DEDUCING TRACE GAS EMISSIONS
USING AN INVERSE METHOD IN THREE-DIMENSIONAL CHEMICAL TRANSPORT MODELS

by

Dana Elizabeth Hartley

B.S., Chemistry, UCLA (1988)

Submitted to the Department of Earth, Atmospheric and Planetary Sciences in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN ATMOSPHERIC CHEMISTRY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1993

©Massachusetts Institute of Technology 1993
All rights reserved

Signature of Author

Certified by

Ronald G. Prinn
Professor of Atmospheric Chemistry
Thesis Supervisor

Accepted by

Thomas H. Jordan
Chairman, Department of Earth, Atmospheric, and Planetary Sciences
DEDUCING TRACE GAS EMISSIONS USING AN INVERSE METHOD IN THREE-DIMENSIONAL CHEMICAL TRANSPORT MODELS

by

Dana Elizabeth Hartley

Submitted to the Department of Earth, Atmospheric and Planetary Sciences on October 13, 1992 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Atmospheric Chemistry

Abstract

We investigate the feasibility of using an inverse method based on a linear Kalman filter to determine regional surface fluxes through comparisons between observations and predictions in a 3-D atmospheric transport model. The ability of the present ALE/GAGE observation sites to quantify the regional fluxes of anthropogenic trace gases is also studied. These investigations are done using CFC13 as the test tracer since its sources are relatively well known. The first of these investigations is done in the low resolution spectral model of Golombek and Prinn (1986) to enable many feasibility tests to be affordably run. Once convinced that the Kalman filter can deduce regional sources, we test the resolving capabilities of a higher resolution model which would be more promising for actually deducing unknown surface fluxes. For this purpose we used the National Center for Atmospheric Research's (NCAR) Community Climate Model (CCM2).

The tests in the low resolution model are first done with the transport model being perfect in the sense that the "observations" were produced by running the model with the CFC13 emissions derived from industry data. The inverse method used is capable in this case of accurately determining regional surface fluxes using the present ALE/GAGE sites and to converge to the correct solution within a year or two even using initial emission guesses very different from the final solution. We also investigate how well the Kalman filter approach works with a less than perfect chemistry-circulation model by using the ALE/GAGE observations of CFC13 for the inversion. The success of this inversion depends largely on the ability of the model circulation to predict observed concentrations of CFC13 since its chemistry is reasonably well understood. The larger the difference between the model and the observed values using the real (industry) emissions then the larger the bias will be in the estimated emissions. Such studies can help to understand the inherent biases in the model when used in an inverse scheme before trying to use the model to estimate unknown surface fluxes such as those for methane, nitrous oxide and carbon dioxide. We also investigate where additional observational stations could be placed to enhance the capability of the present ALE/GAGE network for determining regional net fluxes. It appears that Hateruma (24N,123E) and to a lesser extent Kamchatka (51N,156E) are very promising locations for new stations and that Hateruma is superior to the ALE/GAGE Oregon station in providing information about Asian sources. This type of analysis can aid the process of choosing observation sites by addressing how well each site contributes to the different goals for use of the data.

These tests have shown the Kalman filter can deduce regional sources provided the model is accurate enough. We have chosen NCAR's CCM2 as a promising candidate. The ability of this model to simulate atmospheric CFC13 at the ALE/GAGE observing sites is tested. The model resolves the high frequency events seen in the data in most cases. However, the phasing is not always correct. Next we perform a series of test
inversions using CFCl₃ in the CCM2. We find that due to the high resolution and nonlinearities of the CCM2 there are new considerations for posing inverse problems in such models. Despite the needed changes for the posing of the Kalman filter the results do suggest that the regional source strengths cannot be constrained using the ALE/GAGE observations in this high resolution transport model using calculated dynamics. This may put into question the use of transport models using calculated dynamics to deduce unknowns from observations of tracers. For such questions, a transport model based on observed dynamics may be better suited. Furthermore, it is also apparent that more and better located observation sites could improve our resolving capabilities.

Thesis Supervisor: Dr. Ronald G. Prinn
Title: Professor of Atmospheric Chemistry
Acknowledgements

There are many people to thank for their efforts and support over the last few years. Most of all I would like to thank my advisor, Ron Prinn, for his advice and encouragement on my research, for teaching me a philosophy of science and an independence in my work, and finally for his friendship.

I would like to thank Alan Plumb for enlightening discussions of inverse problems, and Kerry Emanuel for raising interesting questions about my thesis project. I thank Amram Golombek for teaching me the basics of his model and letting me utilize the model for my research.

I also appreciate the help of many people at NCAR. I especially owe a great deal to Dave Williamson and Jerry Olson, who both worked with me in setting up the CCM2 with CFC-11 and continue to be very helpful with information about the CCM2. I also thank Dave and Jerry as well as Phil Rasch and Guy Brasseur for thoughtful conversations and making my visits to NCAR quite enjoyable. I thank Stacy Walters and Linda Bath for helping me with many computer troubles. I also would like to acknowledge SCD at NCAR and especially Sally Haerer for the use of the Cray Y-MP and their help getting my job through as quickly as possible.

During my trips to NCAR there are many other people who also made my trips enjoyable. I would like to thank Gloria Williamson for fun conversations and noon time workouts; Beth Holland and Dave Schimel for sharing their home with me and for their continued hospitality; Donna Sanerib for arranging my trips; Dave Erickson for afternoon coffee breaks. Recently several of my friends from MIT have relocated to Boulder. I want to thank Peter, Lyn, and Cameron Neilley for sharing their home with me and for their friendship. Josh Wurman and Chris Davis both both helped to make visiting Boulder a lot of fun.

Many people at MIT and in Boston have made living in Boston fun. I want to thank my friend Monique Gaffney for much needed lunch breaks. I also want to thank my friend and weight lifting partner Jeannie Sullivan. My officemates Michele Sprengnether and Bob Boldi kept life entertaining on the 13th floor. I thank both of them for their company and Bob for all his computer advice and assistance. Natalie Mahowald, Anne Slinn, Debbie Sykes, Chun-Cheih Wu, Michael Morgan, Dan Reilly, and many more of the meteorology students added a lot to my stay at MIT.

I would like to dedicate this thesis to my family. To my father, Dan Hartley, who has always been my mentor and role model. To my mother, Barbara White, who always encouraged me to be myself and to do my best. To my step-mother, Linda Hartley, who taught me to appreciate art, good food, good books, and to maintain a balanced life. To my sister, Robyn Davis, for being one of my best friends. To my great-grandmother, who immigrated to the United States from Wales since women were not allowed to get an education in Wales. She went on to get her degree in science and started a legacy of belief in education and the sciences.

And most of all, to my husband, Rob Black who makes everyday of my life special, and who has been my day to day support while finishing this thesis. To all of my family for believing in me more than I did at times.
Chapter 1.: Introduction:

Several long-lived trace gases in the atmosphere are growing in concentration and have very significant radiative and chemical effects. The direct greenhouse forcing by these gases has become an important issue for society as well as science since our climate through this forcing could change in a detrimental way (Dickinson and Cicerone(1986), Ramanathan et al.(1985), Wang et al.(1976), Hansen et al.(1989), Cicerone(1988)). To address this issue we need to determine much more definitively why the levels of these gases are increasing and how much their radiative effects will change regional and global climate. Indeed, current uncertainties in the science have produced much debate about the need for societal action (Mathews(1987), Lindzen(1990), Schlesinger and Jiang(1991), and Risbey et al.(1991)). The trace gases of particular interest for greenhouse forcing are carbon dioxide, methane, nitrous oxide, ozone, and chlorofluorocarbons. The chlorofluorocarbons (CFCs) are also involved in stratospheric ozone depletion (Molina and Rowland(1974), Solomon(1988), Molina(1988)). Scientific understanding of the ways in which the chlorofluorocarbons can destroy ozone, thus, allowing more ultra-violet light to reach the earth's surface, is sufficiently good to have led already to major policies to restrict and phase out these compounds.

The positive temporal trends for the trace gases listed above have been documented by a combination of in situ measurements, flask sampling and satellite remote sensing over the globe (Keeling et al.(1989), Steele et al.(1987), Blake and Rowland(1988), Weiss(1981), Prinn et al.(1990, 1992), Cunnold et al.(1986), Penkett(1988)). But just knowing these gases are increasing in the atmosphere is of course not enough. It is also necessary to know the processes governing their production, circulation, and removal. The surface sources and sinks are often the biggest unknown as well as the relative roles of natural and human (and thus to various extents controllable) activity in these processes.

Currently, the methods for determining surface sources and sinks can be placed in the following two general categories. The first category involves in situ flux measurements with examples for methane being reported by Cicerone et al.(1981,1983), Crutzen et al.(1986), and Seiler et al.(1984a,b). The flux measurements are usually taken at a specific site and then extrapolated to larger scales for example by knowing the area of rice paddies, natural wetlands and ruminant populations in the case of methane. There are attempts to use flux measurements from better understood sources to estimate those from areas not as extensively studied (Bingeman and Crutzen(1987), Kvenvolden(1988)).
These measurements for methane have been compiled in this and other ways to produce high-resolution global data sets for methane emissions (Lerner et al. (1988), Matthews and Fung (1987)).

The second category involves using atmospheric chemical transport models which are supposed to simulate the transport and destruction of the gas and then comparing the predicted concentrations to the observed ones to determine which distribution of the sources and sinks (or more specifically the net surface flux) best simulates the observations. This can be done simply by trial and error, synthesis methods (Fung et al.(1991)) or, more objectively, by inverse methods (Cunnold et al.(1983,1986), Prinn et al.(1990), Enting and Mansbridge(1989)). So far the inverse methods have been applied using 2-dimensional models for the atmosphere. They determine emissions as a function of latitude but provide no insight into the variation of the sources with longitude which for most of the gases of interest is substantial. The inverse method in a 3-dimensional model could, if feasible, obviously be a powerful tool for objectively determining the net surface flux of trace gases as a function of both latitude and longitude.

Three-dimensional atmospheric transport models have been developed in recent years to help understand the behavior of trace chemicals in the atmosphere and to determine their sources, transformations, and sinks. There are currently several different models in use for this purpose. The MPI Hamburg Global Atmospheric Tracer Model (Heimann (1992), Heimann et al (1989)) runs using 12 hour average windfields from ECMWF analyses on a 8° latitude by 10° longitude grid. This model has been used mostly for studies of the cycles of carbon dioxide. The advantage of using observed winds is that events seen in observations can be seen in the model. However, the low resolution weakens some of the events that can occur in tracer transport. There are higher resolution models available. However, they currently use model calculated wind fields. The GISS/Harvard model (Prather(1992)) which has a resolution twice that of the MPI Hamburg model has been used for several different studies including the distributions and cycles of carbon dioxide (Fung et al. (1987)), chlorofluorocarbons (Prather et al. (1987)), and methane (Fung et al. (1991)). Other models have since been developed with an even higher resolution. The Lawrence Livermore National Laboratory GRANTOUR model (Penner (1992)) is a Lagrangian transport model with around 50,000 parcels. This gives a resolution of approximately 3° by 3°. This model has been used to study the sources and distributions of tropospheric odd nitrogen (Penner (1991)). There is also the GFDL model which was one of the first three-dimensional tropospheric chemical transport model (Levy et al.(1979)) and now has a resolution of approximately
2.4° in each direction. This model has also been used to study the odd nitrogen distribution (Levy and Moxim (1989)). Recently, the National Center for Atmospheric Research (NCAR) has developed the latest version of their Community Climate Model (CCM2) (Williamson (1992)) which includes tracer transport. The transport is semi-Lagrangian and on a grid of approximately 2.8° in each direction. At a recent World Climate Research Programme (WCRP) workshop on long-range transport of trace gases all the chemical transport models mentioned above ran a CFC$_1$3 (CFC-11) experiment designed to try to compare the transport processes in each of the models. It was at this workshop that it was apparent that the CCM2 possessed a favorable combination of transport climatology and resolution sufficient to provide perhaps the best predictions for CFC$_1$3 in comparison with observations. This fact and the availability of the model made it an attractive candidate for using to attempt inverse problems.

Some of the first work in inverse modeling for atmospheric chemistry was done in two-dimensional models. ALE/GAGE was the first group to start applying inverse methods by using their data along with a simple 9-box zonal transport model to deduce lifetimes and sources of the various gases they measured (Cunnold et al. (1983, 1986), Prinn et al. (1990, 1992)). ALE/GAGE used the linear Kalman filter as their inverse method which gave an objectively estimated source/sink and uncertainty. Other two dimensional studies that have been done have focused on carbon dioxide surface fluxes. Enting and Mansbridge (1989) used what has become know in atmospheric chemistry as "The Green's Function Method". This method as been extensively analyzed by Newsam and Enting (1990). In this method the basic assumption is that the observations can be explained by a series of basis functions of the source distribution. The coefficients of the series are determined in the model and then the fluxes are determined which minimize the difference between the series and the observations. This method is based on the assumption that both the concentration and surface flux fields are periodic which is not always realistic. Tans et al. (1989) did a trial and error analysis to deduce latitudinal distributions of carbon dioxide surface fluxes. They concluded among other things that the longitudinal variations of the surface fluxes in this case are large enough that in order to get credible source distributions it is necessary to deduce these fluxes in a three-dimensional model.

Newsam and Enting (1988) studied the feasibility of a three-dimensional inversion by looking at the error amplification of applying the "Green's Function Method" to a purely diffusive 3-D model. They find the error to cause damping of information from observations. However, it is difficult to extend their results to the real atmosphere where advection plays a much more important role in transport. Fung et al
(1991) have done a synthesis study for methane. In this work they find linear combinations of different source distributions that give the best fit to the observations. However, in this study they do not find a unique solution.

To summarize to date, the methods that have been applied specifically in atmospheric chemistry are trial and error (Tans et al (1989)), synthesis (Fung et al (1991)), Green's Functions (Enting and Mansbridge (1989)), and the Kalman filter (Cunnold et al (1983)). There is also an approach for inverse problems known as the "Adjoint Equation" method which is more general than the Kalman filter. In atmospheric chemistry, when applying this method to source determinations it becomes essentially a weighted least squared fit between the model and the observations. If the error in the emissions estimate is also minimized and the problem linearized, we then arrive at the Kalman Filter.

We will utilize the linear Kalman filter method in this thesis (see Gelb (1988) for a detailed discussion of the Kalman filter). The Kalman filter does not require any explicit assumptions about the periodicity of fields as is required for the Green's function method utilized by Enting and Mansbridge (1989). The Kalman filter properly applied deduces sources objectively and uniquely (uniqueness was found to be a problem with the trial and error and synthesis methods). Also, it utilizes all available a priori information on emission and emission errors and being linear its implementation is not difficult. For this work the linear Kalman filter is applied for the first time in a 3-dimensional model. This enables us to study net surface fluxes with arbitrary distributions and to utilize a sparse (non-periodic) network of observations. Similar to Newsam and Enting (1988) and Enting and Newsam (1990), we are limited by the number of observing sites. However, since we can define the size and number of regions for net flux determination we can choose the number of regions to coincide with the number of observations to help ensure uniqueness.

The goal of this thesis was a study of the feasibility of resolving a 2-dimensional surface regional net flux distribution using a sparse network of observation sites. We first use a low-resolution spectral 3-dimensional transport model (Golombek and Prinn (1986)) which although insufficient to resolve synoptic events is computationally very efficient. This enables us to run many different feasibility tests which do not depend on the model circulation accurately simulating the observed. Next we use a more realistic high-resolution model (NCAR's CCM2) to see if it has an accurate enough circulation to be used to deduce the (known) emissions of the industrial compound CFCl₃ and thus be applicable to determining unknown emissions for compounds such as methane and nitrous oxide.
In Chapter Two the Kalman filter is derived and discussed in detail. Chapter Three describes the emission proxy that was set up for CFCl₃. This proxy defines what we believe to be the true emissions of CFCl₃. We use this as the test of the inversion and also to test the model's transport. Chapter Four discusses inverse tests in a low-resolution chemical transport model, and Chapter Five discusses the results for the higher resolution model (CCM2).
Chapter 2. The Inverse Method:

2.1 General Theory:

The inverse method we use in this study is the optimal estimation scheme known as the linear Kalman filter (Gelb(1988), Cunnold et al. (1983)). This method estimates the surface fluxes from different regions by making use of measurements of the tracer in the atmosphere and a priori knowledge about the fluxes and their distributions (for example, in situ flux studies could be used to constrain the geographical distribution and provide a first guess of their strengths). The goal is to use a model to determine the flux strengths in defined regions that best reproduces the tracer distribution within the error allowed (measurement and a priori emission uncertainty).

This method works by assuming a piece-wise in time linear relationship between the observed tracer mixing ratios ($\chi^{\text{obs}}$) and the true emission scenario ($E$).

$$\chi^{\text{obs}}_t - \chi^{\text{obs}}_0 = PE_t + \sigma^{\text{obs}}$$  \hspace{1cm} (2.1)

Where $\sigma^{\text{obs}}$ = the error in the measurement

$0_t$ = value for time $t$

$0_0$ = initial value

$\chi^{\text{obs}}_t$ = vector of observed mixing ratios where each element represents a different observation site.

$E_t$ = vector of emissions where each element is the source strength from defined regions emitted over time $t$.

In most cases the mixing ratio measured at some site and the emissions at various locations around the globe are not linearly related. Changes in circulation means that $P$ will change with time. Also, even if averaged over time scales long compared to the circulation, $P$ can still change as sources and sinks change. For example, consider methane, whose main sink is the hydroxyl radical (OH). If we assume OH is produced at a set rate, then the more methane emitted, the more OH is consumed, and thus the less there is to destroy methane. So, the methane ends up increasing at a higher rate due to a reduction of its sink. This is, of course, a simplified example, but it illustrates one way that as the emissions change so can $P$. However, for CFC-11 the feedback between emissions and loss is weak (involving ozone loss affecting CFC-11 photodissociation), and for many other long-lived compounds including methane the relation between concentrations and emissions can be considered linear over short periods. In equation
2.1, \( \chi \) and \( E \) are arrays of values, and \( P \), therefore, is a matrix relating each of the sources to each of the observations. By the assumed linearity in equation (2.1), \( P \) is the following time dependent function:

\[
(P_t)_{ik} = \frac{\partial (\chi_{i}^{\text{obs}})}{\partial (E_j)_{k}}
\]

where \( i \) is the observation site and \( k \) is the source region.

However, we do not know \( P \) for the true atmosphere. It is calculated by using a model that simulates the atmosphere. In this study this means that \( P \) is calculated using a three-dimensional chemical transport model. This is further discussed in section 2.3.

The main tool in the inverse method is a transport model of the atmosphere with the known chemistry of the tracer of interest. If in our model we make a guess at the emissions, we can assume

\[
E = E^* + \sigma_E
\]

(2.2)

where \( E \) = the vector of estimated emissions
\( E^* \) = the vector of true emissions
\( \sigma_E \) = the error in the guess

Furthermore, we would want our next guess to be related to our previous one as well as to the observations. This can be expressed as follows:

\[
E_{t+1} = G'E_t + G(\chi_t^{\text{obs}} - \chi_0^{\text{obs}})
\]

(2.3)

where \( G \) and \( G' \) are matrix operators to be defined later.

Extending equation (2.2) gives

\[
E_{t+1} = E^* + \sigma_{E+}
\]

(2.4a)

\[
E_t = E^* + \sigma_{E-}
\]

(2.4b)

where \( \sigma_{E+} \) = error in estimated emissions at time \( t+1 \)
\( \sigma_{E-} \) = error in estimated emissions at time \( t \)

We can now substitute equations (2.1), (2.4a) and (2.4b) into equation (2.3) to get the following relation for the error.
\[ \sigma_{E+} = (G'+GP-I)E + G'\sigma_{E-} + G\sigma_{obs} \quad (2.5a) \]

where I is the identity matrix. If the error is random (i.e., gaussian), the expectation values \( E[] \) of the errors are as follows:

\[ E[\sigma_{E+}] = E[\sigma_{E-}] = E[\sigma_{obs}] = 0. \]

Therefore, by taking the expectation value of equation (2.5a), we find it is necessary that

\[ G'+GP-I = 0, \]

so

\[ G' = I-GP. \quad (2.5b) \]

Substituting this in for \( G' \) in equation (2.3), we have our update equation for our emission estimate.

\[ E_{t+1} = E_t + G_t((X_{t\text{obs}} - X_{0\text{obs}} - PE_t) \quad (2.6) \]

Finally, using the relation in equation (2.1), we assume the model calculated tracers are related to the emission estimates just as in the real atmosphere, so \( PE_t \) is \( \chi_t^{\text{model}}-\chi_0^{\text{model}} \). This implies the transport model represents the true atmospheric transport, which is a major underlying assumption of the Kalman Filter. So, the update equation becomes

\[ E_{t+1} = E_t + G_t((X_{t\text{obs}} - X_{0\text{obs}}) - (X_{t\text{model}} - X_{0\text{model}})) \]

However, in this study the model is initialized to agree with the observations, so \( X_{0\text{obs}} \) and \( \chi_0^{\text{model}} \) cancel giving

\[ E_{t+1} = E_t + G_t(X_{t\text{obs}} - X_{t\text{model}}) \quad (2.7) \]

It can be seen in (2.7) that what the inverse method strives to do is to continually update \( E \) until the residual between the observations and the model calculations reaches an "optimal level". This "optimal level" is defined as being the estimate of the emissions vector for which the estimated emission least square error is minimized. This is the objective when defining \( G \).

In order to determine \( G \), the emission error needs to be further defined. We will use the relation in (2.5b) and substitute this into equation (2.5a). This gives
\[ \sigma_{E+} = (I - GP)\sigma_{E+} + G\sigma_{\text{obs}} \]  \hspace{1cm} (2.8)

We can define a covariance matrix \(C_t\).

\[ C_{t+1} = E[\sigma_{E+}\sigma_{E+}^T] \]  \hspace{1cm} (2.9)

Where \(T\) represents the transpose of the vector. Using (8) and the following definitions:
- \(E[\sigma_{E}\sigma_{E}^T] = C_t\) -- the covariance matrix
- \(E[\sigma_{\text{obs}}\sigma_{\text{obs}}^T] = N\) -- the noise matrix
- \(E[\sigma_{E}\sigma_{\text{obs}}^T] = E[\sigma_{\text{obs}}\sigma_{E}^T] = 0\) -- emission and observational errors uncorrelated

we find

\[ C_{t+1} = (I-GP)C_t(I-GP)^T + GNG^T \]  \hspace{1cm} (2.10)

Now, we can explicitly define \(G\). We want to optimize \(G\) such that we minimize the trace of \(C_t\). (that is we obtain the least squared error). We first define

\[ J = \text{trace}[C_{t+1}] \]

We want to find \(G\) so that

\[ \frac{\partial J}{\partial G} = 0 \]  \hspace{1cm} (2.11)

Using (2.10) and solving (2.11) for \(G\), we find

\[ G_t = C_tP_t^T[P_tC_tP_t^T + N_t]^{-1} \]  \hspace{1cm} (2.12)

where \(G\) is the so-called Kalman Gain Matrix. From here on it will be referred to simply as the Gain Matrix. Now equation (2.7) is complete, and the emissions can be continually updated. With equation (2.10) the error matrix can be updated as well. However, using equation (2.12), equation (2.10) can be further simplified. Rearranging (2.10) we get

\[ C_{t+1} = C_t - G_tC_t - C_tP_t^TG_t^T + G_t(P_tC_tP_t^T+N_t)G_t^T \]
Making use of (2.12), the last two terms cancel, and we are left with the simpler form

\[ C_{t+1} = C_t - G_tC_tC_t \]  

(2.13)

as our update equation for the emission error.

The Kalman filter can also be derived by adjoint equations which minimize both the weighted least squares of the residuals between model and observations and the emissions update. This involves minimizing the cost function \( J \)

\[ J = (\chi^\text{obs} - F(E_{t+1}))^T N^{-1} (\chi^\text{obs} - F(E_{t+1})) + (E_{t+1} - E_t)^T C_0^{-1} (E_{t+1} - E_t) \]

where the 0 subscript refers to the initial value and \( F(E_{t+1}) \) is \( \chi^\text{model} \) as a function of \( E \) (as before we initialize \( \chi_0^\text{model} = \chi_0^\text{obs} \)). If we linearize \( F(E_{t+1}) \) about \( E_t \) (expand in taylor series and only keep linear terms) then \( F(E_{t+1}) = \chi_t + P_t (E_{t+1} - E_t) \). Now by minimizing \( J \) with respect to the emissions we can derive the Kalman filter (Malchow and Whitney (1977) and Daley (1990)). Thus, the adjoint method for the source determination in atmospheric chemistry can with certain assumptions be reduced to the Kalman filter.

If we wanted to express the "Green's Function Method" in terms of the Kalman filter notation, we would define the cost function

\[ J = \int (\sum_k \sum_i (P_{ik} E_k - \chi_i^\text{obs}))^2 dt' \]

where \( k \) refers to the source region and \( i \) to the observation site (Enting (1985)). \( J \) is minimized with respect to the emissions. The minimization is done once for all time so it does not require running the model through again after generating the \( P \) values. However, this version does not give an estimate of the error in the emissions. That would require further error analysis. Also in this form the equation is not weighted by errors in the data.

The equations used in the Kalman filter for the inversion are (2.7), (2.12), and (2.13). These are the update equations for the emissions, Gain Matrix, and covariance matrix respectively. The way these equations work is not transparent. In order to get a more intuitive feel for what these equations do, it is helpful to consider a simplified case. Consider a situation where there is no error in the measurements. This means that \( \sigma_{\text{obs}} = 0 \), and, therefore, the noise matrix, \( N \), is also zero. Equation (2.12), the update equation for the Gain matrix, becomes
\[ G_t = P_t^{-1} = \frac{\partial E}{\partial \chi_{\text{model}}} \]

Substituting this into equation (2.7), the update equation for the emissions simplifies to the following:

\[ E_{t+1} - E_t = \frac{\partial E}{\partial \chi} (\chi_t^{\text{obs}} - \chi_t^{\text{model}}) \]

or

\[ \Delta E = \frac{\partial E}{\partial \chi} (\Delta \chi) \]

This is now a much more intuitive relationship. It becomes apparent when simplifying the problem, that the inverse method is a very simple relation with flexibility added to account for errors in the data. Furthermore, equation (2.13) now becomes

\[ C_{t+1} = C_t - C_t = 0 \]

So, with no error in the measurements the estimated emissions will have no error.

To see how noise enters the solution, consider the case where the noise matrix is very large (i.e., lots of error in the data). From equation (2.12), the Gain matrix is now very small. Assuming \( G \sim 0 \), we find

\[ \Delta E = 0 \]

and

\[ C_{t+1} = C_t \]

So, the inverse method now does nothing to improve on the emissions estimate. For large errors in the observations, the emissions are changed very little. Basically, the inverse method is operating on the fact that there is not useful information in the observations when the error is very large.

Another important aspect of the Kalman Filter is that it assumes formally that the model circulation and chemistry are perfect. This is obviously not a characteristic of any atmospheric circulation model available today (not even those based on observed winds). However, what this means is that the model used for an inversion will have to be very carefully tested using a tracer like CFCl_3 (whose emissions and chemistry are known) to determine to what extent a less than perfect circulation degrades the ability to obtain the correct solution. It may require that the magnitude of the elements in the Noise Matrix (N) be increased to account at least qualitatively for the difference between model and observed circulation. These increased magnitudes will tend to slow the convergence of
the inverse method, but without it the inversion could prove unstable or erroneous. Obviously, high resolution transport models based on high frequency observed wind fields are potentially the best type of model to be used with inverse methods to determine surface fluxes using trace species observations. In this study we specifically use model generated concentrations as our "pseudo-data" to run our tests, so that the perfect model circulation condition is met. We also however, use real data to explore the effects of a much less than perfect circulation on the inversion process.

2.2 Time Averages

The Kalman filter works by comparing the model calculated mixing ratios to the observed values. In order to do this correctly, it must be decided over what time scales these values are comparable. If the model exactly simulated the atmospheric circulation, then conceivably, the comparison to observations could take place at each time there is data available. In the case of ALE/GAGE, this would mean every two hours. However, most models do not accurately simulate the observed circulation. Thus, we must consider over what time scales the statistics of the model and the observed could be comparable. Furthermore, even with the absolute perfect model, inverting every two hours would not be productive, since even nearby observation sites will not feel the effect of a previous update of sources over that period of time. However, you do not want to average over too long of a time period or the importance of the nearby observation sites diminishes. Thus, there needs to be a balance between a time period in which effects of updates of emissions are felt, a time period over which statistics are comparable between the model and observations, and a time period that is short enough so that the sensitive emissions information from nearby observation sites is not lost.

In the first inverse study for this thesis, we use a low-resolution transport model. For this model, due to the low resolution and spectral nature, it is best to analyze on monthly or longer time scales. Furthermore, we wanted to consider a time frame adequate for air parcels to be transported around the hemisphere. This process requires the order of weeks. So we chose to use one month as our time period for defining concentration averages and updating emissions estimates.

In the higher resolution CCM2 model, we could justify looking at model results at higher frequency. However, when comparing to observations this is inappropriate since the synoptic situations do not occur at the same time in both the model and the real atmosphere. Thus we want to look over at least a full synoptic period (e.g. 7-10 days). Again the patterns are different enough in the model and observations to suggest that at
least a couple of synoptic periods are needed. We, therefore, chose again to look at monthly averages. From the studies, it appears that this may not be long enough of an average for a high-resolution climate mean dynamical model such as the CCM2.

2.3 Partial derivative matrix:

The partial derivative is calculated to relate the mixing ratio at sites to sources over the decided time scale (in this thesis as noted above is one month). The Kalman filter assumes that the mixing ratio at an observing site is linearly (or nearly-linearly) related to the emissions. This would imply that the elements $\partial \chi / \partial E_k$ in matrix $P$ are constant. To address this assumption consider the continuity equation for a one box model

$$\frac{d\chi}{dt} = \frac{E}{M} - \frac{\chi}{\tau}$$

where $\tau$ is the chemical lifetime, and $M$ is the global integrated number density $[M]$. For constant $\tau$ and $E$, this can be solved analytically for $\chi$. If we then differentiate $\chi$ with respect to the emissions $E$ to determine the scalar $P = d\chi/dE$ we get

$$P = \frac{\tau (1-e^{-\tau t})}{M}$$

In this study $t$ is only a few years whereas $\tau$ is approximately 40 years for CFCl$_3$. This means $t/\tau$ is very small and therefore $e^{-\tau t}$ can be expanded to $1-t/\tau$ giving

$$P = \frac{t}{M}$$

Therefore, $P$ is linearly increasing with time. However, when actually using the matrix $P$ in the Kalman filter the time ($t$) of interest in (2.16) is the time $\Delta t$ from the last change in emissions. Thus $t$ in our case is a constant as well since the update of the emissions (however small) always occurs once a month.

To implement this method into a 3-D transport model of the atmosphere there are, however, some special considerations. In the real atmosphere and in our a 3-D model the transport changes with time, so $P$ will vary in time as well. Also, since one purpose of using a 3-D as opposed to a 2-D model is to resolve regions in the same latitude belt it is best to invert on time scales similar to the time (days to weeks) for newly injected tracers to go from sources to the nearest observational sites. While this allows air from sources
to reach nearby observation sites during one inversion iteration, more distant observational sites (for example those in the other hemisphere) do not see the newly injected tracers from those sources until much later. For this reason it is necessary to determine the time lag between tracer injection and first appearance at a remote site and not let $\chi^{obs}$ from the distant observation sites contribute to the updating of the distant sources until this time lag has been exceeded.

In order to determine the time lag, each element of the $P$ matrix must be calculated in time. This is done by running the model for a few years once for each source plus a control run. For the control case all sources are at some set value. Then one source is increased by a finite amount $\Delta E_k$ and the differences $\Delta \chi_i$ between the mixing ratios from these two model runs at the observation sites of interest are used to calculate a column of $P$. Figure 2.1 shows as examples the time dependence of the $P$ matrix elements for the Ireland and Tasmania observing sites corresponding to changes in the emissions from Europe. Each element $P_{ik}$ of $P$ is then fit to a linearly piece-wise function. The first is a horizontal line ($P_{ik} = 0$), and the second is a line with a constant slope $(dP_{ik}/dt)$ starting from the end of the first line. The point of intersection of the two lines is the time lag. Figure 2.1 shows examples of these fits. For the Ireland station which is right next to the European source, there is no time lag. However, for the Tasmania station which is in the southern hemisphere the best fit is found with a time lag of 2.8 months; that is the piece-wise fit for Tasmania has $P_{ik} = 0$ for 2.8 months and then $P_{ik}$ increases linearly with time.

Note that the time lag is only of the order of a few months for sources and sites in different hemispheres. We emphasize that these time lags are not the e-folding times used in defining inter-hemispheric transport. Rather these are times that define how long it takes for the first signal of a tracer injected from a distant source to reach a station. In the 2-D model inverse method used by ALE/GAGE (Cunnold et al(1983)) they compared 12-month running means from the model with observations so this time lag was less of an issue.

$P$ is calculated for the $\Delta t$ from which the last change in emissions took place. To deal with this requirement we calculate the monthly residuals $P_{ik}'(m)$ for each month $m$ for the sloping line fit discussed above to account for month-to-month variations in transport. Then as the model progresses the desired $P_{ik}(m)$ can be calculated as

$$P_{ik}(m) = P_{ik}'(m) + (dP_{ik}/dt) \Delta t$$  \hspace{1cm} (2.17)
Figure 2.1: Time series of the partial derivatives for two ALE/GAGE stations representing different hemispheres based on the emissions from Europe for (a) Ireland and (b) Tasmania. Calculated monthly average P values (Grey), piece wise fit (Black).
For stations with time lags from the source of interest, \( \Delta t \) is the difference between the actual time and the time lag once this difference exceeds zero. For example, consider the case shown in Figure 2.1b. Since inversions occur once a month then 3 months time would have had to pass in order for Tasmania's time lag to be exceeded. At that time \( \Delta t \) would be 3-2.8 or 0.2 months.

2.4 Implementing the Kalman Filter

Once a time scale is specified and the partial derivatives are calculated the Kalman filter can be implemented into the transport model. The first step of the inversion model run is to initialize the model mixing ratios to agree with available observations. Next it is necessary to initialize the runs using a priori estimated emissions from the identified geographic regions. We will also need to define how good the a priori guess is in order to initialize the covariance matrix. This is where the in situ measurements extrapolated to regional scales are very helpful. They not only give a reasonable first guess, but the uncertainty in the measurements and extrapolation methods can be used to assess uncertainty in the first guess. In our test studies where we use the Kalman filter to deduce sources of CFC-11, we simply initialize all the sources as equal and make the initial error the same size as the source. Thus, we make a deliberately very poor first guess.

After the initial mixing ratios and emissions are put in the model, it is run for the one month period. At the end of this month the model calls the filter routine. In the filter routine, the average over the one month period of model mixing ratios from the various observation site locations are subtracted from the observed mixing ratios averaged over the same time period to obtain the residuals. Also at this time, the Gain matrix is calculated. The partial derivative for the month is used along with the latest estimate of the emission errors (covariance matrix) and the variance in the observations (noise matrix, specifically the standard deviation in the monthly mean is used as \( \sigma_{\text{obs}} \)). The Gain matrix is multiplied by the residuals to give the necessary change in emissions to improve the emission estimate. This change in emissions is fed back into the transport model which alters the emissions at the surface accordingly and then runs another month with the new emissions and the process repeats itself.
Chapter 3: The Emissions

There are two primary reasons that a compound with known emissions is needed for this study. The first is to evaluate global 3-D chemical transport models on their ability to represent the real transports of the atmosphere. This is done by using a tracer that can be compared to observations. In order to simulate the tracer, it must have a well known distribution of emissions to input into the model. Secondly, for our studies of the inverse method, using a compound with known sources allows an evaluation of how well the inversion worked. For this purpose, CFCs work very well (Golombek and Prinn (1986), Prather et al. (1987)). Not only are there now 12 years of atmospheric measurements through ALE/GAGE, but the CFCs are simpler than other species for determining emissions. That is, we know they are purely anthropogenic. Biogenic sources are much more difficult to pin down (which is the motivation for doing the inverse problem). Furthermore, the only known loss for CFCl₃ and CF₂Cl₂ is UV dissociation in the stratosphere. This means there are less unknowns in the problem since other chemicals (i.e. tropospheric constituents whose distribution and concentrations are not well known) do not have to be explicitly inputted. For this test we have chosen to work with CFCl₃ rather than CF₂Cl₂ since there is less coming from the Eastern Block, so presumably less error in the values we do have for global CFCl₃ releases (CMA(1992), Cunnold et al. (1986)).

The only known sources for CFCl₃ are anthropogenic and technologically based. The Chemical Manufacturers Association (CMA) puts out an annual report on the global production and release of CFCl₃ by reporting countries. This gives a global total (excluding the releases from what used to be the eastern block), but what is needed is a breakdown of this emission to a surface grid. We chose the basic emissions grid to be that compatible with a trapezoidal 42-wave spectral model (T42); approximately 2.8° by 2.8°. This can then be easily integrated to larger grid scales where necessary. Others have already estimated emissions from larger grids (Golombek and Prinn(1986), Prather et al. (1987), Cunnold et al.(1983)).

There is a lack of industry information regarding detailed regional releases which requires that the emissions at each grid point be based instead on a justifiable proxy. Since CFCl₃ is generally a "high" technology related chemical, electricity consumption is most often used to represent a measure of how much CFCl₃ a country might use relative to another (Golombek and Prinn(1986), Prather et al. (1987)). Electricity consumption
by country is reported in the Britannica Book of The Year (1989). The use of electricity for aluminum production (Metal Statistics, 1988) is subtracted off each relevant country's consumption since this process, while in some places a large user of electricity, does not fit our definition of a "high" technology industry. This changes most countries relatively little, with the exception of places like New Zealand and Norway which are using ≥ 10% of their electricity for the production of aluminum. For our purposes, consumptions by China and the USSR were multiplied by a factor of 0.4 based on a statistical analysis of GNP per capita versus electricity usage for all countries. This is also consistent with Cunnold et al. (1986) who found the eastern block not to be a significant source of CFCl₃. After these adjustments, the percentage of world electricity use Fₖ is calculated for each country. This is used as each country's percentage of world emissions of CFCl₃. Only countries with more than 0.1% of the global emissions are considered. Since our proxy implies non-zero emissions from the USSR and China, the implied emissions from these countries are added to the CMA global release numbers. The modified total global emissions are given in Table 3.1 for 1985 to 1988.

Table 3.1 Global emissions of CFCl₃. Units are 10⁶kg/year.

<table>
<thead>
<tr>
<th>Year</th>
<th>Global Emissions CMA</th>
<th>Global Emissions (this paper)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985</td>
<td>280.8</td>
<td>311.9</td>
</tr>
<tr>
<td>1986</td>
<td>295.1</td>
<td>327.8</td>
</tr>
<tr>
<td>1987</td>
<td>310.6</td>
<td>345.0</td>
</tr>
<tr>
<td>1988</td>
<td>314.5</td>
<td>349.3</td>
</tr>
</tbody>
</table>

Within each country, these emissions are divided among land grid points based on the population fᵢ occupying the grid square as determined from The Whole Earth Atlas (1986). The percent of global emissions at each grid point Sᵢj is calculated as

\[
\frac{F_k f_j}{\Sigma_i f_i} = S_{jk} \quad (2)
\]

where the summation is over the grid points within the country of interest. This then gives the high resolution (T42) percentage emission map shown in Figure 3.1.
Figure 3.1: CFCl₃ emissions on a T42 grid based on electrical usage proxy (see text). Shades denote the log(\% global emissions at grid point).
Chapter 4. Inversion Studies in a Low Resolution Chemical Transport Model

4.1 The Model:

To determine the ability to resolve regional sources using a 3-D spectral atmospheric model we chose to work with a computationally fast but low resolution model since it suffices for the required feasibility tests and also enables multiple runs of several years duration. This model (Golombek and Prinn (1986)) has a horizontal resolution of 11.5 degrees latitude by 22.5 degrees longitude. In the vertical there are 26 levels from the surface to approximately 72 km but only the first 40 km are relevant here. CFCl$_3$ is used as the tracer of interest since its sources are known probably better than any other trace gas. The prediction equation for a CFCl$_3$ is

$$\frac{\partial \chi}{\partial t} = -(k \times \nabla \psi) \cdot \nabla \chi - W \partial <\chi>/\partial Z - J[M]\chi - k[O(1D)][M]\chi + E + \frac{1}{(H^2P)}(K \partial \chi/\partial Z)/\partial Z - (1/P)\partial (P<W\chi>)/\partial Z$$  \hspace{1cm} (4.1)

where

- $\chi$ = volume mixing ratio of CFCl$_3$ ([CFCl$_3$]/[M])
- $[M]$ = total molecular number density
- $\psi$ = horizontal stream function
- $W$ = vertical velocity (dZ/dt)
- $Z = -\ln P$
- $P$ = ratio of pressure to surface pressure (1000 mb)
- $H$ = standard atmospheric scale height ($\sim$ 7 km)
- $J$ = photodissociation rate constant calculated in the model
- $k$ = unit vector in the vertical
- $k$ = rate constant for reaction of CFCl$_3$ with O(1D)
- $K$ = vertical eddy-diffusion coefficient
- $<>$ = horizontal average
- $E$ = $\chi$ tendency due to surface source in lowest layers

The photodissociation reaction is the most significant loss process for CFCl$_3$; however, loss due to reaction with O(1D) is also included in equation (4.1) but it is < 2% of the CFCl$_3$ loss. In Golombek and Prinn (1986) the model and its strengths and limitations for tropospheric studies are described in detail. They illustrate in that paper that the large scale circulation in the model is sufficiently accurate to reproduce long-term variations of
the observed mixing ratios of CFCl₃ and CF₂Cl₂ at clean-air sites over the globe. As expected the model is insufficient for study of synoptic-scale phenomena such as pollution events. Because synoptic events are not resolved we will use it to look at averages over time no shorter than a few weeks (specifically monthly averages will be used in this study). Unlike Golombek and Prinn (1986) we use model output from the nearest grid to each ALE/GAGE observing station rather than a spectral transformation to the exact station location. This model can be integrated for one species for one year in approximately eight minutes on the NASA/GSFC Cray Y-MP.

The model is initialized for CFCl₃ for January 1987 with a zonal distribution. The latitudinal gradient is based on a fit to the ALE/GAGE observations for that month. The vertical profile is from a previously spun-up run (Golombek and Prinn (1986)) renormalized to the January 1987 surface values. Thus, at the start of our runs there is close agreement between the model and the observations. Any differences as time proceeds will be assumed to be due to emission (not chemical loss) differences. It can be shown simply using a one box model of the atmosphere that the inverse procedure is quite sensitive to the initial difference between the model and the observations since these differences reflect an accumulated error in model emissions prior to the time of available observations.

4.2 The Sources

For use in the low resolution Golombek and Prinn (1986) model the emissions discussed in Chapter 3 are integrated onto the larger grid (Figure 4.1a). For the purposes of our study we define five geographical source regions (figure 4.1b). When estimating emissions from these five regions the distribution of emissions within each region is maintained as shown in Figure 4.1a. In Figure 4.1b the locations of the nearest model grid point to each of the five ALE/GAGE observation sites are marked by X's. The five stations are Macehead, Ireland (51N, 10W), Cape Meares, Oregon (45N, 124W), Ragged Point, Barbados (13N, 59W), Point Matatula, American Samoa (14S, 171W), and Cape Grim, Tasmania (41S, 145E).

4.3 The Results:

We first ran the model forward for several years using the emissions for CFCl₃ derived from industry data (Figure 4.1a). The output from this run at the five
Figure 4.1: a) The pattern of the percent of global emissions from each grid point within the source regions based on an integration of the emissions shown in Figure 3.1 b) The five source regions (shaded) and ALE/GAGE observation sites (X)
ALE/GAGE observation sites is defined as our "pseudo-data" and used to define the $\chi^{\text{obs}}$ vector in equation (2.7).

The problem is then approached as if the magnitudes of the sources in the five regions in Figure 4.1b are not known. To begin the inverse runs the emissions from the five regions are initially set equal (which is deliberately a very poor guess) and the global total emission is initially set to be the same as used for the "pseudo-data". However, we do not subsequently constrain the global total emission during the inversion. The a priori emission error (in matrix $C$) is set equal to the emission strengths themselves thus giving a factor of two a priori uncertainty. Although this is overly pessimistic for CFCl$_3$, it is appropriate for the more biogenic trace gases such as methane and nitrous oxide. The inversion is then done once a month (of model integrated time) using the monthly averages of both $\chi^{\text{obs}}$ and $\chi^{\text{model}}$. After each inversion the emissions $E$ are updated as well as the estimated error (in $C$). Figure 4.2 shows the progression of the predicted source strengths and their uncertainties through a two year integration. As can be seen the derived source strengths converge rather rapidly to the true emissions used to generate $\chi^{\text{obs}}$ despite the poor first guesses used for these emissions. The dashed lines corresponding to the error in the estimated emissions also converge as the inversion proceeds with convergence being greatest for Europe and the Southern Hemisphere. For example, the European source which started out at $65 \pm 65$ ends up $133 \pm 20$.

It is of interest to note that the North American region converges least rapidly toward the true solution. This may seem surprising considering there is an observing site (Oregon) on the west coast of this source region. Europe in contrast converges quickly despite the fact that it also has an observing site on its western margin (Ireland). The differences are due to a combination of station location and circulation. Both observing sites are climatologically upwind of these source regions but the Oregon station is a significant distance from the maximum North American source grid points (see Figure 4.1) whereas the Ireland station in this model is located at one of the maximum European source grid points. Figure 4.3 shows the 667 mb monthly mean streamlines in the 3-D model for February, May, August, and November. The locations of the nearest grid points to the ALE/GAGE stations are shown as black squares in this Figure. The winds consistently lack a strong easterly component at the Oregon observation site. However, at the Ireland station there are times (e.g. August) when the flow has an easterly component thereby carrying European air over this station. The effect of this circulation is in fact more dramatic for the true location of the Ireland station. To illustrate this further, Figure 4.4 shows the time series of $\partial \chi / \partial E_k$ for Ireland and Oregon for a changing North American source. The evidence of the change in the North American source is
Figure 4.2: Time progression of the estimated source strengths at each source during an inversion run. The dark solid line represents the industry based emissions used to generate the "pseudo-data". The solid grey line is the estimated source strengths from the inversion process initialized with all the sources equal. The dashed grey lines are the estimated range of error in the estimated emissions.
Figure 4.3: Streamlines for 667 mb from the Golombek and Prinn (1986) low-resolution spectral model for (a) February and (b) May. The black squares are the grid point locations nearest to the five ALE/GAGE observing stations and the two possible new stations (refer to text).
Figure 4.3 continued: Streamlines for 667 mb from the Golombek and Prinn (1986) low-resolution spectral model for (c) August and (d) November. The black squares are the grid point locations nearest to the five ALE/GAGE observing stations and the two possible new stations (refer to text).
Figure 4.4: Time series of the partial derivatives for two ALE/GAGE stations in the northern hemisphere related to the North American emissions (a) Ireland and (b) Oregon. Calculated monthly average P values (Grey), piece wise fit (Black).
Figure 4.5: Time progression of the residuals corresponding to the estimated emission time series in Figure 4.2. Also shown (vertical axis on the right) is the corresponding partial derivative relating the model calculated value to the source(s) shown in the right hand corner of each frame.
first seen at Ireland. In this model the Oregon station does not see air from the major 
Eastern North American source until it has circled back around the hemisphere. This 
shows up in Figure 4.4 since the effect of the North American source on Oregon is nearly 
zero after the first month. In Figure 2.1 it was clear that the Ireland station does not have 
this time lag for the European source.

Besides North America, it is evident from Figure 4.2 that convergence for the 
Asian and Japanese/Korean source regions is somewhat slower than for the other regions. 
Of course, the fact that they do eventually converge onto the true values in the two years 
is very promising. Figure 4.5 illustrates some of the key variables in the inversion 
process. The lighter grey line is the residual between the pseudo-data and the calculated 
mixing ratio using the estimated emissions. These values are shown for each observation 
site. The dashed lines bracket the standard deviation of the pseudo-data. The darker 
lines are some of the partial derivatives (defined in the upper right hand corner). In 
considering the Japanese and Asian sources, Oregon is the largest affected in the model 
of the five observation sites. We show the partial derivatives for Oregon due to both 
Japan and Asia in Figure 4.5. Looking at this graph, it is evident that the influence from 
Japan is greater (the partial derivative from Japan is the darker line). Thus, although the 
convergence is slower for both Japan and Asia than the other sources, we would expect 
Japan to converge faster of the two, and this is what we see. Furthermore, in Figure 4.1 it 
is apparent that the sensitivity of the ALE/GAGE network is expected to be less for these 
regions since there is not an observation site near them. Since the Asian and 
Japanese/Korean regions have substantial emissions of many trace gases (e.g. CH$_4$ from 
rice agriculture) it is therefore of considerable interest to ask where a station should be 
located to resolve sources in these regions.

For this purpose we chose first to move the Oregon station (which is in fact no 
longer in operation) to Kamchatka (51N,156E) which is a peninsula off Russia just 
Northeast of Japan. The nearest grid point to this location is shown in Figure 4.3. We 
ran the same inversion procedure as above, only this time rather than using the grid point 
closest to the Oregon station we used the output from the grid point closest to Kamchatka 
to define $\chi^{\text{obs}}$ using the true sources, to determine $\partial \chi / \partial E_k$ values, and to define $\chi^{\text{model}}$ 
during the inverse procedure. The results from this inversion are shown in Figure 4.6.
Comparing these results with Figure 4.2 it is apparent that the predicted Japanese source 
strength now converges much faster to the true value. This is expected due to the 
proximity of Kamchatka to Japan. Furthermore, Figure 4.3 illustrates that there is 
occasionally direct flow from Japan to the Kamchatka grid point. This is especially 
evident in February. However, the predicted Asian source strength converges even
Figure 4.6: Same as Figure 4.2 but the output from Kamchatka rather than from Oregon is used as "pseudo-data".
Using Hateruma for Inversion

Europe

Correct = black
Inverted = grey
Error = dashed

Asia

Japan/Korea

North America

Southern Hemisphere

Million Kg/Year


Figure 4.7: Same as Figure 4.2 but the output from Hateruma rather than from Oregon is used as "pseudo-data".
slower to the true value than when the station was located in Oregon. Looking at the streamlines in Figure 4.3, it is evident that there is not generally a transport of Asian air masses in this model to Kamchatka. So the air from Asia must travel around the hemisphere to reach Kamchatka, whereas it reaches the Oregon station after only travelling over the Pacific. Also note that the North American region now converges to the solution more rapidly than it did when using the Oregon location. This follows from the previous discussion of the North American source. Kamchatka is west of Oregon and therefore sees the effect of the Eastern North American source sooner than Oregon (but later than Ireland).

As seen above, Kamchatka helps resolve the Japanese/Korean source but it does not help with Asian sources. Recently the Japanese government has proposed opening a baseline station in Hateruma (24N, 123E) an island just east of Taiwan. To investigate this site we next replaced the Oregon station with Hateruma and repeated the inversion process to give the results shown in Figure 4.7. Now all of the sources converge rather well. Both Asia and Japan are resolved by the Hateruma location. This is also presumably a very promising location for deducing methane emissions since it is in some months located downwind from the most concentrated Asian region of rice paddies.

The above results show that a sparse (specifically ALE/GAGE) observational network can be used with the linear Kalman filter and a 3-D atmospheric chemistry-transport model to determine regional source strengths when the circulation is exact. However, when inverting using real observational data to determine sources the model circulation will no longer exactly match the real circulation that produced the observational data. An important question is how realistic does the model circulation need to be in order to obtain useful results? We address this question using our low resolution model with its imperfect circulation and the actual ALE/GAGE CFC13 observations. Figure 4.8 shows a time series of the model monthly averages obtained for CFC13 at the five ALE/GAGE stations using the emissions derived from industry data. Also shown are actual ALE/GAGE monthly average observations (with pollution episodes removed) at those locations and the standard deviations for both time series. There is a tendency for the model to underpredict concentrations in the northern hemisphere and overpredict them in the southern hemisphere due to a somewhat too rapid model interhemispheric circulation. Note that the model initiation process ensures that the model and observations both agree in January, 1987.

To address the effect of the differences in the above two time series on inversions, it is again assumed that the magnitudes of the emissions in the five source regions are unknown and as a first guess these magnitudes are set equal. Figure 4.9 shows the
Figure 4.7: Monthly average concentrations and standard deviations from the model run (grey) with the industry based emissions and the ALE/GAGE data (black) with pollution removed for CFCl₃. Units are in parts per trillion by number (ppt).
Figure 4.9: Same as Figure 4.2 only now the real ALE/GAGE data was used as the observations in the inversion process.
Figure 4.10: Time progression of the residuals corresponding to the estimated emission time series in Figure 4.9. Also shown (vertical axis on the right) is the corresponding partial derivative relating the model calculated value to the source(s) shown in the right hand corner of each frame.
progression of the estimated source strengths as the inversion proceeds. The predicted emissions differ significantly from the industry values and the progression is very noisy and appears unstable. This behavior can be understood in part by considering Figure 4.8. Evidently, in order for the linear Kalman filter to get the model to agree with the observations it has to make large adjustments to the emissions leading to instability and ends up replacing emissions by sinks in the southern hemisphere in order to lower the concentrations there. Figure 4.10 shows the diagnostics for the inversion with observational data. By observing the lighter gray solid line it is quite evident that the model adjustments for the residuals are based on a model that does not adequately simulate the atmospheric dynamics. The partial derivatives define how a change in emissions will effect an observation site. If this is different than the way the real atmosphere feels this effect than we can expect to see unstable behavior. One way to deal with instability is to add to the Noise Matrix (Equation 2.12) such that the standard deviation of the data is large enough to include the model predicted values so that they are not statistically different from one another. This can be rationalized in our case as a way to include in the inversion method the uncertainty in $\chi^\text{model}$ caused by the imperfect model circulation. To do this we specifically calculate the average difference between the time series at each station in Figure 4.8 and add this to the corresponding elements of the Noise Matrix. The results of doing the inversion with this added noise is shown in Figure 4.11. The instability evident in Figure 4.9 is no longer present but the derived emissions are still unrealistic due to the imperfect circulation. Furthermore, with the added noise the error in the estimated emissions does not improve much over the two years. Now there is difficulty constraining the sources since our data, having greater uncertainty, contains much less information.

So, how good does the model circulation have to be? As a final investigation, we defined a new "pseudo-data" set by forming a hybrid of the model predictions and actual observations. To form the hybrid we took the average of the ALE/GAGE data and the model values from Figure 4.8 and then continued to take the average between the previously calculated hybrid and the model until the hybrid values were within $1\sigma$ of the model values (with $1\sigma$ defined as the standard deviation of the actual ALE/GAGE data). We ran the inversion process again using these final hybrid values as the "pseudo-data" to define $\chi^\text{obs}$ with the $\sigma$'s of the ALE/GAGE data used to define $N$. The results are shown in Figure 4.12. This time the estimated emissions agree with the industry-based values within the derived error in these estimates. The exception is the southern hemisphere which has a very small estimated emission error due to the very small standard deviations in the Southern Hemisphere data (note on Figure 4.8 that the
Figure 4.11: Same as Figure 4.9 with the noise matrix increased to include the average difference between the data and the model simulation shown in Figure 4.8
Figure 4.12: Same as Figure 4.9 except that a hybrid of the model predictions and ALE/GAGE observations was used as "pseudo-data" at the sites (see text).
concentration scale for the southern hemispheric stations is about half that of the northern hemispheric stations). It appears that for the inversion to be successful the model must be capable of predicting concentrations using the true emissions to within $1\sigma$ or better of the observations. Thus it is clear that when doing an inverse problem in a model, it is imperative to test the model with a tracer with well known sources and chemistry in order to determine how well the model predictions agree with the data. This should be accompanied by a test inversion to establish the expected biases in any estimated source strengths for a less understood tracer.

4.5 Summary

We have investigated the feasibility of using an inverse method based on a linear Kalman filter to determine regional trace gas source strengths through comparisons between observations and predictions in a 3-D atmospheric transport model. The inverse method was successful when the transport model used was perfect. This perfection was achieved because the "observations" were produced by running the model with the true trace gas source strengths. We also showed in this case that the location of the present ALE/GAGE observation sites were sufficient for resolving the primary source regions for CFCl$_3$. The inverse method used is specifically capable of accurately determining regional emissions using the present ALE/GAGE sites and to converge to the correct solution within a year or two. The transport model used in this study is a low resolution model which may underestimate the promise of the inverse method in a perfect circulation model since the large grid areas mean more irreversible mixing and thus loss of information.

The question of possible locations for future observational sites to most enhance the capability for determining regional source strengths was also addressed. It appears that Hateruma and to a lesser extent Kamchatka are very promising locations and that Hateruma is superior to Oregon in providing information about Asian sources. This type of analysis can obviously greatly aid the process of choosing observation sites by addressing how well each site contributes to the different goals for use of the data (in this case to determine sources).

We also addressed the question of how well the Kalman filter approach works with a less than perfect chemistry-circulation model. We find that it is necessary for model predictions of a tracer like CFCl$_3$ (whose sources and chemistry are reasonably well understood) to agree well (specifically within $1\sigma$) with the observed. The larger the
difference between the model and the observed, the more unstable the inversion and the larger the bias in the estimated source strengths. A test inversion of the model using the real CFCl₃ data can help to understand the biases in the model before trying to estimate unknown source strengths such as those for methane, nitrous oxide and carbon dioxide. By doing these tests the dependence of the final determined source strengths on the model circulation and chemistry can be better assessed.
Chapter 5. Using the CCM2 for Inverse Studies

As discussed in the previous chapter, the transport model is an integral part of the inverse method. Therefore, the model should be carefully evaluated. The high-resolution transport model we are working with is NCAR's CCM2 which has been under development over the last two years. In the next two sections, the general characteristics of the model and the CFC-11 experiment done to evaluate the model transport are described. The first experiment was completed in the fall of 1991. At that time the CCM2 was in a version that will be referred to as CCM2x (Holtslag and Boville(1991)). The latest version we will refer to as CCM2. In the fall of 1991 a four year simulation was completed using the CCM2x with CFC-11. In the first half of this chapter that run will be the focus. Recently, a two year simulation was done to begin the inversion work in the CCM2 which is discussed later in this chapter. This later simulation is shown along with the CCM2x time series in order to expedite comparisons.

5.1 The CCM2:

We are using the T42 grid, which has a resolution of approximately 2.8x2.8° on a Gaussian transform grid. There are 18 hybrid levels (sigma and pressure) in the vertical with approximately six levels in the stratosphere and twelve in the troposphere. The nominal levels are 992.5, 970.4, 929.3, 866.4, 786.5, 695.2, 598.2, 501.3, 409.0, 324.8, 251.2, 189.2, 138.7 99.0, 63.9, 32.6, 13.1, and 4.8mb. The model uses a 20 minute time step by dynamically adjusting the spectral resolution of the top layer to maintain a courant number of less than one.

The transport code is a monotonic semi-Lagrangian scheme (SLT, Rasch and Williamson (1989)). The arrival point of a parcel is always assumed to be the Gaussian grid points, and then the departure point is determined by back trajectory. The tracer concentration at the departure point is determined by a shape preserving interpolating scheme (Rasch and Williamson (1990)) based on the grid concentrations at the previous time. To augment the vertical transport there is a stability dependent mass flux representation of moist convection developed by J. Hack (Williamson et al.(1992)). The scheme predicts a mass flux of chemical constituents for both shallow and deep penetrative convection. There is also a parameterization of the boundary layer which diagnoses the PBL height and mixes the chemical constituents as well as the dynamical
variables (Holtslag and Bolville (1991)). The later parameterization plays the dominant role of these two in the vertical transportation of tracers to the free troposphere.

The photodissociation for CFCl₃ is based on the J rates calculated in the Golombek and Prinn (1986) model and used in the previous chapter. The J values are interpolated onto the CCM2 grid once a month. However, there are very few stratospheric levels in the CCM2.

5.2 Initialization/Spinup:

CFCl₃ is used as a diagnostic tracer to test the transport of the CCM2. Besides just knowing the emissions, the CFCl₃ distribution must be initialized in the model. The better this is done the less time it takes the model to adjust. When initializing a 3-D model it is important to consider the trace constituent that is being used. The desired goal is to run the model to a point where the model results are no longer dependent on the initial conditions, and then to begin the desired experiments. One option is to simply run the model for many years to allow it to settle down before starting any experiments with it (Prather et al. (1987)). However, this is a very expensive option. This is especially true of tracers such as CFCl₃ which have long lifetimes (40 years for CFCl₃). CFCl₃ is destroyed in the upper stratosphere. In fact it has a very short lifetime once it is up there. However, to transport the CFCl₃ up can take several years, so it can take a long time to adjust in the model.

Another method which was chosen for this study is to use two vertical profiles with the same latitudinal gradient and run the model until the two profiles are the same (Golombek and Prinn, 1986). The ALE/GAGE data gives a latitudinal gradient and other sources provide some localized vertical profiles (Fabian et al. (1981a,b)). These vertical profiles can then be normalized for the surface values based on the latitudinal gradient. The latitudinal gradient is a linear fit to the ALE/GAGE observations. The initialization is then zonal. Figure 5.1 shows the vertical profiles measured as well as the two profiles chosen for this work. The solid lines are the lower and upper profiles (cfcl and cfch respectively). As the model runs cfch will be quickly adjusted as the destruction is greater in the upper levels. Cfcl on the otherhand will rely move heavily on transport. The model is considered independent of the initial conditions once these two profiles start to converge.

The diagnostic used to observe the progress of the profiles is the global lifetime $\tau$. This is calculated by the following equation:
Initial Profiles of CFC-11

Figure 5.1: Vertical profiles of CFCl₃ (CFC-11) for both data and the two initial model profiles normalized to the surface data values.
Figure 5.2: Time progression of the lifetime $\tau$ in the CCM2 during two years of spinning up.
\[ \tau = - \frac{M}{(dM/dt)_{\text{chem}}} \]

where \( M \) is the global mass of CFCl₃, and \( dM/dt_{\text{chem}} \) is the change in \( M \) due to photodissociation in the stratosphere. Not only is it desired that these two profiles start to converge, but also that they start to converge on the lifetime that is found from the data. Figure 5.2 shows the progression of the lifetime value during two years of spin-up in the CCM2 for the two profiles shown in Figure 5.1. The final lifetime is around 50 years which is higher than that found by ALE/GAGE by about 10 years. This is most likely due to the small number of model stratospheric levels in the CCM2.

The model was initialized for January 1, 1986. The grid point values were renormalized at each time step to the tropospheric mass of January 1, 1986. This process was repeated during the two year spinup time until the two vertical profiles converged as shown in Figure 5.2.

### 5.3 Comparison to data:

The comparison to observations is the crucial aspect of this diagnostic (Hartley and Prinn(1991, 1992), Williamson et al. (1992)). ALE/GAGE has measurements taken every few hours (ALE every 4 hours, GAGE every 2 hours). Ideally, the model results should be written out at the same frequency as the observations have been taken. However, the amount of storage space necessary to save the model results every couple of hours integration time would have been overwhelming. Thus, the model values are saved once every twelve hours. The observations have been averaged into 12 hour periods for comparison. These comparisons will help us to evaluate how well the transport model does represent the real transport.

After the spin up process, the CCM2x was run for four years with the CFCl₃ tracer. Figure 5.3 shows the time series from the CCM2x simulation, the more recent CCM2 simulation and the observations at the five ALE/GAGE stations shown as twelve hour averages for the data and the more recent CCM2 results, and twelve hour instantaneous values for the CCM2x results. Figure 5.3a shows the results for Macehead, Ireland (52N,10W). These values are shifted down by 100ppt for the CCM2x results and 200ppt for the CCM2 results for clarity of the high frequency events. This station is polluted about one third of the time (Prinn et al. (1983)). Most of the time this station receives westerlies carrying background air off the ocean. The flat background is evident in the observations, CCM2x and the CCM2 results. At times the wind direction changes carrying the polluted air off of Europe. In Figure 5.3a these events stand out strongly.
Figure 5.3: (a) Ireland time series of CFCl₃ twelve hour averages from the CCM2 and the ALE/GAGE data and twelve hour instantaneous values from the CCM2x. The model results have been shifted down by the amount in the legend to allow for an easier visual comparison.
Figure 5.3: (b) Oregon time series of CFCl₃ twelve hour averages from the CCM2 and the ALE/GAGE data and twelve hour instantaneous values from the CCM2x. The model results have been shifted down by the amount in the legend to allow for an easier visual comparison.
Figure 5.3: (c) Barbados time series of CFCl$_3$ twelve hour averages from the CCM2 and the ALE/GAGE data and twelve hour instantaneous values from the CCM2x. The model results have been shifted down by the amount in the legend to allow for an easier visual comparison.
Figure 5.3: (d) Samoa time series of CFCl$_3$ twelve hour averages from the CCM2 and the ALE/GAGE data and twelve hour instantaneous values from the CCM2x. The model results have been shifted down by the amount in the legend to allow for an easier visual comparison.
Figure 5.3: (e) Tasmania time series of CFCl$_3$ twelve hour averages from the CCM2 and the ALE/GAGE data and twelve hour instantaneous values from the CCM2x. The model results have been shifted down by the amount in the legend to allow for an easier visual comparison.
against the background values of air coming off the ocean. Thus, these events can be used as flags for the dynamics.

Comparing these events in the model to those seen in the observations it is evident that the pollution events in the model tend to be weaker than those seen in the observations. However, the 1987 observations seem to have anomalously high pollution events as they are much stronger than those in 1988 or 1989. Comparisons of the later years does not imply as dramatic a difference between the CCM2x and the data. Weaker events might be expected just on the basis of model resolution. Although the T42 grid is small enough to resolve synoptic scale events, air masses must be somewhat diluted to fit on the grid. An example of the synoptic pattern during these events is shown in Figure 5.4. Figure 5.4b shows the wind vectors inducing the pollution event on October 2, 1987 in the observations. It is evident that the air off of Europe is being funnelled directly over Ireland. Figure 5.4a shows an example of one of the pollution events in the CCM2x. We can see the air flow is more southerly than easterly for this event than is seen for a similar event in the observations. Thus, we also might expect weaker events in the CCM2x.

Besides just the strength of these events the phase is also important. The ALE/GAGE observations for Macehead show the major events occurring in the winter and late spring. In the CCM2x the dominant events are in the summer time. This brings into question what dynamics are occurring in the model that cause this seasonal difference. Figure 5.5 shows wind roses for the model and observations (ECWMF). Figure 5.5a and 5.5b are the model and the observations respectively at the location of the Ireland station. It is immediately evident that the model does not have easterlies as often as the actual Ireland station during the winter. Figure 5.5c shows the model wind rose about 15 degrees further north. Now we find the wind rose looks much more like that of the observations at 51N. This in fact leads to the conclusion the storm track is shifted northward in the CCM2x. However, we cannot simply use the model output at this higher latitude for our analysis since the sources are at the lower latitudes. Figure 5.6 shows the time series from 65N in the CCM2x along with that at 51N and the observations. As might be expected the pollution events are even smaller since the major sources are not to the east. But the seasonal pattern does look more like the observed. The time series from the latest CCM2 shown in Figure 5.3a, suggest better seasonal agreement.

Figure 5.3b shows the CCM2x results and the observations from the Cape Meares, Oregon ALE/GAGE station (45N, 124W). Opposite from Ireland, the CCM2x now has events that are too intense and seem to drown out a minimum in the winter time.
Figure 5.4: Wind Vectors for 850mb from the CCM2x (September 21, 1987) and the ECMWF analyzed winds (October 2, 1987)
Figure 5.5: Wind roses showing the percentage of time that the wind blows from each direction (pie slices) and the concentrations of CFC13 that are associated with each wind direction (dotted lines) for a) the CCM2x results at Ireland (52N,9W), b) the observed at Ireland (52N,9W), and c) the CCM2x at a higher latitude (66N,9W)
Ireland CFCl₃

Figure 5.6: (a) Ireland time series of CFCl₃ for the CCM2x results at 51N and 65N versus the ALE/GAGE data.
Figure 5.7: Oregon time series of CFCl₃ twelve hour averages from ALE/GAGE data and the CCM2x using the improved emission scenario.
seen in the observations. The dynamics are very similar in the CCM2x and the observed in this location. Therefore, we addressed the sensitivity to the sources to see if there was a local source causing this based on the noisiness of the baseline during these events. By adjusting the population factor for Oregon to a lower estimate (based on a reassessment of the Atlas values) and rerunning the CCM2x with the new emissions for three month periods in both winter and summer, we find that these events disappear. Figure 5.7 shows a time series for one of these winter periods for Oregon, and the pollution events are now much more similar to the magnitude of the observational events. There will obviously be a problem in defining a detailed source distribution on a proxy. The two year simulation with the more recent CCM2 used this improvement in the source distribution. Thus, in the CCM2 time series for that simulation there are not overly large pollution events.

Figure 5.3c shows the time series at Barbados (13N,59W). The ALE/GAGE observations show several rather intense but very short pollution events that are not seen in the CCM2x. The mechanism of these events is due to local inversions occurring on scales of a few kilometers. This is much smaller than the resolution of the model. Thus, we would not expect it to simulate these events. The short decreases in mixing ratios during the fall in the observations are due to influxes of Southern Hemispheric air to this station which is a larger scale process the model should simulate. Indeed, these events are evident in the CCM2x time series. They tend to start earlier in the year in the CCM2x and occur somewhat more frequently. However, the fact that they occur in the model suggests that the dynamical processes in this region of the tropics is reasonably simulated. Figure 5.8 shows the time series for 1988 with the wind components divided into quadrants. Figure 5.8a shows the results for the CCM2x and Figure 5.8b for the observed. In both cases it is apparent that there is a strong easterly component year round. During the time of the decreases there are southerly components. However in the observed these are difficult to extract from the major easterly flow. In the model these southerlies are stronger and more frequent. The more recent CCM2 time series looks very similar to that from CCM2x.

The time series at Samoa is shown in Figure 5.3d. The seasonal patterns in the CCM2x and the observations look very similar. In the Northern Hemispheric winter there is an influx of Northern Hemispheric air to the Southern Hemisphere which is evident in the Samoa observations. Note that in the data this is a time of greater variability in the time series. In the CCM2x this increase due to Northern Hemispheric air is less pronounced, and there is not as striking of a seasonality to the variability in the
Figure 5.8: Time series for CFC-11 from a) CCM2x and b) observed at Barbados with the wind divided into quadrants above for the year 1988.
Figure 5.9: Tasmania time series of CFC\textsubscript{13} twelve hour averages from ALE/GAGE data and the CCM2x using the improved emission scenario.
time series. Again the more recent CCM2 time series at Samoa is very similar to the CCM2x results.

Finally, in Figure 5.3e the time series for Cape Grim, Tasmania is shown. It is immediately evident that pollution events are essentially absent in the CCM2x. Again we address the emissions distribution. After further analysis into the Australian CFC13 emissions it seemed more reasonable to account for all the emissions from Melbourne and Sydney rather than the whole of Australia. This is rerun along with Oregon as described above. The results from a simulation with these new emissions is shown in Figure 5.9. Now there are pollution events in the CCM2x. However, they occur much too often, do not have the characteristic frequencies seen in the observations, and they do not show the seasonality which the observations have with the events occurring predominantly in the boreal winter. The new version of CCM2 used this improvement in the sources as well. Thus, there are many pollution events. We do see some agreement with this version of the CCM2 seasonality of events and that of the observed.

Besides just the high frequency events a lot can be learned from longer time scale phenomena as well. Figure 5.10 shows the monthly average values for both the CCM2x and ALE/GAGE data with pollution episodes removed. Even with the pollution events removed it is obvious that the seasonality of these events affects the seasonal cycle of the background. First note that in Ireland the CCM2x and the data have opposing cycles. This also corresponds to the model having the strongest pollution events during the summer rather than the winter like the data shows. A similar phenomena is seen in the Oregon monthly averages. The data has two minima a year, but in the model the first minima is drowned out by the pollution in the winter which also increases the baseline (Figure 5.3b).

The other most notable feature is the different trends. The linear fits to the data and the CCM2x results are shown in Table 5.1. The trend in the CCM2x time series is always greater than that of the data. The error in the trends (approximately 10%) is of about the right order to be explained by the larger lifetime found in the model (50 years versus 40).
Figure 5.10: Monthly average time series of CFCl₃ from the CCM2x and the ALE/GAGE data with pollution episodes removed.
Table 5.1:

<table>
<thead>
<tr>
<th>Station</th>
<th>CCM2 trend (ppt/yr)</th>
<th>ALE/GAGE trend (ppt/yr)</th>
<th>%error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ireland</td>
<td>10.4</td>
<td>9.0</td>
<td>15.</td>
</tr>
<tr>
<td>Oregon</td>
<td>10.5</td>
<td>10.1</td>
<td>4.0</td>
</tr>
<tr>
<td>Barbados</td>
<td>10.7</td>
<td>9.1</td>
<td>17.6</td>
</tr>
<tr>
<td>Samoa</td>
<td>10.9</td>
<td>10.0</td>
<td>9.0</td>
</tr>
<tr>
<td>Tasmania</td>
<td>10.7</td>
<td>9.9</td>
<td>8.1</td>
</tr>
</tbody>
</table>

5.4 Inverse Studies

Since we have found that it is feasible to deduce regional source strengths using a low resolution model, we can now apply this method in a more realistic high-resolution transport model namely the CCM2. For the inverse studies in the CCM2 we choose to invert once a month since the high frequency analyses of the CCM2 in comparison to the ALE/GAGE observations do not justify doing smaller time averages. Figure 5.11 shows the monthly averages for CFC-11 with pollution events included in both the observations and the latest CCM2 which is the one to be used in the inversions. As can be seen in Figure 5.11 there is rather good agreement between the model and the observations on monthly time scales within the error bars. However, there are also large standard deviations in the data which will increase the noise matrix in the inversion.

The inversions in the CCM2 were set up similarly to the inversions done in the low resolution model described in Chapter 4. The source regions are defined in the same locations. These are shown for the CCM2 in Figure 5.12. We do a series of eight inversion tests summarized in Table 5.2. For the first series of tests we used the data and "pseudo-data" including the pollution events. Similar to the inversions in the low resolution model, we do not include month to month correlations in the noise matrix. The change in the estimated emissions that is allowed at each inversion step is constrained to be no larger than the initial emissions of the source region. This constraint would be much improved upon by a priori information in an inversion to deduce unknown sources. Here we are trying to address very general situations. In the first series we discuss below the weighting of grid points within a source region was done in such a way that caused Australia to be underweighted (and South America and South
Figure 5.11: Monthly average concentrations and standard deviations from the CCM2(grey) with the industry based emissions and the ALE/GAGE data(black) with pollution CFCl₃.
Figure 5.12: Source regions for test inversions in the CCM2
Africa to be correspondingly overestimated) compared to Chapter 3 emissions for the S.H. source region. This is corrected in the later runs discussed, and is found to have little effect.

Table 5.2

<table>
<thead>
<tr>
<th>INVERSION TEST</th>
<th>OBSERVATIONS</th>
<th>INITIAL GUESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>Time series from P-matrix calculation with all sources equal at ALE/GAGE observation sites</td>
<td>Best estimate of five sources</td>
</tr>
<tr>
<td>Pseudo-Data 1</td>
<td>Time series shown in Figure 5.11 from the latest CCM2 at ALE/GAGE sites</td>
<td>All equal</td>
</tr>
<tr>
<td>Pseudo-Data 2</td>
<td>Time series shown in Figure 5.11 from the latest CCM2 at ALE/GAGE sites except Oregon which is substituted by Hateruma</td>
<td>All equal</td>
</tr>
<tr>
<td>Pseudo-Data 3</td>
<td>Time series shown in Figure 5.11 from the latest CCM2 at ALE/GAGE sites except Oregon which is substituted by Kamchatka</td>
<td>All equal</td>
</tr>
<tr>
<td>Pseudo-Data 4</td>
<td>Time series shown in Figure 5.11 from the latest CCM2 at ALE/GAGE sites except Oregon which is substituted by Hong Kong</td>
<td>All equal</td>
</tr>
<tr>
<td>Data</td>
<td>ALE/GAGE Observational data</td>
<td>All equal</td>
</tr>
<tr>
<td>Data + Noise</td>
<td>ALE/GAGE Observational data with added error in the noise matrix to account for differences between the observed values and the model simulated values using the best estimate of the sources</td>
<td>All equal</td>
</tr>
<tr>
<td>Hybrid</td>
<td>The average of the model time series with best sources and the observed time series</td>
<td>All equal</td>
</tr>
</tbody>
</table>
The control test is the complete perfect model test. The "observations" are from the base source scenario used to calculate the partial derivative matrix (Chapter 2). This is the only test that is really a fictional compound since the correct solution is that all the sources are the same. The results from this test are shown in Figure 5.13. The results are noisier than the similar test done in the low resolution model. The high resolution model is behaving differently. Rather than converging, both Japan and the Southern Hemisphere are over adjusted during the Northern Hemispheric summer and then the sources do not change much after that. However, the Southern Hemisphere does show signs of readjusting the following summer. Asia does not adjust at all. In the low resolution model it converged, but slowly. To understand this behavior we have to consider both the nature of the higher resolution transport and the Kalman filter. Consider first the situation in Asia. In the low resolution model the tracer leaves a source and is immediately mixed into a large volume. By the time air from Asia reaches the US (i.e., Oregon) it is well spread out. However, in the high resolution model air parcels maintain more integrity and directionality, so if the air does not head toward Oregon then Oregon naturally senses very little of the change in emissions. In the low resolution model the effects of the Japanese and the Asian source on Oregon are similar as seen in Figure 4.5. However, this is not true in the high resolution model. Figure 5.14 shows various diagnostic information about the model and the Kalman filter for each of the observing sites. The lightest shade of grey solid line is the residual between the "observations" and the model. The light dashed pair of lines bracket the $\sigma_{\text{obs}}$. The darker thinner lines illustrate the values of key partial derivatives (right hand vertical axis). For Oregon two partial derivatives are shown. These are for the sources in Japan and Asia. Japan is the dark thin line, and Asia is the medium shaded line (close to zero). It is apparent comparing these two that Asia has little influence on Oregon, and thus on all the observation sites since Oregon is by far the most influenced. Thus, we can understand why there is essentially no adjustment made to the Asian source.

Using Figure 5.14 we can also address the over-adjustment issues for the Southern Hemisphere and Japan. The key to understanding this is to look at the Kalman filter. In most cases using atmospheric data the Noise matrix is much larger than PCPT which means that the Gain matrix is essentially $\text{CPT}/N$. This is especially true of these runs where we include the pollution events. Thus, large partial derivatives mean large Gain matrix and therefore, large changes in emissions. Thus, we compare the residual to the partial derivative. Oregon is where Japan's influence is most strongly felt. As can be seen in Figure 5.14, the partial derivative has a strong maximum in the Northern Hemisphere summer. It just happens to correspond also to a large residual. This causes
Figure 5.13: Time progression of the estimated source strengths at each source region during the control inversion run of the CCM2 (refer to text).
Figure 5.14: Time progression of the residuals corresponding to the estimated emission time series in Figure 5.13. Also shown (vertical axis on the right) is the corresponding partial derivative relating the model calculated value to the source(s) shown in the right hand corner of each frame.
the emissions to be over-adjusted and since the partial derivative then decreases strongly there is no adjustment in the emissions after that period. Similarly Tasmania and Samoa have the same correspondence between a large partial derivative and a large residual. The following Northern Hemisphere summer we see the Southern Hemisphere source starting to readjust back. Due to the strong seasonality of the partial derivative in the high resolution model we may find these long induced oscillations always occurring but presumably with decreasing amplitude as time progresses. In the low resolution model, the partial derivatives have very little seasonality as seen in Figure 4.5. Thus, these problems were avoided in that model.

Notice that the large partial derivative relating the mixing ratio at Ireland to the European source also has this correspondence. However, it also corresponds to very large N which damps the effect in the Gain matrix.

There are four "pseudo-data" tests that we ran. In each of these the "observations" are from a run using the CCM2 and the FULL emissions (shown in Chapter 3) including the corrections in these emissions discussed in section 5.3. For the inversion we are only considering five regions to be sources while the pseudo-data we use has the full global distribution of sources. The assumption here is that the areas not defined as part of the five source regions are in fact very small sources as compared to the others. This is true for CFC-11. The first of these tests is made using the model generated monthly average mixing ratios at the five ALE/GAGE stations as the observations. Each of the other three tests substitutes the ALE/GAGE Oregon station with the model generated mixing ratios at a possible new observation site. Similar to the low resolution studies in Chapter 3 we address the usefulness of stations at Kamchatka and Hateruma. This time we also address a new station which is currently under negotiation (R. Newell, private communication, 1992) in Hong Kong.

The results from these four pseudo data runs are shown in Figures 5.15 through 5.22. In each of these there are similar oscillations as were seen in the control inversion. Using the stations closer to the far east we are better able to constrain the Asian and Japanese sources. Similar to the studies in the low resolution model, Hateruma improves the constraint on both Asia and Japan. However, Europe is now less constrained. It seems Hong Kong and Hateruma are far enough south to not feel much effect from Europe. Hong Kong and Hateruma behave very similarly in these studies which is not surprising since on the T42 grid they are only spaced a couple of grid points apart. Kamchatka is able to resolve Japan, but it does not contain much information about the Asian source. Figure 5.20 shows the partial derivative at Kamchatka due to both Asia and Japan. The Asian partial derivative is hard to see since it is right along the zero line!
Thus, this station gains us information about Japan, but not Asia similar to what we found in the low resolution model. An observation site at either Hateruma or Hong Kong appears promising from these model runs for deducing Asian and Japanese emissions.

We also included studies with real ALE/GAGE CFC-11 data in this set of experiments. We did a series of three tests in this category. The first is just a straight inversion using the real data. The next is again using the real data, but the error in the real data is augmented by the average difference between the model with the "correct" sources and the observed as shown in Figure 5.11. Finally, another test is performed using an average between the model and the observed real values as the "hybrid" data set. The noise used in this case is the original noise in the real observations. This hybrid data set is less than 1σ away from the model generated values (using the "correct" sources).

Similar to the low resolution model, we find that when comparing the model calculated values to real observed values, we get very noisy results. The inversion results are shown in Figure 5.23 and the corresponding residuals and partial derivatives are in Figure 5.24. Looking at Figure 5.24 the consequence of the model transport being different from the real circulation which produced the observational data is apparent. The partial derivatives correspond to the model transport not to the real atmosphere. Thus, adjustments occur at the specific times that the model transport allows the sites to see the source not the when actual transport would. To attempt to adjust to the observations, the source changes that are made induce large spurious changes in the model mixing ratios. Recall Figure 5.11 compares the model with the "correct" sources to the real observations. Although the agreement does seem reasonable, there are different oscillations in the observed and model calculated values particularly at the Northern Hemisphere stations.

Adding additional noise does help the convergence of the source determination as can be seen in Figure 5.25. We needed to add much less noise in this case than we did in the low resolution model, so we do not lose as much information. That is the emission uncertainties do decrease. The two sources that are not correctly converging - Europe and Asia - are also not decreasing their emission uncertainty. This tells us that we are no more confident in the final source estimation than we were in the beginning. From Chapter 2 we saw that the error in the emissions updates is based on \( C_{t+1} = C_t - GPC_t \). For large \( N \), \( G \) is small, thus the error is updated very little. This is the situation for the European source since Ireland has so much noise with all the pollution events. When \( P \) is small, there is also very little update. This is the situation for Asia. Furthermore, when looking at the residuals shown in Figure 5.26, it is evident that the inversion process is not well behaved. The calculated values at Tasmania diverge from the observed during
Figure 5.15: Time progression of the estimated source strengths at each source region during the pseudo-data inversion run of the CCM2 using the ALE/GAGE stations.
Figure 5.16: Time progression of the residuals corresponding to the estimated emission time series in Figure 5.15. Also shown (vertical axis on the right) is the corresponding partial derivative relating the model calculated value to the source(s) shown in the right hand corner of each frame.
Figure 5.17: Time progression of the estimated source strengths at each source region during the pseudo-data inversion run of the CCM2 using Hateruma rather than the Oregon ALE/GAGE station.
Figure 5.18: Time progression of the residuals corresponding to the estimated emission time series in Figure 5.17. Also shown (vertical axis on the right) is the corresponding partial derivative relating the model calculated value to the source(s) shown in the right hand corner of each frame.
Figure 5.19: Time progression of the estimated source strengths at each source region during the pseudo-data inversion run of the CCM2 using Kamchatka rather than the Oregon ALE/GAGE station.
Figure 5.20: Time progression of the residuals corresponding to the estimated emission time series in Figure 5.19. Also shown (vertical axis on the right) is the corresponding partial derivative relating the model calculated value to the source(s) shown in the right hand corner of each frame.
Figure 5.21: Time progression of the estimated source strengths at each source region during the pseudo-data inversion run of the CCM2 using Hong Kong rather than the Oregon ALE/GAGE station.
Figure 5.22: Time progression of the residuals corresponding to the estimated emission time series in Figure 5.21. Also shown (vertical axis on the right) is the corresponding partial derivative relating the model calculated value to the source(s) shown in the right hand corner of each frame.
Figure 5.23: Time progression of the estimated source strengths at each source region during the data inversion run of the CCM2 using ALE/GAGE observations.
Figure 5.24: Time progression of the residuals corresponding to the estimated emission time series in Figure 5.23. Also shown (vertical axis on the right) is the corresponding partial derivative relating the model calculated value to the source(s) shown in the right hand corner of each frame.
Figure 5.25: Time progression of the estimated source strengths at each source region during the data inversion run of the CCM2 using ALE/GAGE observations with added noise.
Figure 5.26: Time progression of the residuals corresponding to the estimated emission time series in Figure 5.25. Also shown (vertical axis on the right) is the corresponding partial derivative relating the model calculated value to the source(s) shown in the right hand corner of each frame.
Figure 5.27: Time progression of the estimated source strengths at each source region during the hybrid inversion run of the CCM2 using ALE/GAGE observations averaged with the model time series in Figure 5.11 and using the original noise of the data.
Figure 5.28: Time progression of the residuals corresponding to the estimated emission time series in Figure 5.27. Also shown (vertical axis on the right) is the corresponding partial derivative relating the model calculated value to the source(s) shown in the right hand corner of each frame.
Figure 5.29: Time progression of the total global estimated source strengths from each of the inversion tests.
the second half of 1987, and yet no adjustment is made to the emissions. This can be understood by the small error on the emissions after the first half of 1987. This defines the covariance matrix which controls the size of the emissions update. In this case, the emissions cannot change even though the residual is large. This later fact suggests that the transport in the CCM2 is not adequate for deducing sources by comparisons with observations on monthly time scales.

In the hybrid inversion shown in Figure 5.27, we see that there is reasonable movement toward the correct sources based on our previous analyses. We also note that the problems here are now a combination of those in the pseudo-data and real data cases. The Southern Hemisphere source shows the same over-adjustment and then no change. It is starting again to readjust in the summer of 1988 when the partial derivatives become significant again. However, the covariances are so small that the gain matrix is too small to make the adjustments needed to converge rapidly.

Figure 5.29 shows the progression of the global total in each of the eight inversion tests listed in Table 5.2. In all the pseudo-data cases the oscillation seen in the individual sources is seen in the global total as well. Interestingly, despite the difficulty constraining individual sources using the real observational data in the CCM2, the global total is reasonably well constrained.

To ensure that the oscillations seen above were not due to a loss of information due to a large Noise matrix when the pollution events are included, we removed the pollution events from the data and the pseudo-data. The resultant time series for the observations and the pseudo-data are shown in Figure 5.30. We also flagged pollution days predicted in the model during the inversion so that pollution events would also be removed for monthly averages within the model. We then repeated all of the experiments listed in Table 5.2 for several months. The results are shown in Figure 5.31 through Figure 5.40.

The control run looks quite similar to that with pollution shown in Figure 5.13. Europe goes toward the "correct" value faster due to the smaller noise, but it then overshoots. This is easily understood from our analysis of the control case including pollution events. In summer there is a correspondence of the large residual and partial derivative, but now with smaller noise the effect is not damped out. Likewise the same situation is true for Oregon in relation to Japan and in Samoa and Tasmania due to the Southern Hemisphere. Notice the effect of an incorrect initial guess for the source is apparently more strongly felt at Tasmania with pollution removed than with pollution included. This is in fact due to the correction of the Australian weighting to the Chapter 3 value (see section 5.4, second paragraph) and not to the removal of pollution.
Figure 5.30: Monthly average time series of CFCl₃ from the CCM2 and the ALE/GAGE data both with pollution episodes removed.
Figure 5.31: Time progression of the estimated source strengths at each source region during the control inversion run of the CCM2 with pollution removed.
Figure 5.32: Time progression of the residuals corresponding to the estimated emission time series in Figure 5.31. Also shown (vertical axis on the right) is the corresponding partial derivative relating the model calculated value to the source(s) shown in the right hand corner of each frame.
Using Pseudo-Data for Inversion

Figure 5.33: Time progression of the estimated source strengths at each source region during the pseudo-data inversion run of the CCM2 using the ALE/GAGE stations with pollution removed
Figure 5.34: Time progression of the residuals corresponding to the estimated emission time series in Figure 5.33. Also shown (vertical axis on the right) is the corresponding partial derivative relating the model calculated value to the source(s) shown in the right hand corner of each frame.
Using Hateruma for Inversion

Figure 5.35: Time progression of the estimated source strengths at each source region during the pseudo-data inversion run of the CCM2 using Hateruma rather than the Oregon ALE/GAGE station with pollution removed.
Figure 5.36: Time progression of the estimated source strengths at each source region during the pseudo-data inversion run of the CCM2 using Kamchatka rather than the Oregon ALE/GAGE station with pollution removed.
Figure 5.37: Time progression of the estimated source strengths at each source region during the pseudo-data inversion run of the CCM2 using Hong Kong rather than the Oregon ALE/GAGE station with pollution removed.
Figure 5.38: Time progression of the estimated source strengths at each source region during the data inversion run of the CCM2 using ALE/GAGE observations with pollution removed.
Figure 5.39: Time progression of the estimated source strengths at each source region during the data inversion run of the CCM2 using ALE/GAGE observations with added noise with pollution removed.
Figure 5.40: Time progression of the estimated source strengths at each source region during the hybrid inversion run of the CCM2 using ALE/GAGE observations averaged with the model time series in Figure 5.21 and using the original noise of the data with pollution removed.
In the pseudo-data cases the Southern Hemisphere source oscillates a great deal. Due to smaller \( N \), the coincidence of the residual and partial derivative at large values produces large jumps. This is causing oscillations not seen in the control case. Recall, however, that the size of \( \Delta E \) is constrained by \( E_0 \). In the control case the \( E_0 \) is much less than the \( E_0 \) of the other tests. The coincidence of large partial derivatives and residuals is only a problem since the partial derivative is so small the rest of the year giving only a small window for adjustments in the source.

We see similar behavior starting for all the test inversions with pollution excluded that were seen in the series with pollution included. It is apparent that having such strong seasonality to the partial derivatives induces problems since there is only a short period during the year for the model to adjust the emissions. It may be best to smooth the partial derivatives or even use an average value to allow adjustments continuously. This could also induce some oscillations, but they would be on shorter time scales allowing us to run the model for a shorter period to get these to damp than the on year oscillations we are seeing now. This would also prevent large decreases in the emission error which would also help to balance out the oscillations. A further control that could be added to the Kalman filter would be to constrain the change in emissions during each inversion step to be related to the size of the partial derivatives to avoid over-adjusting during the times that the partial derivatives are large.

5.5 Summary

The CCM2 is able to simulate many of the features seen in the ALE/GAGE observations. However, there are some phase differences for pollution events and seasonal patterns at Oregon and pollution events at Ireland and Tasmania. There is reasonable agreement in the latitudinal gradient defined by the ALE/GAGE stations and that defined by the same locations in the CCM2.

However, the inversion studies using real ALE/GAGE observations suggest that the CCM2 transport does not adequately simulate the atmospheric circulation for deducing surface sources from tracer observations using month-to-month varying partial derivatives and the monthly mean concentrations. Although, the issue of induced oscillations with strong seasonality in the partial derivatives will still be an issue, it may be that high resolution transport models using observed winds will be better suited for inversion studies.

The inversion tests in the CCM2 using pseudo-data as observations show that using a higher resolution model may induce spurious oscillations in predicted emissions
due to the stronger seasonality of the partial derivatives. This may be improved by smoothing the partial derivatives so the change from month to month is less, using varying constraints for the size of the update of emissions as related to the partial derivative, doing longer than one month time averages, initializing the inversion run during the time when the partial derivatives are large, or, when the seasonality is different in the various locations, ensemble runs could be done. This latter approach would entail running the model for short periods of times during the various peak partial derivative times in order to constrain the source strengths better before continuing the inverse run.
Chapter 6. Conclusions

As many chemically and radiatively important trace gases increase in concentration in our atmosphere, the need to understand their cycles becomes crucial. Often the hardest aspects of the cycle to quantify are the surface sources and sinks. If the regional surface sources and sinks of compounds such as methane, nitrous oxide, and carbon dioxide could be more precisely quantified, we could clarify both the processes involved and the needed policies to deal with these increases in the future. As it stands we now often do not know the regional source strengths within an order of magnitude (although the global values are usually better constrained).

There have been studies to try to measure fluxes of trace gases from known local sources and then extrapolate to global scale by multiplying by the areas of each type of source. This gives some constraints to the regional and global sources, but the uncertainties are still large. Other attempts to quantify these sources further have involved using models and comparing to observations. Some of the first studies of this type were trial and error. After running the model for a few years with set sources, the results are compared to observations and the emissions adjusted subjectively according to the difference in the two time series. Recently, inverse methods have been getting more attention by atmospheric chemists. These methods allow for objective adjustments of the emissions based on the difference between model results and observations. The goal of this thesis was to investigate one of these methods using three-dimensional global models.

In this work the Kalman filter was chosen as the inverse method. The strengths of this method are its straightforward implementation and the error estimate that is obtained for the optimally estimated state. This latter feature helps define how well the final state is constrained. It appears that the Kalman filter can work for estimating regional source strengths using a three-dimensional chemical transport model and observational data as long as the following assumptions are met:

a) Sources can be defined in specific regions such that areas outside those regions do not contribute significantly to the global input. And, the emissions within each region can be defined with a constant shape.

b) The model simulates the true transport and atmospheric chemistry.

In this thesis, this method was used on CFC-11 since its sources are well quantified and could, therefore, serve as a test of the inverse method. Assumption (a) is appropriate for CFC-11. A priori information about the sources is available from
industry. When the internal shape could not be defined, the region was subdivided into smaller regions. A high resolution source map of CFC-11 was produced in this thesis based on industrial information, population, and electrical usage to define regions and shapes. From this detailed map the predominant source regions were selected for the inversion process.

There is no way to avoid (b). All inverse methods for deducing emissions rely on the model simulating the transport and chemistry correctly. The model is used to generate partial derivatives for the Kalman filter, or the Green's functions, orthogonal functions, and cost functions for the various other inversion methods. All of these generated functions are dependent on the model.

The first step of our study was to determine the feasibility of using the Kalman filter in a three-dimensional transport model to deduce regional source strengths when the transport was the true one. We made use of a low resolution model that could be quickly run on the Cray Y-MP at Goddard. A test inversion was run using a model generated time series as "observations". We found that indeed, under these "perfect model" conditions the sources could be correctly deduced. We also ran a test using this model with ALE/GAGE observational data for which the model circulation was not the true one. This time the inversion was very poorly behaved, and did not converge to the correct solution. We proceeded to illustrate that if the predicted model time series is within $1\sigma$ of the observations then a reasonable solution is found to the source strengths.

The next step was then to look at recently available high-resolution general circulation climate model, the CCM2, that was capable of simulating the atmospheric dynamics well enough to possibly be capable of deducing source strengths at high spatial resolution. The CCM2 had shown itself to be one of the better transport models based on calculated dynamics to simulate important features seen in tracer time series (Williamson et al. (1992)). In running inversion tests in the CCM2 we found that the Kalman filter was no longer as straightforward to implement as it had been in the lower resolution model. In the high resolution model the influence on an observational site due to a source region (i.e., the partial derivative matrix elements) is strongly time dependent. The strong seasonality of this matrix causes there to be only a restricted window of time during the year that the emissions are effectively updated. Thus, it may be necessary to smooth the partial derivatives in time to avoid the oscillations we see induced by their seasonality or to include constraints on the size of the emission updates that are related to the size of the partial derivative. The oscillations could also be avoided by starting the inversion while the partial derivatives are large rather than small so that there is not a backlog of erroneous initial guesses effecting the subsequent evolution.
Curiously, we were able to avoid these problems in the low resolution model. We had thought the low resolution model would be inferior due to the larger dispersion of injected trace gases and what we, therefore, considered to be a greater loss of information. However, it turns out that this simplifies the implementation of the Kalman filter because it depresses the seasonal cycles in the partial derivative matrix. As an air parcel leaves a source region it is more quickly spread out over large areas which smooths the effect of added emissions. Thus, we find that even as the winds change direction during the year, the effect of a source on an observation site is comparatively constant or has high frequency cycles. In contrast in the higher resolution model the air parcels maintain more integrity and directionality. This induces much more time dependence in the influence of a source on an observation site in the high resolution inversion. It will be useful to investigate the effect of various possibly beneficial amendments to the Kalman filter to see if the filter will be better behaved. Some of the possible changes are to smooth the partial derivative matrix (e.g. a multi-month running mean), constraining the update in emissions relative to the size of the partial derivative, invert over longer time averages than monthly, initialize the inversion when the partial derivatives are all large, or do ensemble runs during large partial derivatives to constrain the emissions closer to the true value before starting the several year inversion.

Even with the necessary corrections to the Kalman filter, the results of this thesis suggest that a climate-mean predicted transport represented by the CCM2 (and other similar models) may not be well-suited for deducing unknown tracer emissions by comparison of predictions with observations of these tracers. Although, the emissions are reasonably behaved during the inversion, the behavior of the residual between observed and the predicted mixing ratio suggest that the model is not adequately simulating the transport of the real atmosphere on monthly time scales. It is possible that transport models based on observed winds will be better suited to address these problems. These models will still have to parameterize the sub-grid scale transport (convection and boundary layer transport). Thus, as these models are developed they will need to be carefully assessed to assure that these parameterizations have not introduced other transport errors. With these models the issue of not having enough data may still be a major question.

Finally, the studies with both low and high resolution model inversions showed that inverse methods can be used to identify favorable locations for future observation sites for determining emissions from specific regions. This exercise could greatly enhance the information we can obtain from future networks.
REFERENCES:


