar at high pressure, although growth rates were generally lower than at atmospheric pressure. The results did not support the generally held belief that deep-sea species are more sensitive than shallow-water species because they experience less environmental variability.

- Overall, the nematodes and bacteria showed acute effects when CO₂ concentrations were high enough to effect a pH of 6.0 or lower; above pH 6.0, acute effects were generally not observed.

The nematode mortality and bacteria growth rate decline data of Takeuchi et al. [125] are not used in developing new isomortality curves due to CO₂ exposure for two reasons. First,
CO₂ concentrations were not reported and cannot be estimated without making assumptions regarding the seawater and incubation medium used in the experiments. Second, and more importantly, the CO₂ tolerance of the nematodes and bacteria in this study was apparently greater than that of pelagic copepods and fish, based on a comparison of observed mortalities and pH ranges. Thus, not using these data seems conservative from the standpoint of CO₂-induced mortality assessment.

Additional data on the microbial response to elevated CO₂ can be found in a series of mesocosm experiments reported in Sugimori et al. (2000, [121]), Takeuchi et al. (2002, [126]), and Sugimori et al. (2004, [120]); shifts in bacterial populations were noted for pCO₂ values in the range of approximately 1-10 kPa. For the reasons noted above, these data are not used further in the present analysis.

5.2.4.3 Ishida et al. (2005)

Ishida et al. (2005, [58]) used a benthic chamber to conduct one short (about 3 days) and two long (about two weeks) in situ CO₂ exposure experiments on benthic communities. In each experiment, the abundances of benthic meiofauna, nanofauna, and bacteria were observed for average CO₂ exposures of 5,000 ppm and 20,000 ppm and compared to the control. The results were mixed: meiofauna showed for the most part significantly decreased abundance at the highest CO₂ exposure when compared to the control; nanofauna showed decreased abundance only in one long-term experiment; and bacteria showed a significant increase in abundance in the two long-term experiments. Of the meiofauna, nematodes showed decreased abundance for the two long-term experiments with highest CO₂, and foraminifers showed decreased abundance in one long and the short experiment. The observed sensitivity of the meiofauna to high CO₂ exposure seems consistent with the conclusions of [9], [17], and [128], and the sensitivity of nematodes in particular seems consistent with the findings of [125]. The increase in bacterial growth in the presence of high CO₂ is not consistent with the findings of [125], and is thought to be caused the growth of bacteria adapted to the new environmental conditions, i.e., high CO₂ and reduced feeding pressure from nano and meiofauna. Taking the benthic community as a whole, Ishida et al. point out that the response to CO₂ perturbations was neither simple nor linear; the variable responses to CO₂ among the
different trophic groups led to non-linear effects. Although the data do not lend themselves to isomortality modeling, they do suggest that biological impact calculations using mortality data collected on individual species may be inadequate as these will not capture impacts on the relative abundance of different trophic levels such as those noted herein.

5.2.4.4 Ishida et al. (2006)

Ishida et al. (2006, [57]) augment the above study with additional data from Storfjorden, a Norwegian fjord, using the same experimental setup. Somewhat different trends were observed. Meiobenthos abundance varied greatly, nanobenthos abundance increased, and bacterial abundance was unchanged. As in [58], the response among the various trophic groups was neither simple nor linear.

5.2.4.5 Langenbuch and Pörtner (2004)

Langenbuch and Pörtner [77] studied the CO₂ sensitivity of the eurybathic sediment-dwelling marine worm Sipunculus nudus. They found that although the species can alter its metabolic rates to survive short-term exposure to hypercapnic conditions, long-term mortality is sensitive to sustained high CO₂ levels even when these levels are within the natural range experienced by the organism. Mortality curves are provided, but have not been incorporated into the isomortality analysis because the discharge scenarios studied herein are in the water column and because Pörtner et al. [101] identify Sipunculus nudus as showing exceptional tolerance to acutely elevated pCO₂, at least in the short term.

5.2.5 Combined mortality dataset for all species

The combined mortality dataset for all data (CO₂ and non-CO₂) is plotted in Figure 5-9 as a function of pH. The distinction between reported and estimated mortality statistics is not shown in this figure, although the reader is reminded that most of the adult fish data were estimated in the present study. The data in Figure 5-9 clearly indicate that marine organisms tend to be more sensitive to hypercapnia than the equivalent acidosis caused by another acidifying agent, as has been previously noted for a variety of species.
The dataset for CO$_2$-induced mortality is shown in Figure 5-10. Looking ahead to Section 5.3.3.2.2, it is noted that the adult fish data indicate a higher tolerance to CO$_2$ exposure than zooplankton (copepods). Also, if the adult fish data are excluded, then the remaining dataset is dominated by developmental stage fish for short exposures and by copepods for longer exposures; this has implications for the development of isomortality relationships in Section 5.3.3.2.2.

It should be noted that acute mortality data have apparently been collected for a number of other species as well, including cephalopods (*Sepia lycidas* [73], *Sepioteuthis lessoniana* [73], and the common octopus *Octopus vulgaris* [60]) and decapods (the prawn *Penaeus japonica* [73] and the western rock lobster *Panulirus cygnus* [60]). These species are not included in the present analysis as the data could not be found in the peer-reviewed literature, and the above sources only mention mortality statistics briefly. Moreover, with the exception of the common octopus, these species appear to be significantly more tolerant of CO$_2$ than copepods.

### 5.2.6 Sub-lethal impacts

The focus of the preceding literature review has been to identify sources of toxicity data that are well-suited for the specific objective of the study, namely to update the isomortality analysis provided by Auerbach [4] and Caulfield [18]. In addition, a sampling of closely related studies was reviewed in an attempt to provide a more complete picture of ongoing CO$_2$-related research and a better qualitative picture of the potential biological impacts of ocean sequestration. To that end, it would be remiss of us to not at least mention the substantial body of literature on sub-lethal effects of increased ocean CO$_2$ as well as the existence of a number of comprehensive review articles which attempt to integrate lethal and sub-lethal effects with physiological and ecosystem perspectives. In particular, the reader is referred to IPCC [56], Pörtner *et al.* [101, 102], Ishimatsu *et al.* [59, 61], Siebel and Walsh [114], and Kurihara *et al.* [75].

The potential importance of sub-lethal effects on both the organism and ecosystem level must be taken into consideration, and findings of some recent work in this area are discussed briefly below as an illustration. Kurihara and Shirayama [76] studied the effect of CO$_2$ on the
Figure 5-9: Combined LC$_{50}$ dataset for mortality due to pH depression by CO$_2$ (filled symbols) and other acids (open symbols). The non-CO$_2$ dataset is mainly comprised of zooplankton [5, 145]; the CO$_2$ dataset consists of adult fish (AF, filled diamonds [48, 51, 52, 79, 124]), developmental fish (DF, filled triangles [68, 69, 70, 71]), and copepods (C, filled squares [112, 140]).

early development of the sea urchins *Hemicentrotus pulcherrimus* and *Echinometra mathaei*, looking specifically at fertilization rate, cleavage rate, developmental speed, and pluteus morphology. Also, the effect of CO$_2$ and HCl were compared to differentiate between the effect of increased CO$_2$ and pH depression. CO$_2$ partial pressures of 10,000 μatm were used in the study, resulting in about a 1.2 unit drop in pH. All of the studied factors decreased with increased CO$_2$ concentration, but only fertilization rates showed a greater impact when CO$_2$ rather than HCl is used as the acifidying agent. From this the authors conclude that both reduced pH and increased CO$_2$ can significantly affect the development of these organisms and consequently alter the marine ecosystem. The study suggests, much like Kurihara *et al.* [74] did for pelagic copepods, that there can be deleterious, sub-lethal effects with inter-
Figure 5-10: Combined LC$_{50}$ dataset for mortality due to CO$_2$ exposure. The dataset consists of adult fish (AF, open triangles [48, 51, 52, 79, 124]), developmental fish (DF, open circles [68, 69, 70, 71]), and copepods (C, filled squares [112, 140]).

generational and ecosystem consequences. Such a conclusion is also supported by a longer term investigation of Shirayama and Thornton [113] on the effects of mildly elevated pCO$_2$ (+200 μatm) on the growth rates of gastropods and sea urchins (both calcifying organisms). They demonstrate that adverse effects can be identified even at this low level. Growth rate impacts have also been observed for marine mussels for somewhat higher exposure levels [86]. Although such considerations are beyond the capacity of an isomortality-type analysis, results of the present study are interpreted in this context in the concluding chapter.
5.3 Modeling Approach and Scenarios

The following section describes the approach taken to simulating the acute impact of three discharge scenarios, which are based on the scenarios put forth in Adams and Wannamaker [1], and which seek to maximize near-field dilution of the injected CO$_2$ in different ways. Previous investigations are first reviewed, in particular Auerbach [4], Caulfield [18], and Sato et al. [109], and then these methodologies are adapted to the present study. The specific discharge scenarios considered are also described in this section.

5.3.1 Previous studies

Since the goal of the present study is to update the original work of Auerbach [4, 5] and Caulfield [18, 19] with new biological data and enhanced discharge approaches, these studies are first reviewed in detail. Two additional acute impact modeling studies are also reviewed briefly, Sato et al. [109] and Chen et al. [21]. As discussed in Section 5.3.3, much of the biological impact modeling approach of Sato et al. [109] is ultimately adopted in favor of the Auerbach [4] approach, although many elements of the original framework of [4] and [18] are retained in the present study.

5.3.1.1 Auerbach et al. (1997) and Caulfield et al. (1997)

Together these studies simulated the acute biological impact to clusters of organisms due to a variety of discharges. Their approach consisted of three parts:

- Calculation of the CO$_2$ concentration and pH field due to the discharge.

- Simulation of the exposure history of each organism cluster, i.e., its trajectory through the plume.

- Calculation of cumulative impact to each organism cluster (% mortality).

The overall impact of a discharge scenario was calculated by considering the fate of a large number of organism clusters entering the discharge area at different locations. Each of these steps is described below. Note that the terms organism and organism cluster are
used interchangeably even though the model technically simulates the latter (since a percent mortality does not make sense for a single organism).

5.3.1.1 Discharge scenarios (plume modeling) Although Caulfield et al. [19] reports on only two discharge scenarios, the original work by Caulfield [18] considered four scenarios. A brief outline of each scenario is given below; details of the plume calculations can be found in [18] but will not elaborated on here as the present study considers different scenarios. For each scenario, CO\(_2\) loadings of 130 and 1,300 kg/s were considered, which was taken as the CO\(_2\) produced by 1 and 10 500-MW coal-fired power plants including an energy penalty for capture and storage. Summary statistics of the plumes modeled by [18] are shown in Table 5.1.

**Dry ice release:** Dry ice cubes would be released from a fixed location at the surface of the ocean. For the 3-meter cubes considered, this corresponds to one cube every 5.4 minutes for a single power plant release. Plume calculations assumed that the descending cubes form a two dimensional line source (extending to the sea-floor) spread by a combination of ambient and cube-induced (wake) turbulence.

**Towed pipe release:** Liquid CO\(_2\) would be released from a pipe which is towed by a ship traveling at 5 m/s using a 1-m diameter pipe with diffusers which distribute the CO\(_2\) over a depth range of 1000-1500 m. As in the dry ice scenario, plume calculations assumed a two dimensional line source spread by the combination of ambient and pipe-induced (wake) turbulence.

**Unconfined droplet plume release:** Liquid CO\(_2\) would be released from a fixed multiport diffuser at 1000-1500 m, forming a buoyant plume which distributes the CO\(_2\) vertically. CO\(_2\) enters the surrounding water column through a series of discrete peeling events caused by the entrainment of seawater by the plume as the droplets rise. Here the entrainment dilutes the outer portion of the plume, causing it to detach from the more buoyant inner core of the plume and sink until it reaches neutral buoyancy, forming an intrusion layer. Although it is unclear what spacing between diffuser ports was used in
<table>
<thead>
<tr>
<th></th>
<th>Volume with pH &lt; 7 (km$^3$)</th>
<th>Distance to pH 7 (km)</th>
<th>Min pH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dry Ice, 1 plant</td>
<td>0.001</td>
<td>0.09</td>
<td>6.1</td>
</tr>
<tr>
<td>Dry Ice, 10 plants</td>
<td>1.1</td>
<td>2.2</td>
<td>5.8</td>
</tr>
<tr>
<td>Towed pipe, 1 plant</td>
<td>0.00004</td>
<td>0.2</td>
<td>6.5</td>
</tr>
<tr>
<td>Towed pipe, 10 plants</td>
<td>0.3</td>
<td>14</td>
<td>5.7</td>
</tr>
<tr>
<td>Droplet plume, 1 plant</td>
<td>1.8</td>
<td>23</td>
<td>5.5</td>
</tr>
<tr>
<td>Droplet plume, 10 plants</td>
<td>130</td>
<td>60</td>
<td>5.5</td>
</tr>
<tr>
<td>Dense plume, 1 plant</td>
<td>7.2</td>
<td>94</td>
<td>4.0</td>
</tr>
<tr>
<td>Dense plume, 10 plants</td>
<td>510</td>
<td>690</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Table 5.1: Plume characteristics modeled by Caulfield [18].

In the calculation, a flow of 13 kg/s was specified for each port and plumes from neighboring ports were assumed to interact only in the diffusive regime, i.e., after intrusion. All peeling events were combined into a single intrusion layer, forming a relatively thin but wide two-dimensional plane source. The maximum thickness of the intrusion layer was calculated as 23 m in each case, with widths of 880 m and 8800 m for the 1 and 10 plant releases (reflecting the fact that a longer diffuser was used in the 10 plant case).

**Descending confined dense plume release:** Here seawater would first be enriched with CO$_2$ in a confined vessel, and the resulting dense mixture would be released along the sea floor in a confined trench. The plume would eventually reach a level of neutral buoyancy and form an intrusion layer, which was modeled as a two-dimensional plane in a manner similar to the droplet plume. A single plant release created an intrusion layer 23 m thick (at its maximum) and 520 m wide at 1,000 m depth when released from 855 m; a ten plant release from 755 m reached the same depth and thickness but with a width of 3,000 m.

The relative dilution achieved by each scenario modeled by Caulfield [18] is demonstrated by her plot of centerline dilution as a function of time (equivalent to distance assuming a current velocity of 0.05 m/s or a ship speed of 5 m/s), as shown in Figure 5-11. Caulfield attributes the advantage of the dry ice and towed pipe scenarios to their large plume thicknesses, but also points out the likely higher cost of these approaches. For the two fixed plume scenarios, she notes the advantage of the droplet plume in achieving greater dilution and allowing more
5.3.1.1.2 Organism exposure history  The exposure history of an organism cluster was obtained by simulating its trajectory through the plume. Caulfield [18] considered only planktonic organisms, i.e., floating organisms that move with ambient currents and exhibit little to no ability to influence their horizontal position. Although some planktonic organisms can migrate vertically hundreds of meters per day (diel migration), this effect was conservatively ignored as it was thought to lessen impact by allowing organisms to escape the plume. Thus, organisms were assumed to move only in the horizontal plane.

The fate of organisms drifting through the plume was modeled as a relative diffusion problem between the organism and the plume. Specifically, the separation distance $y$ between the organism cluster and the plume centerline was modeled using a modified version of Richardson’s distance neighbor equation [104]:

$$\frac{\partial p(y)}{\partial t} = \frac{\partial}{\partial y} \left( F(y) \frac{\partial p(y)}{\partial y} \right)$$  \hspace{1cm} (5.1)
where $p(y)$ is the probability that an organism is at $y$ at time $t$ given an initial separation of $y_0$, and $F(y)$ is a scale-dependent distance-neighbor diffusivity given by

$$F(y) = 0.0017y^{1.15} \quad (5.2)$$

where $F$ and $y$ have units of m$^2$/s and m, respectively. The diffusivity defined by (5.2) is based on the data from Okubo [93] on the radial spreading of dye releases in surface layers of coastal waters. Caulfield [18] describes simulating (5.1) with a Monte Carlo technique, although few details are given. Implementation of this modeling approach in the present study is given in Section 5.3.3.1.

### 5.3.1.1.3 Cumulative impact to the organism (isomortality modeling)

The response of the organism to the plume was modeled as consisting of two parts: (1) a % mortality based on observed relationships between mortality, exposure pH, and exposure time; and (2) a population recovery based on a growth rate which varies with exposure pH.

**Isomortality calculations** Auerbach [4] developed the *isomortality* methodology for calculating the cumulative impact to an organism cluster based on its exposure history. He began by developing isomortality curves, which were equations expressing LC$_0$, LC$_{50}$, and LC$_{90}$ as a function of exposure time, from a compilation of available pH mortality data (see Section 5.2). The dataset included zooplankton and benthos data, and in all cases the acidifying agent was not CO$_2$. The derived isomortality curves are shown in Figure 5-12.

Auerbach [4] gives few details of how the source data were merged and manipulated. However, it is noted that some data have been shifted to account for observed adverse reproductive effects, and that all datapoints were shifted “downward” by 0.25 pH units to be conservative.

The basic assumption behind this method is that there is an equivalence between low pH exposure for a short time and higher pH exposure for a longer time. Auerbach [4] discusses the justification of this approach. Although a number of variations of the isomortality calculation are described in [4], the basic method employed involves the following steps:

1. Divide the exposure history into $N$ discrete time intervals ($\Delta t_1$, $\Delta t_2$, ..., $\Delta t_N$), and
define an average exposure for each interval \((pH_1, pH_2, ..., pH_N)\).

2. For interval \(\Delta t_1\), calculate a fractional mortality (population deficit), \(D_1\), by interpolating between the isomortality curves at \(t = \Delta t_1\) to \(pH_1\) (extrapolate if beyond LC\(_{90}\)).

3. Estimate the equivalent cumulative time, \(t_1^*\), that corresponds to \(D_1\) had the organism been exposed to \(pH_2\) instead, i.e., move along the isomortality line that corresponds to \(D_1\). Thus, \(t_1^*\) is defined by the equality

\[
D(pH_2, t_1^*) = D_1. \tag{5.3}
\]

which must be solved iteratively since the curves are nonlinear and not parallel.

4. Calculate the mortality after the second step, \(D_2\), by interpolating/extrapolating between the isomortality lines at exposure time \(t_1^* + \Delta t_2\) to \(pH_2\).

5. Repeat steps (3) - (4) for all remaining time intervals.
Although not explicitly stated in [4] or [18], before reaching non-zero mortality the equivalent cumulative time is incremented by moving parallel to the LC\textsubscript{0} curve.

**Population recovery** Caulfield [18] allowed for species recovery in regions of low stress to account for species reproduction. This was modeled using a logistic equation appropriate for a single species population:

\[
\frac{dN}{dt} = r \frac{N}{K} \left(1 - \frac{N}{K}\right) \tag{5.4}
\]

where \(N\) is the number of organisms, \(K\) is the “carrying capacity” (maximum number of organisms that can exist), and \(r\) is the growth rate with unlimited resources. Thus, \(\frac{N}{K}\) is the fractional population, and \((1 - \frac{N}{K})\) is the population deficit. The growth rate \(r\) was based on laboratory observations of surface copepods and specified as:

\[
r = \begin{cases} 
0 & pH \leq 7.35 \\
0.09 \frac{(pH-7.35)}{0.15} & 7.35 < pH < 7.5 \\
0.09 & pH \geq 7.5 
\end{cases} \tag{5.5}
\]

It should be noted that because a species recovery was included, a finite impact zone could be defined for each plume, i.e., at some distance from the source the plume becomes sufficiently dilute and species populations are allowed to recover to their original levels.

**Acute impact estimates** The combined investigations of Caulfield [18] and Auerbach [4] yielded the acute impact estimates summarized in Table 5.2. Three measures of impact were used. First, *integrated total mortality* is the spatial integration of mortality over the impacted volume, which may be thought of as an equivalent volume of “dead water”. Similarly, the *mortality flux* is the flow rate of “dead water” crossing a plane perpendicular to the plume centerline, which is expected to vary along the centerline. Third, the *maximum spatial deficit* is given by the grid cell with highest average deficit.

The results are qualitatively similar to the pH impact volumes in Table 5.1, i.e., the dry ice and towed pipe scenarios cause the lowest impact, the dense plume causes the most impact by far, and the droplet plume is in between.
5.3.1.2 Sato et al. (2004)

The use of the isomortality method has been extended and somewhat generalized by Sato et al. [109], based on initial work in Sato and Sato [111]. This extended probit model is based on the standard assumption used in the LC$_{50}$ concept, namely that the acute mortality of a species follows a sigmoid function of the log of the concentration of the toxic agent (see Finney [38]). The sigmoid function is made linear by the probit transformation

$$Z = \int_{-\infty}^{Y-5} \frac{1}{\sigma \sqrt{2\pi}} \exp \left(-\frac{(X - \mu)^2}{2\sigma^2}\right) dX \quad (5.6)$$

where $X = \log(x)$ which has mean $\mu$ and standard deviation $\sigma$, $x$ is the concentration of the toxic agent, $Z$ is the probability (mortality), and $Y$ is the probit of $Z$. Applied to $\log(x)$ vs $\log(t)$ plots of mortality (e.g., the data in Figure 5-8), (5.6) can yield a linear relationship between probit transformed mortality $Y$, $\log(x)$, and $\log(t)$:

$$Y = a \log(t) + b \log(x) + c \quad (5.7)$$

where $a$, $b$, and $c$ are regression constants. The cumulative mortality can then be expressed as

$$Y = a \log \left( \int_0^t [x]^{b/a} \, d\tau \right) + c \quad (5.8)$$

Table 5.2: Acute impact estimates from Caulfield [18].
which can be estimated numerically by dividing organism exposure into a series of discrete steps in a manner similar to Auerbach [4, 5]. Thus, the probit mortality model is a generalization of Auerbach’s isomortality approach in that the assumption of a sigmoid function to describe acute toxicity yields a single probit function which describes all isomortality curves, as Auerbach called them. By contrast, Auerbach derived isomortality relationships for each of three different impact levels (LC₀, LC₅₀, and LC₉₀) by independently fitting curves to the data corresponding to these impact levels, and then interpolating linearly between them as needed. In the present study, the term isomortality curve or function is for convenience also used to describe the probit function (5.7) proposed by Sato et al. [109], since this function defines lines of equal mortality for any mortality level.

Sato et al. [109] fitted their probit function to copepod mortality data from the study later reported in Watanabe et al. [140], focusing specifically on *Metridia pacifica* (a species which exhibited relatively high CO₂ sensitivity). Two discharge scenarios were evaluated: (1) a fixed pipe discharging 1 kg/s CO₂ (spray droplets) from each of 100 nozzles spaced 10 m apart over 1,000 - 2,000 m depth, and (2) the same pipe towed by a ship moving at 4 knots (about 2 m/s). The resulting CO₂ plume was calculated using a CFD simulation forced by low-wavenumber ocean eddy velocity data (see [109] for details). Simulations predicted “nontrivial” mortality for the fixed pipe scenario and insignificant mortality for the towed pipe scenario. The latter finding was based on the fact that the target organism, a single copepod traveling with the plume, reached a ΔpCO₂ within 100 μatm of the ambient value before reaching 0.125% mortality, which was taken as the threshold for adverse effects.

The work in Sato et al. [109] was an improvement upon the earlier work by Sato and Sato [111] (where the probit approach was first introduced) and Sato [108]. Here the original isomortality curves by Auerbach were used to derive a probit equation, i.e., the impact analysis was based on non-CO₂ induced mortality. The form of the resulting equation was quadratic in [111] and linear in [108], and it was applied to different discharge scenarios. The main scenario considered in each study was a 200 kg/s descending droplet plume released at 2,000 m depth over 10 hours (from a ship which changed location every 10 hours). Despite reaching pH values as low as 4.2, zero mortality was predicted in Sato and Sato [111] for three test organisms encountering the discharge nozzle, i.e., mortality did not surpass 0.125%
(which was again taken as the threshold for impact). Similarly, the plume was predicted to reach pH as low as 4.7 in Sato [108], but with negligible biological impact. These findings were consistent with an earlier work, Sato and Hama [110], which considered the same injection scenario using Auerbach’s method and data directly.

5.3.1.3 Chen et al. (2004)

Rather than constructing isomortality curves from observed mortality data, Chen et al. [21] used “biological activity” data to develop a relationship for activity as a function of pH and time of the form

$$A_k = \exp\left(\frac{-t^2}{\sigma}\right)$$

where $A_k$ is an activity index which varies from 1 (normal activity) to 0 (no activity), $t$ is the exposure time, and $\sigma$ is a fitting parameter which varies with pH and organism type. Relationships were derived for pelagic zooplankton using the data from Watanabe [139] described in Section 5.2.3, and were implemented in an Eulerian-Eulerian scheme (see [21] for details). It is noted that the activity index method does not take mortality into account, meaning that organisms reaching $A_k = 0$ are allowed to recover. As such, the approach is an attempt to model sub-lethal effects due to pH perturbations, and the authors note that future improvements to the model will incorporate mortality effects.

A test case was studied in which 1 kg/s of CO$_2$ (droplets with lognormally distributed sizes and a mean diameter of 8 mm) was injected about 1,000 m depth into a mean current speed of 2.35 cm/s. The resulting pH field for the buoyant plume was computed with the LES model described in [20], yielding a minimum pH of 5.6 near the nozzle (ambient pH was 7.57). Zooplankton were predicted to be most injured (lowest activity) about 20 m downstream from the nozzle and recovered to normal activity levels about 100 m further downstream (within 2 hours). Thus, the impact of this small injection on zooplankton was confined to a small region.
5.3.1.4 Other investigations

The studies in the prior sections were reviewed in detail because they develop original methods for assessing acute biological impacts. In addition, there are a large number of studies which provide quantitative estimates of water quality impacts of various discharge schemes to the open ocean. For example, near-field modeling studies of stationary sources (bubble, droplet or hydrate plumes) include Liro et al. [81], Golomb et al. [44], Haugan et al. [50], Thorkildsen and Alendal [130], Sundfjord et al. [122], Sato and Hama [110], Alendal and Drange [2], Drange et al. [36], Nihous et al. [90], Chen et al. (2003 [20] and 2005 [23]), Adams and Wannamaker [1], and Wannamaker and Adams [138]. Modeling studies of near-field CO₂ dilution from a towed pipe include Ozaki et al. [95], Chen et al. (2002 [22], 2005 [23], and 2006 [24]), Minamiura et al. [87], Hirai et al. [55], Adams and Wannamaker [1], Jeong et al. [67], and Tsushima et al. ([134], which also applies the Sato et al. [109] framework and concludes negligible biological impact for a 100 kg/s towed pipe discharge). In addition, the far-field water quality impacts of injection schemes have been considered by numerous studies, including Dewey and Stegen [32], Stegen et al. [117], Nakashiki and Ohsumi [89], Xu et al. [144], Dewey et al. (1997 [34] and 2000 [33]), Wickett et al. [143], Caldeira and Wickett [15], Masuda et al. [83], and Magi et al. [82].

5.3.2 CO₂ discharge scenarios & plume modeling

Three CO₂ discharge scenarios are considered in the present study: a sinking CO₂ hydrate plume released from a fixed platform, CO₂ hydrate particles released from a moving ship, and rising CO₂ droplets released from a bottom manifold. The scenarios are based on those developed in Adams and Wannamaker [1], but have been adapted to incorporate enhanced understanding of discharge possibilities. For each discharge scenario, loadings of 10, 100, and 1000 kg/s are evaluated (for reference, a 500-MW coal-fired power plant generates about 95-126 kg/s CO₂ [54]).
5.3.2.1 CO$_2$ hydrates

Two of the scenarios discharge CO$_2$ as clathrate hydrate particles (CO$_2$·5.75H$_2$O), which warrants some discussion. The benefit of hydrate particles is that they are negatively buoyant, meaning that a sinking plume can be generated at shallower depths than those that would otherwise be required using liquid CO$_2$ (which becomes negatively buoyant in the 2000 - 3000 m range, depending on ocean conditions). Although pure hydrate particles have a density of about 1,143 kg/m$^3$ (i.e., about 10% more dense than seawater), such particles have not yet been achieved in laboratory or field studies [79]. As discussed in West et al. [142], present approaches form hydrates in a reactor vessel by creating a slurry of the reactants in a coflow injector. Under a steady flow of CO$_2$ this setup yields an extruded cylindrical hydrate particle which breaks off at the nozzle due to shearing by the ambient current once it reaches some critical length. To date, particle diameters of up to 0.7 cm have been achieved in a laboratory setting. In field studies, CO$_2$ hydrate particles up to 2.2 cm in diameter with a typical length of 30 cm have been observed at about 1,500 m depth with conversions in the range 10-55% [105, 132, 133]. The result of a partial hydrate conversion is a composite particle that is partially pure hydrate and partially unreacted CO$_2$ and water “stuck” to the hydrates. For the present purpose, composites are referred to as hydrate particles even if conversion is less than 100%.

Both the size and conversion percentage of a particle are important in dictating the fate of the CO$_2$ that is locked within the particle. Conversion efficiency dictates the density and thus the settling speed of the discharged particle, while the particle diameter dictates the time required for it to dissolve. Observations of descending hydrate particles indicate a particle dissolution rate of about 6 $\mu$m/s, and that this rate is not strongly dependent on the conversion percentage. Figure 5-13 shows the dependence of the settling depth on conversion percentage ($X$) and diameter for individual 30 cm long cylindrical hydrate particles discharged at 1,500 m depth into a typical stratified ocean of 4,500 m depth (i.e., a descent depth of 3,000 m indicates the particle reaches the sea-floor). Figure 5-13 is calculated using the approach of Riestenberg et al. [105], using a constant particle dissolution rate and a particle density that assumes that the unreacted CO$_2$ and water in the reactor vessel “stick”
Figure 5-13: Descent depth as a function of particle diameter and conversion efficiency for 30 cm long cylindrical CO₂ hydrate particles released at 1,500 m depth into a typical ocean stratification, from Chow [25].

to the converted hydrates. If the particle loading is sufficiently large relative to the ambient crossflow, multiple particles will together form a descending plume which will sink further then the descent depths in Figure 5-13 due to the combination of a group effect and a solute density effect.

Before proceeding to the specific discharge configurations considered herein, it should be noted that the main challenge ahead for improving hydrate formation techniques is to increase the conversion percentage. Ongoing research suggests that creating particles of larger diameter with conversion percentages similar to those created to date should not pose a major problem, although it may require different configurations for introducing the CO₂ into the reactor vessel (e.g., using multiple injection nozzles or a radial injection from around the perimeter). It is thus reasonable to assume that some flexibility will be available in the future to generate larger particles with higher and more controlled hydrate conversion percentages.
5.3.2.2 Stationary CO$_2$ hydrate plume

In this scenario an offshore stationary pipe extends from the ocean surface down to a depth of 1,500 m, discharging hydrate particles which are 30 cm long with a reaction efficiency of 25% (shown schematically in Figure 5-14, which was modified from [1]). The combination of the negatively buoyant particles and the increased density of the seawater surrounding the descending particles causes enhanced sinking. Particle diameter is calibrated to yield a plume which dissolves entirely over the remaining depth of the water column (3,000 m in this case) such that the CO$_2$ plume does not intercept the ocean floor. The discharge is somewhat idealized since with present day hydrate discharge technology the conversion efficiency is difficult to control, but nonetheless it is in theory possible to calibrate particle diameter for a given discharge configuration to generate a plume with the correct average descent depth even if a range of conversion efficiencies were created. For the present purpose this was achieved using a near-field integral plume model, specifically a version of the double plume model described in Adams and Wannamaker [1], which has been modified to handle cylindrical hydrate composite particles by Chow [25]. The same model was applied in Riestenberg et al. [105]. Mass loadings of 10, 100, and 1000 kg/s resulted in plume descent depths near the ocean floor using initial particle diameters of 8.3, 5.8, and 3.1 cm, respectively, for a 25% hydrate conversion. As the plume descends, entrainment of the surrounding water column leads to reduced density in the outer core of the plume which causes detrainment of the outer core in a series of discrete peeling events. Thus, the plume introduces CO$_2$ to the surrounding environment through a number of intrusion layers (one for each peeling event). As in [1], near field plume model results suggest that such a source can reasonably be approximated as a thin two-dimensional plane source extending from 1,500 m to the bottom of the ocean (at 4,500 m) with no particle deposition on the sea-floor. Following [1], the resulting plume is assumed to have an initial width $w$ given by:

$$w = \frac{\dot{m}}{hDIC_0u_a}$$  \hspace{1cm} (5.10)

where $\dot{m}$ is the carbon mass loading rate, $h$ is the plume height (3000 m), $u_a$ is the ambient (mean) current speed, and DIC$_0$ is the average plume concentration of excess dissolved
inorganic carbon (i.e., above background) within the intrusion layers as computed by the near-field model. Since the lateral shape of the source is likely closer to a Gaussian distribution than a rectangular plane source, the downstream lateral diffusion of this depth-averaged plume is modeled as:

\[
DIC(t, y) = \frac{\dot{m}}{\sqrt{2\pi hu_a \sigma_y}} \exp \left[ -\frac{y^2}{2\sigma_y^2} \right].
\]  

(5.11)

where \(\sigma_y = \sigma_y(t)\) is the standard deviation of the Gaussian concentration distribution and \(t\) is the time since discharge, which can be converted to longitudinal distance from the source \(x\) by substituting \(t = \frac{x}{u_a}\). The initial concentration distribution is parameterized by \(\sigma_{y,0} \equiv \sigma_y(t = 0)\), which is taken as

\[
\sigma_{y,0} = \frac{w}{\sqrt{12}}
\]  

(5.12)

since this is the standard deviation of the “top hat” distribution corresponding to a uniform rectangular source of width \(w\). In the interest of being conservative, vertical diffusion is ignored in (5.11). Lateral diffusion, i.e., the growth of \(\sigma_y(t)\), is modeled using data from Okubo [93] on dye patch spreading in the surface layers of coastal waters:

\[
\sigma_y(t) = a (t_0 + t)^n
\]  

(5.13)

where \(a = \frac{0.00071}{\sqrt{10}}\) for \(t\) in s and \(\sigma_y(t)\) in m; \(n = 1.15\); and \(t_0\) is a fictitious plume initial time, i.e., the time required for a point source \((\sigma_{y,0} = 0)\) to reach the actual \(\sigma_{y,0}\) of the plume being considered if it grows according to (5.13). In the interest of being conservative, the value of \(a\) used in the present study is reduced by a factor of \(\sqrt{10}\) to reflect the reduced mixing expected for deeper waters (i.e., diffusivity is thereby reduced by a factor of 10). The nominal value of \(u_a = 0.05\) m/s used by Adams and Wannamaker [1] is also applied here.

5.3.2.3 Moving CO₂ hydrate release (“towed pipe”)

A ship traveling at speed \(u_s\) in a direction perpendicular to the ambient current \(u_a\) releases hydrate particles at 1,500 m depth from a towed pipe (Figure 5-15). The particles are 30 cm long and of a fixed diameter \(d\) with hydrate conversion percentages distributed over the range \(X_{min}\) to \(X_{max}\). Because dissolution rate can be assumed constant for all \(X\), all particles
have the same diameter at any given distance behind the ship, but reach different depths before dissolving completely (higher conversion causes greater sinking). The ship is moving sufficiently fast such that the descending particles do not form a negatively buoyant plume but rather settle as a collection of individual particles (i.e., the particles create a passive plume of CO$_2$, but not a buoyant plume which actively entrains the surrounding seawater). Thus, the descent depth of each particle depends on its diameter and conversion efficiency as was discussed previously (see Figure 5-13). Four different discharge configurations are considered using particle diameters of 2.5, 5, 10, and 15 cm. For each configuration, the discharged particles have the same diameter but varying hydrate fractions. In each case $X_{min}$ was set to 12% (equivalent to a neutrally buoyant particle at this release depth). $X_{max}$ was set to 55% for the 2.5 and 5 cm diameter cases (i.e., the maximum conversion observed in the field) and to 50 and 30% for the 10 and 15 cm cases, respectively (in the latter cases the heaviest particles dissolve completely just before reaching the ocean floor at 3,000 m below
the injection point). In each case the distribution of $X$ values was selected such that the injected CO$_2$ is distributed uniformly over the height of the plume, i.e., the interval between the injection location (1,500 m) and the descent depth of the heaviest particle ($h_{\text{max}}(x)$), where $x$ is the longitudinal distance from the source). Ignoring vertical diffusion, the plume height reaches a steady value $h_{d,\text{max}}$ once the particles in this wedge-shaped “fractionated plume” (Figure 5-16) have dissolved completely. The $h_{d,\text{max}}$ values for particle diameters 2.5, 5, 10, and 15 cm and the $X$ ranges described above were 514 m, 1,409 m, 3,000 m, and 3,000 m, respectively. It is noted that even though the maximum particle diameter observed in the field to date is about 2.2 cm, the formation of larger particles is not expected to be a major hurdle given current technology.

The plume created by this discharge can be modeled using an approach similar to (5.11), but with some modification. First, $u_s$ is substituted for $u_{a_1}$, such that (5.11) describes the concentration distribution trailing the towed pipe in a moving reference frame, i.e., the
scenario is perfectly analogous to the situation where the source is fixed and the CO₂ mass is advected downstream by a current of speed $u_s$. In terms of a fixed reference frame, the longitudinal coordinate $x = u_s t$ is the distance along the plume centerline, which is at an angle $\omega = \tan^{-1} \left( \frac{u_a}{u_s} \right)$ to the negative $u_s$ direction (see Figure 5-17), and the lateral coordinate $y$ is perpendicular to $x$. For the present purpose, the moving reference frame is more relevant, since both CO₂ mass and organisms within the plume move at ambient current $u_a$, and since the relative diffusion as parameterized by Okubo’s relation (5.13) is the same in either reference frame. Second, the concentration of DIC is adjusted for the period of particle dissolution by introducing $f_d(t)$, the fraction dissolved CO₂, into (5.11):

$$DIC(t, y) = \frac{f_d(t) \dot{m}}{\sqrt{2\pi u_s h_{max}(t) \sigma_y}} \exp \left[ -\frac{y^2}{2\sigma_y^2} \right]. \quad (5.14)$$

Here $\dot{m}$ is the mass loading rate, $y$ is the lateral direction perpendicular to the plume centerline (coming out of the page in the right panel of Figure 5-16), $\sigma_y = \sigma_y(t)$ is the
standard deviation of the concentration distribution, \( t \) is the time since discharge, \( u_s \) is the ship speed, \( f_d(t) \) is the fraction dissolved CO\(_2\), and \( h_{\text{max}}(t) \) is the plume depth. The fraction dissolved CO\(_2\) is given by Chow [25]

\[
f_d(t) = 1 - \left( \frac{d(t)}{d_0} \right)^2 \quad (5.15)
\]

where \( d_0 \) is the initial particle size and \( d(t) \) is the particle size at time \( t \), equivalent to a longitudinal (downstream) distance \( x = u_s t \) from the injection, which is given by

\[
d(t) = d_0 - \left| \frac{dd}{dt} \right| t = d_0 - \left| \frac{dd}{dt} \right| \frac{x}{u_s} \quad (5.16)
\]

Here \( \left| \frac{dd}{dt} \right| \) is the particle dissolution rate, which as noted previously has been observed in the field to be roughly constant at 6 \( \mu \)m/s for the range of hydrate conversion efficiencies achieved. The plume depth \( h_{\text{max}}(t) \) is the descent depth of the heaviest particle which is given by

\[
h_{\text{max}}(t) = \begin{cases} 1 & t \leq t_d \\ 1 - \left( \frac{1 - \frac{t}{t_d}}{1 - \frac{t}{t_d}} \right)^{1/2} & t > t_d \end{cases} \quad (5.17)
\]

where \( t_d \) is the time for complete particle dissolution:

\[
t_d = \frac{d_0}{\left| \frac{dd}{dt} \right|}. \quad (5.18)
\]

and the total descent depth \( h_{d,\text{max}} \) is obtained from Figure 5-13. Equations (5.15) to (5.18) are adapted from Chow [25], where they are given as a function of \( x = u_s t \) and \( l_d \) (i.e., the downstream distance of complete particle dissolution). Variables \( f_d(t) \) and \( h_{\text{max}}(t) \) eventually reach steady values of 1 and \( h_{d,\text{max}} \), respectively. Vertical diffusion is conservatively ignored. Lateral diffusion, i.e., the growth of \( \sigma_y(t) \), is as before modeled using (5.13) with the parameter values detailed in Section 5.3.2.2.

For each configuration \((d = 2.5, 5, 10 \text{ or } 15 \text{ cm})\), three loading rates were tested: 10 kg/s, 100 kg/s, and 1,000 kg/s. The ship speed was set to \( u_s = 3 \text{ m/s (6 knots)} \) in each case, which is large enough to prevent a buoyant plume from being achieved even for the highest loading, i.e., treatment of the discharge as a collection of individual particles is appropriate. This is
Figure 5-17: Illustration of towed pipe moving reference frame. The source travels (up) with velocity $u_s$ and there is a prevailing ambient current $u_a$ such that in a fixed reference frame the plume centerline (longitudinal coordinate $x$) would be at an angle $\omega = \tan^{-1} \left( \frac{u_a}{u_s} \right)$ to the $-u_s$ direction.

concluded from the fact that the predicted plume separation depth due to the “crossflow” (ship motion) is much smaller than the depth at which a plume peeling event would be predicted to occur, as per the methodology of Socolofsky and Adams [116]. Ship speed is considered a design parameter, but at this value, assuming as in the other scenarios that the ambient current is 0.05 m/s and unidirectional, the angle between the ship travel direction and the plume centerline is small ($\omega < 1^\circ$). The width ($w$) of the source is also assumed to be a design parameter and was set to 10 m in each case, from which an equivalent Gaussian initial concentration distribution was parameterized using (5.12). The $w$ value of 10 m is of the order of the width of a ship, and was selected based on a reasonable arrangement of injectors to deliver the mass loadings evaluated in the present analysis if a hydrate release velocity of 4 m/s is assumed for a scaled-up discharge (see Chow [25] for more details). Since $u_s$ and $w$ are held constant for each scenario, the shape of the source does not vary with loading but merely increases in intensity.

This version of a “towed pipe” scenario is thought to be more realistic than the one in Adams and Wannamaker [1], where a single stream of pure hydrate particles was assumed to create a thin plume with $\sigma_{y,0} = 0.1$ m. The scenario described here is based on field observations in which a range of conversion percentages was achieved. This range causes a fractionation of the particle trajectories, which in turn allows the discharged CO$_2$ to be
more effectively distributed in the vertical. As hydrate formation technology improves, it is conceivable that it will eventually be possible to custom design a discharge configuration which maximizes vertical dilution by optimizing the distribution of discharged particle sizes and/or conversion percentages.

5.3.2.4 Rising CO\(_2\) droplets from a bottom manifold

A bottom manifold extends along the sea-floor and liquid CO\(_2\) is injected through equally spaced multiport diffusers starting at 800 m depth to form a series of positively buoyant (rising) plumes which are oriented perpendicular to the ambient current \((u_a)\). Assuming a slope \((\theta)\) of 5°, the pipeline would be 13.5 km long in total \((L_0)\), of which approximately the last 4.5 km \((L)\) would be below 800 m (Figure 5-18). Port nozzles with an appropriate diameter would be spaced along \(L\) such that a plume height \(h\) of 250 m would be achieved (a taller plume would be undesirable for sequestration purposes since the CO\(_2\) would form gas bubbles in the depth range 400-500 m). For example, for \(\dot{m} = 100 \text{ kg/s}\), 100 ports of diameter 0.7 cm each discharging 1 kg/s each would achieve such a plume height, where rising droplets would become covered with a thin hydrate film at this depth.

As in [1], the discharge is simplified into a rectangular source of width \(w = \min \left( \frac{h}{\tan \theta}, L \right)\) and height \(h^* = L \sin \theta\) (which is used in favor of the plume trap height \(h\) in order to preserve the cross-sectional area of the source). For consistency with the other two discharge scenarios, downstream diffusion of the source is here computed using the Gaussian solution (5.11) in favor of the rectangular source equation used by [1]. As in the other scenarios, (5.12) is used to compute \(\sigma_{y,0}\), and the nominal value of \(u_a = 0.05 \text{ m/s}\) is used.

In contrast to the other two scenarios in which the CO\(_2\) mass is discharged at 1,500 m depth, this scenario discharges over the depth interval 1,200 - 800 m with a constant plume trap height of 250 m. Thus, from the standpoint of sequestration efficiency (i.e., the length of time the CO\(_2\) is sequestered from the atmosphere), the scenario is less attractive. This can of course be addressed by changing the manifold configuration to reach deeper depths. Likewise, the manifold configuration can be changed to achieve higher dilution by “painting” a larger cross-sectional area with the CO\(_2\). Indeed, it should be possible in theory to select a mass loading and a manifold design that avoids significant acute biological impact.
altogether (e.g., the total mass loading can be restricted, the diffuser can be extended to
deeper depths and the diameter of the diffuser nozzles can increase with depth to achieve
greater plume rise). However, cost/benefit considerations will of course dictate what is
feasible in reality. For the present study we have chosen to simply adopt the configuration
proposed by Wannamaker and Adams [1] as an illustrative example, recognizing that in the
absence of cost constraints the bottom manifold design has a number of degrees of freedom
that could be exploited.

5.3.2.5 Plume representation in the discharge scenarios

The CO$_2$ discharges are modeled in a manner similar to Auerbach and Caulfield in that
an analytical solution is used. By contrast, some previous studies [23, 109, 111] have con-
ducted detailed hydrodynamic modeling near the injection point to resolve small-scale plume
gradients. Simplifying the geometry of the source and using an analytical solution amounts

Figure 5-18: Schematic illustration of the bottom manifold CO$_2$ discharge from Adams and
Wannamaker [1].
to simulating an ensemble plume, i.e., the average plume conditions that might occur over many realizations (see discussion in [63]). This naturally ignores the small-scale gradients that would be present within the plume during any one realization of the discharge, i.e., the CO$_2$ concentrations near the source are “smeared” relative to the concentrations that would be observed in a real plume. This approach has two counteracting effects from a biological standpoint. On the one hand, the inherent patchiness and high concentrations that would occur in the immediate vicinity of the source are not represented, likely under-predicting the peak concentration that an organism might experience during a realization. On the other hand, the ensemble plume introduces elevated concentrations to a larger area, which in turn means that a larger number of organisms are exposed to elevated concentrations. Sources have been simplified, but in a manner which attempts to err on the side of keeping concentrations on the high side. For example, for the towed pipe case, the mass is placed within a relatively narrow source, even though width is a design parameter. For the stationary hydrate plume case, the average concentration within the intrusion layers is applied over the entire depth of the plume even though the CO$_2$ really only leaves the plume at these layers, thus yielding a narrower plume (because $DIC_0$ in (5.10) is higher). In addition, by representing all sources as being Gaussian in shape, peak concentrations are specified near the source so as to reduce the “smearing” effect of the inherent averaging of the plume representation approach.

The approach used here is somewhat analogous to defining a mixing zone as is usually done when setting water quality compliance guidelines for ocean wastewater diffusers. In these cases, water quality standards are typically only enforced at the edge of the mixing zone, i.e., the environmental impact of the discharge is evaluated in terms of its impact on a larger scale rather than on the details of what occurs within the small dynamic mixing zone. Likewise the analysis here focuses on larger scale average impacts as opposed to near-source, small-scale impacts. Although future work might attempt to resolve these effects in the analysis, the reader should note that small-scale patchiness and high concentrations could to some extent be minimized by refining the design of the discharge configurations to homogenize the near-field region.
5.3.2.6 Treatment of carbonate system chemistry

When CO$_2$ is introduced to seawater a series of reactions take place which are collectively referred to as the carbonate system [35, 88]:

\[ \text{CO}_2(g) \rightleftharpoons \text{CO}_2(aq) \] (5.19a)
\[ \text{CO}_2(aq) + \text{H}_2\text{O}(l) \rightleftharpoons \text{H}_2\text{CO}_3(aq) \] (5.19b)
\[ \text{H}_2\text{CO}_3(aq) \rightleftharpoons \text{H}^+(aq) + \text{HCO}_3^-(aq) \] (5.19c)
\[ \text{HCO}_3^-(aq) \rightleftharpoons \text{H}^+(aq) + \text{CO}_3^{2-}(aq) \] (5.19d)

which is here shown for a water sample at equilibrium with the gaseous phase (e.g., at the air-water interface). However, because H$_2$CO$_3$(aq) and CO$_2$(aq) are difficult to distinguish analytically, it is common to combine these terms and designate their sum as H$_2$CO$_3^*$ (aq). Thus, the carbonate system may be expressed as

\[ \text{CO}_2(g) \rightleftharpoons \text{H}_2\text{CO}_3^*(aq) \] (5.20a)
\[ \text{H}_2\text{CO}_3^*(aq) \rightleftharpoons \text{H}^+(aq) + \text{HCO}_3^-(aq) \] (5.20b)
\[ \text{HCO}_3^-(aq) \rightleftharpoons \text{H}^+(aq) + \text{CO}_3^{2-}(aq) \] (5.20c)

For the present study, the approach summarized by Wannamaker [137] is applied to solve for the equilibrium concentration of the carbonate system components. This approach combines the methodology outlined in Morel and Hering [88] with the equilibrium constants defined by Dickson and Goyet [35]. Briefly, the concentrations of the individual species in (5.20) are calculated by solving the following nonlinear equation for the hydrogen ion concentration [H$^+$]:

\[ C - \text{Alk} = -[H^+] + 10^{-pK_w}[H^+]^{-1} + \alpha_1 \text{DIC} + 2\alpha_2 \text{DIC} \] (5.21)

where \( C - \text{Alk} \) is the carbonate alkalinity

\[ C - \text{Alk} \equiv -[H^+] + [\text{OH}^-] + [\text{HCO}_3^-] + 2[\text{CO}_3^{2-}], \] (5.22)
DIC is the total dissolved inorganic carbon

\[ DIC \equiv [H_2CO_3^+] + [HCO_3^-] + [CO_3^{2-}], \]  

(5.23)

and \( \alpha_1 \) and \( \alpha_2 \) are ionization fractions for the bicarbonate \( (HCO_3^-) \) and carbonate \( (CO_3^{2-}) \) ions

\[
\alpha_1 \equiv (1 + 10^{pK_1}[H^+] + 10^{-pK_2}[H^+]^{-1})^{-1}
\]  

(5.24a)

\[
\alpha_2 \equiv (1 + 10^{pK_2}[H^+] + 10^{(pK_1+pK_2)[H^+]^2})^{-1}.
\]

(5.24b)

Here the notation \( pK \equiv -\log(K) \) is applied for brevity to \( K_1, K_2, \) and \( K_w \), which are the dissociation constants of \( H_2CO_3^+ \), \( HCO_3^- \), and water, respectively:

\[
K_1 \equiv \frac{[H^+][HCO_3^-]}{[H_2CO_3^+]} \]  

(5.25a)

\[
K_2 \equiv \frac{[H^+][CO_3^{2-}]}{[HCO_3^-]} \]  

(5.25b)

\[
K_w \equiv [H^+][OH^-]. \]  

(5.25c)

The equilibrium concentrations of the carbonate system species are given by

\[
[H_2CO_3^+] = \alpha_0 DIC \]  

(5.26a)

\[
[HCO_3^-] = \alpha_1 DIC \]  

(5.26b)

\[
[CO_3^{2-}] = \alpha_2 DIC \]  

(5.26c)

where the ionization fraction for \( H_2CO_3^+ \) is given by

\[
\alpha_0 \equiv (1 + 10^{-pK_1}[H^+]^{-1} + 10^{-(pK_1+pK_2)[H^+]^2})^{-1}.
\]

(5.27)

The values for the dissociation constants as a function of temperature and salinity are taken from Dickson and Goyet [35]. Pressure effects on the constants are not included as the present study considers a range of depths and the pressure adjustments are expected to be
minor based on a cursory analysis with the CO2SYS program [80].

Equation (5.21) is solved for \([H^+]\) with given values of \(DIC\) and \(C - Alk\), where \(DIC\) is computed during the plume calculations (see Section 5.3.2) and \(C - Alk\) is a constant specific to the seawater being studied. Note that addition of \(CO_2\) does not change the value of \(C - Alk\). Once \([H^+]\) is determined, all species concentrations can be calculated easily. All pH values are reported on the total hydrogen ion scale (as opposed to the free hydrogen ion scale, see [35]).

In the present application, the source of DIC is not in equilibrium with a gaseous phase as indicated by the first equation in (5.19) and (5.20). Instead, CO\(_2\)(aq) is introduced directly (by dissolving hydrates or droplets). Had the seawater instead been in equilibrium with a gaseous phase, \([H_2CO_3^+]\) would be constrained to be (e.g., [41])

\[
[H_2CO_3^+] = pCO_2 \cdot K_0(T, S) \tag{5.28}
\]

where \(pCO_2\) is the partial pressure of the gaseous \(CO_2\) and \(K_0(T, S)\) is the solubility of \(CO_2\), which is an empirical function outlined in Weiss [141] (the difference between fugacity and partial pressure has here been ignored as it is < 1% for \(CO_2\) in seawater at 1 atm). In the present study, (5.28) is used to convert computed \(H_2CO_3^+\) concentrations to an equivalent \(pCO_2\) at atmospheric pressure. This is motivated by the biological mortality data reviewed in Section 5.2, which were for the most part reported in terms of \(pCO_2\). \(pCO_2\) is a convenient measure because it is easily compared to prevailing atmospheric concentrations. It is usually reported in units of \(\mu\)atm or kPa. In the present study, the latter unit is used, but readers more comfortable with \(\mu\)atm can easily do an approximate conversion by multiplying the \(pCO_2\) values by \(10^4\) (1 kPa \(\approx 9.87 \times 10^2\) \(\mu\)atm).

In the above calculations, the kinetics of the reactions (5.20) are ignored, i.e., equilibrium is assumed to be established instantaneously. Zeebe et al. [148] estimates the timescale to reach equilibrium in seawater to be on the order of 16 s. While this timescale is comparable to that of some near-field plume processes, the present study considers the average conditions caused by simplified representations of the \(CO_2\) sources and thus disequilibrium kinetics are appropriately ignored, consistent with common practice.
5.3.2.7 Ambient ocean conditions

The plume and hydrate particle descent calculations above use an ambient density profile taken from a 1999 survey cruise near Keahole Point, Hawaii, as in previous studies (e.g., see [1, 137]). For ocean chemistry calculations, global average values for depths > 2000 m were selected from Volk and Hoffert [136]: salinity = 35 ppt, DIC = 2306 μmol/kg, and C-Alk2 = 2367 μmol/kg. The temperature used was 3°C for the stationary and towed hydrate releases and 5°C for the bottom manifold, based on the observed temperatures at Keahole Point at release depths of 1,500 and 800 m. For comparison, Volk and Hoffert [136] used a temperature range of 1.5 - 3°C as representative of the deep ocean. The ambient pH computed for these conditions is about 7.94 and 7.90 for temperatures of 3 and 5°C, respectively.

5.3.3 Biological impact analysis approach

The following section details the methodologies employed in simulating the biological impact. Following Auerbach [4] and Caulfield [18], this consists of two parts: (1) simulating the exposure history of planktonic marine organisms and (2) simulating the biological impact of the exposure. The approach implemented herein is a combination of the approaches of Auerbach [4], Caulfield [18], and Sato et al. [109].

5.3.3.1 Organism exposure modeling

The present study simulates organisms in a manner that is highly similar to that of Caulfield [18], i.e., planktonic organisms are advected with the CO₂ plume and sample different concentrations as they undergo a random walk in the lateral direction using a diffusivity that increases with distance from the plume centerline. The following section describes the implementation of this approach, expanding on the work of Caulfield [18] and modifying the methodology as needed.

5.3.3.1.1 Stochastic simulation of organism trajectory  Caulfield [18] discusses various ways in which (5.1) can be employed to estimate the trajectory of an organism through

\[ \text{Computed using CO2SYS [80] from a reported total alkalinity of 2414 \( \mu \text{mol/kg} \) with a phosphate concentrations of 2.21 \( \mu \text{mol/kg} \).} \]
Figure 5-19: Schematic diagram of the simulation of organisms drifting through a CO$_2$ plume. Organisms undergo a random walk relative to the plume centerline.

A plume, including stochastic simulation by a Monte Carlo approach and a number of other approaches aimed at reducing computation time. The subsequent paper Caulfield et al. [19], however, only mentions the stochastic approach, and such an approach is employed here. A schematic diagram of this approach is shown in Figure 5-19, where the paths of drifting organisms are tracked through a CO$_2$ discharge plume.

Although details are not provided, presumably the stochastic simulation was conducted by finding the random walk equivalent of (5.1). Rearranging (5.1) yields

$$\frac{\partial p(y)}{\partial t} + \frac{\partial}{\partial y} \left( p(y) \frac{\partial F(y)}{\partial y} \right) = \frac{\partial^2}{\partial y^2} \left( p(y) F(y) \right). \quad (5.29)$$

Equation (5.29) is similar in form to the one-dimensional Fokker-Planck equation

$$\frac{\partial f}{\partial t} + \frac{\partial A f}{\partial x} = \frac{\partial^2}{\partial x^2} \left( \frac{1}{2} B^2 f \right) \quad (5.30)$$
where $p = p(x)$ is the probability density function of the stochastic variable $f$, and $A$ and $B$ are coefficients to be determined. Since (5.29) and (5.30) are equivalent if $A = \frac{\partial F(y)}{\partial y}$ and $B = \sqrt{2F(y)}$, (5.29) can be simulated by the random walk equation

$$
\Delta y = \frac{\partial F(y)}{\partial y} + \xi \sqrt{2F(y)} \Delta t
$$

where $\xi$ is a random number drawn from a Gaussian distribution with zero mean and unit standard deviation [43, 129], $\Delta t$ is the timestep, and $\Delta y$ is the lateral displacement over $\Delta t$. Equation (5.31) is used to simulate the lateral position of particles (organism clusters) relative to the plume, where both the plume and the organisms move longitudinally with the ambient current $u_a$. Although Caulfield [18] adapts the method to be more computationally efficient by calculating transition probabilities between cells on a spatial grid rather than for individual organisms, initial simulations with the “brute force” approach described above show that computation time is not a major constraint with present computing technology. Thus, the computational domain is initialized with a sufficiently large number of particles which are all placed at $x = 0$ (where $x$ is the longitudinal coordinate) but spaced out evenly across the domain in $y$. The domain is bounded in the lateral direction by a reflecting wall placed sufficiently far away from the centerline so as to make it highly unlikely that a particle will reach the plume centerline and the domain boundary multiple times during a simulation. Because of the symmetry of the plumes being modeled, only half of the domain needs to be simulated with particles, and thus another reflecting wall is placed along the plume centerline. At each timestep, the average $pCO_2$ experienced by a particle is used to update that organism cluster’s % mortality (population deficit) via the isomortality calculation detailed later in this section. A spatial grid is used to bin particles and calculate gridded variables as needed.

### 5.3.3.1.2 Scale-dependent diffusivity

The motion of the organisms relative to the plume centerline is inherently scale-dependent. As organisms move further away from the plume centerline, the scale of eddies that can differentially advect them increases. Since this unresolved motion is parameterized as a diffusivity, this diffusivity must be dependent on the separation distance of the organisms and the plume centerline. This is directly analogous to
using a scale-dependent diffusivity to compute the spreading of the discharged CO$_2$ plume; as the plume becomes larger, it is subject to shearing by eddies of larger scales.

The field data from Okubo [93] on observed diffusion in coastal waters have been reinterpreted to yield a new scale-dependent diffusivity relationship which is applied to simulating the lateral diffusion of organisms relative to the plume centerline. The relationship differs from (5.2), which was used by Caulfield [18]. The derivation of the new equation is outlined below.

Okubo [93] examined the growth of dye patches in coastal waters in two ways. First, he defined the apparent diffusivity $K_a$ implied by the growth of the length scale of dye patches:

$$K_a \approx 0.0103 l^{1.15}$$ (5.32)

where $l \equiv 3\sigma_r = 3\sqrt{2}\sigma_y$ is his length scale, $\sigma_y$ is the one-dimensional standard deviation of the concentration distribution, and $\sigma_r = \sqrt{2}\sigma_x$ is the equivalent radial standard deviation of the distribution. Second, he developed an equation describing the time rate of growth of the radial variance of the dye patches

$$\sigma_r^2 \approx 0.0108 t^{2.34}$$ (5.33)

where $\sigma_r$ is in cm and $t$ is in s. From (5.33) the effective diffusivity $E_r$ and the apparent diffusivity $K_a$ can be calculated:

$$E_r \equiv \frac{1}{4} \frac{d\sigma_r^2}{dt} \approx 0.0063 t^{1.34}$$ (5.34)

$$K_a \equiv \frac{\sigma_r^2}{4t} \approx 0.0027 t^{1.34}$$ (5.35)

where $E_r$ and $K_a$ are in cm$^2$/s. Thus, (5.33) implies that the effective diffusivity is 2.34 times larger than the apparent diffusivity. The difference is that $K_a(t)$ is the constant diffusivity that would achieve $\sigma_r(t)$ in $t$, whereas $E_r(t)$ is the actual scale-dependent diffusivity at time $t$. Thus, multiplying (5.32) by a factor of 2.34 yields the effective diffusivity in terms of the
length scale \( l \):

\[
E_r \approx 0.0241 l^{1.15}. \tag{5.36}
\]

Inserting \( l \equiv 3\sigma_r \) and (5.33) into (5.36) confirms that (5.36) and (5.34) are equivalent. Because the length scale used by Okubo [93] is arbitrary, for the present purpose \( E_r \) is expressed in terms of \( \sigma_r \) or \( \sigma_y \):

\[
E_r \approx 0.0853 \sigma_r^{1.15} = 0.127 \sigma_y^{1.15}. \tag{5.37}
\]

The result in (5.37) is then adapted to the distance-neighbor diffusivity \( F(\gamma) \) needed by (5.1) to simulate organism diffusion relative to the plume centerline. Consider a group of particles spread out according to an arbitrary spatial distribution. Richardson [104] showed that for any distribution, the variance of the separation distance between the particles is twice the variance of the distribution. Thus, if the variance of the particle distribution grows at a certain rate, then the variance of the separation distances must increase twice as fast, i.e., its diffusivity must be twice as large. When applying this logic to (5.37) to simulate the separation between a plume centerline at \( \gamma = 0 \) and an organism located at \( \gamma \), it is not obvious which of \( \sigma_y \) and \( \sigma_r \) is most analogous to the separation distance \( \gamma \). To be conservative, \( \sigma_r \) is selected because it results in a smaller diffusivity. Multiplying (5.37) by a factor of 2 to convert it into a distance-neighbor diffusivity yields

\[
F(\gamma) \approx 0.171 \gamma^{1.15} \tag{5.38}
\]

where \( \gamma \) is in cm and \( F(\gamma) \) is in cm\(^2\)/s. Finally, to be consistent with the assumption of reduced horizontal mixing in the deep ocean used in the CO\(_2\) plume diffusion calculation, the diffusivity \( F(\gamma) \) is reduced by a factor of 10, yielding

\[
F(\gamma) \approx 0.00034 \gamma^{1.15} \tag{5.39}
\]

where the units of \( F(\gamma) \) and \( \gamma \) have been converted to m\(^2\)/s and m, respectively. Overall the diffusivities used to separate organisms from the plume centerline in the present study are
about one fifth of the values used by Caulfield [18].

5.3.3.2 Organism impact modeling

As in Auerbach [4] and Caulfield [18], the general approach of adding exposures via an isomortality model of biological impact is used in the present study. However, the specific approach used is adapted from Sato et al. [109]. The following section outlines the biological impact calculations and also derives the isomortality relationships employed in the present study from a subset of the CO₂ mortality data reviewed in Section 5.2.

5.3.3.2.1 Implementation of isomortality model The implementation of the isomortality model in the present study differs in three main ways from that of the original investigation by Auerbach [4] and Caulfield [18]. First, based on the data reviewed in Section 5.2, pCO₂ is used as the stressor rather than pH. Second, the extended probit model developed by Sato and Sato [111] and Sato et al. [109] is used in favor of the original approach in [4]. Third, the species recovery included in previous modeling efforts is not included in the present study.

As discussed in Section 5.3.1.2, the main difference between Auerbach’s isomortality approach and Sato’s extended probit model approach lies in the form of the isomortality functions. Both methods add up exposures by translating an organism cluster’s accumulated percent mortality from one stress level (pH or pCO₂) to another by defining an equivalent cumulative exposure time. However, while Auerbach [4] developed regression curves based on independent fits of the LC₀, LC₅₀, and LC₉₀ data, Sato [109] used a single function fitted to all the mortality data to derive regression equations. The mathematical elegance of the latter approach makes it more convenient in that a single probit equation is used to describe the biological impact, allowing the process of adding exposures to be done without iteration if the function is linear. More importantly, the form of the equation is based on the probit method commonly used to estimate LC values and is therefore more consistent with the underlying toxicity data. Specifically, the inherent assumption in the probit method for calculating an LC₅₀ (for example) is that the cumulative mortality of organisms follows a sigmoid function of the log of the stressor, and thus a sigmoid curve is fitted to the observed
toxicity data so that any LC can be calculated directly. This imposes a specific form to the LC values (e.g., in log space, $LC_{50} - LC_{10} = LC_{90} - LC_{50}$; $LC_{50} - LC_{40} \neq LC_{40} - LC_{30}$). The extended probit method respects these relationships since it is based on the probit method. In contrast, the Auerbach method does not in that it independently regresses relationships for each mortality level and linearly interpolates between them. Furthermore, to avoid having isomortality lines converge, the form of the curves must be adjusted manually. On the basis of these considerations, the Sato approach to isomortality modeling was selected because it permits less subjectivity on the part of the modeler. As noted previously, for simplicity the probit function of Sato et al. is here referred to as an isomortality function since it accomplishes the same goal as the original curves of Auerbach [4].

Thus, the isomortality function used in the present study is of the form:

$$Y = a \log(t) + b \log(\Delta pCO_2) + c$$  \hspace{1cm} (5.40)

where $Y$ is the probit of mortality (defined by (5.6)), $t$ is the exposure time, $\Delta pCO_2$ is the excess $pCO_2$ over the ambient $pCO_2$, and $a$, $b$, and $c$ are regression coefficients. Excess $pCO_2$ is used instead of the absolute value so that data with different background values can be used together. The specific regression coefficients used in the present study are developed in the next section. It should be noted that the isomortality function does not necessarily have to be linear; the fit could be performed using, for example, a quadratic, such as $Y = a_1 (\log(t))^2 + a_2 \log(t) + b \log(\Delta pCO_2) + c$ (this was the form of the function used in Sato and Sato [111]). For the present purpose, however, a linear function provides a reasonable fit to the data and is convenient for its simplicity and consistency with Sato et al. [109, 112].

The algorithm to step through time is similar to the approach by Auerbach [4] (see Section 5.3.1.1.3), with a few modifications. First, calculations are performed in terms of $Y$ instead of $D$, i.e., using the probit unit instead of fractional mortality. Second, $Y$ is calculated directly from $t$ and $\Delta pCO_2$ using (5.40) without interpolation between specific LC curves, and no special treatment is required for low mortality since a zero mortality does not exist in probit space ($D \to 0$ as $Y \to -\infty$). Consequently, (5.3) becomes $Y(\Delta pCO_2, t^*) = Y_1$
from which \( t^* \) can be solved without iteration. Lastly, the probit value must be converted back into fractional mortality for output (and for insertion into a species recovery function, if applicable):

\[
D = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{Y - 5}{\sqrt{2}} \right) \right]
\]

(5.41)

which was derived by integrating the Gaussian distribution in (5.6) with zero mean and unit standard deviation.

The extended probit method is here modified to include a minimum stress level, \([\text{pCO}_2]_{\text{min}}\), below which mortality is not incurred. This is introduced to avoid an inherent problem, namely that (in the absence of species recovery) an infinitesimal mass of CO\(_2\) would be predicted to eventually kill all organisms. Sato et al. [109] also recognized this effect, and as a result they discounted mortality effects when pCO\(_2\) dropped to within 100 \( \mu \text{atm} \) (\( \approx 0.01 \) kPa) of the ambient value.

The fact that the present study does not include a species recovery as in Caulfield [18] warrants some discussion. Solving equation (5.4) yields the discrete form of the species recovery which was likely implemented in [18]:

\[
D(t + \Delta t) = 1 - \frac{1 - D_{im}(t + \Delta t)}{1 + D_{im}(t + \Delta t) \left( e^{-r\Delta t} - 1 \right)}.
\]

(5.42)

where \( D_{im} \) is the population deficit (fractional mortality) predicted by the isomortality method and \( r \) is the growth rate. Thus the deficit of each particle decreases for each timestep that it encounters a low stress environment. In Caulfield’s case, this corresponded to pH > 7.35. The benefit of including such a recovery formulation is twofold: (1) it seems realistic to expect a population to recover when stress is reduced, and (2) it allows the calculation of a finite footprint for each CO\(_2\) discharge, i.e., the entire population will recover to its carrying capacity at some distance from the source. However, recovery is not included in the present study because the recovery rate is unknown and poorly constrained. Instead, calculations are made as transparent as possible and the issue of species recovery is deferred until later in the analysis.
5.3.3.2.2 Isomortality functions  Listed below are some conclusions from the literature review in Section 5.2 that are relevant to the development of isomortality functions. Since they are based on limited data for a small number of species, they can hardly be taken as generalizations of all marine organisms; rather, they will be used to select the most appropriate data from the existing CO$_2$ mortality dataset.

- Marine organisms are much more sensitive to pH depressions caused by CO$_2$ than by other acidifying agents [70, 52, 140].

- Fish are among the more CO$_2$ tolerant marine species due to their internal compensation mechanisms [101, 112], and there is considerable variability among different species [69, 61]. Tolerance varies over the life cycle, with the lowest tolerance exhibited during the early developmental (egg cleavage) and juvenile stages [69]. The tolerance is highly dependent on the nature of the CO$_2$ exposure; a gradual increase in concentration gives a higher tolerance in developmental and adult fish than a sudden increase, and mortality can be induced by a sudden return to normocapnic conditions from an elevated but sub-lethal CO$_2$ concentration [71, 68]; this implies that they do not fit into the isomortality cumulative approach [112]. As for whether juvenile and adult fish will be able to sense and avoid high CO$_2$ plumes, the data are mixed [127, 135].

- Copepods generally exhibit less CO$_2$ tolerance than fish, apparently lacking similar compensation mechanisms [101, 112]. There is variability between the CO$_2$ tolerance of different species, and between different geographical regions [140]. While some data suggest that deep-sea copepods may be more tolerant of CO$_2$ than their shallow-water counterparts [140], the generality of this conclusion is controversial [56] as it may be limited to the “oxygen minimum zone” that exists at about 1,000 m depth. As for gradual versus sudden changes in CO$_2$ concentration, the response of copepods appears to be more linear, i.e., unlike fish [112].

- CO$_2$ exposure effects are not limited to acute mortality; a wide range of sub-lethal effects have been noted. In particular, a significant effect on reproduction has been noted in copepods and gastropods [74, 76, 75], and in fish [61, 69].
Based on these considerations, the main isomortality function used in the present study is based on the pelagic copepod dataset of Watanabe et al. [140]. Despite the comparatively large dataset, the fish data are not incorporated into the main isomortality function for the reasons noted above, although some of the data are used in a sensitivity analysis. The dataset compiled by Auerbach et al. [5] and the pelagic copepod data from Yamada and Ikeda [145] are not used because they consider mortality due to non-CO$_2$ acidification. The benthic species datasets are not used because (a) there is little in the way of controlled dose-response data, (b) the limited data available suggest higher tolerance than pelagic species at least in the short term, and (c) the discharge scenarios considered herein are dilution strategies most likely to impact pelagic species (since they seek to distribute CO$_2$ over large volumes rather than concentrate it on or beneath the sea-floor). Thus, the present study uses copepods as the main target organisms, which are assumed representative of zooplankton in general. The conservative nature of this choice is supported by the conclusion of Kita and Watanabe [73], who identified copepods as being among the more sensitive species to high-CO$_2$ conditions.

In generating the isomortality functions from the Watanabe et al. [140] copepod dataset, the most sensitive species are selected. Watanabe et al. [140] identified these as being the “surface-living groups”, which they further divided into two groups: (1) subarctic and transitional regions and (2) subtropical region (see Figure 5-20). In each of these groups, the shallow-living copepods exhibited greater sensitivity than their deep-living counterparts. The first group consisted of three single-species experiments (Calanus pacificus, Metridia pacifica, and Euchaeta marina) while the second group consisted of 4 experiments each comprising a mix of epipelagic species (see [140] for a taxonomic breakdown). A comparison of the LT$_{50}$ values in Figure 5-20 suggests that in general the first group (subarctic and transitional regions) displays greater sensitivity. Two datapoints are excluded from the figure because the authors could not calculate an LT$_{50}$. From the subarctic/transitional group, an LT$_{50} >$ 140 hours was observed for a pCO$_2$ of about 0.15 kPa ($\Delta$pCO$_2$ ≈ 0.04 kPa), and from the subtropical group, an LT$_{50} <$ 6 hours was observed for a pCO$_2$ of about 9.9 kPa ($\Delta$pCO$_2$ ≈ 9.8 kPa).
Figure 5-20: Reported LT\textsubscript{50} values for Western North Pacific copepods from Watanabe et al. [140]. Filled squares: shallow-living copepods from the subarctic and transitional regions; hatched squares: shallow-living copepods from the subtropical region; open squares: deep-living copepods. For comparison, the harpacticoid copepod data from Sato et al. [112] are shown (plus signs).

From the raw data plots for these stations (provided by the lead author of [140]), LT\textsubscript{10}, LT\textsubscript{50}, and LT\textsubscript{90} values were estimated using the probit method of sigmoid curve fitting as in [140] (see Finney [38] for an overview). Good agreement was achieved between the estimated and reported LT\textsubscript{50} values after certain assumptions were made during the analysis: (1) the first observed instances of 0% and 100% mortality were assigned values of 1% and 99% during the regressions, and (2) death in the control samples was ignored\textsuperscript{3}. This latter assumption has little impact on the first (subarctic and transitional) group as these exhibited little control sample mortality. However, substantial control mortality was observed in all four experiments in the second group, and as a result these values must be considered rather conservative.

\textsuperscript{3}These assumptions were needed to achieve good agreement with reported LT\textsubscript{50} values; we have been unsuccessful in contacting the lead author of [140] to confirm since receiving the raw data plots.
Multilinear regression was applied to both shallow-living copepod groups to determine the coefficients in a probit-type isomortality function (5.40). The regressions were performed on the estimated $LT_{10}$, $LT_{50}$, and $LT_{90}$ values rather than the raw data because of the required assumptions mentioned above. Prior to regression, the pCO$_2$ values were converted to $\Delta$pCO$_2$, i.e., the control sample pCO$_2$ was subtracted out. The resulting regression lines are plotted together with the underlying datasets in Figures 5-21 and 5-22, and summarized in Table 5.3. Because the reported $LT_{50}$ values are less scattered for the first group (see [140]), and because of the significant control mortality not accounted for in the second group statistics, the isomortality function for the subarctic and transitional region (Figure 5-21) is selected as more realistic. The reader should note that basing the isomortality function on the shallow-living copepods is conservative, since the discharge scenarios considered herein are at or below the Watanabe et al. deep-living sample depths.
Figure 5-22: Isomortality function derived from shallow copepod samples from the subtropical region of the Western North Pacific of Watanabe et al. [140]. Open diamonds: LT$_{10}$; filled squares: LT$_{50}$; open triangles: LT$_{90}$. Solid line: LT$_{50}$ function; dashed lines: LT$_{10}$ (left) and LT$_{90}$ (right) functions.

<table>
<thead>
<tr>
<th></th>
<th>$a$</th>
<th>$b$</th>
<th>$c$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>subarctic/transitional copepods</td>
<td>2.32</td>
<td>1.69</td>
<td>1.36</td>
<td>0.64</td>
</tr>
<tr>
<td>subtropical copepods</td>
<td>1.56</td>
<td>0.741</td>
<td>2.27</td>
<td>0.68</td>
</tr>
<tr>
<td>developmental fish</td>
<td>0.357</td>
<td>3.64</td>
<td>3.40</td>
<td>0.49</td>
</tr>
</tbody>
</table>

Table 5.3: Regression coefficients for various isomortality functions of form (5.40) for time in hours and $\Delta$pCO$_2$ in kPa.

There are two obvious gaps in the copepod mortality dataset which require some discussion, and which future investigations may want to address. The first is the lack of mortality data for low $\Delta$pCO$_2$ exposure levels/long times. This could be due to experimental design or to the fact that the acute toxicity signal is too subtle to measure, i.e., if the LT$_{50}$ approaches the natural lifespan of the organism. The data suggest that the latter is the case. Watanabe et al. [140] noted significant control mortality in the 4-7 day range, and for several lower exposure experiments in the $\Delta$pCO$_2$ range 0.03 - 0.4 kPa there was less than 50% mortality.
during experiments of duration 6 - 10 days. Similarly, Kurihara et al. [74] noted control mortality in excess of 40-60% after 8 days for adult females of two surface-dwelling copepods with no significant added mortality for \( \Delta pCO_2 \) exposures in the range 0.1 - 1 kPa. In addition, the same study indicated that strong impacts on egg production rate, hatching rate, and nauplius survival rate were only observed for \( \Delta pCO_2 \geq 0.5 \) kPa [74], i.e., sub-lethal effects were not observed for lower pCO2.

In the absence of quantitative mortality data for low exposures, the isomortality function is extrapolated to a \([\Delta pCO_2]_{\text{min}}\) of 0.015 kPa (\( \approx 148 \mu \text{atm} \)), i.e., mortality is accrued above this level. Such an extreme low value implies extrapolating well beyond the data range, since the smallest value of \( \Delta pCO_2 \) used in group 1 regression is 0.26 kPa. Anecdotally it is noted that the reported datapoint of \( LT_{50} > 140 \) hours for a \( \Delta pCO_2 \) of 0.04 kPa is not inconsistent with the extrapolation of the \( LT_{50} \) line to 0.015 kPa. The 0.015 kPa criteria is similar to the 100 \( \mu \text{atm} \) (\( \approx 0.01 \) kPa) used by Sato [109], and has been chosen because, with the selected ocean chemistry parameters at 3°C, \( \Delta pCO_2 = 0.015 \) kPa corresponds to \( \Delta pH \approx -0.1 \). As discussed in Section 5.4.2, natural pH variability in the ocean suggests that marine organisms should be capable of tolerating a pH drop of 0.1 without any sub-lethal impacts [8], which is supported by the observations of Kurihara et al. [74]. Furthermore, this value is well below the Predicted No Effect Concentration (PNEC) for CO2 of +500 \( \mu \text{atm} \) estimated for copepods by Kita and Watanabe [73], based on acute and chronic toxicity data. Thus, to allow acute impacts (mortality) to accrue until this level is very conservative within the bounds of the present study.

The other gap in the copepod dataset is the lack of mortality data for short times/high pCO2. For the data used in the regression, the highest \( \Delta pCO_2 \) is 4.6 kPa. The data provided by Watanabe [139] (see Section 5.2) in which copepods were exposed to 20 kPa for short durations are not true toxicity data in the sense that organisms were allowed to recover after appearing dead, and indeed the fact that most of them survived indicates little mortality for exposures of 20 kPa for 30 minutes (and are therefore not inconsistent with the extrapolated curves). Conversely, the datapoint from the subtropical group of \( LT_{50} < 6 \) hours for \( \Delta pCO_2 \approx 9.8 \) kPa suggests that extrapolating the isomortality function into short times and high \( \Delta pCO_2 \) is not conservative.
The latter gap in the data is addressed by a sensitivity analysis in which a second isomortality function is developed from a subset of the developmental fish dataset. The dataset is attractive because mortality was observed for short times (minutes) at concentrations well below those predicted by extrapolating the copepod isomortality curve. It could be argued that they are appropriately considered since a CO$_2$ discharge design should seek to avoid harming any species. However, as noted previously, the response of these organisms is not well-suited to the isomortality approach and thus they are of questionable applicability. Nonetheless, in the interest of being as conservative as possible in the absence of a better dataset, a subset of these data is used. Specifically, the embryo and juvenile Japanese sillago and red sea bream data from Kikkawa et al. [69] are selected because they exhibited the highest sensitivity (Figure 5-23). An isomortality function has been developed in the same manner as for the copepod dataset (Table 5.3), as shown in Figure 5-24 along with the main (copepod) function.

The sensitivity analysis using the developmental fish dataset is conducted in a special manner in the interest of being conservative. At each timestep, the organism’s mortality is evaluated with both the base copepod function and the developmental fish function, and the larger of the two impacts is selected. At the end of the timestep, the cumulative equivalent exposure time using each curve is calculated for the final mortality and exposure pH in preparation for the next step. Thus, the organism’s actual equivalent time is discontinuous as it may vary from timestep to timestep depending on which curve is used. Overall, this approach maximizes impact by allowing a single simulated organism to experience the sensitivity of multiple species$^4$.

$^4$This approach was selected in favor of a single regression line using a combined copepod/developmental fish dataset because a linear fit was too poor to be of value, and experimentation with nonlinear fits did not yield well-behaved functions.
Figure 5-23: LC50 values for developmental fish from Kikkawa et al. [69]. All reported data are plotted. The cleavage stage embryos and juvenile Japanese sillago and red sea bream were selected for the sensitivity analysis isomortality function and are shown as filled symbols.
Figure 5-24: Isomortality function derived from selected subset of developmental fish data from [69], plotted together with the main copepod isomortality function and the selected subset of copepod data. Open diamonds: LT$_{10}$; filled squares: LT$_{50}$; open triangles: LT$_{90}$. Solid line: LT$_{50}$ function; dashed lines: LT$_{10}$ and LT$_{90}$ functions. Note that LC values were reported by the author, but in the isomortality framework the LT and LC concepts are interchangeable.
5.4 Results

Overall, 6 different discharge configurations are studied: a stationary hydrate plume released at 1,500 m and just reaching the ocean bottom at 4,500 m; a bottom manifold releasing CO$_2$ droplets over a 4.5 km distance from the depth range 800-1200 m with a plume rise of 250 m; and four towed pipe scenarios in which single diameter hydrate composite particles (2.5, 5, 10, and 15 cm) are released at 1,500 m with a range of hydrate conversion percentages ($0.12 < X_{hyd} < 0.55$) into an ambient current. For each discharge configuration, CO$_2$ mass loadings of 10, 100, and 1,000 kg/s are evaluated, corresponding roughly to the output of 0.1, 1, and 10 500-MW coal-fired power plants [54].

Calculation results are presented below in the following order: water quality impacts, biological response, and sensitivity analysis. Result interpretation is deferred until the final section of this chapter.

5.4.1 Water quality impacts of discharges

Figure 5-25 shows the pH and pCO$_2$ impact volumes (i.e., the volume with pH below or pCO$_2$ above a given level) for CO$_2$ mass loadings of 100 kg/s and 1,000 kg/s. Only those combinations which cause a ΔpCO$_2$ impact greater than 0.015 kPa (ΔpH ≲ -0.1) are shown. The 10 kg/s mass loading is not shown as only the stationary hydrate plume exceeds the ΔpCO$_2$ criterion.

Some level of impact is predicted (i.e., ΔpCO$_2$ > 0.015 kPa) for all the studied discharge scenarios for the 100 kg/s (and 1,000 kg/s) loading; if the loading is reduced to 10 kg/s, only the stationary hydrate plume exceeds the 0.015 kPa ΔpCO$_2$ impact criterion (not shown). In terms of volume (Figure 5-25), the impact of the towed pipe method decreases with increasing particle diameter, and the towed pipe scenarios demonstrate lower volumes than the stationary hydrate plume at any impact level for any studied combination of loading and particle diameter. The bottom manifold has the greatest volumes at lower impact levels, but its maximum impact level is lower than all but the 10 and 15 cm towed pipe scenarios. In contrast, the stationary hydrate discharge yields a lower volume than the bottom manifold for the lowest impact levels, but higher volumes as the impact level increases. The same is
Figure 5-25: Impact volumes for discharge scenarios in terms of pH (top panels) and pCO$_2$ (bottom panels) for mass loadings of 100 kg/s (left panels) and 1,000 kg/s (right panels). Only those scenarios which achieve an impact of $\Delta$pCO$_2$ $\geq$ 0.015 kPa are shown: sh = stationary hydrate plume; tpX = towed pipe with X cm particle diameter; and bm = bottom manifold. Towed pipe results for d = 15 cm have here been omitted for plot clarity; they are nearly identical to the d = 10 cm results.

true for the towed pipe scenarios with smaller diameters, but as diameter is increased impact volumes for low impact levels can be less than for the bottom manifold (at the spatial scales resolved by this analysis).

The same trends are demonstrated in Figure 5-26, which shows the pH and pCO$_2$ variation along the centerline of the plumes for the same mass loadings. Since the plumes are modeled analytically, the highest concentration is always located at the centerline. It is, however, worth noting that the spatial extent of the plumes differs between scenarios. The horizontal axis of Figure 5-26 is time, i.e., for any point on the centerline the horizontal axis shows the time since that point was co-located with the injection source. As shown in Figure 5-27 and Table 5.4, the axis can be converted to a distance from the source, $x = ut$, where $u$ is the ship speed $u_s$ for the towed pipe scenarios and the ambient current $u_a$ for the
Table 5.4: Plume characteristics of the modeled discharge scenarios. Only those scenarios which achieve a $\Delta p_{CO_2} \geq 0.015$ kPa ($\Delta pH \lesssim -0.1$) are shown.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>CO$_2$ loading (kg/s)</th>
<th>Distance to $\Delta p_{CO_2} = 0.015$ kPa (km)</th>
<th>Final $2\sigma_y$ (km)</th>
<th>Final Plume Height (km)</th>
<th>Area (km$^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stationary hydrate</td>
<td>10</td>
<td>1.06</td>
<td>0.0461</td>
<td>3.00</td>
<td>0.138</td>
</tr>
<tr>
<td>Stationary hydrate</td>
<td>100</td>
<td>8.14</td>
<td>0.458</td>
<td>3.00</td>
<td>1.37</td>
</tr>
<tr>
<td>Stationary hydrate</td>
<td>1,000</td>
<td>61.3</td>
<td>4.57</td>
<td>3.00</td>
<td>13.7</td>
</tr>
<tr>
<td>Towed pipe, 2.5 cm</td>
<td>100</td>
<td>55.0</td>
<td>0.0444</td>
<td>0.514</td>
<td>0.0228</td>
</tr>
<tr>
<td>Towed pipe, 2.5 cm</td>
<td>1,000</td>
<td>479</td>
<td>0.444</td>
<td>0.514</td>
<td>0.228</td>
</tr>
<tr>
<td>Towed pipe, 5 cm</td>
<td>100</td>
<td>18.2</td>
<td>0.0175</td>
<td>1.41</td>
<td>0.0247</td>
</tr>
<tr>
<td>Towed pipe, 5 cm</td>
<td>1,000</td>
<td>192</td>
<td>0.162</td>
<td>1.41</td>
<td>0.228</td>
</tr>
<tr>
<td>Towed pipe, 10 cm</td>
<td>100</td>
<td>6.55</td>
<td>0.00980</td>
<td>3.00</td>
<td>0.0294</td>
</tr>
<tr>
<td>Towed pipe, 10 cm</td>
<td>1,000</td>
<td>94.4</td>
<td>0.0760</td>
<td>3.00</td>
<td>0.228</td>
</tr>
<tr>
<td>Towed pipe, 15 cm</td>
<td>100</td>
<td>6.72</td>
<td>0.00991</td>
<td>3.00</td>
<td>0.0297</td>
</tr>
<tr>
<td>Towed pipe, 15 cm</td>
<td>1,000</td>
<td>94.4</td>
<td>0.0760</td>
<td>3.00</td>
<td>0.228</td>
</tr>
<tr>
<td>Bottom manifold</td>
<td>100</td>
<td>26.8</td>
<td>3.76</td>
<td>0.392</td>
<td>1.48</td>
</tr>
<tr>
<td>Bottom manifold</td>
<td>1,000</td>
<td>362</td>
<td>37.6</td>
<td>0.392</td>
<td>14.7</td>
</tr>
</tbody>
</table>

other scenarios. While the towed pipe scenarios are the shortest in terms of time to reach a $\Delta p_{CO_2}$ of 0.015 kPa, in terms of distance they are longer than the stationary hydrate plume and comparable to the bottom manifold because of the high speed of the ship relative to the ambient current. As a result, the lateral extent of the towed pipe plumes is also much smaller than the other two discharge methods (see Table 5.4).

### 5.4.2 Biological impacts of discharges

A central result of the biological data and modeling analysis is that a subset of the discharge configurations proposed herein yield a prediction of no adverse impact for the region modeled, i.e., outside of the dynamic mixing zone. This conclusion rests on the fact that sufficient dilution is achieved within the dynamic mixing zone to prevent concentrations from exceeding $[\Delta p_{CO_2}]_{min}$, which was set to 0.015 kPa (corresponding approximately to a 0.1 decrease in pH) in an effort to be highly conservative. Since it has been suggested that such a low impact level would likely avoid acute and chronic (sub-lethal) effects alike, the present analysis must conclude “zero” impact. Discussion of the validity and implications
Figure 5-26: Centerline variation in pH (top panels) and pCO\textsubscript{2} (bottom panels) as a function of time from the source for discharge scenarios with mass loadings of 100 kg/s (left panels) and 1,000 kg/s (right panels). Only those scenarios which achieve an impact of ΔpCO\textsubscript{2} ≥ 0.015 kPa are shown: sh = stationary hydrate plume (solid line); tpX = towed pipe with X cm particle diameter (dashdot lines; the larger the particle diameter, the smaller the impact); and bm = bottom manifold (dashed line). [pCO\textsubscript{2}]	extsubscript{min} is shown as a dotted line (ΔpH ≈ −0.1) to indicate the endpoint of the biological impact simulations. Towed pipe results for d = 15 cm have here been omitted for plot clarity; they are nearly identical to the d = 10 cm results.

of this result is deferred to Sections 5.4.4 and 5.5; the remainder of this section summarizes the calculated impacts of each discharge configuration.

Before quantifying the biological impact predicted by the isomortality analysis described in Section 5.3, the nature of these calculations is first illustrated graphically. Figure 5-28 shows the trajectory of an organism cluster traveling along the centerline of the stationary hydrate plume and also the cumulative trajectory which reflects the cumulative exposure predicted by the isomortality method. In the latter case, the horizontal axis is the log of cumulative equivalent exposure time (t\textsuperscript{*} in (5.3)), i.e., the exposure time that embodies cumulative lethal effects had the organism been exposed to a constant pH throughout. As
Figure 5-27: Centerline variation in pH (top panels) and pCO$_2$ (bottom panels) as a function of distance from the source for discharge scenarios with mass loadings of 100 kg/s (left panels) and 1,000 kg/s (right panels). Only those scenarios which achieve an impact of $\Delta$pCO$_2$ $\geq$ 0.015 kPa are shown: sh = stationary hydrate plume (solid line); tpX = towed pipe with X cm particle diameter (dashdot lines; the larger the particle diameter, the smaller the impact); and bm = bottom manifold (dashed line). $[\text{pCO}_2]_{min}$ is shown as a dotted line ($\Delta$pH $\approx$ −0.1) to indicate the endpoint of the biological impact simulations. Towed pipe results for d = 15 cm have here been omitted for plot clarity; they are nearly identical to the d = 10 cm results.

the organism moves along a trajectory that maintains a $\Delta$pCO$_2$ higher than the $[\Delta$pCO$_2]_{min}$, it incurs mortality due to the exposure (the probit function extends to $-\infty$ meaning that even exposures infinitesimally larger than $[\text{pCO}_2]_{min}$ will cause some added mortality). Since the centerline particle is always exposed to the highest plume pH throughout its trajectory, it always experiences the highest mortality of any organism in the plume. Its mortality is entirely due to the predicted plume spreading and the isomortality curve since it experiences no diffusion relative to the plume. Organisms with initial positions not on the centerline experience non-zero relative diffusivity and therefore sample concentrations lower than the centerline as they undergo lateral diffusion.
Figure 5-28: Simulated trajectory of an organism cluster traveling along the centerline of the stationary hydrate plume for a 100 kg/s discharge. The centerline trajectory is a time series of $\Delta pCO_2$ values experienced by the particle. The cumulative trajectory is the centerline trajectory translated to equivalent exposure time via the isomortality method, and thus reflects cumulative exposure. The mortality trajectory shows the mortality incurred by the organism. In this case, the organism cluster incurs about 8.4% mortality before reaching the $[\Delta pCO_2]_{min}$ of 0.015 kPa, after which point no further mortality is incurred.

The predicted mortality of the centerline organism cluster is shown in Figure 5-29 for all discharge configurations and loadings. The reader is reminded that if no mortality is predicted for the centerline organism, then no other organisms experience mortality either. For a 10 kg/s $CO_2$ loading, only the stationary hydrate plume causes any mortality ($<0.1\%$). For 100 kg/s $CO_2$, some mortality is caused by all discharge scenarios, although the levels are less than 0.01% for the towed pipe scenarios. For 1000 kg/s, all discharge configurations yield impact, although the 10 and 15 cm diameter towed pipe scenarios are below 0.1%. For loadings of 100 kg/s or higher, the bottom manifold shows the greatest impact with near complete mortality being achieved for 1,000 kg/s. In each case, the towed pipe scenarios offer the lowest impact, with decreasing impact for increasing hydrate particle diameter. If the $[\Delta pCO_2]_{min}$ is increased from 0.015 kPa to 0.05 kPa to be consistent with the PNEC estimate of Kita and Watanabe ([73], see Section 5.3.3.2.2), then a finding of no impact
would also apply to the 100 kg/s bottom manifold as well as the 10 and 15 cm diameter towed pipe scenarios.

Further insight into the differences in the dilution strategies of the three scenarios can be gained by considering the trajectory of the centerline organism in each case. Figures 5-30 and 5-31 show the trajectories for the bottom manifold and towed pipe (2.5 cm) cases for a 100 kg/s CO$_2$ loading, respectively. The bottom manifold has the lowest initial pCO$_2$, but it persists the longest, reaching a mortality of about 9%. The towed pipe has an order of magnitude higher initial pCO$_2$, but it persists for a much shorter time, reaching a mortality of about 0.0015%. The initial pCO$_2$ for the stationary hydrate plume is yet another order of magnitude higher with a persistence in between that of the other two scenarios, reaching a mortality of 8.4%.

Centerline mortality is a useful indicator of the highest possible impact of a discharge, but it does not provide a measure of predicted impact for the plume as a whole. One such measure is the integrated mortality flux ($Q_M$) as a function of distance downstream ($x$) from the discharge [4, 18]:

$$Q_M(x) = uh \int_{-\infty}^{\infty} D(x, y) dy$$  \hspace{1cm} (5.43)
Figure 5-30: Simulated trajectory of an organism cluster traveling along the centerline of the bottom manifold plume for a 100 kg/s discharge.

Figure 5-31: Simulated trajectory of an organism cluster traveling along the centerline of the towed pipe (2.5 cm diameter) plume for a 100 kg/s discharge.
where \( D(x, y) \) is the fractional mortality, and \( y \) is the lateral coordinate (the centerline is at \( y = 0 \)). Since \( D(y) \) is dimensionless, \( Q_M \) has units \( \left[ \frac{L^3}{T} \right] \) and can therefore be thought of as a flowrate of “dead” water. In the absence of species recovery, this quantity reaches a steady state value at the point where the centerline concentration drops below \( [\Delta p_{CO_2}]_{min} \).

Figure 5-32 shows the integrated mortality flux for each scenario that causes \( \Delta p_{CO_2} \geq 0.015 \) kPa. The same relative trends are noted here, but the separation between the scenarios is multiple orders of magnitude. For example, the mortality flux of the bottom manifold is about 10,000 times greater than the least effective towed pipe scenario \( (d = 2.5 \text{ cm}) \) for a 1,000 kg/s CO\(_2\) loading.

An average fractional mortality for the plume can be calculated by normalizing the integrated mortality flux by the cross sectional area of the plume. The resulting average mortalities are shown in Figure 5-33, where the cross-sectional area is based on a lateral width of \( 2\sigma_y \) of the CO\(_2\) plume at the downstream location where the mortality flux reaches its steady state (see Table 5.4). The same relative trends are again noted, with a maximum average mortality of about 89% being achieved for a 1,000 kg/s loading for the bottom manifold. For a 100 kg/s loading, the maximum average mortality (again for the bottom manifold) is about 4%. The average mortalities are considerably lower for the other scenarios. The towed pipe scenarios yield average mortalities orders of magnitude lower than the bottom manifold and stationary hydrate scenarios (for the same loading).

### 5.4.3 Sensitivity analysis on isomortality function

During the development of the plume calculations detailed in Section 5.3.2, choices were consistently made to be conservative. For example, the observed diffusivity relationship of Okubo [93] was reduced by a factor of ten to reflect deep ocean mixing, vertical diffusion has been ignored, and source representations were crafted to restrict initial plume width to give narrower plumes with higher initial concentrations at the edge of the dynamic mixing zone. While we recognize that more sophisticated simulation of the near field plume formation regime would result in small scale patchiness with higher concentrations near the injection point (e.g., of the type noted by Chen et al. [23]), such impacts would be limited to a small zone. While future work may seek to apply isomortality simulation to high-resolution
Figure 5-32: Integrated mortality of organisms encountering each discharge plume. Only scenarios which cause $\Delta p_{\text{CO}_2} \geq 0.015$ kPa are shown. Note that the towed pipe results for both 10 and 15 cm particle diameters are plotted above, but they are nearly indistinguishable.

Figure 5-33: Average mortality of organisms encountering each discharge plume. Only scenarios which cause $\Delta p_{\text{CO}_2} \geq 0.015$ kPa are shown. Note that the towed pipe results for both 10 and 15 cm particle diameters are plotted above, but they are nearly indistinguishable.
modeling of the mixing zone, this is beyond the scope of the present study. Rather, the mixing zone analogy is maintained and attention is focused on impacts outside of this zone. Because the plume calculations already make a number of conservative assumptions, and because discharge scenarios such as the towed pipe and bottom manifold can be adjusted to further reduce the predicted environmental impact, no additional sensitivity analysis on the plume calculations is offered at this point. The interested reader may, however, refer to Adams and Wannamaker [1] for a discussion of plume sensitivity to current speed, lateral diffusion, and mass loading as their analysis is highly applicable to the plumes considered in the present study.

The sensitivity analysis is instead focused on the isomortality function. As described in Section 5.3.3.2.2, an alternative isomortality function has been developed using the most sensitive subset of the developmental fish dataset (see Figure 5-24) to compensate for a lack of copepod toxicity data for short exposures at high pCO$_2$. The discharge scenarios were simulated again using this additional function in the manner described previously, in which the organism chooses the isomortality function which yields the highest impact at each timestep. To demonstrate the impact of using dual isomortality functions in this manner, the centerline trajectory of an organism in the stationary hydrate plume is shown in Figure 5-34, which can be compared to the previous result in Figure 5-28. In the dual function case, the organism switches from the first (developmental fish) isomortality function to the second (copepod) function at about 0.36 hours into the simulation, where the equivalent cumulative time $t^*$ jumps from 2.5 hours to 16.7 hours. The second function is then used for the rest of the simulation. Here the impact of the more severe isomortality function for short times is substantial; the mortality of the centerline organism goes from 8.4% to 51%.

The centerline mortality and average plume mortality for each discharge scenario are shown in Figures 5-35 and 5-36, where again only those scenarios with $\Delta$pCO$_2 \geq 0.015$ kPa are shown in the latter figure (the other scenarios, as modeled here, have no sensitivity to the isomortality function). The comparison with the base case is summarized in Table 5.5. The stationary hydrate plume shows the greatest absolute increase in the impact parameters because, as noted previously, it causes the highest pCO$_2$ values initially. The bottom manifold shows the least sensitivity because of its low initial pCO$_2$. Most of the stationary
hydrate plume and towed pipe discharges show sensitivity, although the towed pipe impact remains low relative to the stationary hydrate plume. Overall, the bottom manifold impact remains the largest and the towed pipe the smallest.
Figure 5-35: Fractional mortality incurred by an organism traveling down the plume centerline, when the dual isomortality functions in Figure 5-24 are used.

Figure 5-36: Average mortality of organisms encountering each discharge plume, when the dual isomortality functions in Figure 5-24 are used. Only scenarios which cause $\Delta p_{CO_2} \geq 0.015$ kPa are shown.
Table 5.5: Change in predicted centerline and average mortalities (i.e., fractions of organisms killed) when dual isomortality functions are applied.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>CO₂ loading (kg/s)</th>
<th>Change in centerline mortality</th>
<th>Change in average mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stationary hydrate</td>
<td>10</td>
<td>$1.1 \times 10^{-4}$ $\rightarrow$ $2.5 \times 10^{-2}$</td>
<td>$5.1 \times 10^{-6}$ $\rightarrow$ $6.8 \times 10^{-4}$</td>
</tr>
<tr>
<td>Stationary hydrate</td>
<td>100</td>
<td>$8.4 \times 10^{-2}$ $\rightarrow$ $5.1 \times 10^{-1}$</td>
<td>$5.4 \times 10^{-3}$ $\rightarrow$ $1.5 \times 10^{-2}$</td>
</tr>
<tr>
<td>Stationary hydrate</td>
<td>1,000</td>
<td>$8.0 \times 10^{-1}$ $\rightarrow$ $9.7 \times 10^{-1}$</td>
<td>$1.5 \times 10^{-1}$ $\rightarrow$ $1.6 \times 10^{-1}$</td>
</tr>
<tr>
<td>Towed pipe, 2.5 cm</td>
<td>100</td>
<td>$1.5 \times 10^{-5}$ $\rightarrow$ $5.0 \times 10^{-4}$</td>
<td>$1.5 \times 10^{-6}$ $\rightarrow$ $3.5 \times 10^{-5}$</td>
</tr>
<tr>
<td>Towed pipe, 2.5 cm</td>
<td>1,000</td>
<td>$8.9 \times 10^{-2}$ $\rightarrow$ $7.1 \times 10^{-1}$</td>
<td>$4.9 \times 10^{-3}$ $\rightarrow$ $2.1 \times 10^{-2}$</td>
</tr>
<tr>
<td>Towed pipe, 5 cm</td>
<td>100</td>
<td>$9.5 \times 10^{-9}$ $\rightarrow$ $1.2 \times 10^{-7}$</td>
<td>$1.3 \times 10^{-9}$ $\rightarrow$ $1.3 \times 10^{-8}$</td>
</tr>
<tr>
<td>Towed pipe, 5 cm</td>
<td>1,000</td>
<td>$6.8 \times 10^{-3}$ $\rightarrow$ $1.5 \times 10^{-1}$</td>
<td>$4.0 \times 10^{-4}$ $\rightarrow$ $5.8 \times 10^{-3}$</td>
</tr>
<tr>
<td>Towed pipe, 10 cm</td>
<td>100</td>
<td>$1.3 \times 10^{-12}$ $\rightarrow$ $1.5 \times 10^{-10}$</td>
<td>$7.5 \times 10^{-15}$ $\rightarrow$ $1.4 \times 10^{-13}$</td>
</tr>
<tr>
<td>Towed pipe, 10 cm</td>
<td>1,000</td>
<td>$4.1 \times 10^{-4}$ $\rightarrow$ $1.1 \times 10^{-2}$</td>
<td>$3.4 \times 10^{-6}$ $\rightarrow$ $6.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>Towed pipe, 15 cm</td>
<td>100</td>
<td>$1.6 \times 10^{-11}$ $\rightarrow$ $1.7 \times 10^{-11}$</td>
<td>$5.8 \times 10^{-15}$ $\rightarrow$ $1.0 \times 10^{-13}$</td>
</tr>
<tr>
<td>Towed pipe, 15 cm</td>
<td>1,000</td>
<td>$4.6 \times 10^{-4}$ $\rightarrow$ $1.2 \times 10^{-2}$</td>
<td>$3.5 \times 10^{-9}$ $\rightarrow$ $6.7 \times 10^{-8}$</td>
</tr>
<tr>
<td>Bottom manifold</td>
<td>100</td>
<td>$8.9 \times 10^{-2}$ $\rightarrow$ $8.9 \times 10^{-2}$</td>
<td>$4.0 \times 10^{-2}$ $\rightarrow$ $4.0 \times 10^{-2}$</td>
</tr>
<tr>
<td>Bottom manifold</td>
<td>1,000</td>
<td>$9.8 \times 10^{-1}$ $\rightarrow$ $9.8 \times 10^{-1}$</td>
<td>$8.9 \times 10^{-1}$ $\rightarrow$ $8.7 \times 10^{-1}$</td>
</tr>
</tbody>
</table>

5.4.4 Discussion of results

The results of the preceding analysis can be interpreted in two main ways. First, as a means of comparing the various discharge strategies to each other and to those proposed in other studies. Second, as a means to say something about the expected absolute impact of these sequestration schemes. The following section considers each of these, and closes with a brief discussion of the limitations of the present analysis.

5.4.4.1 Relative performance of discharge scenarios

Relative to each other, the preceding results suggest that the towed pipe scenario offers the best performance. As the particle diameter is increased, the predicted impact approaches “zero” for both 100 kg/s and 1,000 kg/s loadings. The second best performer is the stationary hydrate plume on the basis of biological impact, even though it generates volumes with higher levels of pH/pCO₂ perturbation than the other two methods. The bottom manifold exhibits the worst performance as configured here; although it generates lower peak perturbations, it covers a larger area and therefore causes more overall impact.
Conclusions based on these results alone are, however, incomplete. The bottom manifold configuration, which was selected from [1], has more degrees of freedom from a design standpoint than the stationary hydrate plume. The dilution strategy of the manifold is to "paint" as large a region as possible with low concentrations by extending far in the lateral direction and creating a buoyant plume which spreads the CO$_2$ vertically. In contrast, the stationary hydrate plume covers only a small horizontal length scale but takes advantage of the negative buoyancy of the hydrate composite particles to spread the CO$_2$ over the full height of the water column and to dilute it through turbulent entrainment of the surrounding seawater. While the bottom manifold design could be altered to achieve greater dilution by "painting" a larger cross-sectional area (e.g., a longer manifold reaching greater depths with higher plume rise), the stationary hydrate plume does not offer the same flexibility. Although some additional lateral dispersion could be achieved by having a distributed source, (e.g., a platform with many nozzles), it is limited in comparison to the bottom manifold for practical reasons and also because the flow through each nozzle must be large enough to achieve the buoyant plume effect. Caulfield [18] also noted this design flexibility of a bottom manifold approach when considering her droplet plume, in particular looking at the impact of distributing diffuser nozzles in the horizontal as well as the vertical. Thus, in the absence of cost constraints, the length, depth, and spatial loading distribution of the bottom manifold could be configured to reduce environmental impact much the same way the towed pipe particle diameters, reaction efficiencies, and ship speed could be adjusted to reduce impact. Therefore, the lack of design flexibility in the stationary hydrate plume effectively makes it the least attractive option of the three.

5.4.4.2 Comparison to past studies

The discharge scenarios proposed herein seem capable of offering equal or better near-field dilution than comparable strategies in previous studies, some examples of which are reviewed below (only studies which used roughly similar loadings as the present study are considered). Chen et al. [23] studied a 10 x 10 m horizontal platform located 20 m above the seafloor at 878 m with a uniform array of 100 1-kg/s nozzles injecting liquid CO$_2$ droplets with a mean diameter of 0.8 mm. The plume achieved a maximum rise of 190 m but significant interaction
between nozzles apparently caused a descending plume with significant accumulation of CO$_2$ enriched seawater at the seafloor and pH depressions up to 2.6 units. Chen et al. also simulated a towed pipe releasing 100 kg/s of liquid CO$_2$ droplets from an array of nozzles at 1500 m depth moving at 3 m/s. The resulting buoyant droplet plume was reported to dissolve (after 70 minutes) into a passive plume 450 m tall and over 150 m wide, with a maximum pH decrease of 1.7 units. Impact volumes (given in [56]) for the platform release were much larger than those considered in the present study for $\Delta$pH $<$ −1 and slightly higher than the 100 kg/s stationary hydrate plume for $−0.3 < \Delta$pH $< −1$. For the Chen et al. towed pipe scenario, impact volumes are two orders of magnitude larger than the 2.5 cm hydrate particle towed discharge in the present study, which is surprising given the reported extent of the plume. The pH impact volumes predicted by Adams and Wannamaker [1] for their bottom manifold and stationary hydrate plumes are generally smaller than those predicted herein, largely due to the ten-fold decrease in ambient diffusivity employed in the present study as a conservative measure. Employing the same diffusivities, the bottom manifold scenarios are the same, the stationary hydrate plumes are similar ([1] used a shallower release depth which yields slightly greater dilution), and the towed pipe scenarios offer more favorable dilution in the present study due to the increased lateral source width. The impact volumes reported by Caulfield [18] (volume with pH $<$ 7 are reported, see Table 5.1) for the droplet plumes are 0.5-2 orders of magnitude higher than the stationary hydrate and bottom manifold plumes in the present study (Caulfield’s droplet plume scenario is most similar to the bottom manifold scenario considered here). Impact volumes for her towed pipe scenario are comparable to the present 2.5 cm particle towed pipe scenario. This comparison is complicated by the fact that 30% higher loadings were used in [18] (130 kg/s for one plant, 1,300 kg/s for ten plants) along with an order of magnitude higher ambient diffusivity; it is however expected that the present scenarios offer better dilution than the ones considered by Caulfield, save for her dry ice scenario which was dismissed due to cost concerns. The impact volumes for the 25 kg/s stationary droplet release modeled by Drange et al. [36] (5 ports discharging 100 m above the sea-floor in a 0.05 m/s current) are smaller than those of the 100 kg/s stationary hydrate plume considered here. Compared to the 100 kg/s towed pipe 2.5 cm hydrate particle release considered here, the Drange et al. volumes are similar for $\Delta$pH $\geq −0.5$ and larger for $\Delta$pH
Comparisons to the plumes simulated by Sato and Hama [110], Sato and Sato [111] and Sato [108] (a 10-hour 200 kg/s liquid CO$_2$ injection at 2,000 m) and Sato et al. [109] (a 100 kg/s liquid CO$_2$ injection over depth interval 1,000 - 2,000 m moving at speeds of zero and 4 m/s) are difficult in that no impact volumes are reported. Nonetheless, the organism ΔpCO$_2$ experience reported by Sato et al. for the stationary release appears to be greater than the centerline ΔpCO$_2$ for the stationary hydrate plume in the present study, and the ΔpCO$_2$ of the towed pipe scenario of Sato et al. is comparable to the centerline ΔpCO$_2$ of the 2.5 cm diameter towed pipe scenario in the present study (higher for $t < 4$ hours, somewhat lower for $t > 4$ hours). The towed pipe scenario of Minamiura et al. [87] (100 kg/s liquid CO$_2$ discharged through dual horizontal 10-m long diffusers with 50 ports each, towed at 3 m/s) yields a maximum ΔpCO$_2$ of about 600 μatm after 3 hours, which is close to the towed pipe scenario of Sato et al. [109] and the centerline ΔpCO$_2$ of the 2.5 cm towed pipe scenario in the present study. The towed pipe scheme proposed by Hirai et al. [55] and Tsushima et al. [134] (see Section 5.3.1.4) appears to offer the greatest dilution of past studies, with only about $5 \times 10^{-6}$ km$^3$ exceeding a ΔpCO$_2$ of 0.015 kPa for a 100 kg/s loading [134] after about 1 hour; this is far less than the steady state dilution of the 100 kg/s 10 cm towed pipe scenario considered here. Nonetheless, the fact that the ΔpCO$_2$ perturbation predicted by Tsushima et al. [134] becomes negligible after 2 hours suggests that their scheme yields dilution on par with the present towed pipe scenario with 10 or 15 cm diameter hydrate particles.

The predicted biological impact for the towed pipe scenarios and the stationary hydrate plume can be compared to the findings of several of the previous studies listed in Section 5.3.1. In Caulfield [18], a finding of zero impact was reported for her towed pipe scenarios even at loadings of 1,300 kg/s. The present study only predicts zero impact for loadings of 10 kg/s, although the impacts are very small for larger particle diameters. This difference can be attributed to Caulfield’s higher ship speed, more favorable isomortality curve, a source immediately dispersed vertically over 500 m, and more favorable lateral diffusivities. For the droplet plume considered by Caulfield, mortalities reached as high as 11 and 69% for the 130 and 1,300 kg/s cases, respectively, with maximum mortality fluxes of 307 and 27,500 m$^3$/s. The centerline mortalities in the present case are 8.4% (8.9%) and 80% (98%), for
100 and 1,000 kg/s, respectively, with mortality fluxes of 373 (1,890) and 104,000 (417,000) m³/s for the stationary hydrate plume (bottom manifold). Thus, the biological impact of the stationary hydrate plume scenarios considered herein is comparable to Caulfield’s droplet plume despite her more favorable diffusivities, isomortality curve, and diffuser length (which increased with in her study). This is presumably due to the larger vertical extent of the plumes in the present study. Still, the basic conclusion is similar to the one herein, namely that the towed pipe can offer negligible impacts while a fixed plume results in some impact (although some of the impact could perhaps be reduced by optimizing design variables). This is essentially the same conclusion as in Sato et al. [109], which found non-zero but negligible mortality for a ship speed of 2 m/s and a droplet injection distributed vertically over 1,000 m, but “non-trivial” impact when the ship speed was 0. Although Sato and Hama [110], Sato and Sato [111] and Sato [108] found negligible impacts for a 200 kg/s fixed pipe scenario, this conclusion was based on Auerbach’s non-CO₂ isomortality function and their mortality simulations appear to have been truncated before ambient pH was recovered. As discussed in Section 5.3.1.3, Chen et al. [21] found minor biological impact for a stationary discharge using a biological activity approach, but this was for a small loading (1 kg/s). Negligible biological impact was also found using CO₂-induced mortality data (with Sato’s approach) for the 100 kg/s towed pipe scheme of Tsushima et al. [134], which as noted above seems to offer the best dilution of past studies. Lastly, Masuda et al. [83], using an OGCM to predict far-field CO₂ concentrations but ignoring near-field peaks, estimates that a towed pipe discharge of 420, 500, and 270 kg/s in a region of the North Pacific within the approximate depth intervals 1,000 to 1,500 m, 1,500 to 2,000 m, and 2,000 to 2,500 m could avoid exceeding Kita and Watanabe’s [73] PNEC (Predicted No Effect Concentration) of 500 μatm ΔpCO₂.

Thus, overall the present study is consistent with past studies in identifying a towed pipe of some sort as generally being able to avoid significant biological impacts if configured appropriately, and to a greater degree than a fixed descending or ascending plume (e.g., Caulfield et al. [19], Sato et al. [109], Jeong et al. [67]). To the previous assessments we add: (1) the conclusion holds up even considering the most recent CO₂-induced mortality data across a range of species; (2) CO₂ hydrates provide an effective way to achieve greater vertical
and lateral dilution with relatively shallow injection depths, and (3) a bottom manifold can likely also be configured to largely avoid biological impacts (also suggested by [18]), although the design flexibility would likely be limited by cost.

5.4.4.3 Consideration of absolute impact

While the present analysis is a useful way to compare the efficacy of various discharge strategies and to confirm the general findings of previous investigations, the ultimate goal of the investigation is to say something about the absolute impacts of ocean sequestration discharges. Water quality impacts are less controversial in this respect. While the extent of mixing in both the dynamic zone and the passive zone is open to some debate due to the complex fluid mechanics of multi-phase flows and inherent turbulence of the ocean, the chemistry of the carbonate system is well understood and thus bounds can be placed on the expected water quality impacts of a discharge. The challenge of translating the water quality impact into a biological impact is, however, much greater even on an individual species basis, let alone for an entire ocean ecosystem.

The tool employed in the present study is an isomortality-type analysis which is an attempt to integrate acute toxicity data in a mathematically tenable manner such that water quality impacts can be “translated” into a biological impact. Recent research suggests that this may be a reasonably accurate approach for modeling the acute response of zooplankton but not for fish [112] (see discussion in Sections 5.2 and 5.3). Thus, two fundamental approaches can be taken in interpreting the results presented above. First, accepting the isomortality modeling approach as valid and a copepod as an appropriate surrogate for ocean species in general, and second, by attempting to interpret the water quality results directly without the aid of the isomortality analysis. Interpreting the findings of no acute impact for some of the discharge scenarios falls into the second category since it does not rely on adding exposures to come up with a cumulative mortality statistic. Rather, we first attempt to interpret the non-zero mortality results in terms of an absolute impact.

The approach of Sato and Sato [111] and Sato et al. [109] is to consider the mortality of a single test organism and call it significant if it exceeds 0.125% (i.e., a probit value of 2, corresponding to three standard deviations of a tolerance distribution which is assumed to
be Gaussian with the log of the stressor, see Finney [38]). Applying this to the centerline results in Section 5.4.2 implies that there are no impacts for any 10 kg/s discharge, only for the stationary hydrate plume and bottom manifold for a 100 kg/s loading, and for all but the towed pipe with diameter of 10 cm or above for a 1,000 kg/s loading. Using the dual isomortality function sensitivity results: the stationary hydrate plume also yields impact at 10 kg/s; the same conclusions as for the base case apply for 100 kg/s loadings; and at 1,000 kg/s all scenarios yield an impact. Although this criterion could in theory be imposed as a design constraint, its binary nature (impact vs no impact) yields little in the way of overall impact of a plume.

As one step further, Caulfield [18] and Auerbach [4] estimated mortality fluxes and mortality volumes, the latter being the spatial integration in the downstream direction of the mortality flux given by (5.43) to yield a volume of “dead” water. Such a measure is convenient in that it can be normalized by the volume of the ocean, or of some ocean subregion of interest. However, it can only be defined if a species recovery is implemented, e.g., as done by Caulfield and Auerbach via (5.4), because otherwise the volume increases with time indefinitely. As noted earlier, a recovery process is not included in the current calculations because it requires specifying a growth rate which is not well-known and which makes assumptions regarding the response of organisms beyond the considerable ones already made by the isomortality modeling approach. Instead, a different interpretation of the predicted mortality fluxes is offered below.

Simplistically, $Q_M$ can be used to characterize the timescale to kill all the organisms in the ocean, by ignoring all ecosystem effects and regenerative ability of the organisms, i.e., $T = \frac{V_o}{Q_M}$, where $V_o$ is the volume of the ocean. But such a timescale is hardly useful. A more informative interpretation of $Q_M$ can be gained by considering a simple equation governing the balance of a single species of organisms in the ocean:

$$\frac{dN}{dt} = k_g N - k_d N - S_{CO_2}$$  \hspace{1cm} (5.44)

where $N$ is the species population, $k_g$ is the rate of generation or growth, $k_d$ is the rate of death under natural conditions, and $S_{CO_2}$ is the sink of organisms caused by the CO$_2$
discharge. Taking the ambient concentration of organisms to be \( C \), then \( S_{CO_2} = Q_M C \), i.e., \( S_{CO_2} \) has units \( \frac{\text{Num organisms}}{T} \). Perhaps more intuitively, (5.44) can be written in terms of the concentration of organisms:

\[
\frac{dC}{dt} = k_y C - k_d C - S_{CO_2}
\]  

(5.45)

where \( S_{CO_2} \) now has units \( \frac{\text{Num organisms}}{L^3 T} \). In this case, \( S_{CO_2} = \frac{Q_M C}{V} \) where \( V \) is a characteristic volume. Assuming that the concentration of organisms is naturally in a quasi steady state such that \( \frac{dC}{dt} \approx 0 \) and \( k_y \approx k_d \), the impact of the \( CO_2 \) discharge can be characterized by comparing it to the magnitude of \( k_d \), i.e., how much extra mortality can be attributed to the discharge. Thus, the following balance can be considered:

\[
k_d = \frac{Q_M}{V}.
\]  

(5.46)

Taking \( k_d \approx \frac{1}{\tau} \) (where \( \tau \) is the average lifespan of the organism) as a roughly known constant, the magnitude of \( V \) required for the sink due to the \( CO_2 \) discharge to be some fraction of natural death can be calculated. For example, for the \( CO_2 \) discharge to cause a 1% increase in mortality above the natural sink, a critical \( Q_M \) can be defined:

\[
Q_{M,\text{crit}} = 0.01 k_d V_O
\]  

(5.47)

where the \( V \) in (5.46) has been taken as \( V_O \). Since each discharge contributes \( Q_M \), the quantity \( Q_{M,\text{crit}} \) can be translated into a critical number of discharges, \( N_{d,\text{crit}} = \frac{Q_{M,\text{crit}}}{Q_M} \). Based on this analysis alone, one would conclude that introducing less than \( N_{d,\text{crit}} \) of these discharges into the ocean would increase the natural global sink of organisms by less than 1%. Alternately, for a given number of discharges, one could determine the critical volume \( V_{\text{crit}} \) of ocean required to keep the \( CO_2 \) induced sink less than 1% of the natural sink. This latter approach has been used to interpret the previous findings, as shown in Figure 5-37 for the base case isomortality function and in Figure 5-38 for the dual function sensitivity analysis. For context, it has been assumed that there would be about 4,000 100-kg/s discharges, which is based on the Pacala and Socolow [96] estimate of about 175 GtC in avoided emissions over
Figure 5-37: Induced vs natural mortality over various percentages of the ocean volume using the base case isomortality analysis. The analysis assumes 4,000 discharges at 100 kg/s each. Towed pipe results for d = 10 cm and 15 cm are omitted for plot clarity; they are similar to each other and substantially less than the d = 5 cm result.

the next 50 years required to stabilize atmospheric concentrations at 500 ppm. Since the base isomortality function is derived from copepod data, τ was conservatively taken as 1 year as an upper bound estimate on a deep ocean copepod ([3] and references therein). This analysis suggests, for example, that the towed pipe scenario producing 2.5 cm hydrate particles would only cause an increase of 0.0002 to 0.004% in copepod mortality in about 1% of the ocean below 2 km, whereas the bottom manifold would cause about a 3.5% increase in the same volume. Overall, the analysis suggests that significant disruption of the ocean’s copepod population would occur only in a small percentage of the ocean. Of course, the impact may have importance on a regional scale if, for example, a large number of discharges were concentrated in a smaller volume. Furthermore, it suggests that the impact of the discharge is linearly dependent on the average lifespan of the target organism, so if for example the isomortality curves and approach could also be considered realistic for a fish species with a lifespan of 10 years, then its large scale population would be ten times more sensitive to the perturbation caused by the discharges.

The estimates provided by the above analysis are admittedly crude and full of simplifi-
Figure 5-38: Induced vs natural mortality over various percentages of the ocean volume using the dual isomortality function sensitivity analysis. The analysis assumes 4,000 discharges at 100 kg/s each. Towed pipe results for d = 10 cm and 15 cm are omitted for plot clarity; they are similar to each other and substantially less than the d = 5 cm result.

Solutions. For example, they ignore the interaction of multiple sources on each other and the background CO\textsubscript{2} concentration, chronic effects to individual organisms, the relative tolerance of different species, and ecosystem level impacts of the discharge. Nonetheless, the approach gives a crude measure of impact for a given isomortality result.

Referring back to the discussion at the beginning of this section, the other way to interpret the results of the present analysis is to forego the isomortality analysis and interpret the water quality impacts directly. This is attractive because it does not rely on the functional form of an isomortality function and does not require extrapolating toxicity data across species. Such an approach was taken by Barry et al. [8] to determine, in the absence of species specific data, a possible “safe” threshold for avoiding chronic (and acute) biological impacts. Specifically, the pH variability across various zoogeographic regions and bathymetric ranges relevant to ocean sequestration was analyzed. Average pH variability of 0.05 to 0.24 units was observed, from which it was concluded that a pH decrease of 0.1 units may be a reasonably conservative threshold for ecosystem impacts (for reference, if the 175 GtC of required avoided emissions discussed in Section 5.1 were distributed uniformly over the ocean volume below 2,000 m,
the average pH drop would be less than 0.1 \([56, 62]\)). Unwilling to call any level safe, Pörtner et al. [101] concludes that present data indicate that a moderate increase of 200 \(\mu\)atm \((\approx 0.02\text{ kPa})\) may have significant long term effects, i.e., concerning chronic/ecosystem effects rather than acute toxicity. Thus, the value of \([\Delta p\text{CO}_2]_{\text{min}} = 0.015\text{ kPa} (|\Delta p\text{H}|_{\text{max}} = 0.1)\) used in the isomortality analysis seems if anything highly conservative from an acute impact analysis. Since some of the discharge scenarios can meet this strict threshold, overall the present study must conclude that, in the absence of complicating factors such as cost, ocean sequestration schemes can be engineered to largely avoid deleterious environmental impact beyond that which is already occurring under the “business as usual” scenario.

One major caveat to the present analysis of which the reader is reminded is that organism exposure within the dynamic mixing zones of the plumes has not been included in the analysis. Although they did not model hydrate particle plumes or a pure rising droplet plume from a bottom manifold, the detailed fluid mechanics investigations of, for example, Chen et al. [23] and Sato et al. [109] indicate that resolving the dynamic mixing zone will yield small-scale perturbations well in excess of the above threshold. Future work should therefore focus on a combined modeling approach in which near-field active plume mixing and far-field dilution are simulated together by, for instance, coupling a CFD code with a finely resolved ocean general circulation model (e.g., as in Drange et al. [36]), together with an isomortality calculation of the type applied herein. However, it is also noted that because of the design flexibility in the towed pipe hydrate plumes (and the bottom manifold), it seems likely that the proposed discharge scenarios could be further refined as needed to reduce peak concentrations to some extent. For example, hydrate injection nozzles could be distributed laterally at the end of the towed pipe to further increase the initial width of the plume and thereby lessen peak concentrations (analogous to the diffuser used in [134], for example).
5.5 Summary, Conclusions, and Policy Implications

The specific objective of the present study is to provide an updated assessment of the expected acute biological impacts of direct ocean injection. This has been achieved by adapting the methods developed in previous studies (Auerbach et al. [5] and Caulfield et al. [19]) and applying them to new biological data and improved discharge scenarios. An extensive literature review of CO\textsubscript{2}-induced mortality data has been performed, and it clearly demonstrates the need for an update to the initial work of [5] and [19] because (a) a substantial amount of CO\textsubscript{2} toxicity data has since been collected and (b) the sensitivity of marine organisms to CO\textsubscript{2} is greater than the sensitivity to equivalent pH depression by other acids (i.e., the type of data upon which [5] was based). Likewise, advances in, for example, CO\textsubscript{2} hydrate formation techniques have led to the development of new discharge approaches which offer enhanced dilution with less effort. The modeling approaches of the previous studies were updated as deemed appropriate, e.g., the extended probit model of Sato et al. [109] was adopted in favor of the original isomortality approach in [5]. A revised isomortality function has been developed based on pelagic copepod data from the Western North Pacific [140], thus using a similar target organism for quantifying environmental impact as previous studies. The functions were applied to discharge scenarios developed from those proposed in Adams and Wannamaker [1], including a stationary sinking hydrate plume, a towed pipe releasing CO\textsubscript{2} hydrate composite particles, and a rising droplet plume from a bottom manifold. These discharge methods are believed to offer greater or equal dilution than those considered in [18] and subsequent studies. Although the updates to the previous analyses are considerable, the overall conclusion is the same, namely, that ocean discharge scenarios can likely be designed to largely avoid acute impacts. This conclusion is based on two sets of results. First, for some discharge scenarios the peak impact at the edge of the dynamic mixing zone is predicted to be less than a 0.1 unit drop in pH (\(\approx 0.015\) kPa increase in pCO\textsubscript{2}), which is a highly conservative criterion for judging acute impacts since this level has been suggested as a possible “safe” threshold for avoiding chronic/ecosystem effects on the basis of natural pH variability in the deep ocean [8] (and is less than a third of a recent estimate of the Predicted No Effect Level for CO\textsubscript{2} [73]). Second, even when acute impacts are predicted to occur,
discharge scenarios such as the towed pipe yield small impacts relative to the other scenarios and also in an absolute sense by comparison to the expected natural sink of copepods in the ocean. This latter conclusion is, however, crude in the approach used and in the applicability of a copepod as a surrogate for all deep-sea species.

On a practical note, the considerations above coupled with the analysis in the preceding sections suggest that a towed pipe hydrate particle discharge scenario holds the most promise. Although a bottom manifold could in theory be configured to achieve a similar level of dilution, its major drawback is that it is fixed to one location with large up-front capital costs. Site selection would be critical and design alterations would be difficult once constructed. In contrast, towed pipe scenarios would likely have lower up-front costs and are inherently more flexible due to their mobility, allowing the sequestration region to be shifted as necessary to minimize regional hotspots in background pCO$_2$. Although operating costs are expected to be higher [56] than a fixed pipe or platform release, the approach has not been ruled out due to economic infeasibility by past investigators.

The analysis of acute impacts could be strengthened by additional research from the biological community. Additional copepod toxicity data on short-term, high-pCO$_2$ and long-term, low-pCO$_2$ exposure would be helpful to constrain the isomortality function (although as noted in Section 5.3.3.2.2, the latter category may not be accessible through additional data collection). In addition, data on toxicity due to realistic time-variable exposures would be useful in refining and/or confirming the applicability of the isomortality approach to simulate copepod mortality. Equally important would be the collection of toxicity data on a variety of other species in the target depths being considered. Pörtner et al. [101] suggests that the lack of accessibility of such organisms for *in vivo* laboratory analysis could perhaps be remedied by the use of an appropriate model organism such as benthic Antarctic eelpout (*Pachycara brachycephalum*). Such data would need to be reconciled with the data reviewed herein to delineate the applicability of an isomortality-type approach (since fish data do not seem to fit well into this model [112]). New functional models of acute harm could perhaps be developed, e.g., some combination of the isomortality approach and the activity model proposed by Chen *et al.* [21], or perhaps incorporating elements of ongoing work on modeling stress and recovery of fish in thermal plumes [11]. In addition to acute impacts, research to
better understand chronic and ecosystem impact thresholds would be highly useful in this assessment, as discussed later in this section.

Beyond the biological data, additional research into the behavior of the proposed droplet and hydrate particle plumes would be useful to augment the analysis with a consideration of the dynamic mixing zone, both in terms of the resulting shape of the plumes as well as the distribution of excess DIC within the zone. This could partially be addressed by more resolved modeling of the mixing zone but would ultimately require field verification. Also, a thorough characterization of cost constraints to the design options would be necessary to identify economically viable discharge configurations.

Although the focus herein is on acute impacts, we recognize that chronic (sub-lethal) and ecosystem impacts are at least equally important for assessing the viability of ocean sequestration. Chronic impacts to individual organisms (e.g., reduced lifespan or reproduction rate) are important because they can result in intergenerational effects, i.e., population decline of a species over longer timescales. Reduced reproduction rate due to CO$_2$ exposure has been studied for a variety of species, and has been quantified by metrics such as reduced egg production rates, reduced hatching rates, and reduced survival of larvae/nauplii (e.g., [74, 75, 76]). Such data have not been incorporated directly into the analysis because they do not fit well into the isomortality approach used herein. On the one hand, they could be interpreted as acute mortality data in an effort to make the isomortality analysis reflect some intergenerational effects; this approach was taken in the original work by Auerbach [4, 5] where mortality data for adults was shifted in rough accordance with observed declines in reproduction. On the other hand, such a treatment is incomplete because it is not an accurate prediction of the expected population level over many generations, since the equilibrium population is dictated by the balance of many factors of which reproductive rate is only one. Although developmental fish data were used in a sensitivity analysis on the isomortality function for short, high-pCO$_2$ exposure, this treatment was motivated by data gaps in the copepod dataset and was not intended as a compensation for chronic or intergenerational effects. Given the complexity of predicting intergenerational dynamics, the LC$_0$ or LC$_1$ is sometimes adopted as a conservative threshold at which chronic impacts can be expected. The approach taken here has some similarity to this notion in that a $[\Delta$pCO$_2$]$_{\text{min}}$
for the isomortality analysis was set to a very low level (0.015 kPa, based on a $\Delta p\text{H} = -0.1$, or the possibly “safe” threshold identified by [8] to avoid chronic effects). The fact that acute impacts are accrued well below the observed toxicity data and down to a supposed chronic impact threshold does not make the isomortality approach an adequate treatment of chronic impact. However, the fact that at least some of the discharge configurations resulted in impact levels below this threshold suggests that they could be expected to largely avoid chronic impacts to individual species.

Accurate estimation of ecosystem impacts is an even greater challenge. The acute and chronic impacts on population levels are species specific [101, 58], meaning that CO$_2$ injection could cause a shift in the ecological balance of the ocean. Such effects could perhaps be measured in mesocosm experiments, but few such investigations have been done to date and they are complicated by the huge range of species and ecosystems that would need be considered for large scale deployment of direct injection. Long-term observations of population levels due to natural CO$_2$ perturbations could lend some information, but these would not be controlled experiments. Thus, in the absence of adequate data, the “safe” level of a 0.1 pH decrease previously discussed provides an attractive surrogate. Again, the fact that at least some of the discharge scenarios satisfy this constraint at the edge of the dynamic mixing zone suggests that ecosystem effects could be minimized by optimization of these methods.

Given the present state of knowledge, however, the conclusion that ocean discharges can be configured in a way that largely avoids acute and chronic impacts is controversial and subject to a number of substantial caveats. First, the present analysis does not resolve the dynamic mixing zone, where it is very likely that over small distances the 0.1 pH drop threshold would be violated for any practical discharge scenario. While previous studies (e.g., [23, 109]) have modeled the small-scale fluid mechanics near the injection point and found impact levels well above this low threshold, these did not consider the discharge methods proposed herein (descending hydrate particles or rising, non-interacting droplet plumes from a bottom manifold). It is therefore difficult to extrapolate their findings to the present case. While violations within the dynamic mixing zone are likely to occur, we note that the design flexibility offered by the towed pipe hydrate plume and the bottom manifold in
particular suggest that such violations could be limited to small volumes over short durations for moderate loadings of 10-100 kg/s. As noted previously, future research aimed at bridging this gap in the present analysis would be helpful (fluid mechanics modeling, field verification, and design cost estimates).

Another major caveat is the notion that there exists a “safe” threshold at which no impact, be it acute, chronic, or ecosystem, could be expected. Pörtner et al. [101], after an extensive review and discussion of physiological effects across multiple tolerant and intolerant species, warns against this model on the basis that the responses vary dramatically across different types of organisms. While they suggest that the number of organisms likely to suffer from acute CO₂ toxicity is low, the long-term sub-lethal effects on deep-sea fauna may have a significant effect on population structure and species distribution. In particular, the long-term response of calcifying organisms in the surface ocean to ΔpCO₂ of 200 μatm (≈ 0.02 kPa) is provided as an example. It is beyond the scope of the present study to weigh these considerations with the notion that a “safe” threshold can be identified on the basis of natural pH variability, although the latter argument is somewhat persuasive on the basis that the “business as usual” scenario of atmospheric emissions is thought to already have effected a pH change of 0.1 in the upper ocean since the pre-industrial era, and that a further decrease of at least 0.1 and perhaps as high as 0.7 is expected within a century [15, 56].

In light of the above considerations, the overall conclusion is restated in the following manner: if a “safe” threshold can be defined reasonably near the one considered herein, then the present study finds that discharge scenarios could likely be designed and sited to limit violation of the threshold to a small volume. Unfortunately, this leaves a large question unresolved.

If it is assumed that the threshold of a 0.1 decrease in pH (or one like it) can be confirmed, then two important conclusions would logically follow from the present analysis. First, present or near present (for hydrate formation) technology would allow ocean sequestration to meet a substantial part of the required emissions reductions in the short term. The analysis suggests that impacts near the injection points could be minimized by, for example, use of a towed pipe method with large hydrate diameters. If CO₂ can effectively be dispersed over large areas to mostly avoid adverse biological impacts, then the great capacity of the ocean to
act as a sink for CO$_2$ can be exploited. For example, if the ocean is treated as well-mixed, it can be shown that the 175 GtC of required avoided emissions mentioned in Section 5.1 could in theory be stored in the ocean volume below 2,000 m without causing an average pH drop greater than 0.1, which is consistent with other similar calculations [56, 62]. Second, ocean sequestration can only be regarded as a temporary solution to the carbon problem; for any reasonable threshold defined, the storage capacity of the ocean becomes finite if the threshold is to be respected. Ignoring any assimilative capacity on the part of marine organisms over multiple generations, the amount of dilution required of a particular discharge configuration to respect an absolute threshold increases as the background pCO$_2$ increases [62].

From a practical policy standpoint, it is noted that a safe threshold, if one exists, can likely not be confirmed \textit{a priori}. Natural ecosystems are complex and difficult to model due to the interaction of many species of different trophic levels with each other and with their environment. Mild environmental perturbations affecting only a small fraction of species may alter the internal balance of the ecosystem in ways that are subtle and which take long periods of time to manifest themselves on a large scale. Given the escalating sense of urgency to curb atmospheric CO$_2$ emissions that pervades the scientific community, ocean sequestration schemes would have to be enacted on the basis of an incomplete understanding of the biological impacts.

The existence of some uncertainty prior to deployment does not, however, necessarily preclude the use of ocean sequestration as a mitigation strategy. Rather, it places constraints on the implementation of the technology. Widespread deployment would be infeasible until small-scale demonstration projects of substantial duration could be completed to establish the ability of a regional ecosystem to withstand the perturbation, at least on a macroscopic level. Such projects could be implemented with minimal long-term risk since the effectiveness of a dilution strategy is relatively easy to measure and could be determined quickly after discharge is initiated, thus allowing water quality standards based on present-day biological data to be met while the longer term response of the ecosystem is monitored. Maximum flexibility in modifying the discharge configuration and location to find a suitable combination would be important, which further underscores the advantages offered by the towed pipe scenario (even if a stationary hydrate plume were pursued, a ship at rest would be a de-
sirable test apparatus because discharge location could be easily changed). Flexibility would have to be built into the permitting process as well to allow for tweaking of the discharge configuration along the way. Nonetheless, the time to permit a discharge could be lengthy, making the time to large scale deployment lengthy as well.

As a climate change mitigation strategy, ocean sequestration is not presently in favor. Experience has shown opposition to the idea both in the U.S. and in Europe, largely rooted in concern over perceived biological impacts. The effect is highly pronounced in the US, where government funded research into ocean sequestration has dwindled; for example, the National Energy Technology Laboratory’s Program on Carbon Sequestration is no longer actively investigating ocean sequestration [91, 92]. Rather, focus has shifted toward geologic sequestration. Perceived advantages of this storage medium include its lower potential for ecosystem disruption, potentially longer sequestration times, an existing infrastructure with proven economic viability (e.g., enhanced oil recovery projects), and potential for disposal on a regional scale instead of in a global commons (thus reducing the need for multilateral agreements).

Despite the numerous advantages which at present favor geologic sequestration, this strategy is not without problems. First, it is only possible in places with geologic formations favorable for CO$_2$ storage. As an example, a lack of suitable geology appears to be one contributing factor to Japan’s continued research on ocean sequestration [119], as evidenced by the heavy representation of Japanese studies in the recent literature (see Section 5.2). Second, and perhaps most importantly, future leakage of stored CO$_2$ to potable aquifers, to the ocean (if stored beneath the ocean floor), or to the atmosphere is a major concern, meaning that geologic storage implies monitoring indefinitely. As such, the responsibility and risk of adverse impacts of sequestration sites is inherited by future generations. Indeed, in this regard geologic sequestration faces some of the same challenges as nuclear waste storage, albeit to a lesser degree since CO$_2$ is considerably more benign and may be contained by a variety of geochemical trapping mechanisms (see [56]). Unlike ocean sequestration, which can be stopped at any time if it proves undesirable, the consequences of a geologic sequestration can persist long after injection has ceased. Since we, as a society, presumably wish to avoid passing our problems to our descendants, it follows that geologic sequestration is
appropriately regarded as a short-term solution to mitigating climate change, and not as a “blank check” to consume the remaining fossil fuel resources on the planet.

While the present study does not take a stand for or against ocean sequestration, it finds that ocean sequestration should not be dismissed on the basis of environmental impact alone and, in fact, it could be quite benign. No sequestration alternative to “business as usual” is perfect, but they do offer a practical way to curb the buildup of atmospheric CO₂ in the short term while long-term, sustainable energy solutions are developed. Sequestration inherently involves an acceptance of increased risk associated with the storage (e.g., damage to a deep-ocean ecosystem or harm caused by leakage from a geologic formation) so that consequences that are deemed worse can be reduced (e.g., severe climate shifts and the acidification of the surface oceans). In the absence of more definitive evidence of irreparable impacts, no viable sequestration alternatives should be abandoned. As a society, we may need all the options we can muster.
Bibliography


448


452


Chapter 6

Summary and Major Findings

This thesis consists of four separate studies, which are interconnected by two main themes. First, all four studies focused on the use of a Lagrangian framework for simulating mass transport in the environment, with particular emphasis on the application of random walk models in a variety of contexts. Second, two of the studies applied these modeling techniques to the evaluation of deep ocean carbon sequestration by direct injection, i.e., a proposed climate change mitigation strategy in which anthropogenic CO$_2$ is purposefully injected into the deep ocean in order to at least temporarily “sequester” it from the atmosphere. The evaluation considered the remedy from both an efficacy and marine impact standpoint. Taken as a whole, the thesis has sought to provide the reader with a solid introduction to both topics, while developing novel applications and modeling techniques within each.

The first study considered Lagrangian techniques as a means to extend initial mixing calculations beyond the near field of a pollutant discharge. Previous studies have used such techniques either alone or in combination with Eulerian models, where computational efficiency and accuracy are of prime importance in designing “hybrid” approaches. In particular, three relatively simple Lagrangian techniques were characterized in this regard: random walk particle tracking (RWPT), forward Gaussian puff tracking (FGPT), and backward Gaussian puff tracking (BGPT). RWPT is generally the most accurate, capable of handling complexities in the flow field and domain geometry. It is also the most computationally expensive, as a large number of particles are generally required to generate a smooth concentration distribution. FGPT and BGPT offer dramatic savings in computational expense, but their
applicability is limited by accuracy concerns in the presence of spatially variable flow or diffusivity fields or complex no-flux or open boundary conditions. For long simulations, particle and/or puff methods can transition to an Eulerian model if appropriate, since the relative computational expense of Lagrangian methods increases with time for continuous sources. Although the study focused on simple Lagrangian models that are not suitable to all environmental applications, many of the implementation and computational efficiency concerns outlined herein would also be relevant to using higher order particle and puff methods to extend the near field.

The second study was concerned with the performance of RWPT in one-dimensional situations where the diffusivity profile is either a step or piecewise linear function, both of which are common simplifications of physical situations in which diffusivity varies sharply over small spatial scales. Applying RWPT in these cases is challenging because discontinuities in diffusivity or diffusivity gradient can give rise to unphysical particle accumulations. A common approach in these situations is to smooth the diffusivity field and apply a variable timestep as needed, but this has several potential drawbacks including increased computation time, a case-specific and often arbitrary modification of each discontinuity, and the customization of a variable timestepping scheme which may itself introduce error. Consequently, the present study focused on approaches that simulate such profiles without smoothing or variable timestepping. For the step diffusivity case, a correction technique must be used (i.e., timestep reduction does not help), and the present study reviewed, extended, and unified the work of past investigators in this regard, specifically recommending the particle reflection approach of Thomson et al. [3]. For the piecewise linear case, timestep reduction can be used to prevent error but often with a severe increase in computational expense. A novel correction approach was developed for this case, particle reflection with probability translation, which allows larger timesteps to be used with improved accuracy. The approach is generic, i.e., it can be applied to any piecewise linear diffusivity profile. In the test cases studied, the approach allowed computational efficiency to be improved by one to several orders of magnitude relative to an uncorrected constant timestep simulation. Overall, this study provides the modeler with useful tools for handling step and piecewise linear diffusivity profiles in RWPT simulation.
The third study applied the RWPT framework within the context of oceanic transport modeling. The goals of the study were twofold: (1) to create an RWPT implementation that can accurately emulate the tracer transport calculations of an ocean general circulation model (OGCM) using the OGCM’s flow and diffusivity fields, and (2) to demonstrate the utility of the Lagrangian framework in efficiently estimating OGCM residence time statistics. The particular OGCM used was the Lawrence Livermore National Laboratory version of the Modular Ocean Model (LLNL MOM). To achieve the first goal, RWPT equations were developed to mimic the OGCM tracer equation, in which tracer diffusion occurs in the isopycnal and vertical directions. During implementation, the aforementioned particle reflection with probability translation approach for piecewise linear diffusivity was successfully applied to handle the sharply varying vertical diffusivity fields predicted by the Large et al. (1994) KPP mixing scheme. Still, precisely mimicking OGCM calculations proved difficult due to a number of other complicating factors such as specifying the subgrid scale variation of isopycnal slope and diffusivity in steeply sloped or convectively unstable regions, and accurately implementing the Gent-McWilliams eddy-induced transport. As such, further development is recommended to resolve an artificial upwelling problem which occurs in some regions of the ocean (most notably, in the Southern Ocean), resulting in an over-prediction of surface concentrations relative to the OGCM. Nonetheless, the ability of the present RWPT implementation to mimic the OGCM is impressive in many ways and demonstrates good potential for future applications within this context. The second goal of the study, to illustrate the model’s utility, was achieved by using the RWPT model to generate domain-wide estimates of CO$_2$ sequestration efficiency. A novel particle book-keeping method was developed which allows a single particle to be used to diagnose the sequestration efficiency of many OGCM grid cells. The simplified sequestration efficiency metric from Hill et al. [39] was used, i.e., one that assumes a small perturbation to surface pCO$_2$ and ignores “pushback” from the atmosphere. Although the predicted sequestration efficiencies were generally lower than the OGCM’s due to the artificial upwelling problem, a good level of agreement was observed. The global distributions of CO$_2$ sequestration efficiency for releases at 800, 1,500 and 3,000 m were studied and found to be largely consistent with past investigations. Sequestration at 3,000 m is substantially more effective than shallower depths, with a mean efficiency of
44% after 1,000 years and much less spatial variability. For an 800 m injection, deep water formation regions in the North Atlantic and Southern Oceans offer the highest efficiencies, and in general the Atlantic outperforms the Pacific. At 3,000 m, the Pacific outperforms the Atlantic, with efficiencies on par with the deep water formation regions. In addition, the Arctic Ocean generally has high sequestration efficiency. Since the sequestration calculation is but one example of a much larger class of oceanographic investigations involving residence time estimates and source water identification, the model developed herein has potential for a variety of future applications.

The fourth study used RWPT to evaluate the expected impact of several promising schemes for ocean carbon sequestration by direct injection of CO$_2$, and serves as an update to the assessment by Auerbach et al. [1] and Caulfield et al. [2] of water quality impacts and the induced mortality to zooplankton. The study extended the “isomortality” methodology used in these earlier investigations, in particular incorporating the “probit” model framework developed in several more recent studies. A thorough literature review of relevant acute CO$_2$ toxicity data for marine organisms was conducted, and ultimately toxicity datasets for copepods and developmental fish were used in the impact evaluation of marine CO$_2$ discharges. Three promising discharge approaches were considered: a point release of negatively buoyant CO$_2$ hydrate particles from a moving ship; a long, bottom-mounted diffuser discharging buoyant liquid CO$_2$ droplets; and a stationary point release of hydrate particles forming a sinking plume. The discharge plumes were treated in an idealized manner using analytical solutions appropriate for simulating far-field CO$_2$ transport beyond the near-field mixing zone, and the trajectories of passive marine organisms through these plumes were considered using RWPT. Although the analysis would benefit from a more refined treatment of the CO$_2$ discharge plumes, results suggest that it is possible with present technology to engineer discharge configurations which achieve sufficient dilution to largely avoid acute impacts, most notably with the moving ship (“towed pipe”) hydrate discharge method. Sub-lethal and ecosystem effects were discussed qualitatively, but not analyzed quantitatively. Overall, the study suggests that, as a temporary climate change mitigation strategy, ocean carbon sequestration by direct injection should not be dismissed on the basis of environmental impact alone.
Bibliography

