Advanced Aeromagnetic Compensation Models for Airborne Magnetic Anomaly Navigation

by

Albert Reuben Gnadt

B.S. Mechanical Engineering, University of Wisconsin–Madison, 2015
S.M. Aeronautics and Astronautics, Massachusetts Institute of Technology, 2018

Submitted to the Department of Aeronautics and Astronautics
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Aeronautics and Astronautics
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2022

Author . . . .
Department of Aeronautics and Astronautics
May 17, 2022

Certified by . . . .
Alan S. Edelman
Professor of Mathematics
Thesis Supervisor

Certified by . . . .
Youssef M. Marzouk
Professor of Aeronautics and Astronautics
Chair, Thesis Committee

Certified by . . . .
Christopher V. Rackauckas
Research Affiliate, CSAIL
Member, Thesis Committee

Accepted by . . . .
Jonathan P. How
R. C. Maclaurin Professor of Aeronautics and Astronautics
Chair, Graduate Program Committee

Advanced Aeromagnetic Compensation Models for Airborne Magnetic Anomaly Navigation

by

Albert Reuben Gnadt

Submitted to the Department of Aeronautics and Astronautics
on May 17, 2022, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Aeronautics and Astronautics

Abstract

Using the earth's magnetic anomaly field for navigation of aircraft has shown promise as a viable alternative to the Global Positioning System (GPS) and other navigation systems. An airborne magnetic anomaly navigation (MagNav) system collects real-time magnetic field data and uses predetermined magnetic anomaly maps of the earth to estimate location by aiding an inertial navigation system (INS), which continually drifts. MagNav has the benefits of being passive, globally available at all times and in all weather, and not reliant on sight of land or stars. Since the magnetic field strength of a dipole decreases with the inverse cube of distance, MagNav is also nearly unjammable. A corrupting magnetic source must be flying alongside or in the aircraft to be effective.

This magnetic physics has other implications, though. In particular, the magnetic components of the aircraft itself interfere with the desired magnetic measurements that are required to navigate. Magnetic measurements are a linear superposition of multiple magnetic fields. When the measured data contains magnetic signals from both the (desired) earth field and (undesired) aircraft field, it is difficult to separate the two signals. Previous work has proven the viability of MagNav using exceedingly clean magnetic measurements taken by geo-survey aircraft. The most significant outstanding challenge for real-world, operational MagNav is handling corruption of the measured magnetic signal by magnetic sources from aircraft components.

In this thesis, several approaches to enable high-accuracy MagNav, despite receiving corrupted magnetic field measurements, are explored. These approaches can be split into four groups: linear aeromagnetic compensation, nonlinear aeromagnetic compensation, online aeromagnetic compensation, and covariance-adaptive filtering. The first two approaches evaluate different models that aim to improve on the state-of-the-art linear model used for removing aircraft interference. The last two approaches focus on making adjustments within the navigation algorithm in real-time based on the (corrupted) data provided. Performance is compared against the state-of-the-art compensation and navigation approach, which show that these advanced linear and nonlinear models can benefit MagNav when only corrupted magnetic field measurements are available. Each model and additional tools for aeromagnetic compensation and airborne magnetic anomaly navigation are publicly available in the MagNav.jl Julia software package.

Thesis Supervisor: Alan S. Edelman
Title: Professor of Mathematics

Acknowledgments

I have been at MIT for seven years now and worked on four (too many) different research projects, allowing me to meet many wonderful people. Thank you to Beth Marois for helping me get through each difficult advisor/lab transition. I am extremely grateful to my final research advisor, Prof. Alan Edelman, for welcoming me into the Julia Lab and encouraging me to finish my PhD. Your guidance and non-overbearing, but available, approach to advising was exactly what I needed to be successful. I am similarly grateful to Chris Rackauckas, who has provided great feedback throughout this project. Finally, thank you to my academic advisor, Prof. Youssef Marzouk, for your feedback in committee meetings and overall kindness.

I would like to thank the entire AIA MagNav team, but particularly Allan Wollaber and Joseph Belarge for our many one-on-one conversations. MagNav.jl would not have progressed as it did without your constructive feedback. Thank you to Jonathan Taylor and Mike O'Keeffe for leading the project and Capt Kyle Palko for your comradery. Thank you to our AFIT collaborators, Aaron Nielsen, Maj Joseph Curro, and Maj Aaron Canciani, who greatly helped ramp up the project. Additional team members I would like to thank include Maj David Jacobs, Maj Kyle McAlpin, TSgt Chasen Milner, Glenn Carl, Linh Pham, Michael Yee, and Lauren Conger.

Thank you to Prof. Steven Barrett, Jayant Sabnis, and Raymond Speth for your guidance during my Master's research, a project that I much enjoyed. I would also like to mention my mentors during my time as an undergraduate. Thank you to Kristen Seashore, Jack Page, Strider Hunt, and Steve Lambert for guiding me in real-world engineering during my various internships. Thank you to Prof. Tim Osswald, Prof. Sage Kokjohn, and Prof. Melih Eriten for introducing me to academic research. All of you were great role models and I appreciate the letters of recommendation you

have written to help me move onto the next stages of my academic career. Thank you to Todd Braun for pushing me during high school and for staying in touch.

I am blessed to have so many friends in Wisconsin, Massachusetts, and spread across the country. Most, if not all, of you have inspired and encouraged me in good times and bad. I certainly have left some people out, but thank you to:

- The 125ers, Ian Mosbrucker, Matt (and Kelly) VandeHei, Bill (and Kristen) Hussey, Adam Luepke, and Mike Keim, for the many good memories over the past decade, especially Badger games and ski trips

- My other former roommates, especially Harris Fisher (and Mark Endrizzi by association), Luke Haug, and Mike Zanoni, for all of the shenanigans

- The Bows brothers, Alex (and Chloe) Buesing and Hans (and Aleece) Helland, for taking dumb pictures with our elbows in the air. Thank you to the rest of the Helland family as well, Eric, Mary, and Jozi (who just happen to be standing 10 ft away as I write this), for your support since my sister met Jozi in kindergarten.

- G^3 and my other close high school friends, specifically Patrik Gorun, Nate Gibson, Chad Marthaler, Alex Greiff, Garrett Madland, and Aaron Stroede, for the now decades of friendship. Thank you to your families as well, especially Richard and Cindy Marthaler, for your support.

- My other high school friends, including Devin Winnes, Ethan Freel, Michael Holzem, Abbi Houghton, Karin Clemens, Hunter Martin, Justin Ward, and the entire WDHS soccer and swim teams, for the many experiences we have been through together

- Boy Scout Troop 66, especially Tom, Cindy, and Kevin Jones, for helping me become a well-rounded person
- The UW–Madison ASME members, especially the ones I have stayed in touch with, including Kara Novotny, Chad Van De Hey, Kasey Boxleitner, and Abe Spindel, for the many events we have done together as we became engineers
- The AeroAstro Ballers, including Chris Bradley, John Graham, Kevin Sabo, Matthew Moraguez, Yonatan Tekleab, Blake Bequette, and Jasper Arneberg, for the fun times on and off the court
- The Esteemed Squad, especially Marlis Denk-Lobnig, Arny Leroy, Quantum Wei, Max Khatsenko, Gerardo Bledt, Cal Brooks, and Joe Lake, for the many good times
- My Edgerton House friends, including Divya Ramamoorthy, Cameron McBride, Rachel Kurchin, Jerry Wang, Katina Ross, and Lee Thurston, for helping form a great community
- The Potlucks group, including Sarah Fay, PJ Santos, Jenny Shen, Drew Meyers, Sarah Walker, Daniel Duane, Affi Maragh, Hugo Uvegi, Carolyn Joseph, and Xiaoyu Wu for the tasty food and good conversation
- My other MIT friends, including Alec Kushner, Jinwook Lee, Cory Frontin, Kris Frey, Prashanth Prakash, Carla Grobler, Akshat Agarwal, Parker Vascik, Eric Hinterman, Berk Ozturk, Mohammad Islam, and Diogo Castilho, for the nerdy chats and cold beverages, among other things
- Chrissy Eckman for the best dates EVER and the reminder

Finally, and most importantly, I am deeply grateful for the love and encouragement from my family during graduate school and all of the years beforehand. Thank you to my late grandparents, Ray and Gertrude Scheid and Reuben and Emily Gnadt for treating me like the angel I was not. Thank you to my many aunts, uncles, and cousins, particularly Aunt Char and Uncle Leroy Heider and cousins Robert Giese and Matt Nourmohamadian, for the fun family times. Thank you to my sister Katie Hartman and brother-in-law Taylor Hartman (and your little guy, Harrison Huxley Hartman aka Triple H) for being great role models and now parents. Thank you to my mother, Jane Scheid-Gnadt, and late father, Richard Gnadt, for your endless support throughout my life. My successes would not have been possible without you.

This thesis is dedicated to the many strong women in life who have (sometimes aggressively) pulled me back to reality when needed.

Research was sponsored by the United States Air Force Research Laboratory and the United States Air Force Artificial Intelligence Accelerator and was accomplished under Cooperative Agreement Number FA8750-19-2-1000. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the United States Air Force or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

Contents

List of Figures

List of Tables

Nomenclature

Acronyms and Abbreviations

AGL	above ground level
DFT	discrete Fourier transform
DRMS	distance root mean square
EKF	extended Kalman filter
FFT	fast Fourier transform
FOGM	first-order Gauss-Markov
GPS	Global Positioning System
HAE	height above ellipsoid
IGRF	International Geomagnetic Reference Field
IMU	inertial measurement unit
INS	inertial navigation system
IR	improvement ratio
L-BFGS	limited memory Broyden–Fletcher–Goldfarb–Shanno algorithm
LLA	latitude, longitude, and altitude
LSTM	long short-term memory
MAD	magnetic anomaly detection
MagNav	airborne magnetic anomaly navigation

ML	machine learning
MPF	marginalized particle filter
MSE	mean squared error
NAMAD	North American Magnetic Anomaly Database
NN	(artificial) neural network
PLSR	partial least squares regression
RBPF	Rao-Blackwellized particle filter
RLS	recursive least squares
RNN	recurrent neural network
SciML	scientific machine learning
SGL	Sanders Geophysics Ltd.
SLAM	simultaneous localization and mapping
SNR	signal-to-noise ratio
UKF	unscented Kalman filter
USAF	United States Air Force
USGS	United States Geological Survey
UTM	Universal Transverse Mercator
WDMAM	World Digital Magnetic Anomaly Map
WGS	World Geodetic System
WMM	World Magnetic Model

Chapter 1

Introduction

Using local variations in the earth's magnetic anomaly field for navigation of aircraft has shown promise as a viable alternative to the Global Positioning System (GPS) and other navigation systems. The intent is not to replace GPS entirely, which is often accurate to less than 10 m and available worldwide [1], but to instead have an alternative navigation method in case GPS becomes unavailable. Airborne magnetic anomaly navigation (MagNav) has been shown to achieve accuracies on the order of 10's of meters, which almost reaches typical GPS position accuracy [2]. MagNav has the benefits of being passive, globally available at all times and in all weather, and not reliant on sight of land or stars.

Furthermore, MagNav is nearly unjammable, which is particularly important for military applications, such as United States Air Force (USAF) missions in GPS-denied environments [3]. GPS jamming is done by simply overpowering the GPS signal with an adversarial signal at the same frequency as GPS. The jam-proof ability of MagNav stems from the basic physics of a magnetic dipole. The strength of a magnetic field decreases proportionally with the inverse cube of the distance from the magnetic source [4]. This means a corrupting magnetic field would have to be near an aircraft to be effective. This physics has other implications – the magnetic components on an aircraft itself can interfere with the desired magnetic measurement.

An airborne magnetic anomaly navigation system collects real-time magnetic field data and uses predetermined magnetic maps of the earth to estimate location by aiding an inertial navigation system (INS). However, the received magnetic measurement is actually a linear superposition of multiple magnetic fields. The most significant outstanding challenge for operational MagNav is handling corruption of the earth's magnetic signal by magnetically active aircraft components. When the measured data contains magnetic signals from both the (desired) earth field and (undesired) aircraft field it is difficult to separate the two signals. This work explores several approaches for high-accuracy navigation despite receiving corrupted magnetic field measurements.

This chapter provides background on the earth's magnetic field and how it is measured. The central challenge this thesis aims to address is then discussed. Finally, the thesis contributions and outline are presented.

1.1 Earth's Magnetic Field

The total geomagnetic field of the earth is a linear superposition of fields from several sources. For this work, the three sources considered are the core field, anomaly field, and temporal variations, each of which are discussed individually in this section.

1.1.1 Core Field

The dominant source is the core field, or main field, which is shown in Figure 1-1. Typical values at the surface of the earth are between 20,000 and 60,000 nT – more than 95% of the total geomagnetic field. Thus, the magnetic north direction that a compass points is almost entirely due to the core field. The core field is caused by the rotation of conductive fluids deep within the earth. This fluid rotation generates electrical fields, which in turn generates magnetic fields.

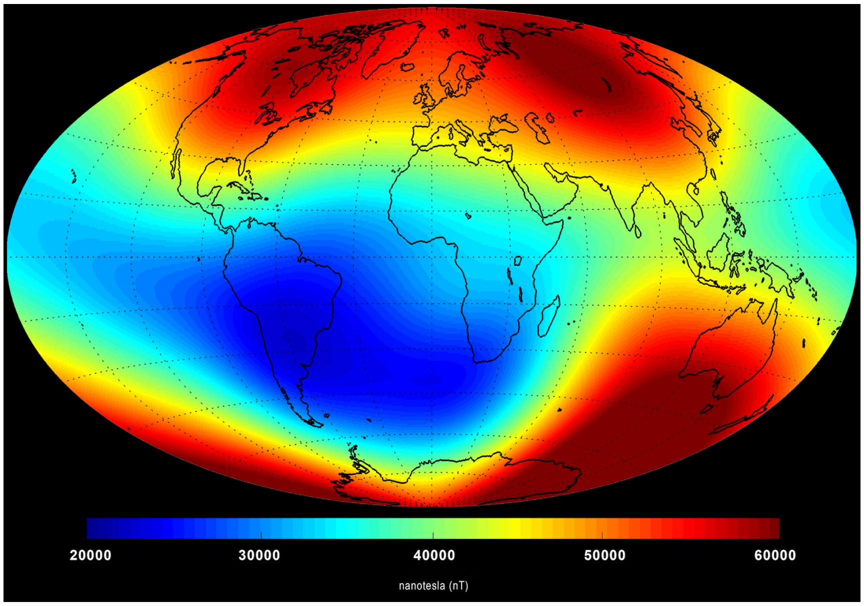


Figure 1-1: Earth's core field. Modified from [5].

A widely used model of the core field is the International Geomagnetic Reference Field (IGRF) [6]. More specifically, the IGRF represents a set of coefficients that are used in a spherical harmonic model. The spherical harmonic coefficients span from degree 1 to 13, indicating that the model covers the first 13 wavenumbers and is only able to model long wavelengths of over 3000 km. Thus, the frequency content of the core field is not suitable for MagNav with 10's or even 100's of meters of accuracy.

Note that the core field changes on timescales of months to millions of years, so the IGRF is typically revised every five years to account for this variation. Using the IGRF model, the core field is known to a high degree of accuracy for a given position and time. An alternative to the IGRF is the World Magnetic Model (WMM), which is used by the United States Department of Defense and other government agencies. Only the IGRF is used in this work.

1.1.2 Anomaly Field

The second geomagnetic source is the anomaly field, also known as the crustal field. The magnetic fields generated in the anomaly field are due to permanent or induced magnetization of rocks in the earth's crust, which is the upper portion of the lithosphere, the rigid outer layer of Earth [7]. Portions of the crust at a temperature less than the Curie point of magnetite or other magnetic minerals can be magnetized by the core field through induction. This is the primary source of magnetization in the crust [8]. Permanent magnetization, otherwise known as remanent magnetization, results when minerals are magnetized by induction and then cool below the Curie temperature, but remain magnetized.

Local variations in the anomaly field, as shown in Figure 1-2, can provide positional information. The anomaly field is also stable over geologic time spans, only varying by 1 nT or so per year. Thus, the anomaly field may be used for navigation, even at high altitudes of more than 10 km, though navigation accuracy is decreased. However, it is important to note that the anomaly field is approximately 2 orders of magnitude weaker than the core field, which must be accounted for along with temporal variations.

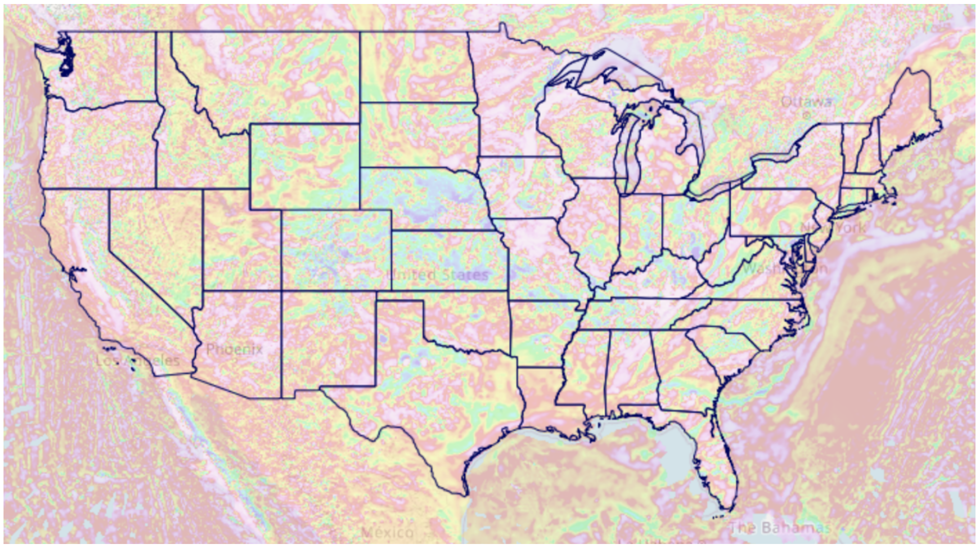


Figure 1-2: Magnetic anomaly field over the continental United States. Modified from [9].

1.1.3 Temporal Variations

While the core and anomaly geomagnetic fields are generated internally, temporal variations are generated externally. This component of the total magnetic field is due to contributions from the ionosphere, magnetosphere, and coupling currents between the two, as shown in Figure 1-3. These three contributions are often collectively termed "space weather."

When the ionosphere is ionized by solar radiation, an electrically conducting plasma (charged particles) is created in which electrical currents can flow [4]. Atmospheric tidal winds, mostly caused by day-night differential solar heating or the gravitational attraction of the moon, cause the plasma to move relative to earth's magnetic field. This movement creates electrical currents, which creates magnetic fields. Additional electrical currents are caused by equatorial and polar electrojets, which are discussed in [11] in greater depth.

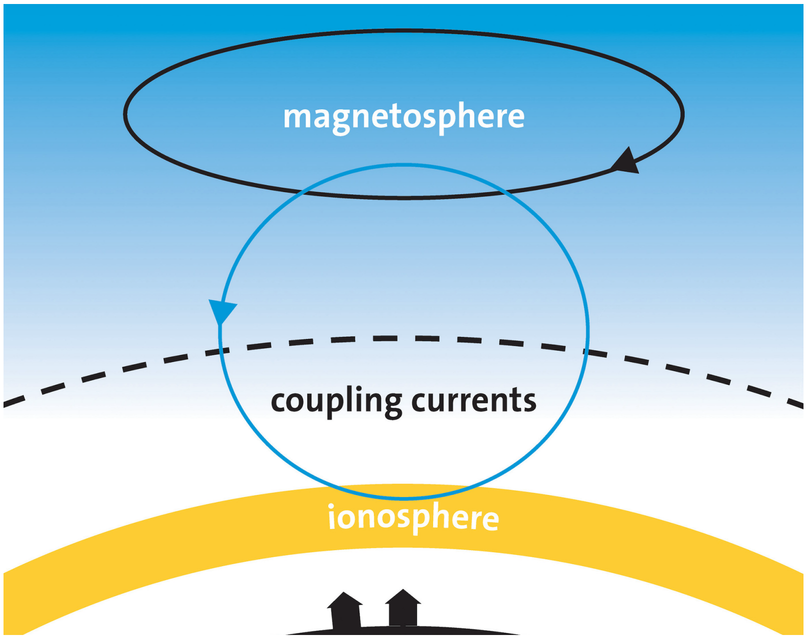


Figure 1-3: Upper atmosphere magnetic sources. Modified from [10].

The magnetosphere is a region in space where plasma (i.e. solar wind from the sun) interacts with the earth's magnetic field. This interaction causes aurora near the northern or southern poles. A high amount of interaction is referred to as a geomagnetic storm. Similar to the ionosphere, relative movement of the plasma and earth's magnetic field creates flowing electrical currents, which creates magnetic fields. Electrical currents, and thus additional magnetic fields, are also created by the coupling of the ionosphere and magnetosphere.

Temporal variations are often called the "diurnal" component of the total magnetic field, as they change from day to day. This component can be estimated at a given position and time using ground-based reference measurements taken at a stationary base station within the flight region [11]. Alternatively, if base station measurements are unavailable, temporal variations can be modeled within the navigation algorithm itself, which is discussed in section 5.3.

1.1.4 Summary of Components

Table 1.1: Nominal magnitudes of earth's magnetic field components. These values were estimated from flight data recorded near Ottawa, Ontario, Canada during the summer of 2020 at an altitude of 800 m. The spatial variation is 2σ during a 1 hr flight covering 240 km (150 mi).

Component	**Magnitude**	**Spatial Variation**	**Temporal Variation**
Core field	50,000 nT	100 nT	1 nT/week
Anomaly field	100 nT	200 nT	1 nT/year
Temporal variations	10 nT	5 nT	10 nT/day

Shown in Table 1.1 are nominal values for each of the components of the earth's magnetic field. Clearly any measurement of the total field is dominated by the core field. However, there is a high amount of spatial variation from anomaly field sources. Not shown is the spatial frequency content of each component, which is generally much greater for the anomaly field than the temporal variations or core field. The anomaly field is also stable over time, while the core field and temporal variations vary on shorter time scales, though these can be accounted for, as discussed in section 5.3.

1.2 Magnetic Measurements

Magnetic instruments were first used in the mid-1800s to discover (magnetic) ore deposits [12]. Later during World War II, magnetic anomaly detectors were developed for aerial use in detecting submarines [4]. Geologists adapted these geomagnetic survey instruments to map magnetic fields in the earth's crust as an aid for understanding sub-surface geology and providing guidance for mineral exploration and mining, magnetic anomaly detection (MAD), which continues to be done today. This has led to the creation of magnetic anomaly maps of various sizes with various cell spacings in areas across the world, which is discussed in section 2.3.

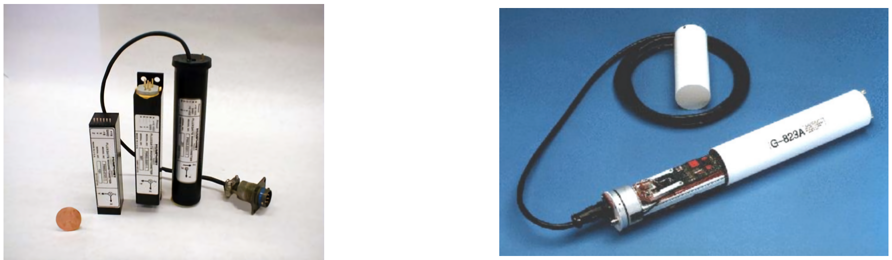
This work considers two types of magnetic measurements and their associated measurement sensors, scalar and vector, which are described in this section. There are advantages and disadvantages to each, and together they make high-accuracy MagNav possible.

1.2.1 Vector Magnetometers

The earth's total magnetic field $\vec{B}_t$ is a vector field,

$$\vec{B}_t = B_x\hat{\imath} + B_y\hat{\jmath} + B_z\hat{k} \tag{1.1}$$

where B_x, B_y, and B_z are magnitudes in the $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$ orthogonal directions of the chosen reference frame. Typically north, east, and down are used for airborne navigation. A measurement of the total field vector is made by a vector magnetometer. For airborne applications, fluxgate magnetometers are used, such as that shown in Figure 1-4a. These instruments have absolute accuracies of ±100 nT or worse typically [13]. The absolute error is due to the combined effect of calibration, temperature, orthogonality, and alignment errors. Due to their low accuracy, fluxgate magnetometers are not suitable as the primary magnetic measurement for navigation. However, they are important for MagNav since they are used for aeromagnetic compensation, as discussed in section 3.1.



(a) Goodrich 2801 fluxgate (vector) magnetometer [14].

(b) Geometrics Model G-823A scalar magnetometer [15].

Figure 1-4: Examples of vector and scalar magnetometers.

1.2.2 Scalar Magnetometers

The magnitude of the total magnetic field is

$$|\vec{B}_t| = \sqrt{B_x^2 + B_y^2 + B_z^2}. \tag{1.2}$$

A measurement of the total field magnitude $|\vec{B}_t|$ is made by a scalar magnetometer. This measurement has lost some information, namely the direction of the total field. However, modern atomic scalar magnetometers are sensitive and accurate. Airborne geomagnetic surveys, along with MAD, extensively use optically pumped, split-beam cesium vapor magnetometers [16, 17], such as that shown in Figure 1-4b. These scalar magnetometers have high sensitivity, high absolute accuracy, and rapid reading capability. For example, the Geometrics Model G-823A has a sensitivity of approximately 0.02 nT at a 10 Hz sample rate and an absolute accuracy of ±3 nT or less at any sample rate [15].

1.3 Thesis Contributions and Outline

Airborne measurements made by scalar and vector magnetometers contain a linear superposition of the earth field magnetic components (core field, anomaly field, and temporal variations). However, they also contain magnetic interference from the aircraft itself, which acts as corruption. The goal of aeromagnetic compensation, discussed in detail in Chapter 3, is to remove this undesired magnetic source. This is the central topic of this work.

Moreover, operational MagNav requires compensation of sources that have not traditionally been compensated, such as control surfaces, lights, radios, etc. Measuring small fluctuations in the magnetic anomaly field, which is used for navigation, becomes difficult when nonlinear and non-stationary magnetic noise is caused by the platform. There have been some recent attempts to use machine learning to improve aeromagnetic compensation [18, 19]. Additionally, online aeromagnetic compensation has recently been introduced [20], but not fully explored, especially with machine learning. With this in mind, the overall objective of this thesis is to evaluate multiple approaches for dealing with moderate to heavily corrupted magnetic measurements. Two overarching strategies are presented in this work, which are shown in Figure 1-5.

The first, a pre-processing strategy, is to improve the aeromagnetic compensation model. The state-of-the-art linear model does not adequately handle stochastically corrupted magnetic data as well as necessary for operational MagNav. This strategy does not require a navigation algorithm, though one is still useful to see the resulting navigation performance with a compensated magnetic signal. The second, an end-to-end strategy, modifies the baseline navigation algorithm and possibly the aeromagnetic compensation model as well.

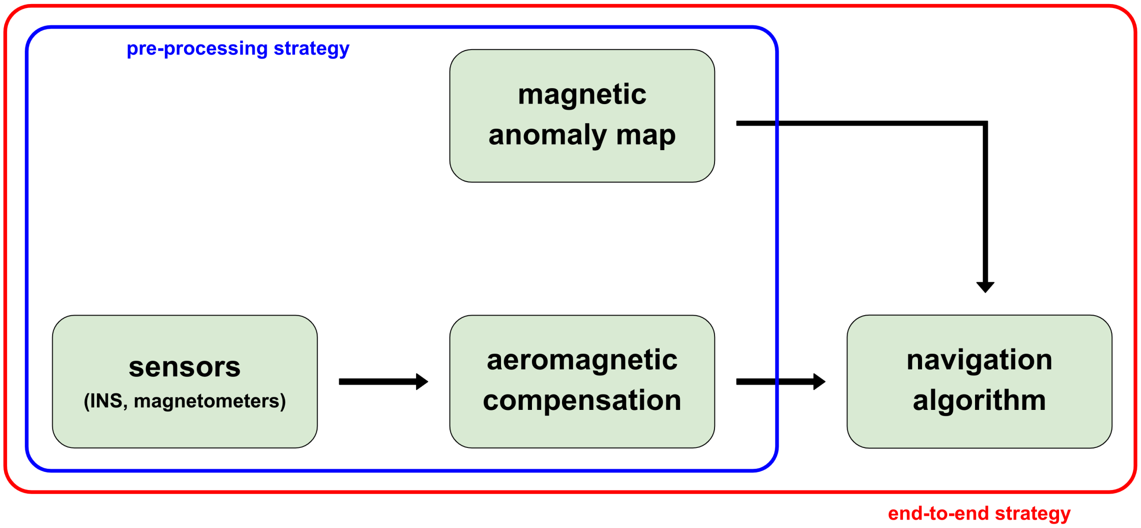


Figure 1-5: Two high-level strategies.

The overall research question is, "Which aeromagnetic compensation models and/or navigation algorithms enable an operational aircraft to perform high-accuracy airborne magnetic anomaly navigation while using corrupted magnetic field measurements?" The specific contributions of this work toward answering this question are:

1. MagNav.jl: a full suite of open-source tools for aeromagnetic compensation and airborne magnetic anomaly navigation implemented in the high-performance Julia programming language.
2. Development of linear aeromagnetic compensation models that outperform the classical model in compensating magnetic sources that have traditionally not been compensated.
3. Formation of neural network-based aeromagnetic compensation models that outperform the classical model in compensating magnetic sources that have traditionally not been compensated. This includes an assessment of which non-

magnetometer sensors provide useful additional compensation information.

4. Introduction of online neural network-based aeromagnetic compensation models with integrated machine learning-based online learning.
5. Evaluation of the performance impact of using adaptive filtering for real-time measurement noise covariance updating.

The remainder of this thesis follows the order of the contributions defined above. Chapter 2 provides an overview of the publicly available Julia-based MagNav software package, which was the primary software package used for this work. The flight data and magnetic anomaly maps used in this work are also described here, which are available in the package as well. Chapter 3 describes in detail the classical aeromagnetic compensation model, then introduces additional linear aeromagnetic compensation models. Chapter 4 presents multiple neural network-based aeromagnetic compensation models. Additionally, multiple analyses of neural network design and feature selection for these models are discussed. The models in both Chapters 3 and 4 are pre-processing strategies. Chapter 5 covers online aeromagnetic compensation and adaptive filtering, both of which are end-to-end strategies. The effect of specific aircraft maneuvers on filter observability is also evaluated. Finally, Chapter 6 compares navigation performance across the different approaches, summarizes the thesis findings and limitations, and proposes future work.

Chapter 2

MagNav.jl and Data Sources

This chapter first describes the primary software package used for this work. This is followed by a description of the flight data and magnetic anomaly maps used for this work, which includes a description of the required map processing – an essential step for high-accuracy airborne magnetic anomaly navigation (MagNav).

2.1 MagNav.jl

MagNav.jl is an open-source [21] software package for aeromagnetic compensation and airborne magnetic anomaly navigation written in the Julia programming language. It was the primary package used for this work. It was developed almost entirely by the author, but it was largely based on work by Aaron Canciani at the Air Force Institute of Technology [13]. This section describes the high-level functionalities of the package, then provides a brief performance comparison with the original MATLAB-based software from which it was developed.

2.1.1 Functionalities

MagNav.jl contains a full suite of open-source aeromagnetic compensation and airborne magnetic anomaly navigation tools, which were developed out of necessity. At the time of writing, no other open-source package of this type is available in any programming language. It is largely designed to be robust for different flight datasets and magnetic anomaly maps, though it is based on the dataset and maps described in sections 2.2 and 2.3.

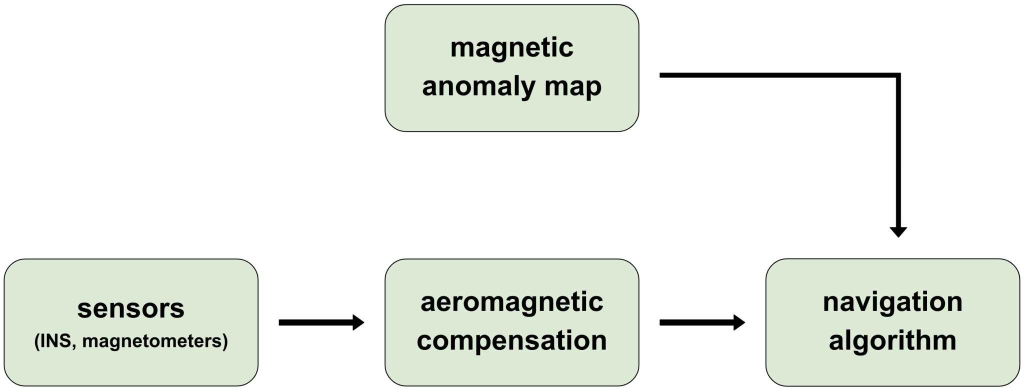


Figure 2-1: Four essential components of MagNav.

The MagNav.jl functionalities can be divided into the four essential components required for MagNav, which are repeated from Figure 1-5 in Figure 2-1. The first essential component, sensors, is essentially obtaining high-quality flight data, including magnetic and inertial data. This is largely well-established, though even higher performing hardware, such as NV-based diamond magnetometers, are actively being developed [22]. The flight data contained within MagNav.jl is described in section 2.2. Likewise, the magnetic anomaly maps and navigation algorithms available within MagNav.jl are described in sections 2.3 and 2.4, respectively. Finally, the aeromagnetic compensation models, the primary contribution of this work, are described in detail in Chapters 3 and 4.

MagNav.jl also includes unit tests, which use the SafeTestsets Julia package [23]. This puts each set of tests into a separate module so that they do not effect each other. Each test set can be evaluated completely independently.

2.1.2 Performance

Many of the functionalities in MagNav.jl are unable to be compared to the original MATLAB-based software from which it was developed in a fair comparison, as the structure of the packages have substantially diverged. However, some of the basic functions are essentially the same, and compared in Table 2.1.

Table 2.1: Benchmark runtime comparison between MagNav.jl and MATLAB-based software functions.

MagNav.jl Function	Function Description	MagNav.jl Runtime [s]	MATLAB Runtime [s]
map_itp	create linear map interpolation	0.19	0.19
map_itp	create cubic map interpolation	0.73	0.19
create_ins	create simulated INS data	5.2	6.6
crlb	compute Cramér–Rao lower bound	5.9	14.4
ekf	run extended Kalman filter	6.3	15.8

These timing benchmarks were determined with the @*btime* function in Julia and the *timeit* function in MATLAB using a 64 GB 2019 MacBook Pro. The data used is simulated flight data with 144668 samples (instances) and a 3195 × 7279 magnetic anomaly map. The function to create a cubic map interpolation takes longer in MagNav.jl, but this is typically only used once prior to evaluating navigation performance. When evaluating the more computationally-intensive functions, the high performance of Julia shows with approximately 50% lower runtimes. Note that neither package was particularly optimized for speed.

2.2 Flight Data

This work uses data from two sets of data collections by Sanders Geophysics Ltd. (SGL) using a Cessna 208B Grand Caravan, which is typically used by SGL for geophysical surveys [24]. The first data collection occurred during the summer of 2020 near Ottawa, Ontario, Canada. This is the primary data source used in this work, and the majority of the data is available within MagNav.jl via data artifact. A total of 34 hr and 46 min of flight data was collected over the course of 9 flights. Due to differences in available data fields, only 18 hr and 49 min of data from flights 1003-1007 is used. The second data collection occurred during the winter of 2021, again near Ottawa with another Cessna 208B Grand Caravan. Some of this data is used to test model performance on a different aircraft as well as for an analysis of the effect of using specific aircraft maneuvers for filter observability in section 5.8. This section provides more details on the flight data, which is followed by a description of the training and testing data splits used in this work.

2.2.1 Flight Data Details

The flight data itself is rather unique, in that it contains a nearly perfect magnetic signal (minimal aircraft interference) from tail stinger measurements, as well as four noisy magnetic signals (varying degrees of aircraft interference) from in-aircraft measurements. These scalar measurements of the total field were generated from five optically pumped, split-beam cesium vapor magnetometers. Four fluxgate magnetometers, one at the base of the tail stinger and three inside the aircraft, were also used for vector measurements of the total field. The specific locations of the magnetometers for flights 1003-1007 are listed in Table 2.2. The tail stinger scalar measurement with classical Tolles-Lawson compensation, which is described in detail in section 3.1, is considered to be the truth magnetic signal. This is essentially the value that would be used to generate a magnetic anomaly map during a geomagnetic survey. Additionally, many portions of the flight data have a corresponding magnetic anomaly map, which can be used with the Global Positioning System (GPS) position data to form an alternative truth magnetic signal. These values are typically very close (within 5 nT), since the anomaly field only changes roughly 1 nT/year.

Table 2.2: Flights 1003-1007 magnetometer locations. The reference point is the front seat rail. x is positive in the aircraft forward direction, y is positive to port (left facing forward), and z is positive upward. Geometrics optically pumped, split-beam cesium vapor scalar magnetometers are used. Bartington Mag-03 and Billingsley TFM100 fluxgate vector magnetometers are used.

Sensor Name	Location	x [m]	y [m]	z [m]
Scalar Magnetometers				
Mag 1	Tail stinger	-12.01	0	1.37
Mag 2	Front cabin, aft of cockpit	-0.60	-0.36	0
Mag 3	Mid cabin, near INS	-1.28	-0.36	0
Mag 4	Rear cabin, floor	-3.53	0	0
Mag 5	Rear cabin, ceiling	-3.79	0	1.20
Vector Magnetometers				
Flux A	Mid cabin, near fuel tank	-3.27	-0.60	0
Flux B	Tail, base of stinger	-8.92	0	0.96
Flux C	Rear cabin, port	-4.06	0.42	0
Flux D	Rear cabin, starboard	-4.06	-0.42	0

To provide a sense of the magnitude of corruption in the available scalar magnetometers listed in Table 2.2, an example of the uncompensated (raw) measurements is shown in Figure 2-2. Mag 1 has nearly no corruption, and Mag 5 follows the trend of Mag 1 well. Mag 4 is worse than Mag 5, but still largely matches the trend of Mag 1. Both Mags 2 and 3 have significant corruption and could not be used for MagNav without aeromagnetic compensation.

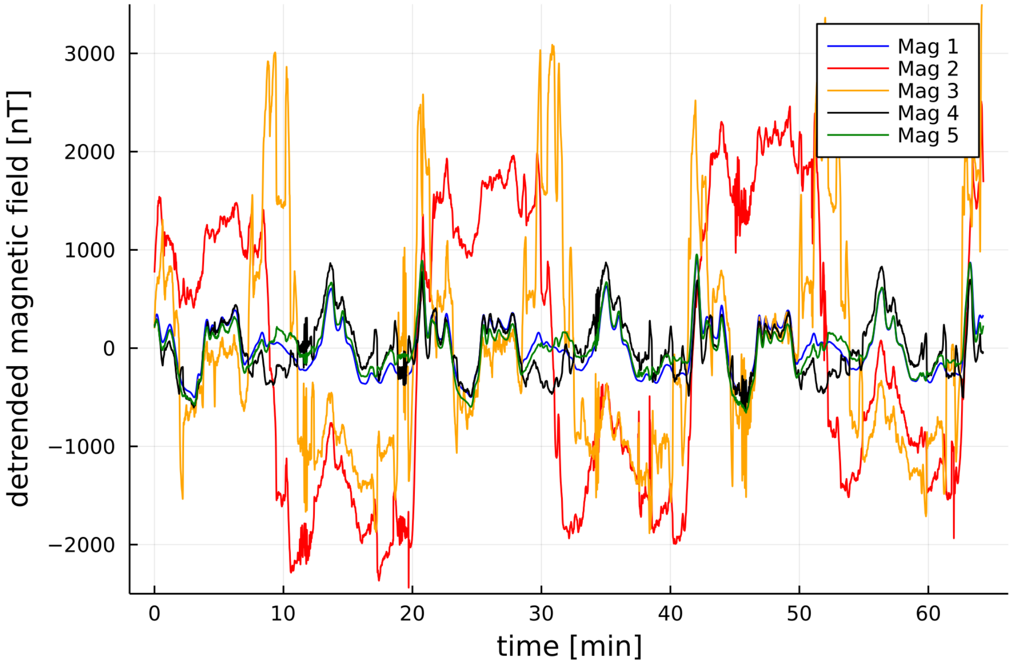


Figure 2-2: Uncompensated scalar magnetometers from flight line 1007.02. Due to their positions, the decreasing order of corruption of the uncompensated scalar magnetometers is: Mag 1, Mag 5, Mag 4, Mag 3, Mag 2.

The scalar magnetometers can of course be compensated using a vector magnetometer, which is shown in Figure 2-3. Here, classical Tolles-Lawson aeromagnetic compensation (described in section 3.1) was performed on each scalar magnetometer using Flux A and the first calibration box of flight line 1006.04. Now Mags 3, 4, and 5 follow Mag 1 quite well, but decreasingly significant errors still remain for the respective magnetometers. Mag 2 performs *worse* after compensation, and is thus not shown for clarity.

To emphasize the magnitude of corruption for each scalar magnetometer, the standard deviations of the magnetic errors are provided in Table 2.3. These errors are in reference to the magnetic anomaly map values along the flight path to give a

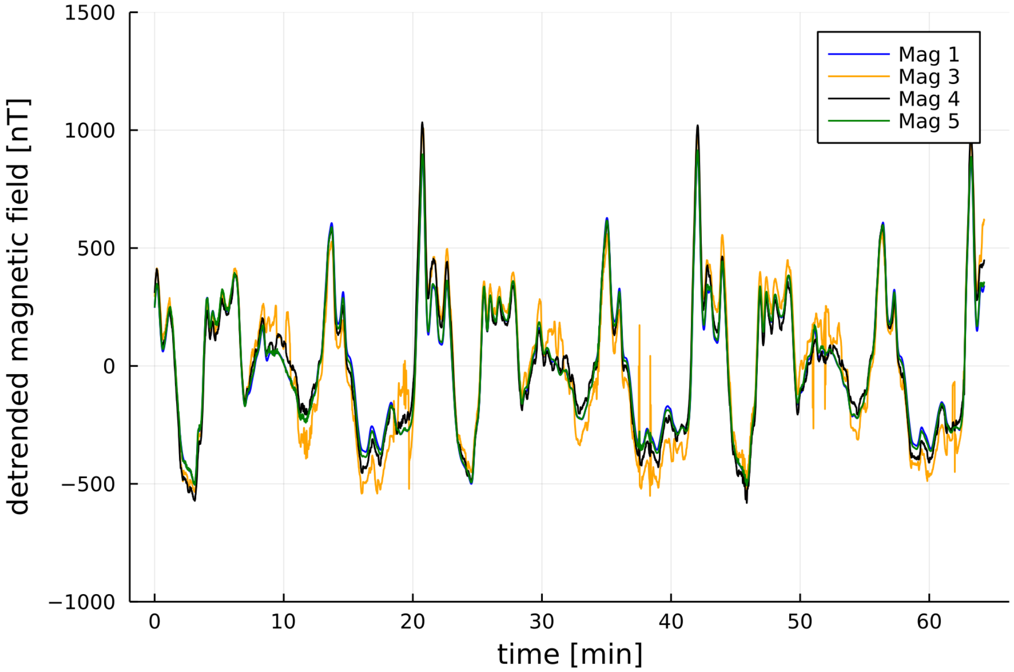


Figure 2-3: Compensated scalar magnetometers from flight line 1007.02. Due to their positions, the decreasing order of corruption of the compensated scalar magnetometers is: Mag 1, Mag 5, Mag 4, Mag 3, Mag 2. Mag 2 is not shown due to significant corruption, more so than in Figure 2-2.

sense of the errors that would be fed into the navigation algorithm. Comparing to the map values is not done in most of this work, since the tail stinger measurement is considered to be a better "truth" measurement as it was measured at the same time and avoids map errors. Mag 1 needs minimal compensation, while Mag 2 performs worse with compensation. The magnetic signal errors for Mags 3, 4, and 5 significantly decrease with compensation. Note that compensation performance is fairly insensitive to which vector magnetometer is used.

The flight data also contains supplemental sensor data from the inertial navigation system (INS), GPS position data, and auxiliary data from additional sensors, such as voltages and currents. This additional data reflects some of the temporal changes

Table 2.3: Aeromagnetic compensation errors for flight line 1007.02. The Tolles-Lawson model is applied to each scalar magnetometer with each available vector magnetometer. Errors [nT] are in reference to the map values along the flight path.

Scalar Mag	Uncompensated	Flux A	Flux B	Flux C	Flux D
Mag 1	6.2	5.8	5.8	5.8	5.8
Mag 2	1389	46109	46038	45316	45330
Mag 3	1032	102	108	111	102
Mag 4	203	54	53	52	52
Mag 5	104	13	12	12	13

in the virtual magnetic dipole of the aircraft, which causes error in the magnetic data that is not removed with classical Tolles-Lawson compensation. Currents were measured using induction sensors (placed around wiring) rather than inline sensors due to additional approvals that would be required. This means the currents sensors may have picked up induced changes to wires' electromagnetic fields that could have been caused by other sources. The full listing of the available data fields is shown in Appendix A. Note that when fitting or training any model, some data should not be used, especially GPS position and tail stinger measurements. In general the following data fields may be used: in-cabin magnetometer measurements and their gradients, diurnal and IGRF magnetic estimates, INS data, voltages, and currents.

During the data collection flights, various events were purposely carried out to cause temporal magnetic field disturbances during some flight lines. This included control surface movements (e.g. flaps up/down), fuel pump on/off, radio use, and movement of magnetic materials within the cabin. The flight patterns and altitudes were also varied from flight to flight to provide a diverse dataset. The flight altitudes were generally either at constant height above ellipsoid (HAE) or constant height above ground level (AGL), otherwise known as flying "on drape." The former uses the altitude above an assumed perfect ellipsoid model of earth, while the latter uses the altitude on a drape surface above surface features, such as hills.

2.2.2 Fitting/Training Flight Data

As previously mentioned, primarily data from flights 1003-1007 is used in this work. Up to approximately 11 hr and 52 min of this data was used for model training, which is shown in Figure 2-4. This includes all flight data from flights 1003-1006, with the exception of flight line 1003.05, which had a data anomaly, and the held out testing data, which is discussed in section 2.2.3. For the neural network-based approaches, typically 3/17 of the training data was used for validation check-pointing.

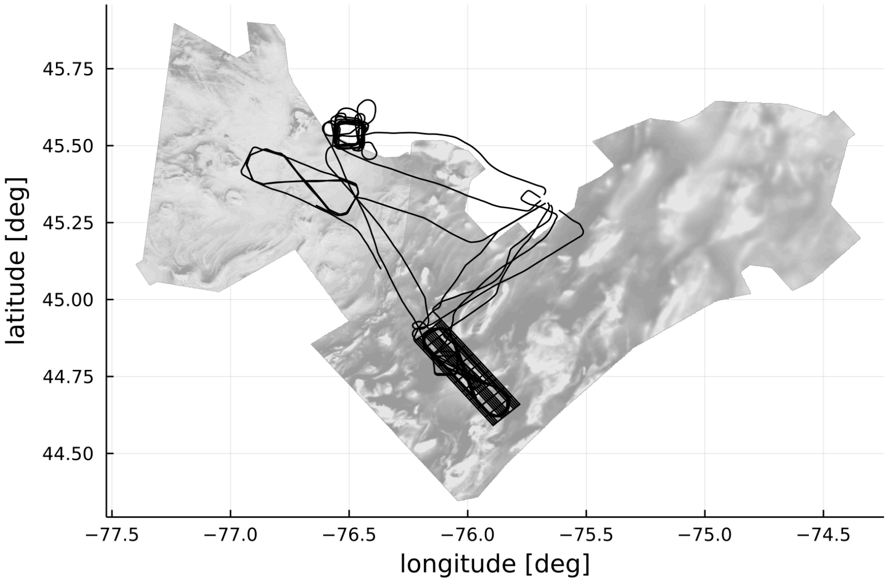


Figure 2-4: SGL training data flight lines. The total flight time was approximately 11 hr and 52 min.

However, not all of the flight data used for training is available for evaluating navigation performance, which is shown in Figure 2-5. This is due to some flight portions occurring outside of the available magnetic anomaly map areas, transitioning between the available magnetic anomaly maps, and significant (50 m or greater) altitude variations. These issues can be resolved by expanding the map areas with further geomagnetic surveying, merging the maps, and using a 3D map grid, respectively. However, given that 7 hr and 23 min of the flight data used for training remained navigation-capable, these solutions were not pursued and are considered out of scope for this work. Note that merging magnetic anomaly maps is not necessarily straightforward. The geomagnetic surveys carried out to create the maps were likely at different altitudes, and upward continuation causes errors, which are largest near the map edges, as explained in section 2.3.

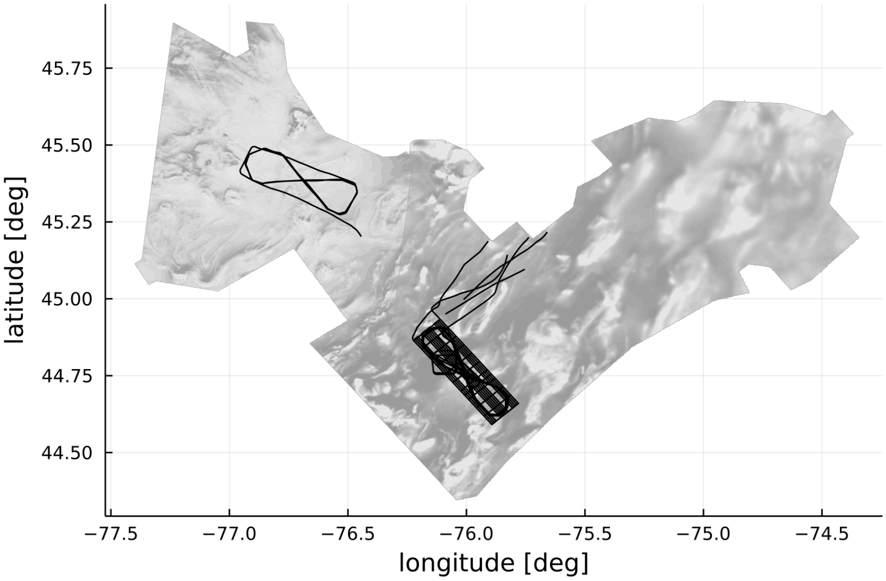


Figure 2-5: Navigation-capable SGL training data flight lines. The total flight time was approximately 7 hr and 23 min. Some flight data may not be used for navigation due to the lack of a map, crossing between maps, or large altitude variations.

Note that there are occasional dropped magnetic signals within the training data. It was determined that these did not have a significant impact on performance, except for flight line 1003.05, and were kept in the training data for this work. Additionally, a large portion of the training data contains straight and level flight lines from a mini-survey that was conducted during flights 1004 and 1005. It was thought that the training data may be biased, but performance was typically better keeping all available flight data.

2.2.3 Testing Flight Data

As previously mentioned, some data was held out of the training data for testing. This includes flights lines 1003.10, 4014.00, 4013.00, 4006.00, 4005.00, 1006.02, and 1007.04, which totalled 3 hr and 39 min. All of these are navigation-capable, except 1006.02, which was flown outside of the available map areas. The standard testing data in this work, a total of 3 hr and 23 min, does not include 1006.02. The flight data for these flight lines is not publicly released within MagNav.jl or otherwise, as this data is used for evaluation of performance for the Signal Enhancement for Magnetic Navigation Challenge Problem [25, 26, 27]. Thus, these flights lines are purposely not shown here, but it is worth noting they were all flown at nominally constant altitudes of less than 1000 m HAE.

For reproducibility, an alternative publicly released flight line, 1007.06, is used in various parts of this work, which is shown in Figure 2-6. This 1 hr and 27 min flight line is navigation-capable and was flown at nominally 400 m HAE in the Renfrew flight area, which is described in section 2.3.3.

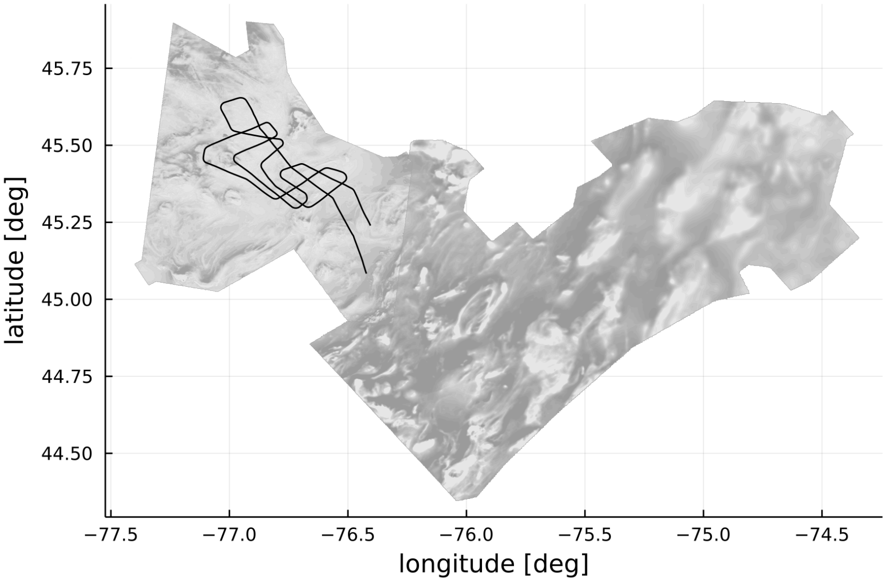


Figure 2-6: 1007.06 flight line.

2.3 Magnetic Anomaly Maps

A high-quality magnetic anomaly map is essential for high-accuracy MagNav. Many commercial and government organizations have performed geomagnetic surveys to varying degrees of precision and created magnetic anomaly maps. For example, a combination of magnetic data from near-surface, oceanic, and satellite measurements was used to create the World Digital Magnetic Anomaly Map (WDMAM), which was first published in 2007 [28]. Another publicly available map is the North American Magnetic Anomaly Database (NAMAD) [9]. However, both of these maps have a 1 km or greater cell spacing, which is not sufficient for high-accuracy MagNav. Instead, smaller maps with a tighter cell spacing are required.

The magnetic field strength decreases with the inverse cube of the distance from the magnetic source to the measurement location. At higher altitudes, magnetic features in the crust tend to blur together, which is analogous to low-pass filtering the magnetic anomaly field. Thus, to detect small magnetic variations in the crust, geo-surveys are typically flown as close to the ground as possible [4]. This also means navigation performance is altitude dependent due to physics that cannot be avoided [29, 30]. Similarly, faster flights generally perform better, since magnetic information is received at a higher rate.

After a low-altitude geomagnetic survey is completed, the data is compiled into a "drape" map, which contains the measurements at the survey altitude for each sample point, possibly with additional smoothing. It is often desirable to have the magnetic anomaly map at a higher, possibly constant, altitude where a flight would take place. This is possible using upward continuation, which is described next. This is followed by a description of downward continuation and a discussion of the maps available in MagNav.jl and used for this work.

2.3.1 Upward Continuation

Magnetic anomaly fields are potential fields, which obey Laplace's equation [31]. Using upward continuation, a potential field can be transformed from the measured surface to a potential field that would be measured on another surface at a higher altitude, further from the underlying magnetic sources [32]. Using Green's Identities, the upward continuation integral can be derived as

$$U(x, y, z_0 - \Delta z) = \frac{\Delta z}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{U(x', y', z_0)}{[(x - x')^2 + (y - y')^2 + \Delta z^2]^{3/2}} \, dx' dy' \tag{2.1}$$

where U is the potential field, x and y are horizontal position (typically east and north directions, respectively), and z is vertical position in the *downward* direction.

See [32] for the full derivation. For upward continuation, Δz is positive. (2.1) is only valid for a constant vertical offset Δz. Recall that the surveyed magnetic anomaly map data is not at a constant altitude, but instead on a drape over the earth's surface. The Cordell-Hildenbrand "chessboard" method may be used to obtain a level map at a higher altitude [33]. This is done by first upward continuing the original map to a series of maps at successively higher altitudes. Downward continuation, described in the next section, may also be needed for a series of maps at successively lower altitudes. The map at the desired altitude is then found by interpolating the series of maps in the vertical direction at each point on the horizontal grid.

The double two-dimensional (2D) integral in (2.1) indicates that each point on an upward-continued 2D map is dependent on each point from the original 2D map. This can be computationally-intensive for even small maps, so instead a more efficient frequency domain approach is used in practice. The Fourier-domain representation of (2.1) is

$$\mathcal{F}[U_{z_0-\Delta z}] = \mathcal{F}[U_{z_0}]\ \mathcal{F}[\psi] \tag{2.2}$$

where $\mathcal{F}[U]$ is the Fourier transform of the potential field and $\mathcal{F}[\psi]$ is an upward continuation (frequency) filter given by

$$\mathcal{F}[\psi] = e^{-\Delta z|k|} \tag{2.3}$$

where $|k|$ is the magnitude of the radial wavenumber (spatial frequency),

$$|k| = \sqrt{k_x^2 + k_y^2}. \tag{2.4}$$

Thus, an upward-continued map is determined by Fourier transforming the original map, multiplying by the upward continuation filter, and inverse Fourier transforming the product. Typically a discrete Fourier transform (DFT), or more specifically a fast Fourier transform (FFT) is used. From (2.3), it is apparent that the upward continuation filter attenuates high frequency (short wavelength) features more so than low frequency (long wavelength) features, acting similar to a low-pass filter.

As indicated by (2.1), upward continuation makes the assumption that the map is infinite in each of the horizontal directions. This assumption is clearly violated in practice when a finite size map is used. This results in errors, particularly along the edges of the map, which are dependent on unavailable data from outside the map area during upward continuation. Additionally, the Fourier transform assumes a periodic map, which is violated by the abrupt edge discontinuities of the map.

Both of these edge effects can be partially mitigated by carefully padding the original map prior to upward continuation. More specifically, each edge of the map is appended with interpolated data between each edge so that it "wraps-around." This is shown in Figure 2-7, where the original map is outlined in the center and the contours fan outward. The appended map can be interpolated to an arbitrary size, since the padded data is only used during upward continuation then discarded. The padding of each horizontal direction should be at least 10× the upward continuation distance to limit edge effects. It is computationally beneficial to pad the map in each direction to a smooth number (highly composite) size, allowing the use of the Cooley–Tukey FFT [34], which is done in MagNav.jl by default.

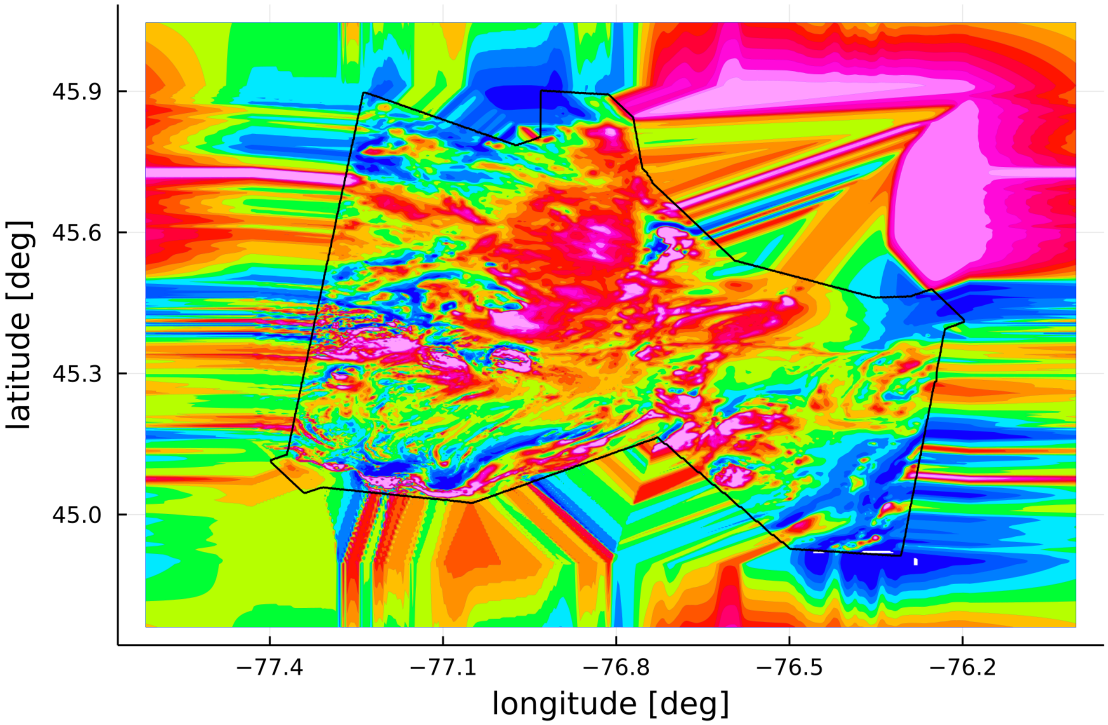


Figure 2-7: Magnetic anomaly map with padded edges.

2.3.2 Downward Continuation

Unlike upward continuation, which is a smoothing operation, downward continuation is an "unsmoothing" operation and is mathematically unstable [35]. This can be understood from (2.3), where negative Δz (for downward continuation) causes extremely high filter values for large radial wavenumbers $|k|$. Rather than attenuating, downward continuation *amplifies* high frequency features (i.e. noise) and leads to a non-unique solution.

One approach to reduce this instability issue is to use a low-pass filter during downward continuation. However, care must be taken when low-pass filtering to prevent unnecessarily removing too much of the higher frequency content. Tikhonov regularization [36], also known as ridge regression, is one commonly used approach for creating a low-pass filter. Modifying (2.2) and (2.3), the downward continuation Fourier transform is

$$\mathcal{F}[U_{z_0-\Delta z}] = \mathcal{F}[U_{z_0}]\frac{e^{-\Delta z|k|}}{1+\alpha k^2 e^{-\Delta z|k|}} \tag{2.5}$$

where α is a regularization parameter and Δz is negative for downward continuation [37]. The optimal α can be found with an L-curve approach [36]. First, downward continuation is performed for a geometric sequence of α. The optimal α is the local minimum on the characteristic curve of the L_∞ norm of the difference between successive downward continuations at each α. An extension of this is using a cutoff wavenumber that divides the potential field into low-pass filtered and unfiltered portions [38].

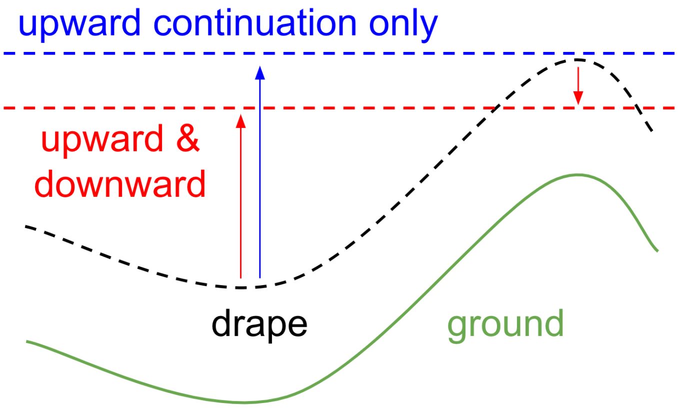


Figure 2-8: Upward and downward continuation. Some constant altitude (HAE) maps only require upward continuation, while others may require both upward and downward continuation.

Note that most geomagnetic surveys are performed near the minimum flying altitude that is able to maintain terrain clearance [39]. Any magnetic anomaly map needed for MagNav should be at or above the survey altitude, so there is limited need for downward continuation. The primary use of downward continuation is to generate a low-altitude constant altitude map where portions of the map need to be slightly downward continued as part of the Cordell-Hildenbrand “chessboard” method, which is shown in Figure 2-8. For this purpose, $\alpha = 200$ is appropriate for downward continuation of 10’s of meters for the maps used in this work. Finally, in addition to the errors introduced with upward and/or downward continuation, other maps errors may be present due to under-sampling and/or smoothing.

2.3.3 Ottawa Area Maps

The flights described in section 2.2 were flown in three different regions, which are shown in Figure 2-9. The northernmost region is the figure of merit (FOM) flight area, which is the typical location for SGL calibration flights. The easternmost region is known as the Eastern Ontario flight area, while the westernmost region is known as the Renfrew flight area. The magnetic anomaly maps for the Renfrew and Eastern Ontario regions were obtained from publicly available Ontario Geological Survey data, with some further processing by SGL [40, 41]. In particular, SGL re-sampled these maps from approximately 400 m (Eastern Ontario) and 200 m (Renfrew) to 30 m cell spacing using minimum curvature interpolation.

However, as previously discussed, “drape” maps were provided by SGL. Many processing steps were required to prepare these maps, Eastern Ontario and Renfrew, for MagNav. All of these steps were completed using MagNav.jl. First, the edges of the map without any data were removed for less memory use and faster upward continuation. Next, the IGRF (core field) was subtracted off, based on the original map survey date (October 20, 2013), leaving only the nearly static anomaly field (temporal variations were previously subtracted off during geomagnetic survey processing).

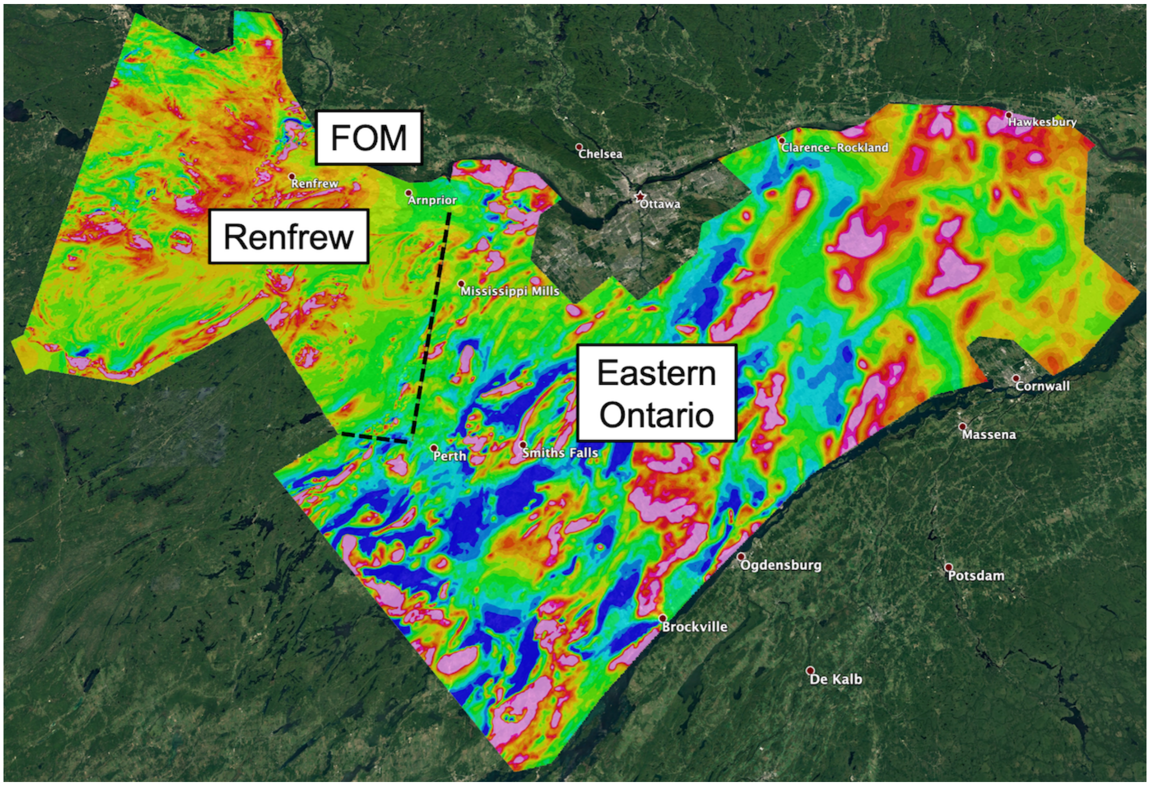


Figure 2-9: Magnetic anomaly maps near Ottawa, Ontario, Canada. The western region is the Renfrew flight area. The eastern region is the Eastern Ontario flight area. The northern region is the figure of merit (FOM) flight area.

Missing data areas (i.e. gaps) were filled using a k-nearest neighbors (KNN) algorithm. This results in better upward continuation and interpolation performance. The Cordell-Hildenbrand "chessboard" method was then carried out, which included upward and slightly downward continuing the maps to multiple levels and then vertically interpolating. During upward (and downward) continuation, the map was temporarily padded using the process described in section 2.3.1 to limit edge effects. Finally, the map was re-gridded from Universal Transverse Mercator (UTM) to latitude, longitude, and altitude (LLA) coordinates.

Table 2.4: Ottawa area maps. The WGS-84 coordinate system is used.

Name	Description
Eastern_395	Eastern Ontario at 395 m HAE
Eastern_drape	Eastern Ontario on drape
Eastern_plot	Eastern Ontario on drape, no fill
Renfrew_395	Renfrew at 395 m HAE
Renfrew_555	Renfrew at 555 m HAE
Eastern_drape	Renfrew on drape
Eastern_plot	Renfrew on drape, no fill

The end result is several magnetic anomaly maps, which are used for this work and available in MagNav.jl. These maps, listed in Table 2.4, are suitable for MagNav with any of the flight data discussed in section 2.2. As with GPS, these maps use the World Geodetic System (WGS) coordinate system, specifically WGS-84. Note that the highest available map should be used when possible, such as Renfrew_555 rather than Renfrew_395, as Renfrew_395 was generated using mild downward continuation and is slightly less accurate. Renfrew_555 only used upward continuation.

It may be of interest to understand how well these maps work with the available flight data. Shown in Figure 2-10 are the magnetic anomaly field errors plotted against the altitude errors for each navigation-capable flight line in the 2020 SGL flight dataset. Each flight line is compared to the corresponding map that would be used for navigation. For the HAE flight lines, the map was upward continued to the mean flight line altitude. The pilot was able to keep the aircraft at constant altitude or on the drape surface within approximately 10 m in most cases. This resulted in roughly 5 nT of error for the HAE flights and 10 nT of error for the drape flights. Of course, the drape flights were at lower altitudes, where the magnetic field strength is larger and thus the errors are larger.

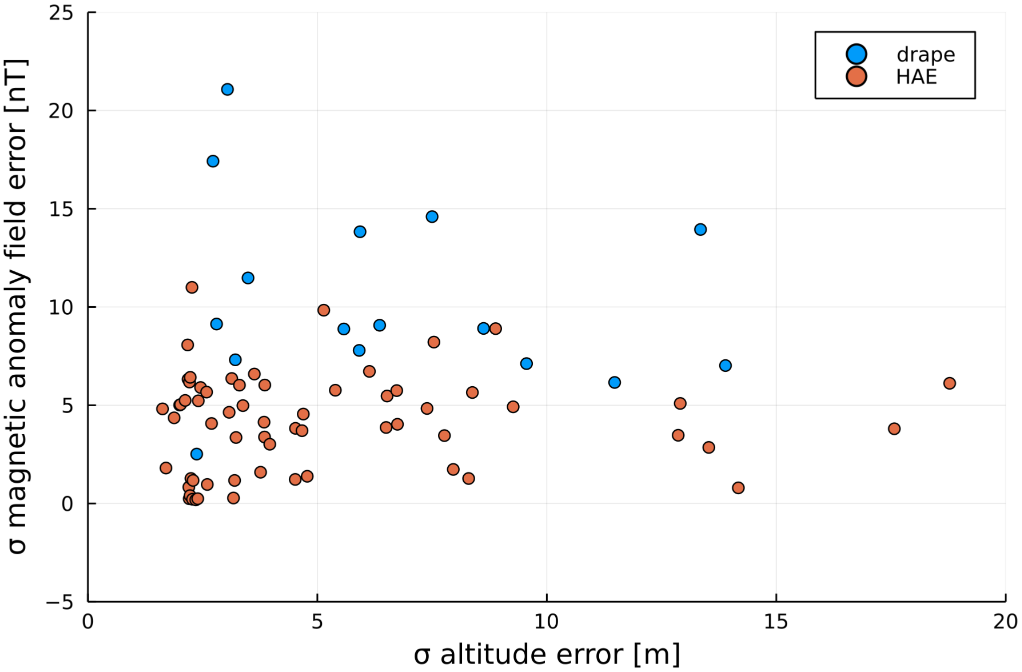


Figure 2-10: HAE and drape map error comparison.

Note that, again due to edge effects, there is a general rule of thumb that magnetic anomaly maps should only be upward continued approximately 1/50× the map size. The maps used in this work were created from flight data measured at nominally 275 m (Eastern Ontario) and 314 m (Renfrew). The map areas are approximately $1.24 \times 10^{10}\,\mathrm{m}^2$ and $6.16 \times 10^{9}\,\mathrm{m}^2$ for Eastern Ontario and Renfrew, respectively. Together, this means the highest altitude for reliable map data is approximately 2500 m for the Eastern Ontario map and 1900 m for the Renfrew map.

2.4 Navigation Algorithms

A magnetic navigation system collects magnetic field data using a magnetometer and uses magnetic anomaly maps to determine the current location by correcting the drift from an INS. MagNav.jl contains multiple navigation algorithms for this purpose. The simplest is an extended Kalman filter (EKF), which is described in section 5.2.2. A Rao-Blackwellized particle filter (RBPF), a type of marginalized particle filter (MPF), developed by Canciani [13] is also available, but is not used in this work. Multiple recently introduced and still in development algorithms are also available. These include an EKF with online learning of Tolles-Lawson coefficients (section 5.5), an EKF with online learning of neural network weights (section 5.6), and a measurement noise covariance-adaptive neural extended Kalman filter (section 5.9.3).

Chapter 3

Linear Aeromagnetic Compensation

Magnetic measurements are only capable of measuring the total magnetic field. The desired magnetic anomaly field must be extracted from the total magnetic field measurement, which contains four main components,

$$\vec{B}_{\text{measured}} = \vec{B}_{\text{core}} + \vec{B}_{\text{anomaly}} + \vec{B}_{\text{tv}} + \vec{B}_{\text{aircraft}}. \tag{3.1}$$

As previously discussed, the earth field components (core field, anomaly field, and temporal variations) of the total field measurement are approximately known for a given position and time. However, the total field measurement is also corrupted with an aircraft interference field. The goal of aeromagnetic compensation is to remove this undesired magnetic source.

This chapter first describes the classical Tolles-Lawson aeromagnetic compensation model in detail. This is followed by the discussion of two variations of the Tolles-Lawson model and two additional linear aeromagnetic compensation approaches. The available fitting or training output targets and compensation performance metrics are then introduced. This chapter concludes with a comparison of the compensation performance of these five models using only data from a single scalar and vector magnetometer pair.

3.1 Classical Tolles-Lawson Aeromagnetic Compensation

Figure 3-1: Magnetometer measurements on a tail stinger.

State-of-the-art aeromagnetic compensation is performed using the Tolles-Lawson model [42, 43, 44], which is a linear model that uses measurements from a vector magnetometer to remove aircraft magnetic field contributions to scalar magnetometer measurements, which are used for navigation. For a geo-survey aircraft, such as that shown in Figure 3-1, this compensation procedure is sufficient to remove nearly all of the corruption due to the aircraft magnetic field if the aircraft is flying in a "magnetically quiet" mode (i.e. limited control surface movements, radio transmissions, etc.). Additionally, a geo-survey aircraft often uses magnetic noise reduction methods, including coil or ring compensation, magnetic shielding, degaussing, and/or optimal placement or removal of magnetic sources [45, 46, 47, 48]. The measurements on a geo-survey aircraft are taken on a tail stinger, far away (nominally at least 3 m) from the greatest corruption sources [49]. Flights using this type of aircraft and the Tolles-Lawson model are able to obtain navigation accuracies of 10's of meters [2].

However, using tail stinger measurements on operational aircraft is impractical, as is having the navigation system impose a "magnetically quiet" flight limitation. The magnetic dynamics of operational aircraft exceed the capabilities of the Tolles-Lawson model, meaning this model does not produce a compensated signal with sufficiently accurate results when the magnetometer is close to the aircraft interference sources (e.g. engine), as shown in Figure 3-2. Thus, the largest remaining challenge for operational airborne magnetic anomaly navigation is the compensation of magnetic measurements on operational aircraft.

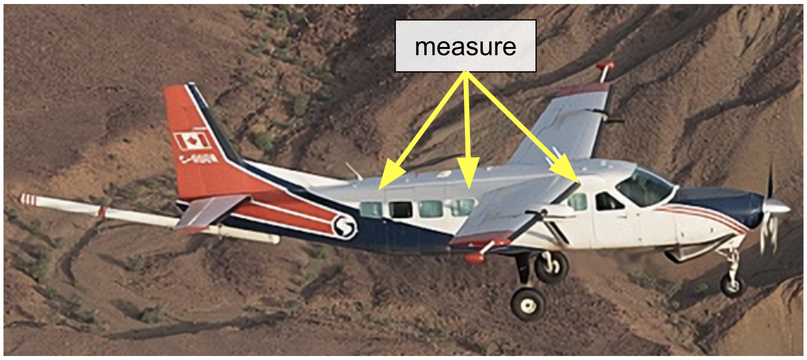


Figure 3-2: Magnetometer measurements within the cabin.

Though the Tolles-Lawson model is not sufficient for aeromagnetic compensation on operational aircraft, it is useful to understand how the model works in order to improve on it. In particular, the vector magnetometer (***A*** matrix) terms are heavily used in the later chapters of this work as features for various models. The remainder of this section describes the origin of the Tolles-Lawson model and how the model coefficients are determined.

3.1.1 Derivation of the Tolles-Lawson Model

Tolles and Lawson first reported their linear aeromagnetic compensation model for scalar magnetometers in 1950 [42], though it was developed earlier during World War II. Tolles was later issued patents for the hardware involved in airborne magnetometer compensation and the primary initial use case was magnetic anomaly detection (MAD) [43, 44]. Leliak later proposed performing sinusoidal maneuvers during a calibration flight to increase observability of the terms in the Tolles-Lawson model [45]. This has been the state-of-the-art for decades, though numerous improvements to this method have been proposed over the years, as discussed in section 3.5. The basic idea of the Tolles-Lawson model [50] is to use magnetic measurements from a vector magnetometer to calibrate a scalar magnetometer, the latter of which is used for navigation. This model provides a means for removing a corrupting aircraft magnetic field from a scalar total magnetic field measurement, yielding the earth magnetic field used for navigation.

An airborne vector magnetometer measures the vector sum of two primary magnetic fields,

$$\vec{B}_t = \vec{B}_e + \vec{B}_a \tag{3.2}$$

where $\vec{B}_t$ is the total field, $\vec{B}_e$ is the earth (external) field and $\vec{B}_a$ is the unknown aircraft (interference) field. Note that here earth field refers to all components – core field, anomaly field, and temporal variations. A vector magnetometer measures $\vec{B}_t$, but for navigation the unknown, desired signal is $|\vec{B}_e|$, the magnitude of $\vec{B}_e$. A scalar magnetometer measures $|\vec{B}_t|$, the magnitude of $\vec{B}_t$. These terms can be related as follows:

$$|\vec{B}_e|^2 = \vec{B}_e \cdot \vec{B}_e = (\vec{B}_t - \vec{B}_a) \cdot (\vec{B}_t - \vec{B}_a) \tag{3.3}$$

$$|\vec{B}_e|^2 = \vec{B}_t \cdot \vec{B}_t - 2\vec{B}_t \cdot \vec{B}_a + \vec{B}_a \cdot \vec{B}_a \tag{3.4}$$

$$|\vec{B}_e| = \sqrt{|\vec{B}_t|^2 - 2\vec{B}_t \cdot \vec{B}_a + |\vec{B}_a|^2} \tag{3.5}$$

$$|\vec{B}_e| = |\vec{B}_t|\sqrt{1 - 2\frac{\vec{B}_t \cdot \vec{B}_a}{|\vec{B}_t|^2} + \frac{|\vec{B}_a|^2}{|\vec{B}_t|^2}} \tag{3.6}$$

To get the Tolles-Lawson model, it is assumed that $|\vec{B}_a|^2/|\vec{B}_e|^2$ is negligible, i.e. the aircraft field is small compared to the earth field. This leaves

$$|\vec{B}_e| \approx |\vec{B}_t|\sqrt{1 - 2\frac{\vec{B}_t \cdot \vec{B}_a}{|\vec{B}_t|^2}} \tag{3.7}$$

which can be linearized using the series expansion

$$\sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} - \dots \tag{3.8}$$

to first order to get

$$|\vec{B}_e| \approx |\vec{B}_t| - \frac{\vec{B}_t \cdot \vec{B}_a}{|\vec{B}_t|}. \tag{3.9}$$

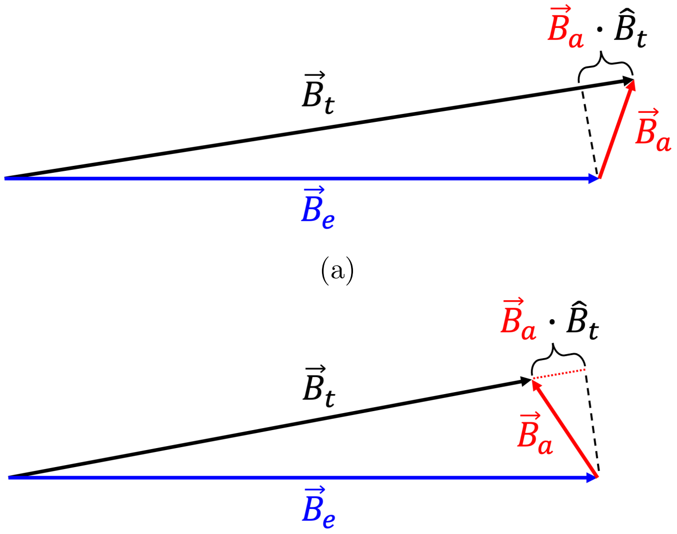
The magnetic field components from the vector magnetometer are used to compute the total field direction cosines,

$$\hat{B}_t = \frac{\vec{B}_t}{|\vec{B}_t|}. \tag{3.10}$$

Using this definition, (3.9) becomes

$$|\vec{B}_e| \approx |\vec{B}_t| - \vec{B}_a \cdot \hat{B}_t \tag{3.11}$$

where $|\vec{B}_e|$ is the magnitude of the earth field (desired signal for navigation), $|\vec{B}_t|$ is the (measured) total field, and $|\vec{B}_a|$ is the aircraft field. Note that $\vec{B}_a \cdot \hat{B}_t$ is a corruption term, i.e. the effect of the aircraft field projected onto the total field. This can be visualized as in Figure 3-3.



(a)

(b)

Figure 3-3: Total, earth, and aircraft magnetic field vectors. The same aircraft magnetic field (magnitude) can be (a) additive or (b) subtractive when projected onto the total magnetic field, depending on the relative orientation between the aircraft and earth magnetic fields during aircraft maneuvers.

Up to this point, no physics knowledge has been incorporated. The derivation comes solely from manipulating vectors and making an assumption about the magnitude of those vectors. In order to get to the final model, Tolles and Lawson [42] assumed that the aircraft field is comprised of permanent, induced, and eddy current magnetic moments

$$\vec{B}_a = \vec{B}_{\text{perm}} + \vec{B}_{\text{ind}} + \vec{B}_{\text{eddy}} \tag{3.12}$$

which is further assumed to take the form of

$$\vec{B}_a = \boldsymbol{a} + \boldsymbol{b}\vec{B}_t + \boldsymbol{c}\dot{\vec{B}}_t \tag{3.13}$$

where coefficient vector $\boldsymbol{a}$ and coefficient matrices $\boldsymbol{b}$ and $\boldsymbol{c}$ are all unknown. The permanent magnetic moment terms,

$$\vec{B}_{\text{perm}} = \boldsymbol{a} = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}^T, \tag{3.14}$$

contain 3 unknown coefficients. These permanent magnetic moment terms represent nearly constant, permanent magnetization of various ferromagnetic aircraft components, including both the aircraft itself and items within the aircraft [50]. These terms do not change unless the aircraft configuration or contents are modified. The induced magnetic moment terms,

$$\vec{B}_{\text{ind}} = \boldsymbol{b}\vec{B}_t = |\vec{B}_t| \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \hat{B}_t, \tag{3.15}$$

contain 9 unknown coefficients. These induced magnetic moment terms represent the earth field inducing a secondary magnetic field in magnetically susceptible aircraft components. The relative orientation of the aircraft and earth field determines the magnitude and direction of the induced magnetization. Since much of the aircraft structure is comprised of non-magnetic aluminum alloys, the primary source of induced magnetic fields are the aircraft engines [17]. The eddy current terms,

$$\vec{B}_{\text{eddy}} = \boldsymbol{c}\dot{\vec{B}}_t = |\vec{B}_t| \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \dot{\hat{B}}_t, \tag{3.16}$$

contain 9 unknown coefficients. These eddy current terms represent electrical current loops caused by the time-varying earth field (relative to the aircraft) interacting with electrically conductive aircraft components. Unlike the permanent and induced fields, eddy currents depend on the time rate of change of earth's magnetic flux through these components, such as the aircraft skin [42, 45]. Magnetic fields created by eddy currents obey Lenz's Law, opposing the magnetic field that created them [13]. This is similar to how current is produced in a coil rotating in a uniform magnetic field [50]. The form of the aircraft generated moments (corruption term) in (3.11) becomes

$$\vec{B}_a \cdot \hat{B}_t = (\boldsymbol{a} + |\vec{B}_t| \ \boldsymbol{b} \ \hat{B}_t + |\vec{B}_t| \ \boldsymbol{c} \ \dot{\hat{B}}_t) \cdot \hat{B}_t. \tag{3.17}$$

There are a total of 21 coefficients in $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$, but due to symmetry in the induced magnetic moment matrix $\boldsymbol{b}$, the repeated off-diagonal terms are removed resulting in 3 fewer coefficients. Thus, there are 18 total unknown coefficients in the standard Tolles-Lawson model,

$$|\vec{B}_e| \approx |\vec{B}_t| - \left(\hat{B}_t^T \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + |\vec{B}_t| \ \hat{B}_t^T \begin{bmatrix} \beta_4 & \beta_5 & \beta_6 \\ \cdot & \beta_7 & \beta_8 \\ \cdot & \cdot & \beta_9 \end{bmatrix} \hat{B}_t + |\vec{B}_t| \ \hat{B}_t^T \begin{bmatrix} \beta_{10} & \beta_{11} & \beta_{12} \\ \beta_{13} & \beta_{14} & \beta_{15} \\ \beta_{16} & \beta_{17} & \beta_{18} \end{bmatrix} \dot{\hat{B}}_t \right), \tag{3.18}$$

which has unknowns on both sides of the equation, $|\vec{B}_e|$ and $\boldsymbol{\beta}$, since only the total field $\vec{B}_t$ and $|\vec{B}_t|$ can be directly measured. However, using a "trick" (3.18) can be modified into a solvable form [50].

3.1.2 Solving for the Tolles-Lawson Model Coefficients

Rewriting (3.18), first a length 18 row vector of direction cosine terms, which are calculated from a vector magnetometer measurement, is created as

$$\vec{\delta} = \begin{bmatrix} \hat{B}_{\;3\times1} \\ \mathrm{vec}(|\vec{B}|\ \hat{B}\hat{B}^T)_{6\times1} \\ \mathrm{vec}(|\vec{B}|\ \hat{B}\dot{\hat{B}}^T)_{9\times1} \end{bmatrix}^T \tag{3.19}$$

where again only 6 of induced magnetic moment terms are taken from $|\vec{B}|\ \hat{B}\hat{B}^T$ due to symmetry. Explicitly, the 18 direction cosine terms are

$$\vec{\delta} = \begin{bmatrix} \hat{B}_x \\ \hat{B}_y \\ \hat{B}_z \\ |\vec{B}|\hat{B}_x\hat{B}_x \\ |\vec{B}|\hat{B}_x\hat{B}_y \\ |\vec{B}|\hat{B}_x\hat{B}_z \\ |\vec{B}|\hat{B}_y\hat{B}_y \\ |\vec{B}|\hat{B}_y\hat{B}_z \\ |\vec{B}|\hat{B}_z\hat{B}_z \\ |\vec{B}|\hat{B}_x\dot{\hat{B}}_x \\ |\vec{B}|\hat{B}_x\dot{\hat{B}}_y \\ |\vec{B}|\hat{B}_x\dot{\hat{B}}_z \\ |\vec{B}|\hat{B}_y\dot{\hat{B}}_x \\ |\vec{B}|\hat{B}_y\dot{\hat{B}}_y \\ |\vec{B}|\hat{B}_y\dot{\hat{B}}_z \\ |\vec{B}|\hat{B}_z\dot{\hat{B}}_x \\ |\vec{B}|\hat{B}_z\dot{\hat{B}}_y \\ |\vec{B}|\hat{B}_z\dot{\hat{B}}_z \end{bmatrix}^T \tag{3.20}$$

where $\hat{B}_x$, $\hat{B}_y$, and $\hat{B}_z$ are the direction cosines. A time series of $\vec{\delta}$ can be composed into an $N \times 18$ matrix

$$\boldsymbol{A} = \begin{bmatrix} \vec{\delta}_1 \\ \vdots \\ \vec{\delta}_N \end{bmatrix} \tag{3.21}$$

where each row is one of N time steps. The column vector of Tolles-Lawson coefficients to learn is $\boldsymbol{\beta}$, as taken from (3.18). Rearranging and substituting,

$$\boldsymbol{B}_{\text{scalar}} - |\vec{\boldsymbol{B}}_e| = \boldsymbol{A}\beta \tag{3.22}$$

where $|\vec{\boldsymbol{B}}_e|$ and $\boldsymbol{\beta}$ are both still unknown and $\boldsymbol{B}_{\text{scalar}}$ is scalar magnetometer measurements that represent $|\vec{\boldsymbol{B}}_t|$. The "trick" is to use a bandpass finite impulse response filter (bpf) on (3.22) [51],

$$\text{bpf}(\boldsymbol{B}_{\text{scalar}} - |\vec{\boldsymbol{B}}_e|) = \text{bpf}(\boldsymbol{A}\boldsymbol{\beta}). \tag{3.23}$$

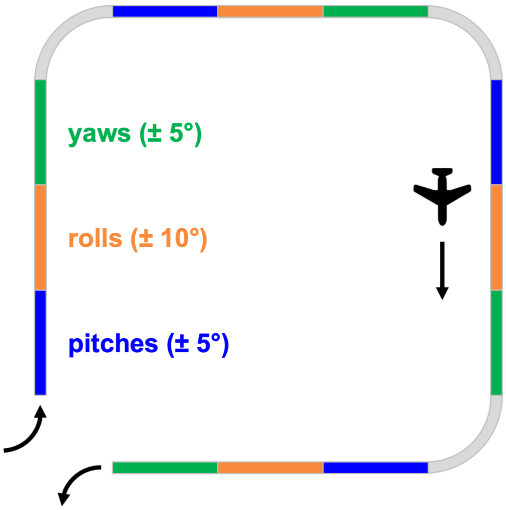


Figure 3-4: Typical aeromagnetic calibration flight pattern. A box-like flight path with tight turns is flown with pitch, roll, and yaw maneuvers (in any order) performed along each leg [52].

The passband frequency range for the bandpass filter is carefully selected in order to remove nearly all of the earth field while keeping much of the aircraft field. In practice, a passband of 0.1-0.9 Hz has been found to perform well, since in this range the frequency content of the aircraft dominates the magnetic signal. The measurements themselves are taken during a specific set of roll, pitch, and yaw aircraft maneuvers during a calibration flight, as shown in Figure 3-4. Roll, pitch, and yaw, as shown in Figure 3-5, are the Euler angles that describe the aircraft orientation in reference to the earth.

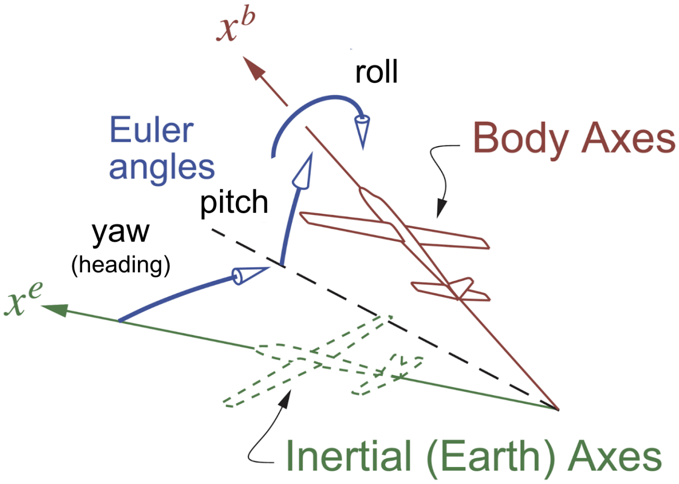


Figure 3-5: Euler angles. Modified from [53].

These maneuvers, are meant to "inject" aircraft magnetic field content into the passband frequency range. The calibration flight is performed at a high altitude over a region with a small magnetic gradient, as shown in Figure 3-6, to reduce the uncertainty imparted by the earth field [54, 52].

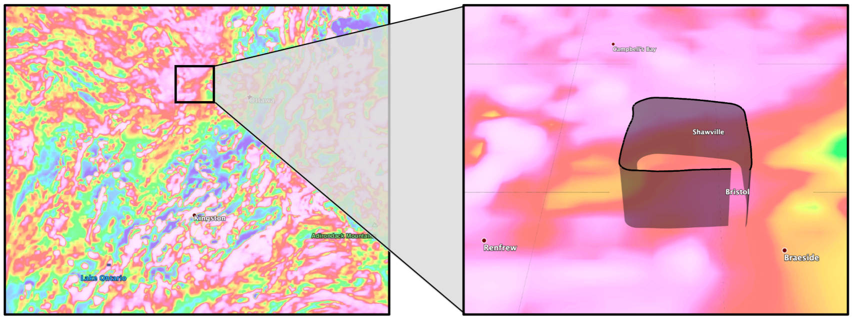


Figure 3-6: SGL calibration flight region. A magnetic anomaly map from [9] is shown with an enlarged region that contains a relatively small magnetic gradient. This is where Sanders Geophysics Ltd. (SGL) typically conducts calibration flights at 10,000 ft (3048 m), an example of which is illustrated with a black flight path.

With $\mathrm{bpf}(|\vec{\boldsymbol{B}}_e|) \approx 0$, (3.23) becomes

$$\mathrm{bpf}(\boldsymbol{B}_{\mathrm{scalar}}) = \mathrm{bpf}(\boldsymbol{A})\boldsymbol{\beta} \tag{3.24}$$

where $\boldsymbol{B}_{\mathrm{scalar}}$ is known from scalar magnetometer measurements and $\boldsymbol{A}$ is known from vector magnetometer measurements, as previously described. The Tolles-Lawson coefficients $\boldsymbol{\beta}$ can then be solved for with linear least squares regression,

$$\boldsymbol{\beta} = (\boldsymbol{A}_f^T \boldsymbol{A}_f)^{-1} \boldsymbol{A}_f^T \boldsymbol{y} \tag{3.25}$$

or ridge regression,

$$\boldsymbol{\beta} = (\boldsymbol{A}_f^T \boldsymbol{A}_f + \lambda \boldsymbol{I})^{-1} \boldsymbol{A}_f^T \boldsymbol{y} \tag{3.26}$$

where $\boldsymbol{y} = \mathrm{bpf}(\boldsymbol{B}_{\mathrm{scalar}})$, $\boldsymbol{A}_f = \mathrm{bpf}(\boldsymbol{A})$, and λ is a ridge parameter. Ridge regression is useful when $\boldsymbol{A}_f^T \boldsymbol{A}_f$ is poorly conditioned due to correlations among the Tolles-Lawson coefficients [55]. A ridge parameter of $\lambda = 0.00025$ is appropriate for this type of problem [56]. For the calibration flights used in this work, $\lambda = 0.025$ is appropriate, as discussed in the next section. Though the coefficients are determined using bandpass filtered measurements, they can be applied to unfiltered measurements. During a measurement or navigation (non-calibration) flight, compensation is then performed as

$$|\vec{\boldsymbol{B}}_e| = \boldsymbol{B}_{\text{scalar}} - \boldsymbol{A}\boldsymbol{\beta} \tag{3.27}$$

where the $\boldsymbol{\beta}$ Tolles-Lawson coefficients represent the average aircraft magnetic field contributions predetermined during a calibration flight and $\boldsymbol{B}_{\text{scalar}}$ and $\boldsymbol{A}$ come from real-time scalar and vector magnetometer measurements, respectively.

3.2 Tolles-Lawson Ridge Parameter and Calibration Flight Selection

As mentioned in the previous section, ridge regression is commonly used with the Tolles-Lawson model, but the ridge parameter must be determined. Fortunately, the flight dataset used in this work has 22 calibration flight lines that can be leveraged to determined an appropriate ridge parameter. However, two of these flight lines were flown at low altitude with a high magnetic gradient and are excluded. An example of this real calibration flight data is shown in Figure 3-7.

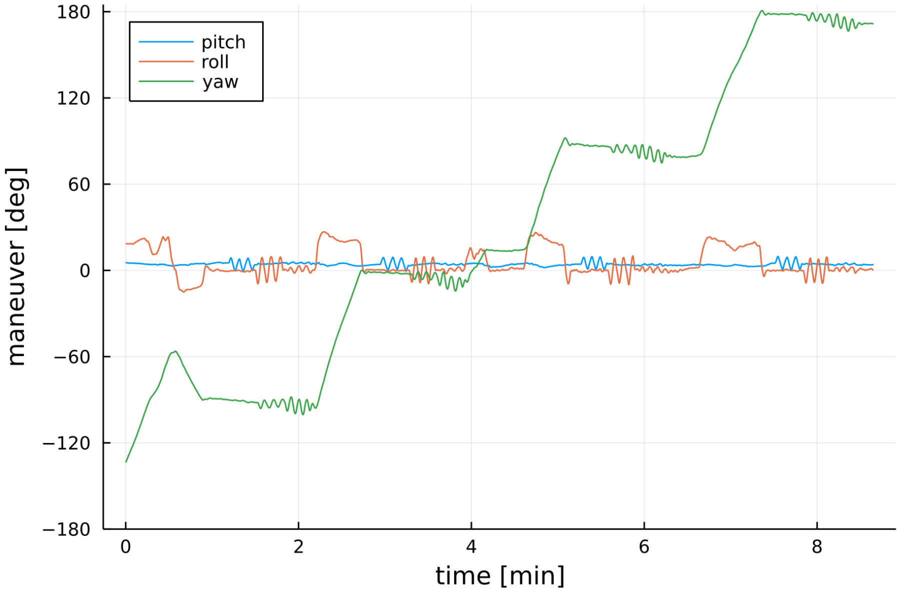


Figure 3-7: Flight line 1002.02 calibration maneuvers. The maneuver amplitudes are approximately 4.5° pitch, 9° roll, and 5° yaw. The period of each is around 6 sec.

Table 3.1: Comparison of Tolles-Lawson coefficients without ridge regression. Only the permanent and induced terms are used here for simplicity.

Flight Line	Box #	cosX	cosY	cosZ	cosXX	cosXY	cosXZ	cosYY	cosYZ	cosZZ
1001.01	4	1974	111	-1223	-1013	32	-167	-1177	-18	-1564
1002.02	2	1798	-159	-964	-787	-189	-137	-656	180	-1024
1006.04	1	1243	424	-937	-1132	348	147	-317	190	-921

As shown in Table 3.1, the Tolles-Lawson coefficients without ridge regression can vary widely. One approach to determine an appropriate ridge parameter is to use the coefficient of variation,

$$c_v = \frac{\sigma}{\mu} \tag{3.28}$$

where σ is the standard deviation and μ is the mean. The coefficient of variation for each Tolles-Lawson coefficient across the various calibration flights can be calculated using (3.28). The average absolute value of these is plotted against the ridge parameter used to calculate the Tolles-Lawson coefficients in Figure 3-8. Here it can be seen that large ridge parameters are very stable, while small ridge parameters can be unstable, depending on the Tolles-Lawson terms that are included. The spikes in Figure 3-8 are due to some Tolles-Lawson coefficients, particularly eddy current `z` terms, switching signs. Note that Flux D is used here, while Flux A is used in most other sections of this work. Flight 1002 had a technical issue with Flux A. Flight 1002 also contained eight calibration flight lines, so it was decided to use Flux D instead so that more calibration flight lines could be used. As explained in section 2.2.1, compensation performance is similar for all available vector magnetometers.

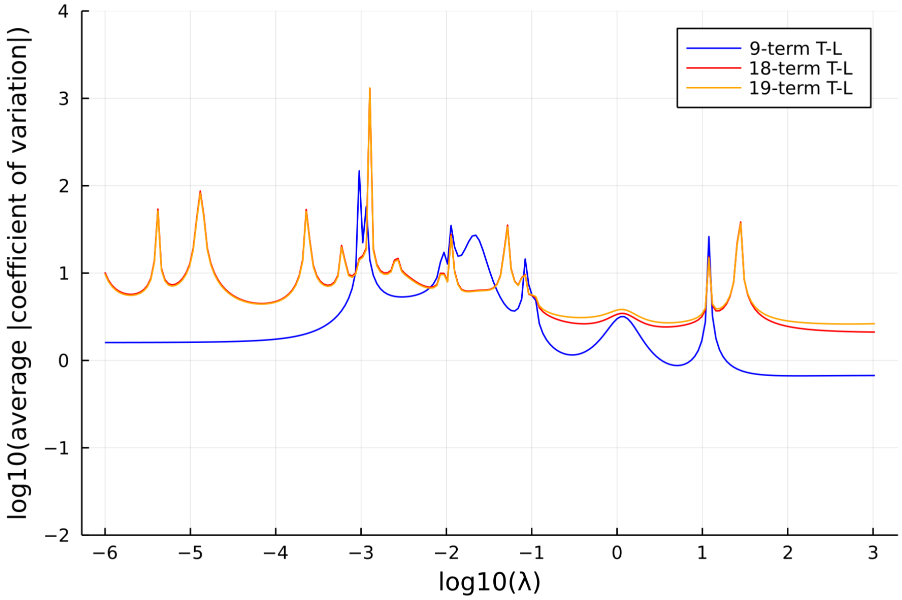


Figure 3-8: Coefficient of variation of Tolles-Lawson coefficients vs ridge parameter. Mag 4 and Flux D are used here with 20 calibration flight lines. "9-term T-L" uses permanent and induced terms, "18-term T-L" further includes eddy current terms, and "19-term T-L" further includes a bias term.

Now looking at the compensation error, as shown in Figure 3-9, using too large of a ridge parameter can result in excessive compensation errors as well. Thus, there is a sweet spot for the ridge parameter at around 10^{-1}. Specifically, $\lambda = 0.025$ was selected as the baseline ridge parameter in this work. As shown in Table 3.2, the Tolles-Lawson coefficients still vary with $\lambda = 0.025$, but the magnitudes have decreased and the values across the flight lines are more similar than in Table 3.1.

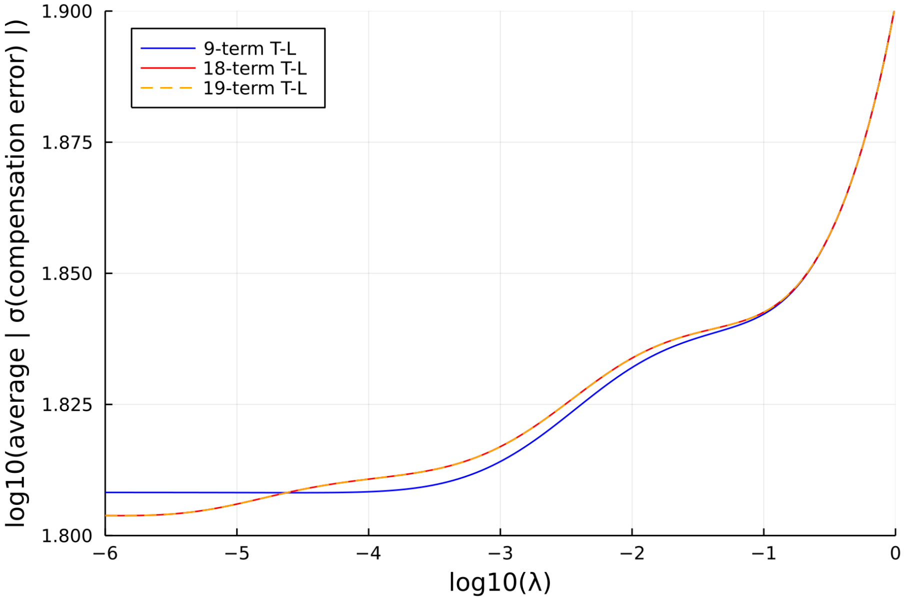


Figure 3-9: Compensation error vs ridge parameter. Mag 4 and Flux D are used here with 20 calibration flight lines. "9-term T-L" uses permanent and induced terms, "18-term T-L" further includes eddy current terms, and "19-term T-L" further includes a bias term.

Navigation is not covered until Chapters 5 and 6, but for completion, using different sets of Tolles-Lawson coefficients (determined using ridge regression) results in only small changes in performance for a well-tuned navigation filter. For example, using the coefficients in Table 3.2 results in navigation errors of 43-49 m DRMS on flight line 1007.06. Similarly, there was no change (1 m or less) in navigation performance

Table 3.2: Comparison of Tolles-Lawson coefficients with ridge regression. Only the permanent and induced terms are used here for simplicity. A ridge parameter of $\lambda = 0.025$ was used.

Flight Line	Box #	cosX	cosY	cosZ	cosXX	cosXY	cosXZ	cosYY	cosYZ	cosZZ
1001.01	4	150	74	-832	-341	23	275	322	0	-19
1002.02	2	305	-48	-865	-462	-56	261	392	178	42
1006.04	1	333	84	-857	-600	41	255	514	168	55

when decreasing the ridge parameter from $\lambda = 0.025$ to $\lambda = 0.00025$, as was done in [56], while holding all other parameters constant. Increasing the ridge parameter to $\lambda = 0.25$ marginally increased the navigation error by 2-3 m. As a final related note, typically the magnetometer data is downsampled to 10 Hz, as shown in Figure 3-10. This can introduce smoothing errors, particularly in the vector magnetometer components (e.g. 10 nT in B_x). Using the raw, downsampled data without smoothing had a negligible impact on compensation and navigation performance and is not considered further.

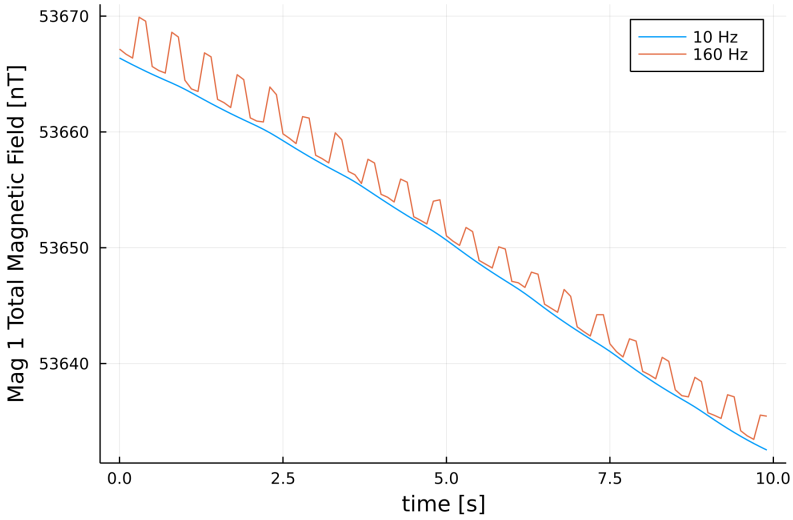


Figure 3-10: 10 Hz vs 160 Hz uncompensated scalar magnetometer data.

3.3 Modified Tolles-Lawson Aeromagnetic Compensation

The first variation of the classical Tolles-Lawson model is termed modified Tolles-Lawson aeromagnetic compensation [20]. The small change from the original model is the scaling of the total field direction cosines,

$$\hat{B}_t = \frac{\vec{B}_t}{B_{\text{scalar}}} \tag{3.29}$$

where the magnetic field components from a vector magnetometer, $\vec{B}_t$, are now scaled by a scalar magnetometer measurement, B_{scalar}, rather than the vector norm shown in (1.2), which would typically come from a vector magnetometer. Again, these direction cosines are used to form the $\boldsymbol{A}$ matrix, an essential part of the Tolles-Lawson model(s).

3.4 Map-Based Tolles-Lawson Aeromagnetic Compensation

The second, more involved variation of the classical Tolles-Lawson model is termed map-based Tolles-Lawson aeromagnetic compensation [20]. Here, rather than using a bandpass filter, the earth field "truth" is determined from a magnetic anomaly map. With the earth field known, (3.22) can be modified as

$$\boldsymbol{B}_{\text{scalar}} - \boldsymbol{B}_{\text{map}} = \boldsymbol{A}\boldsymbol{\beta} \tag{3.30}$$

where $\boldsymbol{B}_{\text{scalar}}$ is scalar magnetometer measurements, $\boldsymbol{B}_{\text{map}}$ is magnetic anomaly map values at known positions, $\boldsymbol{A}$ is formed from vector magnetometer measurements, and only $\boldsymbol{\beta}$ is unknown. (3.30) can be solved with linear least squares regression or ridge regression as with the classical model, as shown in (3.25) and (3.26), respectively. The map-based Tolles-Lawson model allows for direction computation of the aircraft field by using the known total and earth fields. This is only an approximation, as the magnetic anomaly (earth) field may have changed, albeit slightly, between when the map was created and when the calibration flight occurred. Note that the modified Tolles-Lawson variation can also be used here.

3.5 Elastic Net-Based Aeromagnetic Compensation

Due to multicollinearity, the Tolles-Lawson model is often ill-conditioned. Various approaches have been put forth to resolve this, including the use of ridge regression [55], singular value decomposition [57], and principal component analysis [58]. Ridge regression has been the dominant method to resolve multicollinearity, though the optimal ridge parameter must be determined [58]. As previously discussed in section 3.2, a ridge parameter of 0.025 has been found to be appropriate for the flight data in this work. Chen et al. used the ridge regression approach with two boom-mounted magnetometers. Using a transfer function to relate the two signals, the magnetometers were compensated simultaneously [59]. Given the improvements made by these approaches, two linear models are explored in this section and the next, namely elastic net and PLSR.

The elastic net is another regularized regression method that uses both the L_1 and L_2 penalties [60]. Using only the L_2 penalty is ridge regression, which was previously described in section 3.1.2. The L_2 norm is useful for generalization, i.e. to avoid overfitting. Using only the L_1 penalty corresponds to the least absolute shrinkage and selection operator (Lasso). The L_1 norm is useful for sparsity, i.e. variable selection [61]. The elastic net linearly combines these two penalties,

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}}(||\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}||_2^2 + \lambda_2||\boldsymbol{\beta}||_2^2 + \lambda_1||\boldsymbol{\beta}||_1) \tag{3.31}$$

where $\boldsymbol{\beta}$ is the regression coefficients, $\boldsymbol{y}$ is the output (observed) data, $\boldsymbol{X}$ is the input data, λ_2 is the L_2 or ridge parameter, and λ_1 is the L_1 or Lasso parameter. Specifically for aeromagnetic compensation the elastic net takes the form of

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}}(||\boldsymbol{B}_{\text{target}} - \boldsymbol{A}\boldsymbol{\beta}||_2^2 + \lambda_2||\boldsymbol{\beta}||_2^2 + \lambda_1||\boldsymbol{\beta}||_1) \tag{3.32}$$

where $\boldsymbol{A}$ is formed from a vector magnetometer measurement as previously described, and $\boldsymbol{B}_{\text{target}}$ can be one of multiple magnetic field options, as explained in section 3.7. The elastic net-based aeromagnetic compensation model was implemented using the GLMNet and MLJLinearModels Julia packages [62, 63, 64]. More specifically, cross-validation is performed using GLMNet to determine the best ridge and Lasso parameters, followed by determining the final elastic net coefficients using MLJLinearModels, as this was found to result in the best performance.

3.6 PLSR-Based Aeromagnetic Compensation

A more advanced linear model is partial least squares regression (PLSR). The general idea is to find low-dimensional latent variables that are the most correlated with the output (observed) data [65]. To start, the input (and output) data is typically standardized,

$$\tilde{\boldsymbol{X}} = \frac{\boldsymbol{X} - \boldsymbol{\mu}}{\boldsymbol{\sigma}} \tag{3.33}$$

where $\boldsymbol{X}$ is the input data, $\boldsymbol{\mu}$ is the mean for each column (i.e. feature), $\boldsymbol{\sigma}$ is the standard deviation for each column (i.e. feature), and $\tilde{\boldsymbol{X}}$ is the standardized (Z-score normalized) input data. Similarly, $\tilde{\boldsymbol{Y}}$ is the standardized output data. Using the singular value decomposition (SVD), a matrix factorization technique,

$$(\boldsymbol{U}, \boldsymbol{S}, \boldsymbol{V}) = \mathrm{svd}(\boldsymbol{C}_{XY}), \tag{3.34}$$

the left singular vectors $\boldsymbol{U}$, singular values $\boldsymbol{S}$, and right singular vectors $\boldsymbol{V}$ of the input-output covariance matrix $\boldsymbol{C}_{XY}$ can be determined. The first left singular vector, $\boldsymbol{u}$, and right singular, $\boldsymbol{v}$, are then used to determine the input score

$$\boldsymbol{z} = \tilde{\boldsymbol{X}}\boldsymbol{u} \tag{3.35}$$

and output score

$$\boldsymbol{r} = \tilde{\boldsymbol{Y}}\boldsymbol{v}. \tag{3.36}$$

Next, the input loading vector

$$\boldsymbol{p} = \frac{\boldsymbol{C}_{XX}\,\boldsymbol{u}}{\boldsymbol{u}^T\,\boldsymbol{C}_{XX}\,\boldsymbol{u}} \tag{3.37}$$

and output loading vector

$$\boldsymbol{q} = \frac{\boldsymbol{C}_{YX}\,\boldsymbol{u}}{\boldsymbol{u}^T\,\boldsymbol{C}_{XX}\,\boldsymbol{u}} \tag{3.38}$$

are determined, where $\boldsymbol{C}_{XX}$ is the input covariance matrix. At this point, the first left singular vector $\boldsymbol{u}$, input loading vector $\boldsymbol{p}$, and output loading vector $\boldsymbol{q}$ can be used to make output predictions $\hat{\boldsymbol{Y}}$. However, only the first most significant component of the input-output covariance matrix has been used and there is likely more information that can provide better predictions. To use the next most significant components, the input and output data and covariance matrices must be "deflated," which removes the component already used, leaving a residual. The input covariance matrix, input-output covariance matrix, input data, and output data are deflated as

$$\boldsymbol{C}_{XX} = (\boldsymbol{I} - \boldsymbol{p}\,\boldsymbol{u}^T)\,\boldsymbol{C}_{XX}, \tag{3.39}$$

$$\boldsymbol{C}_{YX} = \boldsymbol{C}_{YX}\,(\boldsymbol{I} - \boldsymbol{u}\,\boldsymbol{p}^T), \tag{3.40}$$

$$\tilde{\boldsymbol{X}} = \tilde{\boldsymbol{X}} - \boldsymbol{z}\,\boldsymbol{p}^T, \tag{3.41}$$

and

$$\tilde{\boldsymbol{Y}} = \tilde{\boldsymbol{Y}} - \boldsymbol{z}\,\boldsymbol{q}^T. \tag{3.42}$$

This process is repeated up to the desired number of components, which is limited by the number of columns (features) in the input data. For each iteration, the left singular vector $\boldsymbol{u}$, input loading vector $\boldsymbol{p}$, and output loading vector $\boldsymbol{q}$ are stored as columns in $\tilde{\boldsymbol{U}}$, $\tilde{\boldsymbol{P}}$, and $\tilde{\boldsymbol{Q}}$, respectively. The regression coefficients of the standardized data, $\tilde{\boldsymbol{\beta}}$, can then be calculated as

$$\tilde{\boldsymbol{\beta}} = \tilde{\boldsymbol{U}}(\tilde{\boldsymbol{P}}^T\tilde{\boldsymbol{U}})^{-1}\tilde{\boldsymbol{Q}}^T. \tag{3.43}$$

These regression coefficients are used with standardized input data to predict standardized output data, which are then de-standardized for true output predictions.

3.7 Fitting or Training Output Target Options

There are multiple options for the output target B_{target}, as shown in Figure 3-11.

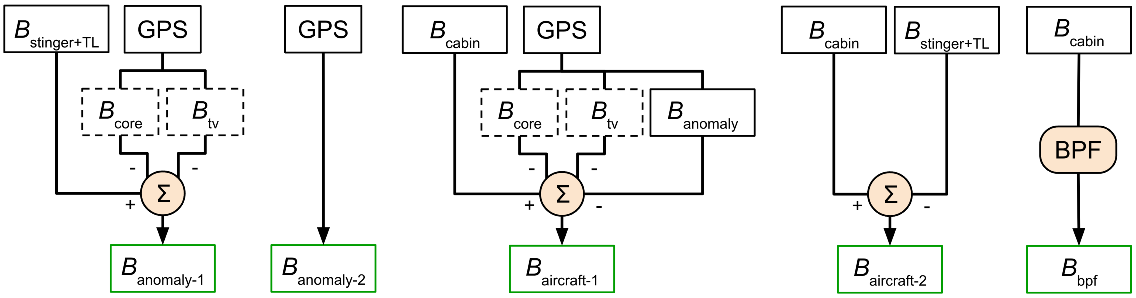


Figure 3-11: Fitting or training output target options.

All of these are some sort of magnetic value. More specifically, the options are:

(a) Anomaly field #1: compensated *tail stinger* total field scalar magnetometer measurements

(b) Anomaly field #2: interpolated magnetic anomaly map values

(c) Aircraft field #1: difference between uncompensated *in-cabin* total field scalar magnetometer measurements and interpolated magnetic anomaly map values

(d) Aircraft field #2: difference between uncompensated *in-cabin* and compensated *tail stinger* total field scalar magnetometer measurements

(e) BPF'd total field: bandpass filtered uncompensated *in-cabin* or *tail stinger* total field scalar magnetometer measurements

The first option (anomaly field #1) requires tail stinger data, while the second option (anomaly field #2) is a direct way to get the magnetic anomaly field, though error is potentially introduced due to position inaccuracy (both GPS measurement and map interpolation). Also, the magnetic anomaly map was generated using data collected at a different time than when the flight data was collected, which also may introduce error. The third option (aircraft field #1) attempts to isolate the aircraft field by removing the anomaly field component from a raw magnetometer measurement. The fourth option (aircraft field #2) isolates the aircraft field by finding the difference between uncompensated cabin and compensated tail stinger measurements. The third and fourth options are representative of typical aeromagnetic compensation, which represents the aircraft corruption that is removed. Finally, the fifth option (BPF'd total field) bandpass filters total field scalar measurements, as in the typical Tolles-Lawson procedure. Optionally, the core field (i.e. IGRF) and/or temporal variations (i.e. diurnal) can be removed from the total field scalar magnetometer measurements to truly isolate the anomaly or aircraft field.

3.8 Compensation Performance Metrics

When the ability to compensate a magnetic signal is being quantified, it is appropriate to use the standard deviation of the magnetic signal error,

$$\sigma_{\Delta_{\text{mag}}} = \sqrt{\frac{1}{N}\sum_{k=1}^{N} \Delta_{\text{mag,k}} - \mu_{\Delta_{\text{mag}}}} \tag{3.44}$$

where Δ_{mag} is the error between the compensated magnetic signal and truth (tail stinger or map) magnetic signal and $\mu_{\Delta_{\text{mag}}}$ is the mean of these errors over N samples (instances). Per-flight standard deviation rather than root mean square (RMS) is used here, since the mean (DC) offset of the magnetic signal has negligible impact on the filter-produced position estimates. Mild bias errors can be accounted for in the temporal variation filter state. Two signals can be compared using a common metric for evaluation of aeromagnetic compensation, the improvement ratio,

$$\text{IR} = \frac{\sigma_1}{\sigma_2} \tag{3.45}$$

where σ_1 and σ_2 are two standard deviations being compared [52]. Typically for aeromagnetic compensation, σ_1 is the uncompensated magnetic signal and σ_2 is the compensated magnetic signal. In general, σ_1 is some baseline and σ_2 is some modification to the baseline. Similarly, the signal-to-noise ratio,

$$\text{SNR} = \frac{\sigma_{\text{mag}_{\text{truth}}}}{\sigma_{\Delta_{\text{mag}}}} \tag{3.46}$$

can be used to quantify the quality of the expected magnetic measurements, $\text{mag}_{\text{truth}}$, in comparison to the magnetic measurement errors, Δ_{mag}. An SNR of greater than 10 often results in accuracies of 10's of meters with a high-quality map, data, etc.

3.9 Linear Model Performance Comparison

Five different linear aeromagnetic compensation models have been presented in this chapter. There are many comparisons that could be made with the various magnetometer options (see Table 2.2) and Tolles-Lawson terms (permanent, induced, eddy current, and/or bias). The comparison made here uses Mag 4, the third noisiest scalar magnetometer, and Flux A. The standard permanent, induced, and eddy current terms are used to create the Tolles-Lawson $\boldsymbol{A}$ matrix, which is an input to all linear models. For the elastic net-based and PLSR models, Mag 4 is an additional input feature. The fitting output target for the Tolles-Lawson models is as previously described, while for the elastic net-based and PLSR models, the fourth option, (aircraft field #2) is used.

Table 3.3: Linear aeromagnetic compensation performance using a single calibration flight pattern. The first calibration box of flight line 1006.04 with Mag 4 and Flux A is used here. The typical value of $\lambda = 0.025$ for the Tolles-Lawson ridge parameter was used, and $k = 2$ (of 19) components was found to work best for PLSR. Testing dataset errors [nT] are reported in comparison to the compensated tail stinger.

Flight	Flight Line	Classical Tolles-Lawson	Modified Tolles-Lawson	Elastic Net	PLSR
1003	1003.10	125	123	125	110
1004	4014.00	51	50	242	79
1004	4013.00	54	53	264	101
1005	4006.00	38	38	161	65
1005	4005.00	30	30	156	62
1007	1007.04	105	101	283	137
1007	1007.06	94	91	153	98

Looking at Table 3.3, it can be seen that the best performing linear aeromagnetic compensation model is modified Tolles-Lawson. However, this model only performs marginally better than classical Tolles-Lawson. Both the elastic net-based and PLSR-based models perform worse, significantly worse for some flight lines. Thus, Tolles-Lawson works best when using a typical limited calibration flight dataset. Note that changing the number of PLSR components used can significantly change the per-flight line performance. For the results in Table 3.3, 2 of 19 components were used, as this resulted in the lowest overall error. Using 18 or 19 components resulted in significantly lower error for the 1003, 1004, and 1005 flight lines, but significantly higher error for the 1007 flight lines. This may be due to the training data containing flight data from flights 1003-1006, but not flight 1007. Specifically, the respective PLSR errors with 18 components are: 34, 53, 58, 35, 36, 173, and 118 nT.

In Table 3.4 the linear aeromagnetic compensation results have completely changed when using a large training dataset compared to Table 3.3, which only used a typical calibration flight. Now, the elastic net-based and PLSR-based models perform

Table 3.4: Linear aeromagnetic compensation performance using a larger training dataset. A Tolles-Lawson ridge parameter of $\lambda = 0.25$ and $k = 17$ (of 19) PLSR components were found to work best for this dataset. Testing dataset errors [nT] are reported in comparison to the compensated tail stinger.

Flight	Flight Line	Classical Tolles-Lawson	Modified Tolles-Lawson	Elastic Net	PLSR
1003	1003.10	123	109	74	73
1004	4014.00	62	58	32	32
1004	4013.00	57	51	36	37
1005	4006.00	39	34	25	26
1005	4005.00	31	28	24	23
1007	1007.04	97	74	42	41
1007	1007.06	90	72	44	44

significantly better than the Tolles-Lawson models. The improvement of modified Tolles-Lawson over classical Tolles-Lawson is now more evident, while the elastic net and PLSR perform similarly. Using more data generally improved classical Tolles-Lawson aeromagnetic compensation performance, but not nearly as much as with the elastic net-based and PLSR-based models. Tolles-Lawson works well for modeling *maneuvering* magnetic interference, especially with small amounts of calibration flight data. Elastic net and PLSR seem to better model *non-maneuvering* magnetic interference with larger amounts of flight data. These models effectively use the additional data to learn better compensation coefficients than Tolles-Lawson.

Note that map-based Tolles-Lawson is not compared in Tables 3.3 or 3.4, since flight line 1006.04 and many of the flight lines in the standard training dataset (see section 2.2.2) were flown over areas without map coverage. Additionally note that the results presented here only include magnetic data from a single scalar and vector magnetometer pair. The elastic net and PLSR models can readily include additional data, which is explored in detail in section 4.9 in comparison with nonlinear models.

Chapter 4

Neural Network-Based Aeromagnetic Compensation

Coupling together machine learning-based data-driven models and physics-based models has the potential for better performance than either individually. This combination can allow for lower data requirements and smaller models (e.g. neural networks), thus greater computational efficiency.

In 1993, Williams first proposed using a neural network for aeromagnetic compensation of geo-survey aircraft [66]. The model was trained to predict the total magnetic field, which was assumed to be a linear combination of 3 components, namely local (earth/crustal), diurnal (temporal variations), and aircraft interference effects. These components correspond with 6 position, 2 time, and 12 attitude input features, respectively. Williams showed that the high frequency portion of the magnetic field was dependent almost entirely on attitude terms. It was also suggested to use a fluxgate vector magnetometer to directly measure the earth to aircraft direction cosines and use these as input features, as in the Tolles-Lawson model. Performance was similar to the classical Tolles-Lawson model.

More recently, Ma et al. used a variational Bayesian neural model to predict the standard deviation of the aeromagnetic compensation error in addition to the compensation magnitude itself [67]. Yu et al. used a denoising autoencoder with a single hidden layer to resolve the previously mentioned multicollinearity issue in estimating the aeromagnetic compensation coefficients for a rotary-wing UAV [68]. Xu et al. used deep learning for MAD, primarily through simulations [69].

Hezel trained 240 different neural network models using magnetometer, accelerometer, gyroscope, and voltage data from a Geometrics micro-fabricated atomic magnetometer development kit (MFAM) [18]. The best performing model was able to decrease the RMSE from approximately 47 nT to 5 nT on a validation dataset. Emery developed a neural network-based aeromagnetic compensation model for an F-16 that used the Tolles-Lawson $\boldsymbol{A}$ matrix terms from (3.19)–(3.21), along with additional voltage and current sensor data, as input features [19]. The hyperparameters from [18] were mostly used without further consideration, though extremely deep neural networks of 7 to 97 layers were examined. None of the models outperformed classical Tolles-Lawson aeromagnetic compensation, which was attributed to a lack of model experimentation and optimization. Evaluating shallower neural network models was recommended.

This chapter first provides background on neural networks, then introduces five neural network-based aeromagnetic compensation models. These models vary from a standard neural network, similar to Hezel and Emery [18, 19], to more involved scientific machine learning (SciML) approaches that combine the Tolles-Lawson model with a neural network. An evaluation of neural network size and feature selection is then discussed. Finally, compensation performance is compared against the Tolles-Lawson and PLSR-based models for a variety of magnetometer selections.

4.1 Neural Networks

An (artificial) neural network (NN) is a collection of connected nodes or units called (artificial) neurons, which is analogous to biological neurons in the brain. The nodes are put into layers to build the neural network, as shown in Figure 4-1. Here, a single-layer neural network is illustrated, but more hidden layers can be used to create a composition of functions. When a multi-layer approach is used, it is called a deep neural network.

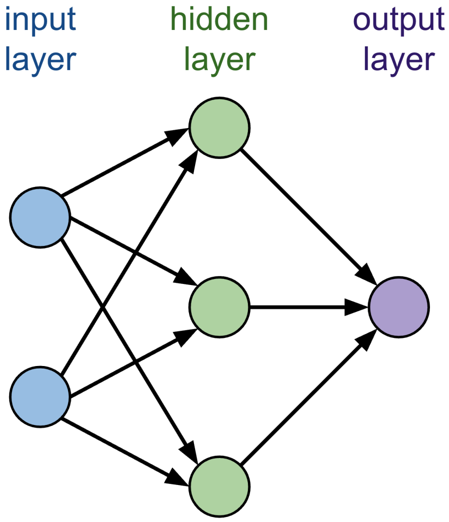


Figure 4-1: Neural network.

A neural network with even just a single hidden layer is a universal approximator [70]. This means that, with enough nodes and proper selection of model parameters, an accurate approximation can be constructed for any arbitrary nonlinear function. There are other universal approximators, but neural networks have additional "order of approximation" properties, meaning that the error can be significantly less for a given number of nodes [71].

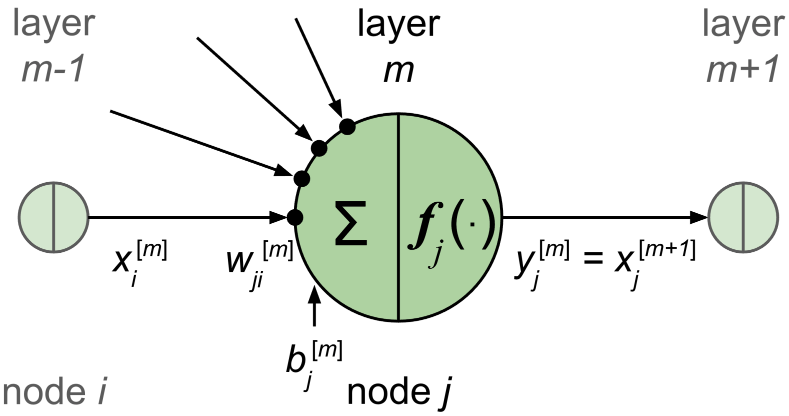


Figure 4-2: Individual node within a neural network.

An individual neural network node is shown in Figure 4-2. Input x from nodes in the previous layer is fed into the node and multiplied by weights w with an added bias b, also known as offset or threshold. The resulting value is the pre-activation

$$z_j = \sum_{i=1}^{n} w_{ji} x_i + b_j \tag{4.1}$$

which passes through a nonlinear activation function f to produce the node's output

$$y_j = f_j(z_j). \tag{4.2}$$

There are many options for the activation, with the primary requirement being differentiability. A few commonly used activation functions are shown in Figure 4-3, along with their derivatives in Figure 4-4. Note that the derivatives are important for training the neural network to learn w and b, which is explained next.

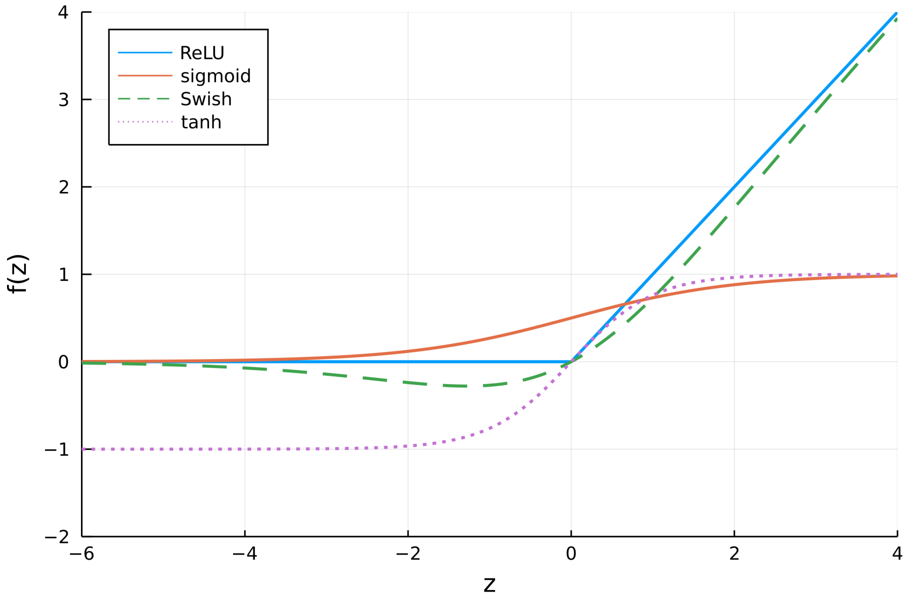


Figure 4-3: Common activation functions. The standard slope parameter of 1 is used for sigmoid and Swish.

The activation functions shown in Figure 4-3 and Figure 4-4 are provided by (4.3)–(4.6). A rectified linear unit (ReLU) is a type of ramp function, taking only the positive part of its argument.

$$\mathrm{ReLU}(z) = \max(0, z) \tag{4.3}$$

The sigmoid function σ, a logistic function specifically, uses a slope parameter β (typically 1) and varies between 0 and 1.

$$\sigma(z, \beta) = \frac{1}{1 + e^{-\beta z}} \tag{4.4}$$

Swish is a relatively new activation function that has been shown to outperform ReLU and others on multiple neural network models [72]. Similar to ReLU, it uses a

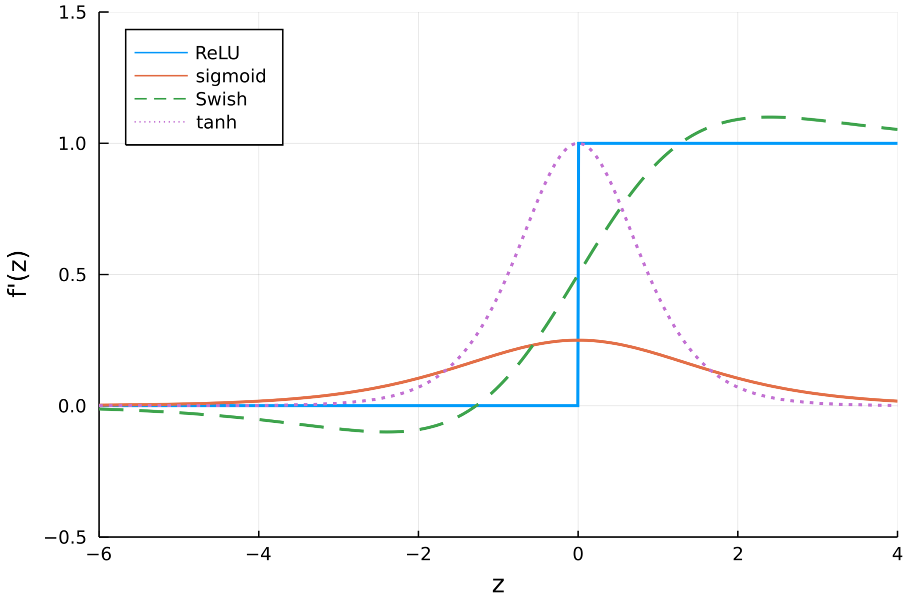


Figure 4-4: First derivatives of common activation functions. The standard slope parameter of 1 is used for sigmoid and Swish.

slope parameter β that is typically 1. This is the primary activation function used in this work, as it was found to perform best compared to the other options mentioned here, as discussed further in section 4.7.

$$\mathrm{Swish}(z, \beta) = z\ \sigma(z, \beta) = \frac{z}{1 + e^{-\beta z}} \tag{4.5}$$

The hyperbolic tangent function tanh varies between -1 and 1.

$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \tag{4.6}$$

The squared loss for input data x with output target data y is

$$L(x, y|w, b) = \frac{1}{2}(y - \hat{y}(x|w, b))^2 \tag{4.7}$$

where $\hat{y}$ is the neural network estimated output. Using the error backpropagation algorithm, the weights of the neural networks can then be updated with

$$\Delta w_{ji} = \rho \; f_j'(z_j) \; (y - \hat{y}(x|w,b)) \; x_i \tag{4.8}$$

where ρ is the learning rate, which controls how much to update the weights of the neural network in response to the calculated loss. It is a small, positive value, typically between 0 and 1.

4.2 Neural Network-Based Aeromagnetic Compensation, Model 1

One of the goals of this work was to determine if nonlinear aeromagnetic compensation was able to outperform Tolles-Lawson with tail stinger and/or in-cabin magnetic measurements. Thus, the goal of neural network-based aeromagnetic compensation is to use machine learning in order to create a nonlinear aeromagnetic compensation model. Additionally, it was questioned if additional sensors, beyond a single scalar and vector magnetometer, could improve compensation performance. Onboard electronics, such as strobe lights, are largely not able to be compensated using the classical Tolles-Lawson model, so it has been previously suggested to specifically model them [73, 74].

The approach taken here is to include additional data along with the typical magnetometer data. An example of possible data fields is provided in Appendix A and discussed in section 2.2.1. The idea here is that this additional data allows for greater observability of the aircraft magnetic field, and machine learning can learn the relationship between the data and aircraft field. A baseline model for this approach is to use a neural network (described in section 4.1) to represent the compensation model in a pure data-driven approach, as shown in Figure 4-5.

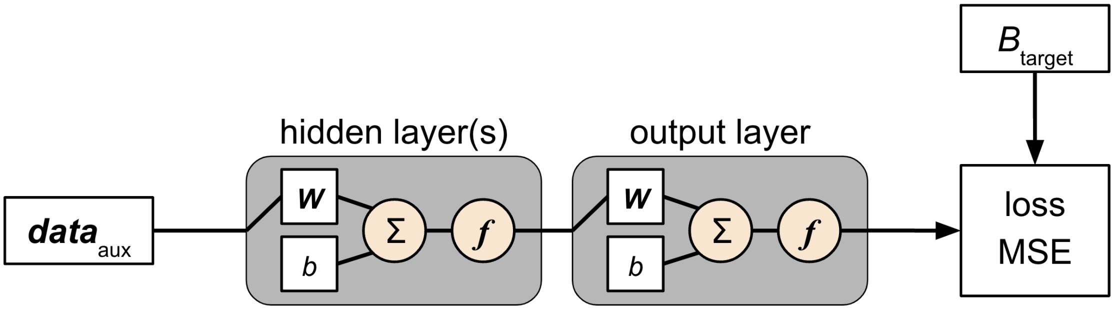


Figure 4-5: Neural network-based aeromagnetic compensation, model 1.

Here, the input data $\boldsymbol{data}_{\text{aux}}$, also known as features, can be raw measurements and/or the Tolles-Lawson $\boldsymbol{A}$ matrix terms from (3.19)–(3.21). Using the Tolles-Lawson terms is a form of scientific machine learning (SciML), as these terms represent prior scientific knowledge. This can potentially reduce the amount of data required for training. Note that during training, the mean squared error (MSE) of the aeromagnetic compensation output is compared with the output target, as described section 3.7. During testing (evaluation), the aeromagnetic compensation output just prior to the "loss MSE" in Figure 4-5 is used for navigation.

4.3 Neural Network-Based Aeromagnetic Compensation, Model 2a

Model 1 can be extended into even more of a SciML approach using the physics of aeromagnetic compensation. Model 2a, shown in Figure 4-6, more closely resembles the classical Tolles-Lawson model, now with

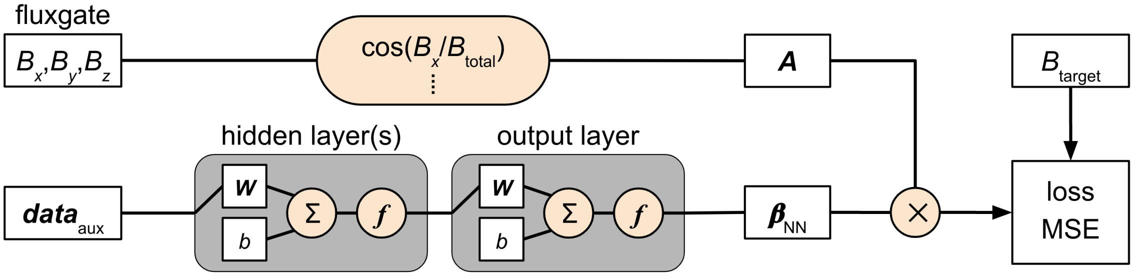


Figure 4-6: Neural network-based aeromagnetic compensation, model 2a.

$$B_{\text{target}} = \boldsymbol{A}\ \boldsymbol{\beta}_{\text{NN}} \tag{4.9}$$

where B_{target} is some target magnetic value, $\boldsymbol{A}$ is the Tolles-Lawson $\boldsymbol{A}$ matrix, and $\boldsymbol{\beta}_{\text{NN}}$ is a vector of varying coefficients.

Rather than directly determining some magnetic value, the output of the neural network is instead a set of varying coefficients. These are then multiplied by the Tolles-Lawson $\boldsymbol{A}$ matrix to produce a magnetic value. The general idea here is that this model is easier to train and more interpretable, since the coefficients have meaning related to the permanent, induced, and eddy-current terms in the Tolles-Lawson model. This model uses the same input data $\boldsymbol{data}_{\text{aux}}$ as previously described, but also separately requires vector magnetometer measurements (B_x, B_y, B_z). The output target B_{target} can again be any of those listed in section 3.7. Note that if the fifth option (BPF'd total field) for the output target is used, the $\boldsymbol{A}$ matrix in Figure 4-6 (or any of the models in sections 4.3–4.6) must also be bandpass filtered.

4.4 Neural Network-Based Aeromagnetic Compensation, Model 2b

Another SciML approach that uses the physics of aeromagnetic compensation is model 2b, which is shown in Figure 4-7. Rather than predicting the Tolles-Lawson coefficients with a neural network, the Tolles-Lawson coefficients for this approach are pre-determined, using any of the classical, modified, or map-based Tolles-Lawson models. Tolles-Lawson aeromagnetic compensation is then carried out, but large errors still remain for compensation of in-cabin magnetometers. An additive correction is then made to the Tolles-Lawson compensated value to further reduce the errors. This additive correction is the output of a neural network, similar to model 1 presented in section 4.2, except now the neural network does not need to learn the Tolles-Lawson (linear) portion of the desired compensation and instead primarily learns the higher-order nonlinear portion.

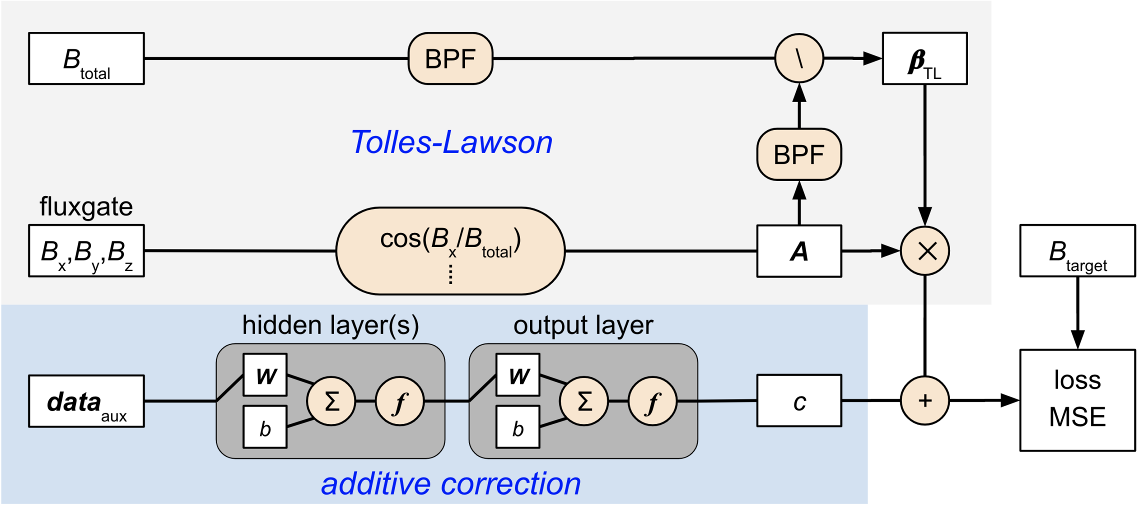


Figure 4-7: Neural network-based aeromagnetic compensation, model 2b.

4.5 Neural Network-Based Aeromagnetic Compensation, Model 2c

An immediate extension of model 2b is model 2c, which is shown in Figure 4-8. These models look very similar, but there is one key, subtle difference. In model 2c, the Tolles-Lawson coefficients are updated by including them as parameters during training of the neural network, rather leaving them as static values. The Tolles-Lawson coefficients are not actually part of the neural network, but the Flux machine learning library [75, 76] and Julia programming language [77] provide the flexibility to backpropagate the error to the Tolles-Lawson coefficients just as with the weights of the neural network. This would be difficult to implement in many other programming languages, but it was rather straightforward with Julia.

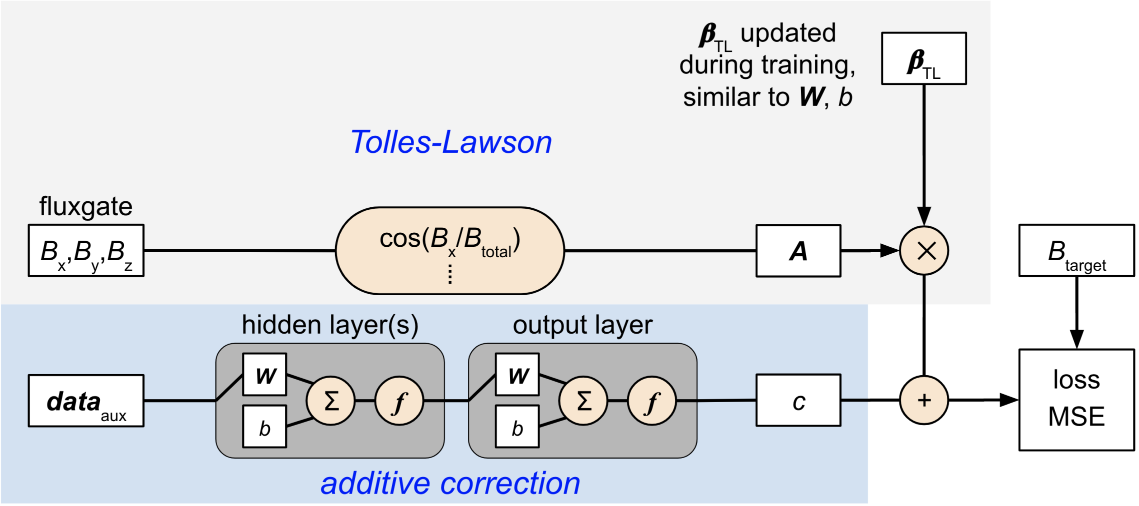


Figure 4-8: Neural network-based aeromagnetic compensation, model 2c.

4.6 Neural Network-Based Aeromagnetic Compensation, Model 2d

The final neural network-based approach is model 2d, which is shown in Figure 4-9. This model combines the general idea behind model 2a and model 2b in sections 4.3 and 4.4, respectively. An additive correction is now made to each individual (static) Tolles-Lawson coefficient. This model was developed as it was thought that making the additive correction in a higher dimensional space may improve performance further.

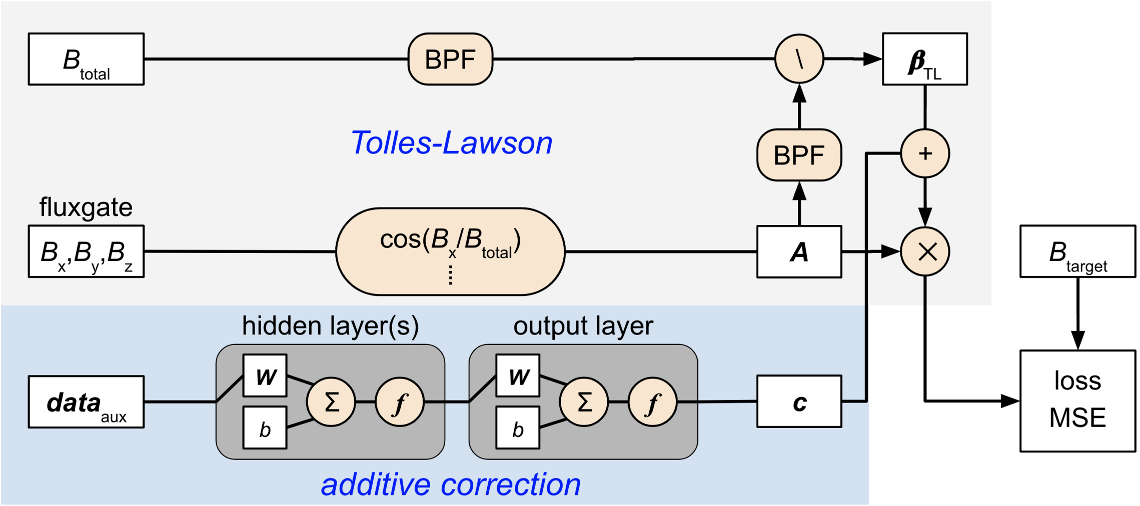


Figure 4-9: Neural network-based aeromagnetic compensation, model 2d.

4.7 Neural Network Size Evaluation

A key unknown for the models presented in sections 4.2–4.6 was the neural network design, including the number of hidden layers, number of nodes per hidden layer, and activation function. To gain insight into these unknowns, the design of experiments (DOE) shown in Table 4.1 was completed. One, two, and three hidden layer neural networks with up to 128 nodes per hidden layer were evaluated. To a lesser extent, the Swish and ReLU activation functions were compared as well. Initial testing prior to this DOE indicated that other activation functions, namely sigmoid and tanh performed consistently worse than Swish and ReLU and were not considered further. For this DOE, batch sizes of 2048, 500 training epochs, the ADAM optimizer (optimizing algorithm) [78], and a mean squared error (MSE) loss function were used. The number of training epochs was held at 500, as this was found to prevent early stopping in preliminary analysis. The batch sizes of 2048 and the ADAM optimizer were also previously found to work well. The training data was split into a true training portion and a validation portion in a 14:3 ratio. Check-pointing, based on validation loss, was used to save the neural network weights during training when a new minimum loss was found. Using this method prevents the model from overtraining and allows for better generalization.

Table 4.1: Initial Neural Network Design of Experiments.

Hidden Layers	Nodes	Activation	Training Error [nT]	Testing Error [nT]
1	4	Swish	3.7	4.1
1	**8**	**Swish**	**3.5**	**4.6**
1	**16**	**Swish**	**3.1**	**5.5**
1	16	ReLU	3.2	6.2
1	32	Swish	2.9	5.9
1	64	Swish	2.7	6.8
1	128	Swish	2.6	8.1
2	4, 4	Swish	3.8	4.0
2	**8, 4**	**Swish**	**3.4**	**5.2**
2	**8, 8**	**Swish**	**3.3**	**5.6**
2	**16, 8**	**Swish**	**3.0**	**5.9**
2	16, 16	Swish	2.9	6.0
2	16, 16	ReLU	3.4	9.1
2	32, 16	Swish	2.5	7.7
2	32, 32	Swish	2.4	9.4
2	64, 32	Swish	2.0	11
2	64, 64	Swish	2.0	10
2	128, 64	Swish	1.7	12
2	128, 128	Swish	1.8	11
3	4, 4, 4	Swish	3.8	4.4
3	8, 8, 8	Swish	3.3	5.4
3	16, 16, 16	Swish	2.7	7.7
3	16, 16, 16	ReLU	3.5	16
3	32, 32, 32	Swish	2.4	8.0

Model 1, presented in section 4.2, was trained on flight data with some of the fields listed in Appendix A using the Flux machine learning library [75, 76], which uses the Julia programming language [77]. More specifically, the in-cabin magnetometers (four scalar and three vector), INS data, barometer reading, and all available current and voltage sensors were used. The vector magnetometer data was transformed into the Tolles-Lawson $\boldsymbol{A}$ matrix terms from (3.19)–(3.21), and the INS roll, pitch, and yaw data was transformed into direction cosine matrix terms. Both of these provide orientation information. A total of 108 features were used. The output target B_{target}, as in Figures 4-5 and 4-6, was the compensated tail stinger total field scalar measurements (anomaly field #1 in section 3.7) with temporal variations and the core field removed, though using the interpolated magnetic anomaly map values resulted in similar performance. The training data itself was selected from flights 1003-1006 (of the available 9 flights performed by SGL), as described in section 2.2.2.

The DOE results, shown in Table 4.1, are fairly consistent across the variety of neural network sizes. In general, larger neural networks have larger out-of-sample (testing) errors, which indicates overfitting to the training data. The Swish activation function clearly has superior performance compared to the ReLU activation function. With this in mind, 5 neural networks were chosen for further evaluation, which are shown in bold in Table 4.1. These were selected based on the tradeoff between model size, training error, and testing error.

A secondary DOE was carried out with a reduced feature set for a preliminary look at feature importance and the impact on compensation performance. Iterative hard thresholding (IHT) was carried out using the MendelIHT Julia package [79, 80]. IHT is a method to develop a sparse approximation to a linear system of equations, thus it is not the best approach for nonlinear feature importance, which is explored in detail in section 4.8. Cross-validation IHT was performed on the training data, which indicated that the best sparse, linear solution used 20 of the 108 features for **linear** compensation. Among these features were several terms from the Tolles-Lawson $\boldsymbol{A}$ matrices, one term from the direction cosine matrix, the barometer reading, multiple magnetometer readings, one INS acceleration, and 5 current and voltage sensors.

The most important current sensors were: `cur_ac_hi` (air conditioner fan high), `cur_strb` (strobe lights), and `cur_heat` (INS heater). The most important voltage sensors were: `vol_bat_1` (battery 1) and `vol_block` (block). Based on these features being selected, it was decided to keep several other related features, including the `cur_com_1` (aircraft radio 1), `cur_ac_lo` (air conditioner fan low), `cur_tank` (cabin fuel pump), and `cur_flap` (flap motor) current sensors and the `vol_bat_2` (battery 2) voltage sensor. Additionally, only one eddy current term from the Tolles-Lawson $\boldsymbol{A}$ matrices was present in the IHT results, which also had a relatively small coefficient. With this in mind, it was decided to only keep the permanent and induced terms. Other work has observed the relative unimportance of the eddy current terms and ignored them as well [81]. In summary, only 10 of the 31 current and voltage sensors and 27 of the 54 terms from the Tolles-Lawson $\boldsymbol{A}$ matrices were kept, reducing the feature count from 108 to 60.

Table 4.2: Secondary Neural Network Design of Experiments. All runs used the Swish activation function, batch sizes of 2048, and 500 training epochs.

Hidden Layers	Nodes	Model 1 Train [nT]	Model 2a Train [nT]	Model 1 Test [nT]	Model 2a Test [nT]
1	**8**	**3.9**	**5.5**	**4.8**	**4.7**
1	16	3.6	6.8	4.5	5.7
2	8, 4	3.9	6.1	5.2	6.9
2	8, 8	3.8	7.6	4.7	6.0
2	16, 8	3.4	6.4	4.7	8.0

The results of the secondary DOE with the reduced feature set are shown in Table 4.2. Batch sizes of 2048, 500 training epochs, and a mean squared error loss function were again used. The ADAM optimizer was again used for model 1. However, model 2 used 5 epochs with the ADAM optimizer for initialization of the neural network weights, then 500 epochs with the limited memory Broyden–Fletcher–Goldfarb–Shanno algorithm (L-BFGS) [82] to test the impact of a different optimizer. This decision was based on previous testing and performance of various optimizers with the models.

The output target B_{target}, as in Figures 4-5 and 4-6, was the compensated tail stinger total field scalar measurements with temporal variations and core field removed. The shallow model with only a single hidden layer containing eight nodes was selected for further analysis as it retained low error (approximately 5 nT) on both the training and testing datasets despite being relatively small and fast. Note that both DOEs were evaluated using an older version of MagNav.jl. Thus, these results are unable to be exactly replicated, but the overall results should be similar.

4.8 Feature Selection

The importance of the individual data fields within $\boldsymbol{data}_{\text{aux}}$ was generally unknown for all of the neural network-based approaches. The linear approach (IHT) previously discussed in section 4.7 may not agree with nonlinear approaches for feature selection. To potentially develop simpler, more interpretable, and faster models, feature (data field) selection was evaluated in-depth. It was anticipated that some of the features, particularly some of the current and voltage sensors, were not providing useful information for neural network-based aeromagnetic compensation. Eliminating these features from the models increases interpretability and decreases the training and evaluation runtimes. To fully explore feature selection, also known as feature importance or reduction, several approaches were taken. These are explained in the remainder of this section, followed by feature selection results.

4.8.1 Drop-Column Feature Importance

The first feature selection approach is drop-column feature importance. This is a “wrapper” type of feature selection. For this approach, the model is trained with each individual feature excluded (i.e. the feature column is dropped) one at a time. This means the model must be trained for as many times as the number of features. For a large number of features, this is computationally expensive. Additionally, as seen in section 4.8.5, the feature importance results are not meaningful if the features are highly correlated, since no individual feature is particularly important. Relative performance of the trained models indicates feature importance, i.e. lower performance without a given feature corresponds to higher feature importance and vice versa.

4.8.2 Permutation Feature Importance

The second feature selection approach is permutation feature importance. This is also a “wrapper” type of feature selection. First, a model is trained with all features included. The model is then evaluated with the data for each individual feature shuffled (i.e. the feature column is randomly permutated) one at a time. Similar to drop-column feature importance, relative performance of the evaluated model indicates feature importance, i.e. lower performance with a given feature shuffled corresponds to higher feature importance and vice versa. This approach has the benefit of only requiring the model to be trained once, and then simply evaluated for as many times as the number of features.

4.8.3 Sparse Group Lasso

The third feature selection approach uses the sparse group Lasso penalty,

$$\lambda \sum_{j=1}^{p} (1-\alpha)\,||\theta_{1,\cdot,j}||_1 + \alpha\,||\theta_{1,\cdot,j}||_2 \tag{4.10}$$

where λ is the sparse group Lasso parameter, α is a Lasso ($\alpha = 0$) vs group Lasso ($\alpha = 1$) balancing parameter, and $\theta_{1,\cdot,j}$ are the first-layer neural network weights for each input feature j among p total features [83]. Sparse group Lasso is a nonlinear sparse-input regularization approach. The penalty uses the L_1 and/or L_2 penalties, similar to the elastic net presented in section 3.5. Again, the L_1 is useful for sparsity, while the L_2 norm is useful for generalization.

To carry out this feature selection approach, a neural network model is trained with the above penalty included in the loss function (alongside the standard MSE). Thus, this is an "embedded" feature selection approach, different from the others discussed in this section. The trained model should have features with small first-layer weights, the number of which is dependent on the λ and α hyperparameters. These hyperparameters and the Lasso portion of (4.10), $||\theta_{1,\cdot,j}||_1$, drive sparsity and appropriately scale the penalty. The ranking of features is based solely on the group Lasso portion of (4.10), $||\theta_{1,\cdot,j}||_2$, after training.

4.8.4 Shapley Values

The fourth and final feature selection approach used in this work is to determine stochastic Shapley values for each feature, another "wrapper" type of feature selection. This approach comes from cooperative game theory and determines how much each feature contributes to a prediction in comparison to the average prediction

[84]. Shapley values are a model-agnostic way to quantify global feature importance, meaning they are valid for any model, linear or nonlinear. Shapley values were determined using the ShapML Julia package [85].

4.8.5 Feature Selection Result Comparison

Note that prior to the feature selection results presented here, many different features were examined with the various neural network-based models. In some cases, the "same" data can be included in multiple ways. For example, the aircraft attitude can be included with the raw values, i.e. roll, pitch, and yaw, or the sine and/or cosine of each value. Alternatively, the individual terms from the 3×3 direction cosine matrix (formed from the roll, pitch, and yaw values) can be used, which was found to perform best after examining all three options mentioned. This is called feature engineering or feature extraction – creating features from functions of the original data using domain knowledge. Additional features that were examined, but not included here due to worse performance, include INS accelerations in the navigation frame (rather than the body frame), yaw, pitch, and roll rates, and lagged scalar magnetometer measurements.

Similarly, for these neural network-based models, there are more than 20 hyperparameters that can be selected. Clearly an exhaustive search for each optimal hyperparameter is not practical. Additionally, the best hyperparameters are not necessarily the same for every flight data subset. Instead, various trials were carried out prior to the results presented here to find well-performing, but not overly optimized, hyperparameters. Slightly better performance could result by further optimizing the hyperparameters, but not enough of an improvement to invalidate the takeaways formed here. The selected hyperparameters are shown in Table 4.3.

Table 4.3: Feature selection hyperparameters.

Hyperparameter	Value	Description
y_type	a	$\boldsymbol{y}$ output target option (see section 3.7)
use_mag	mag_4_uc	scalar mag to use with y_type = c, d, or e
use_vec	flux_d	vector mag to use for "external" Tolles-Lawson $\boldsymbol{A}$
terms	[p,i,e]	terms for Tolles-Lawson $\boldsymbol{A}$ within $\boldsymbol{data}_{\text{aux}}$
terms_A	[p,i,e]	terms for "external" Tolles-Lawson $\boldsymbol{A}$
sub_diurnal	true	if true, subtract diurnal from scalar mag meas
sub_igrf	true	if true, subtract IGRF from scalar mag meas
bpf_mag	false	if true, bpf scalar mag meas in $\boldsymbol{data}_{\text{aux}}$
norm_type_A	none	normalization for "external" Tolles-Lawson $\boldsymbol{A}$
norm_type_x	standardize	normalization for $\boldsymbol{data}_{\text{aux}}$
norm_type_y	standardize	normalization for $\boldsymbol{y}$
TL_coef	TL_d_4	Tolles-Lawson coefficients
η_adam	0.001	learning rate for ADAM optimizer
epoch_adam	500	number of epochs for ADAM optimizer
epoch_lbfgs	0	number of epochs for L-BFGS optimizer
hidden	[8]	nodes per hidden layer
activation	swish	activation function
batchsize	2048	mini-batch size
frac_train	14/17	training data fraction (remainder for validation)
α_sgl	1	Lasso (0) vs group Lasso (1) balancing parameter
λ_sgl	0	sparse group Lasso parameter
k_pca	-1	# components for PCA + whitening, -1 to ignore

Looking at Table 4.3, the first hyperparameter is `y_type`, which is discussed in section 3.7. For this analysis, the first option (anomaly field #1) was selected, though the first four options generally performed similarly. `use_mag` was not relevant for this analysis, and `use_vec` = `flux_d` was selected arbitrarily among the three non-stinger options. The standard Tolles-Lawson terms, permanent (`p`), induced (`i`), and eddy current (`e`), were selected for both the Tolles-Lawson $\boldsymbol{A}$ matrices within $\boldsymbol{data}_{\text{aux}}$ and the "external" matrix used with models 2a-2d. `sub_diurnal` and `sub_igrf` were both set to true, meaning the scalar magnetometer measurements had the core field and temporal variations removed. This may be thought to be "cheating" since these are position dependent, but even with the position uncertainty during navigation, these are known to within approximately 0.1 nT. No normalization was performed on the "external" Tolles-Lawson $\boldsymbol{A}$ matrix, while standardization (Z-score normalization) was found to work well on the input and output data. Note that min-max normalization and scaling by the maximum absolute value (no bias) were also evaluated on the input and output data, and both were found to provide similar or worse results. The selected Tolles-Lawson coefficients, `TL_coef` = `TL_d_4` were calculated using the classical Tolles-Lawson model with Mag 4 and Flux D flight data from two calibration boxes performed during flight line 1002.02. Performance was similar using a different calibration flight line.

As far as neural network-specific parameters, 500 epochs with the ADAM optimizer were used, with a learning rate of η_adam $= 0.001$. Various optimizers were evaluated, and stochastic gradient descent generally performed poorly with a noisy, oscillating solution. L-BFGS, a quasi-Newton method, works as well, but was found to often overfit as MagNav.jl is currently unable to use validation check-pointing with L-BFGS. Thus, the commonly used optimizer, ADAM was exclusively used for generating the results in the remainder of this work. A single hidden layer with 8 nodes and the swish activation function were used, as these were found to work well, as described in section 4.7. Similarly, `batchsize` $= 2048$ and `frac_train` $= 14/17$ were again used. The sparse group Lasso parameters, α_sgl and λ_sgl, as described in section 4.8.3, were set as to not use sparse group Lasso. Similarly, `k_pca` $= -1$ does not use PCA with whitening by default. This is an optional pre-processing step for highly correlated input data that can reduce computation time, as discussed in section 4.9.

With the setup explained, the results can now be examined. Two neural network-based aeromagnetic compensation models were examined, models 1 and 2c. These were selected as model 1 is the least SciML-like approach, while model 2c, with the "external" Tolles-Lawson $\boldsymbol{A}$ matrix and updating Tolles-Lawson coefficients, was deemed the most SciML-like approach and it performed the best at the time of this analysis. The standard training dataset described in section 2.2.2 was used for training here.

Table 4.4: Top 25 features for model 1. “Drop” is drop-column feature importance, “Perm” is permutation feature importance, “SGL” is sparse group Lasso, and “Shap” is for Shapley values, as described in sections 4.8.1–4.8.4. “Train” and “Test” are the standard training and testing datasets described in sections 2.2.2 and 2.2.3, respectively. Sorted by the average of the four right columns.

Feature	Drop Train	Drop Test	Perm Test	SGL Train	Shap Train	Shap Test
TL_A_flux_d_Z	4	79	2	1	3	1
TL_A_flux_c_Z	94	26	1	2	1	2
TL_A_flux_d_XZ	45	22	3	3	2	3
mag_5_uc	36	60	4	5	5	5
TL_A_flux_c_XZ	33	11	5	6	4	4
TL_A_flux_a_YY	40	97	9	4	16	6
TL_A_flux_d_ZZ	14	52	7	7	7	8
TL_A_flux_a_XX	108	18	6	11	9	7
mag_4_uc	34	15	8	8	6	9
mag_4_5_uc	53	53	10	9	8	11
TL_A_flux_c_ZZ	57	81	12	10	26	10
mag_3_uc	52	89	11	18	10	12
TL_A_flux_a_YZ	82	33	14	15	11	13
TL_A_flux_c_X	10	106	17	14	29	14
TL_A_flux_d_XX	29	10	15	13	20	17
dcm_1	63	27	16	22	13	16
TL_A_flux_a_XY	35	78	23	20	23	20
mag_3_5_uc	9	94	21	24	18	19
TL_A_flux_d_X	38	107	24	16	34	24
TL_A_flux_c_XY	44	48	22	23	19	22
TL_A_flux_c_XX	77	8	29	12	32	26
TL_A_flux_c_YZ	69	32	13	44	12	15
TL_A_flux_d_YZ	72	40	19	32	17	21
TL_A_flux_d_YY	47	9	20	29	14	25
TL_A_flux_d_XY	32	59	31	19	22	27

Table 4.5: Top 25 features for model 2c. "Drop" is drop-column feature importance, "Perm" is permutation feature importance, "SGL" is sparse group Lasso, and "Shap" is for Shapley values, as described in sections 4.8.1–4.8.4. "Train" and "Test" are the standard training and testing datasets described in sections 2.2.2 and 2.2.3, respectively. Sorted by the average of the four right columns.

Feature	Drop Train	Drop Test	Perm Test	SGL Train	Shap Train	Shap Test
TL_A_flux_c_Z	12	49	1	2	1	1
mag_5_uc	35	53	3	1	6	4
TL_A_flux_c_XZ	1	1	2	5	2	2
TL_A_flux_d_Z	6	6	4	6	3	3
TL_A_flux_d_XZ	101	89	6	7	5	6
mag_4_uc	16	78	5	8	4	7
TL_A_flux_d_ZZ	2	11	7	10	8	5
mag_4_5_uc	45	62	8	9	7	10
TL_A_flux_c_ZZ	17	33	9	11	12	8
TL_A_flux_a_YY	34	85	16	3	30	11
TL_A_flux_a_YZ	68	90	11	15	9	9
TL_A_flux_c_XX	102	24	13	4	11	19
TL_A_flux_c_X	14	16	12	14	19	12
TL_A_flux_d_YZ	72	67	10	21	10	13
TL_A_flux_a_XY	3	2	15	17	14	14
TL_A_flux_a_XX	60	48	17	18	20	17
TL_A_flux_a_XZ	57	87	14	28	13	15
dcm_1	49	84	19	22	15	16
TL_A_flux_c_XY	74	3	18	24	17	18
TL_A_flux_c_YZ	36	96	20	20	16	20
TL_A_flux_d_XY	77	99	22	16	21	23
TL_A_flux_d_XX	24	22	26	12	31	27
TL_A_flux_d_Y	27	10	21	27	26	21
TL_A_flux_c_YY	28	45	23	25	18	24
TL_A_flux_a_X	99	70	29	19	25	25

Tables 4.4 and 4.5 show the top 25 features for models 1 and 2c, and the full results are in Appendix B. There are multiple clear trends visible in these results. First, the most important features are the magnetometers, which is not at all surprising. The best available scalar magnetometers, Mag 4 and 5, and their differences appear in both Tables 4.4 and 4.5. Additionally, many of the Tolles-Lawson terms from Flux A, C, and D appear. Note that only permanent (e.g. `_Z`) and induced (e.g. `_XZ`) terms are present among the top 25 features. Additionally, the direction cosine terms seem to provide some information, especially `dcm_1 = cos(pitch) cos(yaw)`. Just outside of the top 25 features `dcm_4 = cos(pitch) sin(yaw)` also appears. One additional thing to note is the dissimilarity of the drop-column feature importance results compared to the others. When a single feature is dropped, the overall impact on model performance is minor, since many of the features are correlated. This produces spurious results, as the best performing model during drop-column feature importance is only marginally better than the worst performing model.

Tables 4.6 and 4.7 show the bottom 25 features for both models, and the full results are again available in Appendix B. Contrary to the top 25 features in Tables 4.4 and 4.5, there are nearly no magnetometer features here, the only exceptions being some Tolles-Lawson eddy current terms (e.g. `_YZ_dot`). Nearly all of the unimportant features are INS accelerations, currents, and voltages. This is not necessarily surprising, as some of these are nearly static values. This can be quantified, by looking at Table 4.8. As an example, `vol_res_n` has both a low standard deviation and low skew, and it appears near the bottom of the feature selection lists. However, this introduces a potential issue. A feature may not be important a majority of the time, but particularly important for short periods. For this reason, a feature such as `cur_flap` should be kept, as the skew indicates this may be an “on/off” type of feature. Again note the dissimilarity of the drop-column feature selection results compared to the others. In this case, some of the most important features according to drop-column feature importance are completely unimportant for the other approaches.

Table 4.6: Bottom 25 features for model 1. “Drop” is drop-column feature importance, “Perm” is permutation feature importance, “SGL” is sparse group Lasso, and “Shap” is for Shapley values, as described in sections 4.8.1–4.8.4. “Train” and “Test” are the standard training and testing datasets described in sections 2.2.2 and 2.2.3, respectively. Sorted by the average of the four right columns.

Feature	Drop Train	Drop Test	Perm Test	SGL Train	Shap Train	Shap Test
TL_A_flux_a_XZ_dot	97	43	75	82	82	74
vol_back_n	30	77	95	56	80	85
TL_A_flux_c_YY_dot	74	16	68	92	73	79
vol_block	71	42	89	62	96	88
cur_acpwr	78	21	76	94	58	70
vol_acc_p	16	105	100	46	78	95
TL_A_flux_a_ZX_dot	61	93	73	88	71	82
TL_A_flux_d_ZX_dot	3	100	74	87	94	84
cur_com_1	1	38	88	65	90	94
cur_ac_hi	7	49	94	47	88	106
TL_A_flux_a_ZY_dot	106	44	72	108	77	68
TL_A_flux_c_YZ_dot	104	61	71	102	72	78
TL_A_flux_d_YX_dot	6	101	78	89	76	90
vol_res_p	102	66	98	63	104	100
vol_gyro_2	26	51	107	57	101	97
vol_back	88	39	96	67	97	99
vol_fan	58	13	99	69	102	98
vol_gyro_1	51	1	108	74	93	87
cur_ac_lo	60	57	109	59	99	101
cur_strb	80	4	97	68	108	108
cur_srvo_m	107	45	101	78	107	102
vol_res_n	85	12	106	80	100	103
cur_srvo_i	79	34	104	91	105	104
ins_acc_z	73	55	105	96	103	107
cur_srvo_o	56	68	102	105	106	105

Table 4.7: Bottom 25 features for model 2c. "Drop" is drop-column feature importance, "Perm" is permutation feature importance, "SGL" is sparse group Lasso, and "Shap" is for Shapley values, as described in sections 4.8.1–4.8.4. "Train" and "Test" are the standard training and testing datasets described in sections 2.2.2 and 2.2.3, respectively. Sorted by the average of the four right columns.

Feature	Drop Train	Drop Test	Perm Test	SGL Train	Shap Train	Shap Test
cur_tank	107	88	109	46	91	77
vol_gyro_2	44	81	101	53	95	80
TL_A_flux_a_ZY_dot	85	23	60	102	73	74
vol_cabt	65	100	85	67	90	84
TL_A_flux_d_ZY_dot	52	65	72	97	84	68
vol_acc_n	84	25	89	51	80	97
vol_acc_p	73	52	94	45	74	98
TL_A_flux_c_XY_dot	94	79	64	103	61	71
TL_A_flux_d_YY_dot	79	60	63	100	70	78
TL_A_flux_d_YX_dot	26	15	73	94	78	79
TL_A_flux_d_YZ_dot	19	58	77	82	75	87
cur_com_1	4	13	92	63	96	99
cur_flap	93	4	96	54	105	105
vol_res_p	21	20	93	68	97	96
TL_A_flux_a_XY_dot	98	94	78	92	86	89
vol_fan	63	54	97	75	98	91
vol_back	75	39	98	71	99	95
vol_gyro_1	53	76	99	72	100	101
vol_block	92	106	102	80	104	104
cur_strb	46	64	108	70	108	108
cur_srvo_m	100	93	106	81	103	100
cur_srvo_i	66	77	100	85	106	103
cur_srvo_o	76	14	107	87	102	102
ins_acc_z	95	66	104	90	107	107
vol_res_n	104	69	105	98	101	106

Table 4.8: Standard deviation and skew of current and voltage features.

Features	Std. Dev.	Skew
cur_flap	0.019	21.93
cur_ac_lo	0.041	8.863
vol_bat_1	0.243	6.553
cur_strb	0.037	5.231
cur_com_1	0.016	4.731
cur_ac_hi	0.102	3.275
cur_bat_1	0.003	1.799
cur_tank	0.040	1.352
vol_block	0.026	0.726
vol_res_p	0.014	0.717
vol_cabt	0.307	0.645
vol_back	0.002	0.553
vol_fan	0.014	0.546
vol_acc_p	0.012	0.481
cur_outpwr	1.798	0.476
vol_back_n	0.008	0.473
cur_acpwr	1.784	0.467
vol_outpwr	0.163	0.384
vol_acpwr	0.147	0.323
vol_back_p	0.009	0.311
cur_bat_2	0.004	0.273
vol_res_n	0.012	0.258
vol_gyro_1	0.033	0.193
vol_acc_n	0.006	0.158
cur_srvo_o	0.124	0.132
vol_bat_2	0.264	0.129
cur_srvo_i	0.062	0.098
vol_srvo	0.085	0.095
cur_srvo_m	0.038	0.073
vol_gyro_2	0.029	0.054
cur_heat	1.330	0.021

4.9 Linear and Nonlinear Model Performance Comparison

Following the feature selection analysis in section 4.8.5, the full 108 feature set was down-selected. The features that were removed were entirely based on the feature selection rankings, with the exception of a select few current and voltage features that were kept based on the standard deviation and skew statistics presented in Table 4.8. The specific removed current features were `cur_tank`, `cur_srvo_o`, `cur_srvo_m`, `cur_srvo_i`, `cur_bat_1`, and `cur_bat_2`. All of the voltage features except `vol_bat_1` were removed, as were the INS accelerations (`ins_acc_x`, `ins_acc_y`, `ins_acc_z`). Other than the magnetometer terms, this only left the direction cosine matrix terms (`dcm_1` – `dcm_9`), barometer reading (`baro`), `vol_bat_1`, and the following current features: `cur_com_1`, `cur_ac_hi`, `cur_ac_lo`, `cur_flap`, `cur_strb`, `cur_heat`, `cur_acpwr`, and `cur_outpwr`.

A few features were also added based on further independent trials. These include scalar magnetometer derivatives and 4^{th} order central differences, which is based on an aeromagnetic compensation noise metric used in industry [86]. An IGRF (core field) term and INS latitude, longitude, and altitude were also added. The Tolles-Lawson terms were modified to use the 3 permanent, 6 induced, and 3 *derivative* terms, which are similar to the typical eddy current terms, but reduce the feature count. In general, the eddy current terms were shown to be unimportant, so this choice was based on the desire to retain derivative information, but reduce the (unimportant) feature count. The final 77 feature set used in the following results comparisons is listed in Appendix C.

Table 4.9: Down-selected hyperparameters.

Hyperparameter	Value	Description
y_type	d	$\boldsymbol{y}$ output target option (see section 3.7)
use_mag	mag_4_uc	scalar mag to use with y_type = c, d, or e
use_vec	flux_d	vector mag to use for "external" Tolles-Lawson $\boldsymbol{A}$
terms	[p,i,d]	terms for Tolles-Lawson $\boldsymbol{A}$ within $\boldsymbol{data}_{\text{aux}}$
terms_A	[p,i,e,b]	terms for "external" Tolles-Lawson $\boldsymbol{A}$
sub_diurnal	false	if true, subtract diurnal from scalar mag meas
sub_igrf	false	if true, subtract IGRF from scalar mag meas
bpf_mag	false	if true, bpf scalar mag meas in $\boldsymbol{data}_{\text{aux}}$
norm_type_A	none	normalization for "external" Tolles-Lawson $\boldsymbol{A}$
norm_type_x	standardize	normalization for $\boldsymbol{data}_{\text{aux}}$
norm_type_y	standardize	normalization for $\boldsymbol{y}$
TL_coef	TL_d_4	Tolles-Lawson coefficients
η_adam	0.001	learning rate for ADAM optimizer
epoch_adam	500	number of epochs for ADAM optimizer
epoch_lbfgs	0	number of epochs for L-BFGS optimizer
hidden	[8]	nodes per hidden layer
activation	swish	activation function
batchsize	2048	mini-batch size
frac_train	14/17	training data fraction (remainder for validation)
α_sgl	1	Lasso (0) vs group Lasso (1) balancing parameter
λ_sgl	0	sparse group Lasso parameter
k_pca	-1	# components for PCA + whitening, -1 to ignore

In addition to down-selection of the features, the hyperparameters were modified based on further analysis. The updated set of hyperparameters are shown in full in Table 4.9. It was determined that `y_type = d` performed the same or better than `y_type = a` with `sub_diurnal = false` and `sub_igrf = false`. This is the fourth output target option (aircraft field #2), as discussed in section 3.7. `sub_diurnal` and `sub_igrf` refer to removing the diurnal and IGRF magnetic fields from the scalar magnetometer measurements, respectively. Using this combination is similar to conventional compensation, where the total field is used and the desired output is the compensation, which represents the aircraft magnetic field. Additionally, the Tolles-Lawson $\boldsymbol{A}$ matrices within $\boldsymbol{data}_{\text{aux}}$ used a *derivative* (`d`) term instead of an eddy current (`e`) term, as previously explained, while the "external" Tolles-Lawson $\boldsymbol{A}$ matrix used with models 2a-2d added a bias (`b`) term.

With the down-selected feature set and updated hyperparameters explained, the performance of each neural network aeromagnetic compensation model is compared in Figure 4-10. There are a few takeaways that can be formed. First, 500 epochs is adequate for all models to prevent early stopping and overfitting. As few as approximately 200 epochs could have been used without a significant change in performance. Next, the models perform similarly overall, but model 2c had the slight edge over the others at fewer epochs, prior to slight overfitting. As a reminder, model 2c adds an additive correction to updating Tolles-Lawson coefficients. Finally, all models are able to achieve around 4 nT of error (standard deviation, as defined in section 3.8) on held out testing data.

Despite the large training dataset used, the models in Figure 4-10 were trained quickly, approximately 5-6 min each using a 64 GB 2019 MacBook Pro. However, it may be of interest to decrease the training time even further, without decreasing performance significantly. One method to do so is by pre-processing the standardized data using principal component analysis (PCA) with whitening [87]. PCA is a way to do a change of basis on the data by finding the principle components, i.e. directions of maximum variance [88]. Whitening involves re-standardizing the PCA-transformed data. The results after performing these two steps are shown in Figure 4-11.
Comparing Figure 4-11 to Figure 4-10, the curves corresponding to the training data are again overlapped, but stabilize at a slightly lower error (roughly 2 nT instead of 2.3 nT). Model 2c performs better than the others on held out testing data, though model 1 reaches a similar performance at 500 epochs. Note that model 2c also

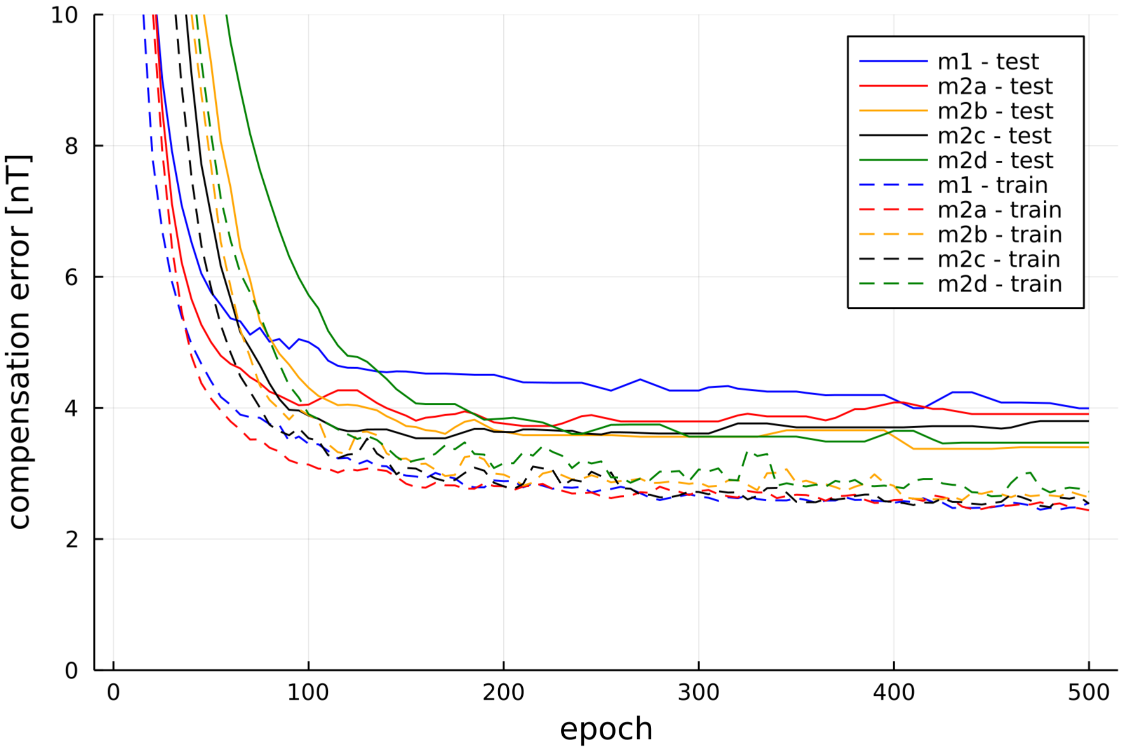


Figure 4-10: Neural network-based model comparison.

tended to perform better than the other models in additional evaluations that are not included in this work. Model 2c (as well as model 2b) additionally has the benefit of containing both a linear and a nonlinear portion, which provides reliability and interpretability. Thus, model 2c is deemed the marginally best neural network-based aeromagnetic compensation approach. Approximately 200 epochs are appropriate for model 2c when using PCA with whitening.

Using the PCA with whitening approach, 200 epochs, and the down-selected feature set, a large sweep of magnetometer combinations was performed. Three approaches are compared here, classical Tolles-Lawson, PLSR-based, and neural network-based (model 2c) aeromagnetic compensation. For classical Tolles-Lawson,

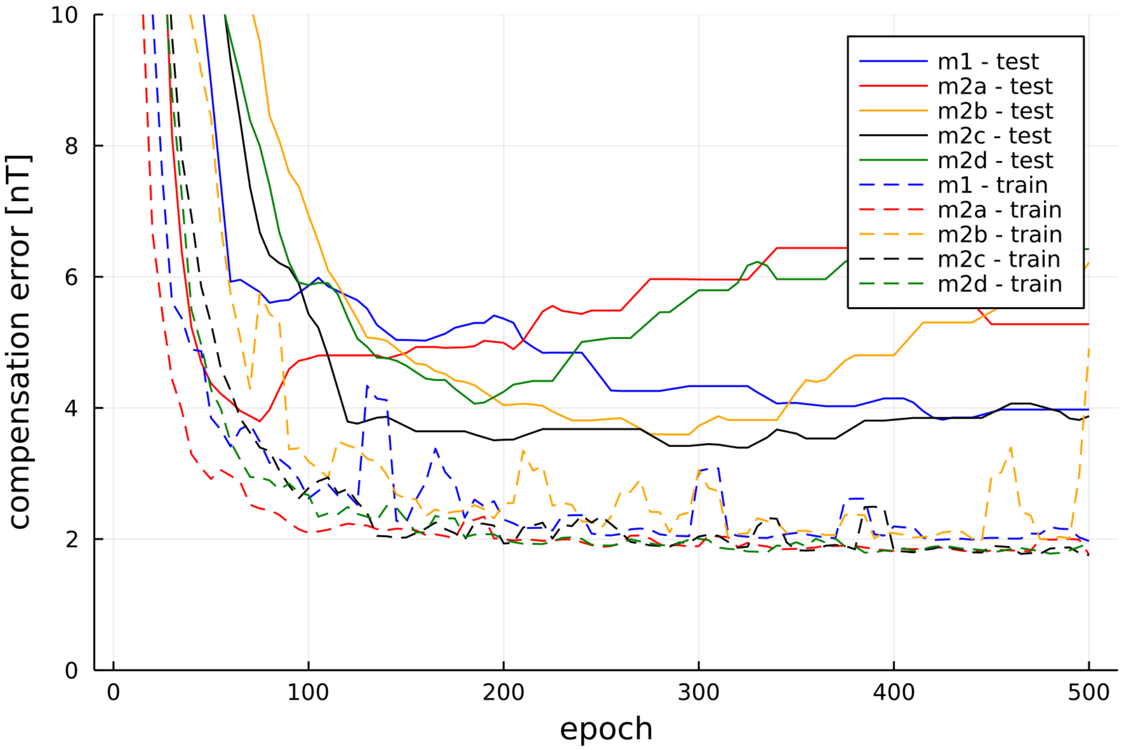


Figure 4-11: Neural network-based model comparison using PCA with whitening. The number of components kept with PCA was 68 (of 77), which retained 99.99% of the input data variance.

the first calibration box of flight line 1006.04 was again used, as in section 3.9. This was done to provide a baseline for the state-of-the-art method. For PLSR and model 2c (the marginally best performing linear and nonlinear aeromagnetic compensation models) the standard training dataset, described in section 2.2.2, was used. The results from using Mag 3, 4, or 5 with up to two of Flux A, C, or D are shown in Table 4.10. This sweep was actually performed three separate times. In the first sweep, the goal was to see which vector magnetometer(s) performed best in combination with which scalar magnetometer(s). In the second sweep, the best available vector magnetometer was used to determine the classical Tolles-Lawson coefficients, which are also used with model 2c. During the first two sweeps the number of PCA components was set to not lose *any* information, i.e. 100% of the input data variance was retained. In the third sweep, the number of PCA components was reduced to retain at most 99.99% of the input data variance, which was 1 to 3 less than the total number of components. This final sweep is reported here.

Table 4.10: Single scalar magnetometer sweep performance on testing data. Mag 2 is evaluated separately.

Scalar Mag	Vector Mag(s)	Tolles-Lawson Error	PLSR Error	Model 2c Error
3	A,C	91	25	22
3	A,D	91	18	16
3	C,D	87	39	24
3	A	91	57	30
3	C	103	68	44
3	D	87	57	49
4	A,C	102	10	8.6
4	A,D	105	16	15
4	C,D	102	16	13
4	A	105	23	26
4	C	102	70	43
4	D	106	54	32
5	A,C	18	4.7	5.4
5	A,D	19	5.3	4.9
5	C,D	18	5.3	5.7
5	A	19	9.5	4.9
5	C	18	26	13
5	D	20	13	12

Looking at Table 4.10, first note that the results with Mag 5 are better than all of the results without Mag 5 (for the same vector magnetometers), which is clearly the best choice for MagNav. However, Mags 3 and 4 also show promising results with Flux A at around 30 nT of error, less than a third of the error from the state-of-the-art approach. It appears that PLSR and model 2c are more sensitive to the vector magnetometer selection than classical Tolles-Lawson. Additionally, classical Tolles-Lawson is worse than PLSR or model 2c in nearly all cases. Finally, model 2c outperforms PLSR in most, but not all, cases. The cases of worse performance for model 2c may potentially be due to the PCA approach taken here, where the results are sensitive to the last included or excluded component. However, the goal here was not to rigorously optimize each model of the sweep, but to instead see the trends across the magnetometer combinations. Note that using the PCA with whitening approach decreased the training time by over 50%, yet sub-5 nT out-of-sample testing errors are still achieved.

The results from using up to two of Mag 3, 4, or 5 with up to two of Flux A, C, or D are shown in Table 4.11. Again note that the performance when Mag 5 is included is consistently better than when Mag 5 is excluded (for the same vector magnetometers). However, a somewhat surprising result is that Mag 5 alone, as seen in Table 4.10, often performs approximately the same or better than in combination with another scalar magnetometer. Thus, getting the least corrupted magnetic signal possible should be the top priority, rather than using multiple noisier magnetometers. However, without Mag 5, 15 nT or less error can be achieved if two scalar and/or two vector magnetometers are used. This could be applied in an operational scenario, in which a scalar and vector magnetometer pair is placed in a pod on each wing. PLSR and model 2c again outperform classical Tolles-Lawson in nearly all cases. With two scalar magnetometers, model 2c outperforms PLSR overall.

Table 4.11: Dual scalar magnetometers sweep performance on testing data. Mag 2 is evaluated separately.

Scalar Mag	Vector Mag(s)	Tolles-Lawson Error	PLSR Error	Model 2c Error
3,4	A,C	102	9.9	10
3,4	A,D	105	16	18
3,4	C,D	102	16	14
3,4	A	105	23	21
3,4	C	102	67	34
3,4	D	106	46	31
3,5	A,C	18	4.7	5.9
3,5	A,D	19	5.4	7.0
3,5	C,D	18	5.3	5.6
3,5	A	19	8.9	5.5
3,5	C	18	19	11
3,5	D	20	13	9.8
4,5	A,C	18	4.4	4.2
4,5	A,D	19	5.3	4.3
4,5	C,D	18	5.1	4.0
4,5	A	19	9.5	4.8
4,5	C	18	17	10
4,5	D	20	11	8.4

The results from using Mag 2 (the worst scalar magnetometer) and/or Mag 5 (the best scalar magnetometer) with up to two of Flux A, C, or D are shown in Table 4.12. Mag 2 alone performs poorly, and is unsuitable for airborne magnetic anomaly navigation unless multiple vector magnetometers are used. When used in combination with Mag 5, the performance can be worse than Mag 5 alone. In general, Mag 2 is too corrupted to be used as a source of magnetic information.

Table 4.12: Magnetometer 2 sweep performance on testing data.

Scalar Mag	Vector Mag(s)	Tolles-Lawson Error	PLSR Error	Model 2c Error
2,5	A,C	18	4.7	3.9
2,5	A,D	19	5.4	5.0
2,5	C,D	18	5.3	4.6
2,5	A	19	9.5	4.6
2,5	C	18	26	19
2,5	D	20	13	7.0
2	A,C	33281	29	41
2	A,D	33281	18	34
2	C,D	32957	58	56
2	A	33281	132	118
2	C	32749	82	156
2	D	32957	129	143
5	A,C	18	4.7	5.4
5	A,D	19	5.3	4.9
5	C,D	18	5.3	5.7
5	A	19	9.5	4.9
5	C	18	26	13
5	D	20	13	12

To set the stage for online neural network-based aeromagnetic compensation, as described in section 5.6, it is worthwhile to look at *even smaller* feature sets. The reason for this is that an online neural network stores each weight (and bias) as a state in the navigation algorithm, so a large feature set becomes a computational burden. In an operational scenario, it would also be beneficial to not rely on many sensors, any of which could experience signal dropouts. Two small feature sets are examined with each of Mags 3, 4, and 5. Mag 2 is not analyzed here or in the remainder of this work, since it is not suitable for MagNav, as shown by the results in Table 4.12. Flux A is solely used, as it provided the best overall performance with these scalar magnetometers. A reminder from section 2.2.1 that Flux A was located in the middle of the cabin near the fuel tank, a location that would be expected to have significant magnetic corruption from the aircraft itself.

Small feature set 1 uses the reduced set of currents and one voltage, as described at the beginning of this section. To repeat, this includes `cur_com_1`, `cur_ac_hi`, `cur_ac_lo`, `cur_flap`, `cur_strb`, `cur_heat`, `cur_acpwr`, `cur_outpwr`, and `vol_bat_1`. It also includes the most important direction cosine matrix terms, specifically `dcm_1 = cos(pitch) cos(yaw)` and `dcm_4 = cos(pitch) sin(yaw)`. For the scalar magnetometer, the measurement itself, the derivative, and the 4^{th} order central difference are used. For the vector magnetometer, the typical permanent and induced Tolles-Lawson terms and the *derivative* terms are used. Small feature set 2 includes only the magnetometer terms. Small feature set 1 contains a total of 26 features, while small feature set 2 contains a total of 15 features.

Prior to evaluating the small feature sets, it was wondered if the neural network size could also be decreased from 8 nodes to even fewer nodes. First, the full 77-term version of model 2c was re-evaluated to see the sensitivity of performance due to neural network size and epochs, which is shown in Figures 4-12 and 4-13. First looking at the training results in Figure 4-12, it can be seen that training performance improves with more nodes. However, the improvement diminishes with size, i.e. there is a substantial improvement between 1 and 2 nodes, but only a marginal improvement between 32 and 64 nodes. Beyond 750 epochs, there is no improvement in training performance for any neural network size.

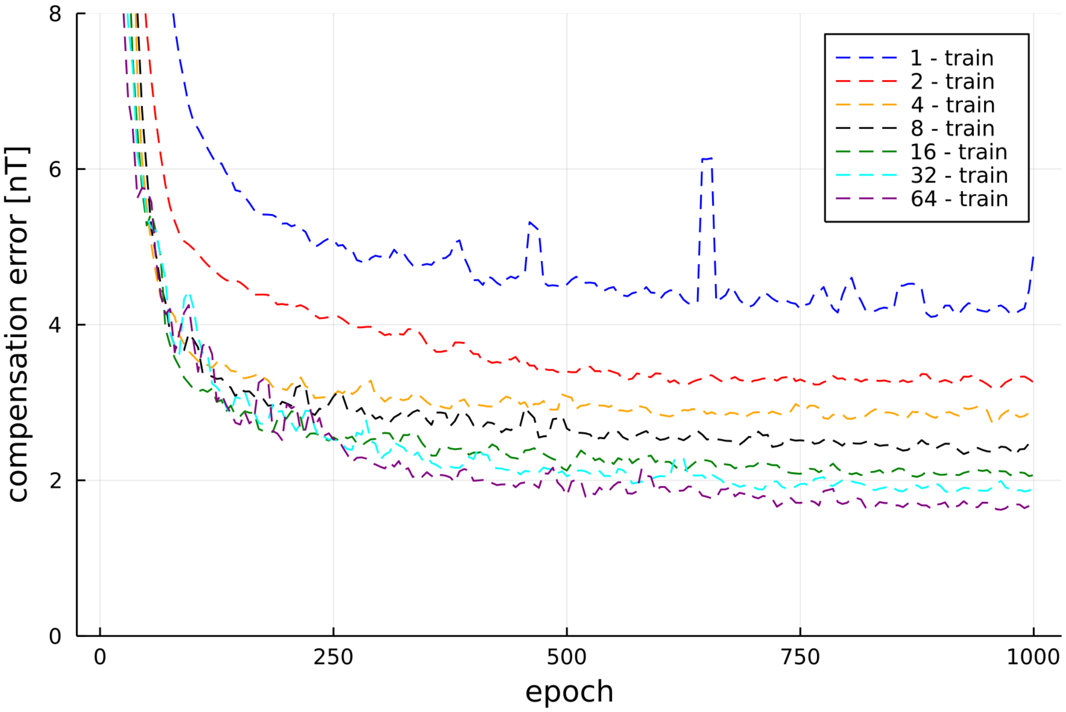


Figure 4-12: Model 2c training performance with neural network size, 77 features. 1, 2, 4, 8, 16, 32, and 64 nodes in a single hidden layer are compared. The feature set is listed in Appendix C. Mag 4 and Flux A are used here.

Now looking at the testing results in Figure 4-13, it can be seen that testing performance does not necessarily improve with more nodes. The lowest compensation error is achieved with 16 nodes at around 500 epochs. However, 8 nodes also performs well at 500 epochs, confirming the decision to use 8 nodes and 500 epochs for much of the previously discussed work. Significantly worse performance occurs with 1 node, and the second worst performance occurs with 64 nodes (which also took more than 2× as long to train). 2, 4, 8, 16, and 32 nodes all have similar performance at around 4 nT of compensation error. With the exception of 1 node, 500 epochs is appropriate for model 2c. Note that previous work in this chapter used a larger feature set with 108 features, which may have benefited from using 8 nodes rather than 2 or 4 nodes.

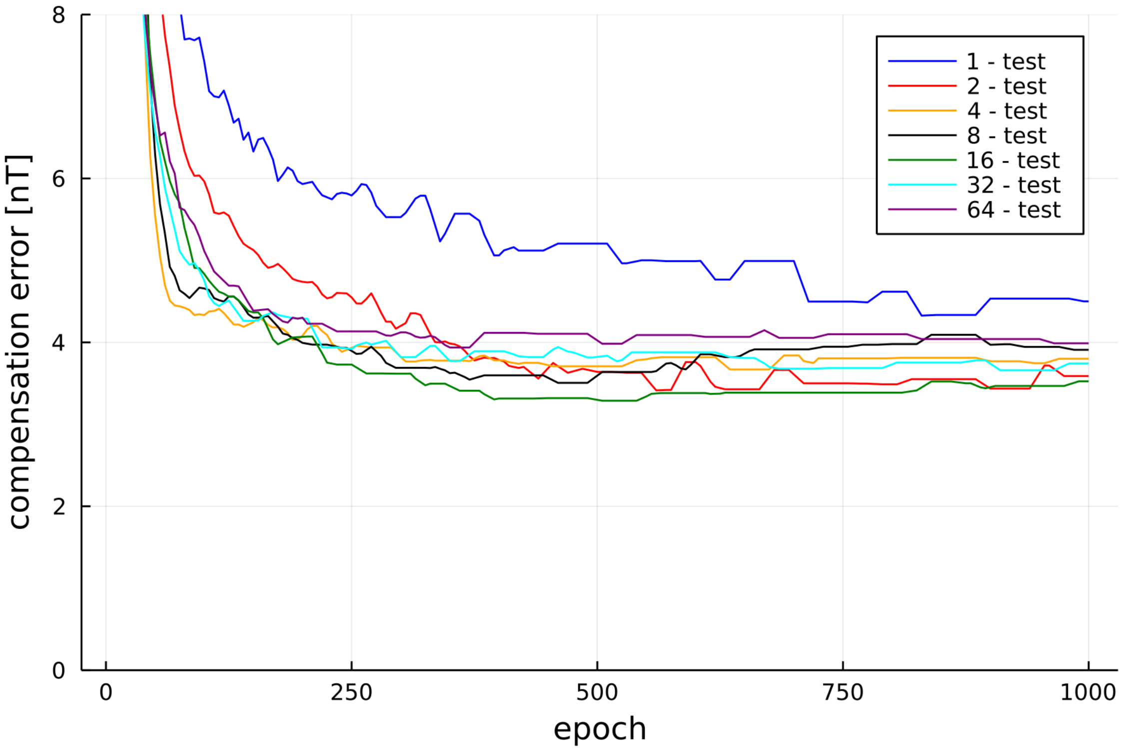


Figure 4-13: Model 2c testing performance with neural network size, 77 features. 1, 2, 4, 8, 16, 32, and 64 nodes in a single hidden layer are compared. The feature set is listed in Appendix C. Mag 4 and Flux A are used here.

Using small feature set 2, an additional analysis of neural network size and number of epochs was completed, as shown in Figure 4-14. Different from the results with the larger feature set, approximately 750 epochs is more appropriate. Using 2 or 4 nodes resulted in similar performance to 8 nodes, while clearly 1 node should not be used. The overall performance was far worse compared to the results with 77 features, with approximately 37 nT of error on held out testing data when using 2, 4, or 8 nodes. This was expected, since far less information, especially magnetometer measurements, was provided to the neural network. Given these results, 2 nodes was deemed the best choice for the small feature sets.

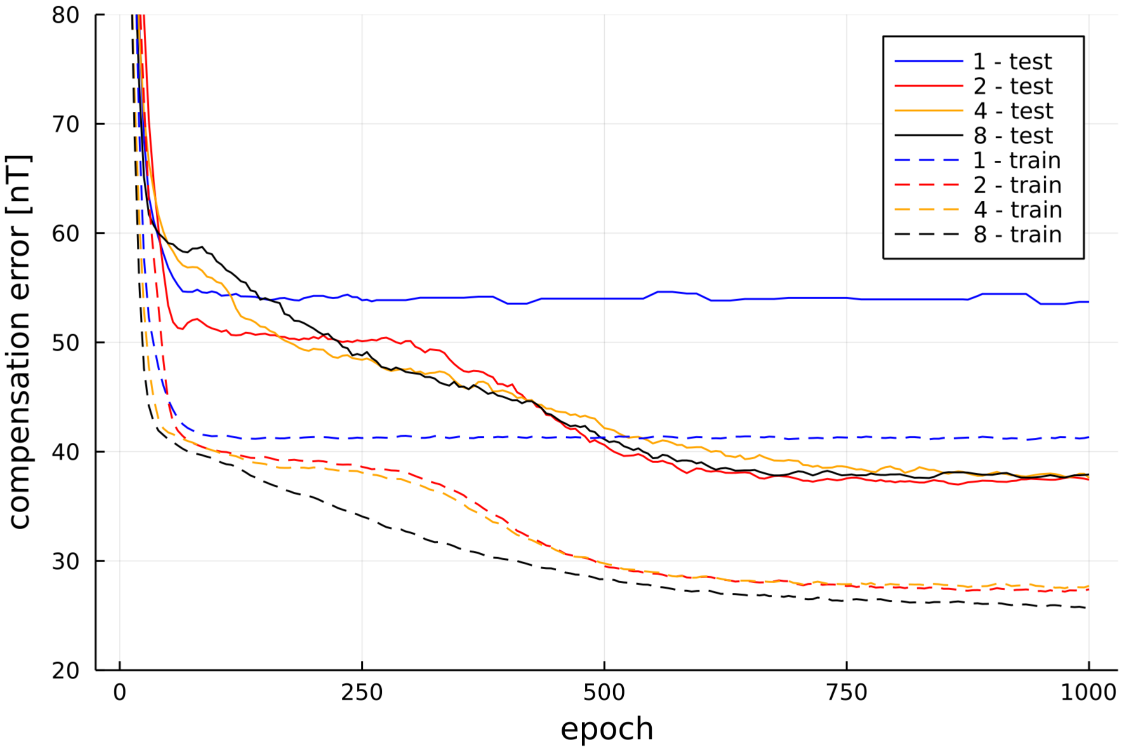


Figure 4-14: Model 2c performance with neural network size, small feature set 2. Small feature set 2 includes only scalar and vector magnetometer data for a total of 15 features. Mag 4 and Flux A are used here.

Similar to Figures 4-10 and 4-11, small feature set 2 was evaluated with each neural network-based aeromagnetic compensation model, as shown in Figure 4-15. As before, the models perform fairly similarly, with the exception of model 2d, which clearly performed worse in this case. Model 2c achieves the lowest compensation error at 750 epochs, while models 1, 2a, and 2b achieve similar performance at 1000 epochs. This further confirms the choice to evaluate model 2c in greater depth.

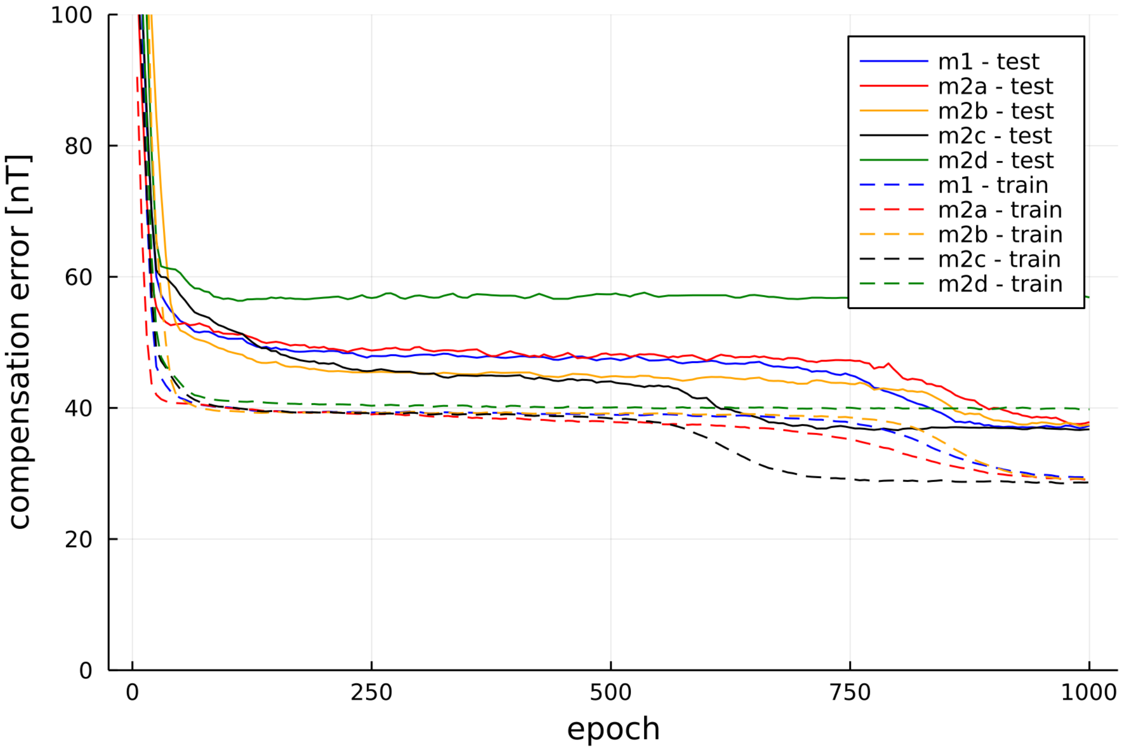


Figure 4-15: Neural network-based model comparison, small feature set 2. Small feature set 2 includes only scalar and vector magnetometer data for a total of 15 features. Mag 4 and Flux A are used here.

Using 2 nodes, 750 epochs, and Flux A, but otherwise using the hyperparameters listed in Table 4.9, the small feature sets were evaluated. PCA with whitening was not used here, since there is uncertainty in the best number of components to include and training runtime was not a concern. As can be seen in Tables 4.13 and 4.14, the neural network-based aeromagnetic compensation, model 2c, performs best for each magnetometer. Using the classical Tolles-Lawson model can only achieve 19 nT of error, while both the PLSR and model 2c can achieve under 10 nT with Mag 5. The more corrupted magnetometers, Mags 3 and 4, can achieve around 50 nT of error with PLSR and/or model 2c, which can still be used for MagNav, as explained in section 6.1. Model 2c performs the same or better than PLSR is all cases. Finally, feature set 1 outperforms feature set 2, indicating that the two terms from the direction cosine matrix and limited current and voltage sensor data does indeed improve aeromagnetic compensation.

An interesting side result is the effect of including the 4^{th} order central difference term in small feature set 2. Without this term, the PLSR results are nearly exactly the same. However, the model 2c results are noticeably worse with errors of 74, 57, and 12 nT for Mags 3, 4, and 5, respectively. The neural network may be using this term, which represents the magnitude of magnetic noise, effectively when the magnetometer has high levels of corruption.

Table 4.13: Single magnetometer testing performance, small feature set 1. Small feature set 1 includes magnetometer data, two terms from the direction cosine matrix, and limited current and voltage sensor data for a total of 26 features.

Scalar Mag	Vector Mag	Tolles-Lawson Error	PLSR Error	Model 2c Error
3	A	91	60	45
4	A	105	26	22
5	A	19	9.7	5.5

Table 4.14: Single magnetometer testing performance, small feature set 2. Small feature set 2 includes only scalar and vector magnetometer data for a total of 15 features.

Scalar Mag	Vector Mag	Tolles-Lawson Error	PLSR Error	Model 2c Error
3	A	91	65	52
4	A	105	54	38
5	A	19	11	11

Chapter 5

Magnetic Navigation

A magnetic navigation system collects magnetic field data using a magnetometer and uses magnetic anomaly maps to determine the current location by correcting the drift from an inertial navigation system (INS). Magnetic navigation is enabled by variations in earth's magnetic anomaly field. However, magnetic measurements can only measure the total field and thus contain the core field and temporal variations in addition to the anomaly field. In practice, the core field at a given (estimated) position and time can be accurately estimated using a core field model, such as the IGRF model [6]. The temporal variations can also be estimated at a given (estimated) position and time using ground-based reference measurements taken at a stationary base station within the flight region [11]. Alternatively, the temporal variations can be modeled within the navigation algorithm, which is described in section 5.3.

Magnetic navigation relies on an INS, which uses accelerometers and gyroscopes to measure specific forces and angular velocities, respectively [89]. The INS provides position, velocity, and attitude estimates. However, velocity is calculated with a single integral and position is calculated with a double integral. Thus, small errors in the specific force and angular velocity measurements are integrated into progressively larger errors in velocity and even larger errors in position. Matching magnetic measurements to a magnetic anomaly map is one method to correct this drift.

Recall that the strength of a static magnetic dipole field decays with the inverse of cubic distance from the source [4]. This high drop-off rate in magnetic fields means that it is difficult for disturbances to affect magnetic sensors from a distance without exhorting a significant amount of power, making it difficult to jam the magnetic signal from ground stations or anywhere other than near the measurement location. However, this also means that magnetic measurements are extremely sensitive to nearby magnetic disturbances. The predominant issue with magnetic navigation comes from magnetic interference generated by the aircraft itself, which is typically handled through aeromagnetic compensation.

The end goal of this work is improving aeromagnetic compensation to enable high-accuracy airborne magnetic anomaly navigation. Thus, understanding, evaluating, and possibly improving the magnetic navigation algorithm is essential. To this end, this section describes previous work on magnetic navigation, then provides background on navigation algorithms and the baseline model used for airborne magnetic anomaly navigation. This is followed by a description of two approaches for improving the navigation algorithm, namely online aeromagnetic compensation and covariance-adaptive filtering. The effect of specific aircraft maneuvers on filter observability is also evaluated. Finally, navigation performance with various navigation algorithms is compared on a common flight line.

5.1 Previous Work on Magnetic Navigation

Magnetic navigation is an active area of research for not only aircraft, but also underwater, space, ground, and indoor vehicles. These other platforms have been more thoroughly investigated than airborne magnetic anomaly navigation over the past decades and can provide insight on the general capabilities of magnetic navigation. Thus, this section describes previous work in the area of magnetic navigation for different types of vehicles.

5.1.1 Underwater Magnetic Navigation

One of the earliest to report on the potential for magnetic anomaly navigation was Tyren in 1987 [90]. Tyren discussed the use of magnetic measurements to calculate submarine ground speed using two offset magnetometers, as well as the superposition of vehicle magnetic field onto the earth magnetic field of the earth. Also mentioned was the potential to use the magnetic anomalies from underwater structures as natural beacons for navigation. Jie used a Kalman Filter for underwater magnetic navigation and found that navigation is improved with mean diurnal correction compared to using no diurnal correction at all [91]. This has applicability for airborne magnetic anomaly navigation, where only an estimate of the diurnal may be known rather than the specific value for the entire flight line.

5.1.2 Space Magnetic Navigation

Shorshi and Bar-Itzhack [92] and Psiaki et al. [93] used an extended Kalman Filter (EKF) to relate magnetometer measurements taken on a satellite to the International Geomagnetic Reference Field (IGRF) core magnetic field model. Both were able to obtain accuracies of less than 10 km. Note that at satellite altitudes the magnetic anomaly field is negligible. Using a batch filter with both magnetometer and sun sensor data, Psiaki was able to achieve 500 m accuracy for inclined low Earth orbits [94]. These examples provide an approximate upper bound on the navigation errors than can be obtained using airborne magnetic anomaly navigation, assuming a clean magnetic signal can be received.

5.1.3 Ground Magnetic Navigation

Shockley was able to show the viability of road navigation using magnetic field variations with a ground vehicle [95, 96]. Magnetometer measurements were used to estimate position using three different likelihood functions. Additionally, a magnetic map-matching particle filter was used to correct INS drift. Meter-level position accuracy was demonstrated when sufficient magnetic field features were available. However, different from airborne magnetic anomaly navigation is the availability of local magnetic field perturbations from roadways, buildings, and other structures. Road-based vehicles also have the benefit of navigating in a "road reference frame" that has fewer degrees of freedom than a world reference frame [97].

5.1.4 Indoor Magnetic Navigation

Similar to road navigation, indoor navigation is able to use magnetic signals from often known, man-made magnetic sources. Li et al. discuss the use of these local anomalies, referred to as a magnetic "fingerprint," as well as the challenges posed by moving magnetic objects and electronic devices [98]. Successful global self-localization for both a mobile robot and human with a wearable magnetometer was demonstrated using Monte Carlo Localization (MCL) by Haverinen and Kemppainen [99]. Sub-meter accuracy was demonstrated for a mobile robot using multiple magnetic sensors, three magnetic field maps, and a particle filter system [100]. Storms was able to obtain sub-meter position errors in a laboratory hallway environment using an INS with aiding from three-axis magnetometer data and a Kalman filter [101, 102].

5.1.5 Airborne Magnetic Anomaly Navigation

In 2006, Goldenberg reviewed terrain navigation efforts (using land topography and geomagnetic maps) and suggested that magnetic sensing technologies were accurate enough for airborne magnetic anomaly navigation as well [14]. Goldenberg also provided one of the first in-depth descriptions of an INS system aided with a best fit between real-time magnetic measurements and a corresponding magnetic map. Finally, Goldenberg addressed issues in obtaining good magnetic measurements in an aircraft environment, as well as the limitation of magnetic anomaly map availability.

Airborne magnetic anomaly navigation is relatively unexplored and there is not a considerable amount of experimental results. Wilson et al. were able to obtain position accuracies of 600-1200 m for 1 hr or longer flights using United States Geological Survey (USGS) maps in 2006 [103]. In this case, the magnetic measurements were used to aid an airspeed dead-reckoning system. A key limitation was the use of a less accurate vector magnetometer as the primary magnetic sensor.

More recently, Canciani achieved 13 m distance root mean square (DRMS) error during a 1 hr flight near Louisa, Virginia, demonstrating the viability of airborne magnetic anomaly navigation [13, 11, 2]. A marginalized particle filter (MPF), more specifically a Rao-Blackwellized particle filter (RBPF) [104, 105], was used as the navigation algorithm. However, the aircraft used was a magnetically optimized geo-survey aircraft, which used magnetic measurements from a tail stinger. Lee used a simultaneous localization and mapping (SLAM) approach to resolve the issue of map availability [106, 107]. Less than 20 m DRMS error was demonstrated for flights at 150 m above ground level (AGL), though again a geo-survey aircraft was used for data collection. Finally, Canciani achieved 59 m DRMS errors on 1.5 hr flights with an F-16 aircraft using online Tolles-Lawson aeromagnetic compensation [20]. This approach is described in section 5.5.

5.2 Navigation Algorithms

The goal in navigation is to move from a start to an end location. Knowing the position at any given time is not necessarily required, but it is often desired. The ability to obtain accurate navigation state estimates, such as position, facilitates accurate navigation [108]. A navigation filter processes measurements and prior information into navigation states.

5.2.1 Kalman Filters

One of the most widely used navigation filters is the Kalman filter [109]. A Kalman filter is analogous to the recursive least squares algorithm in that state estimates are updated recursively based on the prediction error between the previous state and new input data [110]. The Kalman filter is an optimal estimator for a linear system with zero-mean white Gaussian noise. The discrete state dynamics equation for a Kalman filter is

$$\boldsymbol{x}_{k+1} = \boldsymbol{\Phi}_k \boldsymbol{x}_k + \boldsymbol{B}_k \boldsymbol{u}_k + \boldsymbol{w}_k \tag{5.1}$$

where $\boldsymbol{x}$ is the state vector, $\boldsymbol{\Phi}$ is the state transition matrix, $\boldsymbol{B}$ is a matrix that determines the change in the state based on the system input $\boldsymbol{u}$, and $\boldsymbol{w}$ is Gaussian noise with $\boldsymbol{w} \sim \mathcal{N}(0, \boldsymbol{Q})$. Thus, the process is corrupted by process noise $\boldsymbol{w}$ with covariance $\boldsymbol{Q}$. The measurement update equation is

$$\boldsymbol{z}_k = \boldsymbol{H}_k \boldsymbol{x}_k + \boldsymbol{v}_k \tag{5.2}$$

where $\boldsymbol{z}$ is the measurement, $\boldsymbol{H}$ is a matrix that determines the change in the measurement based on the state, and $\boldsymbol{v}$ is Gaussian process noise with $\boldsymbol{v} \sim \mathcal{N}(0, \boldsymbol{R})$ that is uncorrelated with $\boldsymbol{w}$. Thus, the measurement is corrupted by measurement noise $\boldsymbol{v}$ with covariance $\boldsymbol{R}$. The time-varying prediction (a priori) equations for the Kalman filter are

$$\hat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{\Phi}_k \hat{\boldsymbol{x}}_{k-1} + \boldsymbol{B}_k \boldsymbol{u}_k \tag{5.3}$$

$$\boldsymbol{P}_{k|k-1} = \boldsymbol{\Phi}_k \boldsymbol{P}_{k-1} \boldsymbol{\Phi}_k^T + \boldsymbol{Q}_k \tag{5.4}$$

where $\hat{\boldsymbol{x}}$ and $\boldsymbol{P}$ are the state (mean) vector and covariance matrix, respectively. Note that $\hat{\boldsymbol{x}}$ represents a state estimate, while $\boldsymbol{x}$ represents the true state. An initial condition must first be provided for both the state vector and covariance matrix. After a new measurement is received, the state vector and covariance matrix are updated using the time-varying update (a posteriori) equations for the Kalman filter,

$$\boldsymbol{K}_k = \boldsymbol{P}_{k|k-1} \boldsymbol{H}_k^T (\boldsymbol{H}_k \boldsymbol{P}_{k|k-1} \boldsymbol{H}_k^T + \boldsymbol{R}_k)^{-1} \tag{5.5}$$

$$\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k (\boldsymbol{z}_k - \boldsymbol{H}_k \hat{\boldsymbol{x}}_{k|k-1}) \tag{5.6}$$

$$\boldsymbol{P}_k = (I - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{P}_{k|k-1} \tag{5.7}$$

where $\boldsymbol{K}$ is the Kalman gain, which represents the level of trust in the measurement $\boldsymbol{z}$. If the Kalman gain $\boldsymbol{K}$ is small (e.g. due to a large measurement noise covariance $\boldsymbol{R}$), then the measurement is largely ignored and the state vector is highly dependent on the prediction equations. Conversely, if the Kalman gain $\boldsymbol{K}$ is large, then the state vector is updated such that the estimated measurement approaches the actual measurement.

5.2.2 Extended Kalman Filters

The Kalman filter algorithm may be extended to systems with nonlinear dynamics via the extended Kalman filter (EKF) [111, 112]. The basic idea is to linearize the state-space model at the most recent state estimate. The state update equation becomes

$$\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_{k|k-1}) \tag{5.8}$$

where

$$\hat{\boldsymbol{x}}_{k|k-1} = f(\hat{\boldsymbol{x}}_{k-1}, \boldsymbol{u}_k, k) \tag{5.9}$$

and

$$\hat{\boldsymbol{z}}_{k|k-1} = h(\hat{\boldsymbol{x}}_{k|k-1}, k) \tag{5.10}$$

which are nonlinear state transition and measurement models. Additionally,

$$\boldsymbol{\Phi}_k = \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\hat{\boldsymbol{x}}_{k-1},\boldsymbol{u}_k} \tag{5.11}$$

and

$$\boldsymbol{H}_k = \left.\frac{\partial h}{\partial \boldsymbol{x}}\right|_{\hat{\boldsymbol{x}}_{k|k-1}} \tag{5.12}$$

are the state transition and measurement Jacobian matrices, respectively.

5.3 Baseline Model

The baseline navigation algorithm used for this work is an extended Kalman filter (EKF), as shown in Figure 5-1. Here, there is no system input $\boldsymbol{u}$, as magnetic navigation is passive. However, INS position (latitude, longitude, altitude), velocities (north, east, down), specific forces (north, east, down), and attitude (direction cosine matrix) are required for the dynamics model. The INS by itself allows for dead-reckoning, but it drifts over time unless corrected via some measurement. The measurement $\boldsymbol{z}$ is the compensated total field measurement from a scalar magnetometer. The baseline aeromagnetic compensation is performed using the classical Tolles-Lawson model, described in section 3.1.

Other than an initial state vector $\boldsymbol{x}_0$ and initial covariance matrix $\boldsymbol{P}_0$, in general the required components for the EKF are the process noise covariance $\boldsymbol{Q}$, measurement noise covariance $\boldsymbol{R}$, nonlinear state transition model f and its Jacobian matrix $\boldsymbol{\Phi}$, and nonlinear measurement model h and its Jacobian matrix $\boldsymbol{H}$. However, for this particular application, only the measurement equation is nonlinear, not the linearized state dynamics. The state transition model is based on a 9-state Pinson error model, which is a standard 1st-order dynamics model for an INS-based system [113].

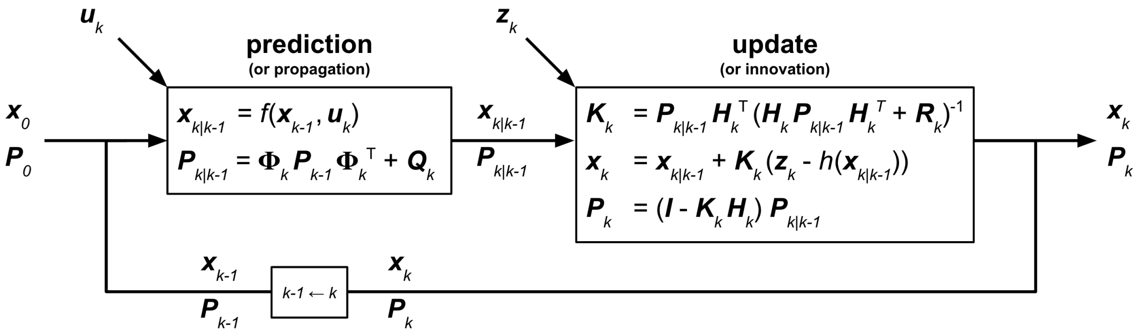


Figure 5-1: Extended Kalman filter.

The state vector for this model is

$$\hat{\boldsymbol{x}} = \begin{bmatrix} \delta \boldsymbol{p} \\ \delta \boldsymbol{v} \\ \boldsymbol{\epsilon} \end{bmatrix} \tag{5.13}$$

where $\delta \boldsymbol{p}$, $\delta \boldsymbol{v}$, and $\boldsymbol{\epsilon}$ are the three-dimensional INS position (latitude, longitude, altitude), velocity (north, east, down), and tilt (north, east, down) errors, respectively. However, additional states are added to this model for improved performance. The accelerometer bias $\boldsymbol{a}$ and gyroscope bias $\boldsymbol{g}$ can be estimated with six additional states, one for each dimension. Two barometer states, aiding altitude error δh_a and vertical acceleration error $\delta \hat{a}$, constrain the vertical channel, which is otherwise unstable for an INS [13]. Generally the barometer is precise, but not accurate, which is why this aiding is necessary. Finally, a world-frame time-correlated magnetic bias state S models temporal variations (e.g. space weather) and the map bias via a first-order Gauss-Markov (FOGM) random process. This final state is only effective if the frequency of temporal variations is outside the magnetic anomaly field frequency band, which is often the case for low frequency space weather effects (or a map bias).

This results in an 18-state model with the state vector

$$\hat{\boldsymbol{x}} = \begin{bmatrix} \delta\boldsymbol{p} \\ \delta\boldsymbol{v} \\ \boldsymbol{\epsilon} \\ \boldsymbol{a} \\ \boldsymbol{g} \\ \delta h_a \\ \delta\hat{a} \\ S \end{bmatrix}. \tag{5.14}$$

The remainder of this section describes the dynamics and measurement models in more detail.

5.3.1 Dynamics Model

The continuous state transition model uses a 18 × 18 matrix, which can be broken down by block as

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{F}_{pp} & \boldsymbol{F}_{pv} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \boldsymbol{F}_{pb} & \mathbf{0}_{3\times1} \\ \boldsymbol{F}_{vp} & \boldsymbol{F}_{vv} & \boldsymbol{F}_{v\epsilon} & \boldsymbol{F}_{va} & \mathbf{0}_{3\times3} & \boldsymbol{F}_{vb} & \mathbf{0}_{3\times1} \\ \boldsymbol{F}_{\epsilon p} & \boldsymbol{F}_{\epsilon v} & \boldsymbol{F}_{\epsilon\epsilon} & \mathbf{0}_{3\times3} & \boldsymbol{F}_{\epsilon g} & \mathbf{0}_{3\times2} & \mathbf{0}_{3\times1} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \boldsymbol{F}_{aa} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times2} & \mathbf{0}_{3\times1} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \boldsymbol{F}_{gg} & \mathbf{0}_{3\times2} & \mathbf{0}_{3\times1} \\ \boldsymbol{F}_{bp} & \mathbf{0}_{2\times3} & \mathbf{0}_{2\times3} & \mathbf{0}_{2\times3} & \mathbf{0}_{2\times3} & \boldsymbol{F}_{bb} & \mathbf{0}_{2\times1} \\ \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times2} & \boldsymbol{F}_{ss} \end{bmatrix}_{18\times18} . \tag{5.15}$$

The eight blocks in the upper left corner of $\boldsymbol{F}$ in (5.15) form the state transition matrix for the original 9-state Pinson error model with the states in (5.13). These relate position to position errors

$$\boldsymbol{F}_{pp} = \begin{bmatrix} 0 & 0 & -\frac{v_n}{R^2} \\ \frac{v_e \ \tan L}{R \ \cos L} & 0 & -\frac{v_e}{R^2 \ \cos L} \\ 0 & 0 & -k_1 \end{bmatrix}_{3\times3} \tag{5.16}$$

velocity to position errors

$$\boldsymbol{F}_{pv} = \begin{bmatrix} \frac{1}{R} & 0 & 0 \\ 0 & \frac{1}{R \ \cos L} & 0 \\ 0 & 0 & -1 \end{bmatrix}_{3\times3} \tag{5.17}$$

position to velocity errors

$$\boldsymbol{F}_{vp} = \begin{bmatrix} -v_e(2\Omega \ \cos L + \frac{v_e}{R \ \cos^2 L}) & 0 & \frac{v_e^2 \ \tan L - v_n v_d}{R^2} \\ 2\Omega(v_n \ \cos L - v_d \ \sin L) + \frac{v_n v_e}{R \ \cos^2 L} & 0 & -v_e \frac{v_n \ \tan L + v_d}{R^2} \\ 2\Omega v_e \ \sin L & 0 & \frac{v_n^2 + v_e^2}{R^2} + k_2 \end{bmatrix}_{3\times3} \tag{5.18}$$

velocity to velocity errors

$$\boldsymbol{F}_{vv} = \begin{bmatrix} \frac{v_d}{R} & -2(\Omega \ \sin L + \frac{v_e \ \tan L}{R}) & \frac{v_n}{R} \\ 2\Omega \ \sin L + \frac{v_e \ \tan L}{R} & \frac{v_n \ \tan L + v_d}{R} & 2\Omega \ \cos L + \frac{v_e}{R} \\ -\frac{2v_n}{R} & -2(\Omega \ \cos L + \frac{v_e}{R}) & 0 \end{bmatrix}_{3\times3} \tag{5.19}$$

tilt to velocity errors

$$\boldsymbol{F}_{v\epsilon} = \begin{bmatrix} 0 & -f_d & f_e \\ f_d & 0 & -f_n \\ -f_e & f_n & 0 \end{bmatrix}_{3\times3} \tag{5.20}$$

position to tilt errors

$$\boldsymbol{F}_{\epsilon p} = \begin{bmatrix} -\Omega\ \sin L & 0 & -\frac{v_e}{R^2} \\ 0 & 0 & \frac{v_n}{R^2} \\ -\Omega\ \cos L - \frac{v_e}{R\ \cos^2 L} & 0 & \frac{v_e\ \tan L}{R^2} \end{bmatrix}_{3\times3} \tag{5.21}$$

velocity to tilt errors

$$\boldsymbol{F}_{\epsilon v} = \begin{bmatrix} 0 & \frac{1}{R} & 0 \\ -\frac{1}{R} & 0 & 0 \\ 0 & -\frac{\tan L}{R} & 0 \end{bmatrix}_{3\times3} \tag{5.22}$$

and tilt to tilt errors

$$\boldsymbol{F}_{\epsilon\epsilon} = \begin{bmatrix} 0 & -(\Omega\ \sin L + \frac{v_e\ \tan L}{R}) & \frac{v_n}{R} \\ \Omega\ \sin L + \frac{v_e\ \tan L}{R} & 0 & \Omega\ \cos L + \frac{v_e}{R} \\ -\frac{v_n}{R} & -(\Omega\ \cos L + \frac{v_e}{R}) & 0 \end{bmatrix}_{3\times3} \tag{5.23}$$

where the variables in (5.16)–(5.23) are listed in Table 5.1. Beyond the eight blocks forming the nine-state Pinson state transition matrix are nine additional blocks that include the accelerometer, gyroscope, barometer, and temporal variation errors.

Table 5.1: State transition matrix variables.

Variable	Units	Description
L	rad	latitude
v_n	m/s	north velocity
v_e	m/s	east velocity
v_d	m/s	down velocity
f_n	m/s^2	north specific force
f_e	m/s^2	east specific force
f_d	m/s^2	down specific force
C_b^n	-	direction cosine matrix (body to navigation)
τ_b	s	barometer time constant
τ_a	s	accelerometer time constant
τ_g	s	gyroscope time constant
τ_{tv}	s	temporal variation time constant
k_1	1/s	barometer aiding constant
k_2	1/s^2	barometer aiding constant
k_3	1/s^3	barometer aiding constant
R	m	WGS-84 radius of Earth (6378137)
Ω	rad/s	rotation rate of Earth ($7.2921151467 \times 10^{-5}$)

The accelerometer to velocity errors are related using the *body to navigation* direction cosine matrix,

$$\boldsymbol{F}_{va} = C_b^n. \tag{5.24}$$

The accelerometer errors are governed by a time constant τ_a,

$$\boldsymbol{F}_{aa} = \begin{bmatrix} -\frac{1}{\tau_a} & 0 & 0 \\ 0 & -\frac{1}{\tau_a} & 0 \\ 0 & 0 & -\frac{1}{\tau_a} \end{bmatrix}_{3\times 3}. \tag{5.25}$$

The gyroscope to tilt errors are related using the direction cosine matrix,

$$\boldsymbol{F}_{\epsilon g} = -C_b^n. \tag{5.26}$$

The gyroscope errors are governed by a time constant τ_g,

$$\boldsymbol{F}_{gg} = \begin{bmatrix} -\frac{1}{\tau_g} & 0 & 0 \\ 0 & -\frac{1}{\tau_g} & 0 \\ 0 & 0 & -\frac{1}{\tau_g} \end{bmatrix}_{3\times 3}. \tag{5.27}$$

The barometer is used for third order altitude aiding with

$$\boldsymbol{F}_{pb} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ k_1 & 0 \end{bmatrix}_{3\times 2}, \tag{5.28}$$

$$\boldsymbol{F}_{vb} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -k_2 & 1 \end{bmatrix}_{3\times 2}, \tag{5.29}$$

$$\boldsymbol{F}_{bp} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & k_3 \end{bmatrix}_{2\times 3}, \tag{5.30}$$

and

$$\boldsymbol{F}_{bb} = \begin{bmatrix} -\frac{1}{\tau_b} & 0 \\ -k_3 & 0 \end{bmatrix}_{2\times 2}. \tag{5.31}$$

The temporal variation errors are governed by a time constant τ_{tv},

$$\boldsymbol{F}_{ss} = -\frac{1}{\tau_{\text{tv}}}. \tag{5.32}$$

The continuous state transition model may then be put into a discrete form as

$$f(\hat{\boldsymbol{x}}) = \boldsymbol{\Phi}\hat{\boldsymbol{x}} \tag{5.33}$$

with

$$\boldsymbol{\Phi} = e^{\boldsymbol{F}\Delta t} \tag{5.34}$$

where Δt is the time step between measurements. The process noise covariance $\boldsymbol{Q}$ is

$$\boldsymbol{Q} = \text{diag}(\begin{bmatrix} \mathbf{0}_{1\times 3} & \mathbf{VRW}_{1\times 3} & \mathbf{ARW}_{1\times 3} & \boldsymbol{D_a}\,_{1\times 3} & \boldsymbol{D_g}\,_{1\times 3} & D_b & 0 & D_{\text{tv}} \end{bmatrix})_{18\times 18} \tag{5.35}$$

where $\mathbf{VRW} = \sigma^2_{\text{VRW}}$ is zero-mean velocity random walk, $\mathbf{ARW} = \sigma^2_{\text{ARW}}$ is zero-mean angular (tilt) random walk, $\boldsymbol{D_a} = 2\sigma^2_a/\tau_a$ is accelerometer driving noise, $\boldsymbol{D_g} = 2\sigma^2_g/\tau_g$ is gyroscope driving noise, $D_b = 2\sigma^2_b/\tau_b$ is barometer driving noise, and $D_{\text{tv}} = 2\sigma^2_{\text{tv}}/\tau_{\text{tv}}$ is temporal variation driving noise.

5.3.2 Measurement Model

As previously described in section 1.2, a magnetometer can only measure the total magnetic field, which contains multiple magnetic sources. The actual magnetic measurements ($\boldsymbol{z}$ in (5.6), (5.8), and Figure 5-1) are compensated to remove aircraft interference. The classical model for this is again Tolles-Lawson, which uses a set of 18 coefficients and vector magnetometer measurements to compensate the scalar measurements. Ideally, after compensation the measurements only contain earth-related magnetic signals. The nonlinear measurement model is thus

$$h(\hat{\boldsymbol{x}}) = f_{\mathrm{IGRF}}(lat, lon, alt, t) + f_{\mathrm{map}}(lat, lon, alt) + S \tag{5.36}$$

where f_{IGRF} is the IGRF core field model, f_{map} is a predetermined magnetic anomaly map interpolation, and S is the temporal variation filter state. The position lat, lon, and alt is the filter-estimated (INS-corrected) latitude, longitude, and altitude, respectively. As described in section 1.1.4, the spatial magnetic variation that allows for navigation is contained in the magnetic anomaly field. The measurement Jacobian matrix is

$$H(\hat{\boldsymbol{x}}) = \begin{bmatrix} \frac{\delta h(\hat{\boldsymbol{x}})}{\delta lat} & \frac{\delta h(\hat{\boldsymbol{x}})}{\delta lon} & \frac{\delta h(\hat{\boldsymbol{x}})}{\delta alt} & \mathbf{0}_{1\times 15} \end{bmatrix}^T \tag{5.37}$$

which is simply the gradient of the IGRF core field model and magnetic anomaly map. Note that the gradient of the core field is much smaller than that of the magnetic anomaly map.

5.4 Online Aeromagnetic Compensation

The aeromagnetic compensation models in Chapters 3 and 4 follow the pre-processing strategy described in section 1.3. This means that compensation is completed independent of navigation. However, these may be more tightly integrated using online aeromagnetic compensation as an end-to-end strategy. This is not a new idea. It was previously suggested to use a recursive algorithm for adaptively updating compensation coefficients en-route to a geo-survey area [114, 73]. Additionally, online compensation of a vector magnetometer was described by Beravs et al. [115] and Siebler et al. [116] using an unscented Kalman filter (UKF) and Rao-Blackwellized particle filter (RBPF), respectively. This section describes two approaches to online aeromagnetic compensation.

5.5 Online Tolles-Lawson Aeromagnetic Compensation

Recently a navigation algorithm with integrated aeromagnetic compensation has been introduced by Canciani [20]. Instead of compensation pre-processing of the scalar magnetic measurements, the raw, uncompensated values are brought directly into the navigation filter. More specifically, an EKF with additional Tolles-Lawson coefficient and vector magnetometer states is used, which allows for the tracking of and adapting to the changing aircraft magnetic field. The original state vector, shown in (5.14), is modified to become

$$\hat{\boldsymbol{x}} = \begin{bmatrix} \delta\boldsymbol{p} \\ \delta\boldsymbol{v} \\ \boldsymbol{\epsilon} \\ \boldsymbol{a} \\ \boldsymbol{g} \\ \delta h_a \\ \delta\hat{a} \\ S \\ \boldsymbol{\beta}_{\text{TL}} \\ \boldsymbol{V} \end{bmatrix}. \tag{5.38}$$

where $\boldsymbol{\beta}_{\text{TL}}$ is the Tolles-Lawson coefficient vector and $\boldsymbol{V}$ is a three-dimensional vector magnetometer measurement. The Tolles-Lawson coefficients are length 19, which includes the typical 18 terms and a bias term. This results in a total of 40 states. Note that in the formulation presented by Canciani [20], the δh_a and $\delta\hat{a}$ terms are not included, resulting in 38 states.

To use this augmented EKF, the dynamics and measurement models must both be modified. The continuous state transition matrix is nearly the same as (5.15), except for additional $-\infty$ terms on the diagonal for the vector magnetometer states,

$$\boldsymbol{F}_{40\times40} = \begin{bmatrix} \boldsymbol{F}_{18\times18} & \boldsymbol{0}_{18\times19} & \boldsymbol{0}_{18\times3} \\ \boldsymbol{0}_{19\times18} & \boldsymbol{0}_{19\times19} & \boldsymbol{0}_{19\times3} \\ \boldsymbol{0}_{3\times18} & \boldsymbol{0}_{3\times19} & -\infty I_{3\times3} \end{bmatrix}_{40\times40} \tag{5.39}$$

where $\boldsymbol{F}_{18\times18}$ is the baseline state transition matrix and $\boldsymbol{F}_{40\times40}$ is the augmented version. Conceptually, the $-\infty$ terms represent overriding the vector magnetometer measurements every time step. This overriding is done through use of

$$\boldsymbol{B} = \begin{bmatrix} \mathbf{0}_{37\times 3} \\ \boldsymbol{I}_{3\times 3} \end{bmatrix}_{40\times 3} \tag{5.40}$$

where $\boldsymbol{B}$ is a matrix that determines the change in the state based on the system input, which is a vector magnetometer measurement in this case. This is used as shown in (5.3). The process noise covariance $\boldsymbol{Q}$ also has additional diagonal terms for the added states, which are modeled as Brownian motion (integrated white noise) with linearly growing variances. The Tolles-Lawson variances can be estimated by observing how much the coefficients change when calibrating the classical Tolles-Lawson model at multiple points in time. The vector magnetometer measurement variances are highly dependent on the magnetometer used and the level of aircraft interference, but should in general be greater than 100 nT. The measurement model is similar to (5.36), but now includes a compensation term,

$$h(\hat{\boldsymbol{x}}) = f_{\text{IGRF}}(lat, lon, alt, t) + f_{\text{map}}(lat, lon, alt) + S + f_{\text{TL}}(\boldsymbol{\beta}_{\text{TL}}, \boldsymbol{V}) \tag{5.41}$$

where f_{TL} is the Tolles-Lawson model using coefficients $\boldsymbol{\beta}_{\text{TL}}$ and vector magnetometer measurement $\boldsymbol{V}$. The measurement Jacobian matrix is similar to (5.37), except derivatives of the Tolles-Lawson coefficient and vector magnetometer states must be included. The derivatives for the Tolles-Lawson coefficient states are simply the Tolles-Lawson $\boldsymbol{A}$ matrix terms, while the derivatives for the vector magnetometer states are more involved. See [20] for details of these derivatives.

5.6 Online Neural Network-Based Aeromagnetic Compensation

An extension of the online Tolles-Lawson aeromagnetic compensation approach described in section 5.5 is online neural network-based aeromagnetic compensation. Instead of learning 19 varying Tolles-Lawson coefficient states, weights of a neural network are learned. The state vector in this case is

$$\hat{\boldsymbol{x}} = \begin{bmatrix} \delta \boldsymbol{p} \\ \delta \boldsymbol{v} \\ \boldsymbol{\epsilon} \\ \boldsymbol{a} \\ \boldsymbol{g} \\ \delta h_a \\ \delta \hat{a} \\ S \\ \boldsymbol{w}_{\mathrm{NN}} \end{bmatrix} . \tag{5.42}$$

where $\boldsymbol{w}_{\mathrm{NN}}$ are the neural network weights and biases. This model allows a nonlinear compensation model, as discussed in Chapter 4, to be integrated into the online compensation approach.

It was unknown prior to this work if this approach would provide improved performance compared to the classical Tolles-Lawson or online Tolles-Lawson models. It was also unknown if this approach is computationally tractable. Thus, a simple version of online "neural network-based" aeromagnetic compensation was first implemented, which is shown in Figure 5-2. This is essentially the same as the online Tolles-Lawson approach if a linear activation function is used. However, using this model verified that online neural network-based compensation can be implemented. The results are nearly identical to that of online Tolles-Lawson compensation, as expected. Differences in performance are seemingly due to different initializations of navigation filter parameters for the two model types.

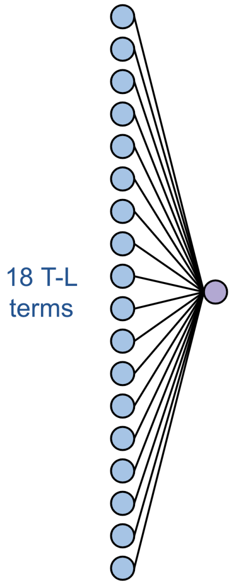


Figure 5-2: Simple online "neural network-based" aeromagnetic compensation.

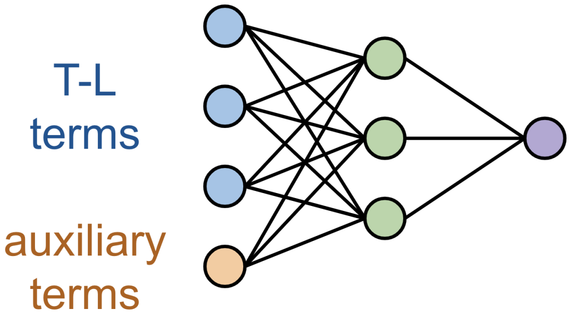


Figure 5-3: Online neural network-based aeromagnetic compensation. The weights (and biases) of the neural network are states in the EKF.

The complexity of the neural network can of course be increased, as shown in Figure 5-3, by using auxiliary data and/or a nonlinear activation function. As described in [110], the neural network can be trained directly within the EKF using weights (and biases) as states. This is similar to online Tolles-Lawson compensation, as previously described, with the notable difference being the use of backpropagation for the neural network weights in the measurement Jacobian H calculation.

A key aspect for implementation is initializing the covariance matrix corresponding to the neural network weights (and biases). The approach taken by Canciani [20] for online Tolles-Lawson aeromagnetic compensation does not work for online neural network-based aeromagnetic compensation. Instead, a recursive least squares (RLS) algorithm can be used with the weights and biases flattened into a coefficient vector. In the RLS algorithm, the coefficient vector β is updated as

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{K}_t \left(y_t - \hat{y}(t|\boldsymbol{\beta}_{t-1}) \right) \tag{5.43}$$

with

$$\boldsymbol{K}_t = \frac{\boldsymbol{P}_{t-1}\,\boldsymbol{\beta}_t}{1 + \boldsymbol{\beta}_t^T\,\boldsymbol{P}_{t-1}\,\boldsymbol{\beta}_t} \tag{5.44}$$

and

$$\boldsymbol{P}_t = \boldsymbol{P}_{t-1} - \frac{\boldsymbol{P}_{t-1}\,\boldsymbol{\beta}_t\,\boldsymbol{\beta}_t^T\,\boldsymbol{P}_{t-1}}{1 + \boldsymbol{\beta}_t^T\,\boldsymbol{P}_{t-1}\,\boldsymbol{\beta}_t} \tag{5.45}$$

where y is the output target, as in section 3.7, and $\hat{y}$ is the model output. The initially unknown covariance matrix $\boldsymbol{P}$ can be initialized with the identity matrix. Using the RLS algorithm with a small subset of training data (e.g. 100 samples), a fairly steady state covariance matrix can be determined and used for initialization within the online neural network-based aeromagnetic compensation algorithm. Similarly, the differences between successive RLS coefficient (weight and bias) updates can be used to estimate the standard deviations for the weights and biases, which are used for the driving noise of the respective states.

5.7 2021 SGL Flight Data

Though the primary data used in this work was collected during the summer of 2020 by SGL, an additional flight dataset was collected during the winter of 2021, again by SGL. For this work, the primary use of this secondary dataset is to look at the effect of specific aircraft maneuvers on filter observability, which is covered in the next section. First, there are some higher-level questions about the data itself and the applicability of models across different aircraft of the same type that this section addresses. Note that both sets of data collection flights used a Cessna 208B Grand Caravan, but it was not the same aircraft, i.e. a different "tail." Different magnetometers were also used, but they were the same types and placed in approximately the same locations, as described in Table 2.2.

As in section 2.2.1, first an example of the uncompensated (raw) scalar magnetometer measurements is shown in Figure 5-4. Note that flight line 2005.36 has nominally the same flight path and altitude as flight line 1007.02, data from which is shown in Figures 2-2 and 2-3. Once again, Mag 1 has nearly no corruption, but now Mag 4 follows the trend of Mag 1 *very* well instead of Mag 5. Mags 2 and 5 are noticeably worse than Mag 4, but both still largely match the trend of Mag 1. Mag 3 has significant corruption and could not be used for MagNav without aeromagnetic compensation. Overall, Mag 1 performs as before, while Mags 3 and 5 have much more corruption and Mags 2 and 4 have much less corruption, which is quantified in Tables 2.3 and 5.2.

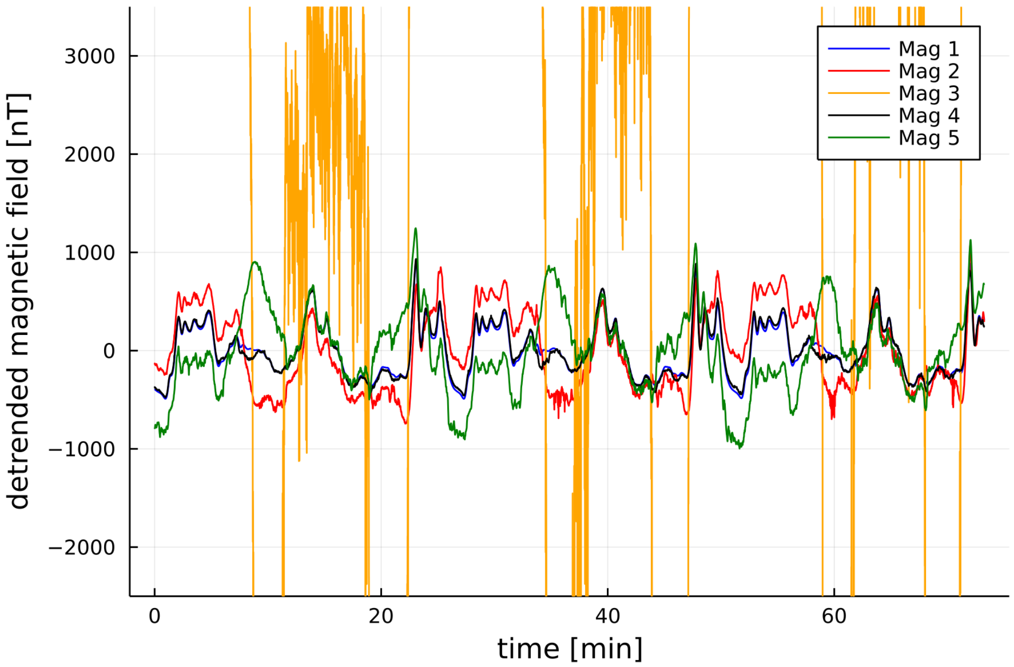


Figure 5-4: Uncompensated scalar magnetometers from flight line 2005.36. Due to their positions, the decreasing order of corruption of the uncompensated scalar magnetometers is: Mag 1, Mag 4, Mag 2, Mag 5, Mag 3.

The compensated scalar magnetometers from flight line 2005.36 are shown in Figure 5-5. Here, classical Tolles-Lawson aeromagnetic compensation with Flux A was again performed on each scalar magnetometer, but now using calibration flight line 2001.22. For flight line 2005.36, Mags 4 and 5 follow Mag 1 quite well, though slight errors remain. Mags 2 and 3 are improved with compensation, but still retain significant error compared to the others. Overall, Mags 1 and 5 perform as before, while Mag 3 has much more corruption and Mags 2 and 4 have much less corruption, which is again quantified in Tables 2.3 and 5.2.

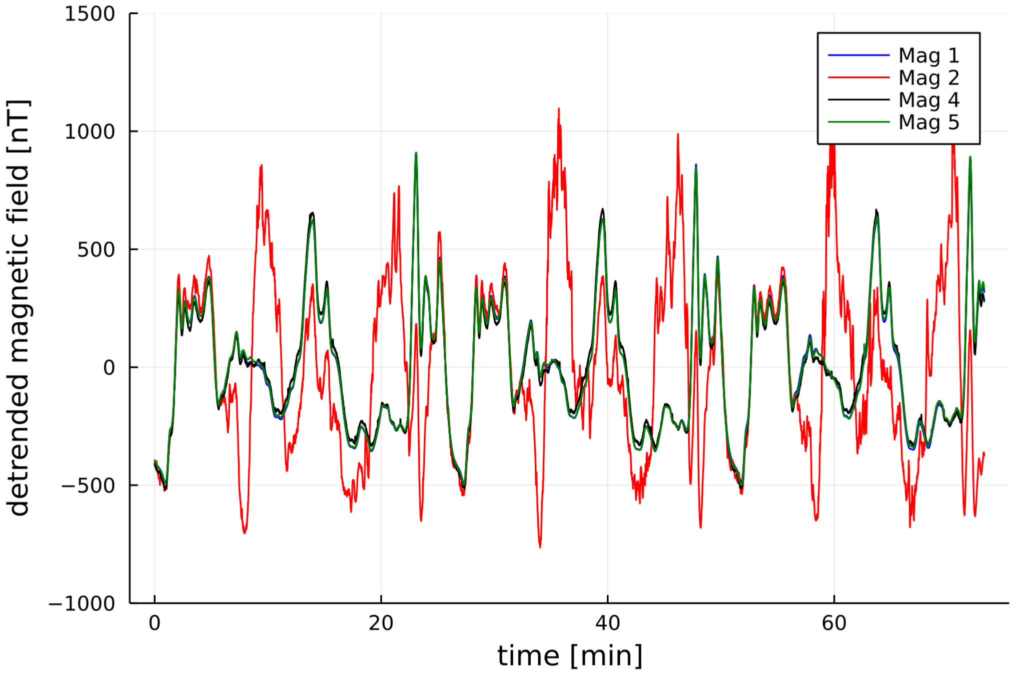


Figure 5-5: Compensated scalar magnetometers from flight line 2005.36. Due to their positions, the decreasing order of corruption of the compensated scalar magnetometers is: Mag 1, Mag 5, Mag 4, Mag 2, Mag 3. Mag 3 is not shown due to significant corruption.

The standard deviations of the scalar magnetometer errors from flight line 2005.36 are provided in Table 5.2. Mag 1 needs minimal or no compensation. The magnetic signal errors for Mags 2, 3, 4, and 5 significantly decrease with compensation. Note that, as in Table 2.3, compensation performance is fairly insensitive to which vector magnetometer is used.

Table 5.2: Aeromagnetic compensation errors for flight line 2005.36. The Tolles-Lawson model is used with each available vector magnetometer. Errors [nT] are in reference to the map values along the flight path.

Scalar Mag	Uncompensated	Flux A	Flux B	Flux C	Flux D
Mag 1	7.3	7.3	7.4	7.4	7.3
Mag 2	272	124	124	126	125
Mag 3	12077	7559	7443	7678	7460
Mag 4	37	22	22	22	22
Mag 5	370	11	11	11	10

Based on the scalar magnetometer differences between the two SGL datasets, it seems clear that aeromagnetic compensation models developed using the 2020 dataset do not readily work on the 2021 dataset. This has important implication for real-world, operational use. A model developed using a specific aircraft is not immediately applicable to another aircraft of the same type. Nonetheless, it is worthwhile to evaluate and quantify performance, first while attempting to use the same model and then from using a re-fitted or re-trained model.

Shown in Table 5.3 is a comparison of aeromagnetic compensation performance on flight line 1007.02 from the 2020 SGL dataset, and flight lines 2005.36, 2005.18, and 2005.38 from the 2021 SGL dataset. As previously explained, flight line 2005.36 follows the same nominal flight path at the same nominal altitude as flight line 1007.02. Additionally, flight lines 2005.18 and 2005.38 follow the same nominal flight path, but at different altitudes, than flight line 1007.02. All flight lines are evaluated using the models presented at the end of section 4.9. Specifically, the classical Tolles-Lawson, PLSR-based, and neural network-based (model 2c, small feature set 2) aeromagnetic compensation models were used with Mag 4 and Flux A. Model 2c was first trained an additional 250 epochs on the 2020 SGL training data (for a total of 1000 epochs), as this resulted in better performance on the 2021 SGL testing data. As expected, the flight lines from the 2021 SGL dataset perform extremely poorly, since this was a different aircraft with a different magnetic field. The performance here on the 2021 flight lines is absolutely not suitable for MagNav. This has important operational implications – each aeromagnetic compensation model must be aircraft-specific.

Table 5.3: Compensation performance on repeated flight line, no re-training.

SGL Dataset	Flight Line	Flight Altitude [m]	Tolles-Lawson Error [nT]	PLSR Error [nT]	Model 2c Error [nT]
2020	1007.02	800	54	32	34
2021	2005.36	800	170	268	228
2021	2005.18	2000	184	288	305
2021	2005.38	drape	170	271	242

However, these models can be re-trained using flight data from the 2021 SGL dataset, which is a form of transfer learning. Typically, transfer learning involves only re-training (i.e. tuning) the final couple of layers of a neural network [117]. However, for the single hidden layer models in this work, that is clearly not possible. Instead the entire neural network is re-trained using new flight data. The original neural network model weights completely change, so this is similar to starting with randomly initialized weights. Specifically, the following flight lines are used for re-training: 2001.00, 2001.01, 2001.02, 2001.03, 2001.04, 2001.05 (flight 2001, 13-Dec-2021), 2002.00, 2002.01, 2002.02, 2002.03, 2002.04 (flight 2002, 14-Dec-2021), 2001.06, 2001.07, 2001.08, 2001.09, 2001.10, 2001.11, 2001.12, 2001.13, 2001.14, 2001.15, 2001.16, 2001.17 (flight 2004, 19-Dec-2021). However, several different sub-datasets are used to evaluate the effect of reduced flight data availability.

First, all of the listed flight lines for flights 2001, 2002, and 2004 were used to re-train the models, the results from which are shown in Table 5.4. For the classical Tolles-Lawson model, only calibration flight line 2001.22 was used. The best performance for PLSR resulted when all components were kept. Model 2c from section 4.9 (trained 750 epochs) was retrained 1000 epochs. Here it can be seen that both PLSR and model 2c can achieve around 10 nT of error per flight line.

Table 5.4: Compensation performance on repeated flight line, re-training with 3 flights from the 2021 SGL dataset. Flight data from flights 2001, 2002, and 2003.

Flight Line	Flight Altitude [m]	Tolles-Lawson Error [nT]	PLSR Error [nT]	Model 2c Error [nT]
2005.36	800	21	5.8	8.4
2005.18	2000	25	13	8.3
2005.38	drape	22	8.3	11

The results from re-training with flight data from only flights 2001 and 2002 are shown in Table 5.5. Model 2c from section 4.9 (trained 750 epochs) was again retrained 1000 epochs with the 2021 SGL training data. Both PLSR and model 2c still outperform classical Tolles-Lawson, but the performance on flight line 2005.18 is of concern. The training data in this case not only contained less data, but also less data at higher altitudes, which may have resulted in worse performance with the higher altitude flight line. The performance on the drape flight line noticeably improved, which may have been for a similar reason – a larger portion of the training data was at lower altitudes.

Table 5.5: Compensation performance on repeated flight line, re-training with 2 flights from the 2021 SGL dataset. Flight data from flights 2001 and 2002.

Flight Line	Flight Altitude [m]	Tolles-Lawson Error [nT]	PLSR Error [nT]	Model 2c Error [nT]
2005.36	800	21	8.1	6.4
2005.18	2000	25	23	17
2005.38	drape	22	6.8	5.7

The results from re-training with flight data from only flight 2001 are shown in Table 5.6. In this case, the best performance for model 2c resulted from first training an additional 250 epochs on the 2020 SGL training data (for a total of 1000 epochs) then re-training for significantly longer epochs. Similar to the previously discussed results in Table 5.5, performance is worse overall when less training data is used. In particular, the performance on flight line 2005.18 is roughly the same for all of the models. This indicates that the classical model should not be abandoned when very limited flight data is available. Additionally, embedding the classical model in a SciML approach, such as model 2c, may help it remain accurate.

Table 5.6: Compensation performance on repeated flight line, re-training with 1 flight from the 2021 SGL dataset. Flight data from flight 2001.

Flight Line	Flight Altitude [m]	Tolles-Lawson Error [nT]	PLSR Error [nT]	Model 2c Error [nT]
2005.36	800	21	11	13
2005.18	2000	25	26	22
2005.38	drape	22	11	15

In generating the model 2c results shown in Table 5.6, it was noticed that substantially more training epochs were required. This could have been mitigated by increasing the learning rate, but it was left unchanged for consistency with all of the other results in this work. Shown in Figure 5-6 are the training and testing errors at each epoch. It can be seen that the best results are not achieved until around 10000 epochs. Beyond this point there is negligible improvement, then over-training. However, there is noticeable benefit, for both training and testing, in initializing the neural network weights with the model 2c weights that were determined when originally training on the 2020 SGL data rather than a random initialization.

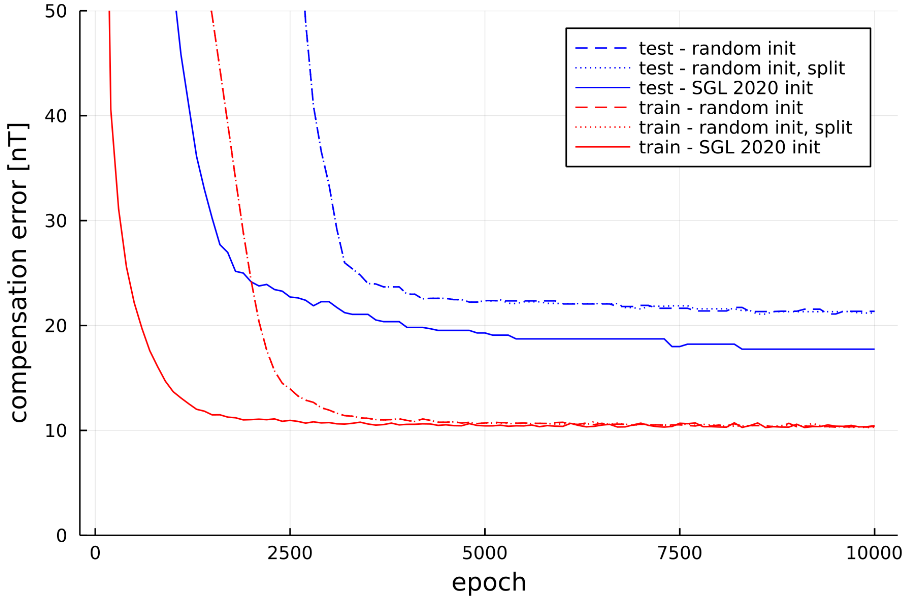


Figure 5-6: Model 2c comparison with different neural network weight initializations. Using the weights that were previously determined when training on the 2020 SGL training data resulted in better performance than randomly initialized weights.

5.8 Aircraft Maneuvers for Filter Observability

Unlike traditional aeromagnetic compensation, the concept of operations for online compensation is to have the filter compensation states updated in real-time. Thus, it may be beneficial to perform specific aircraft maneuvers that increase filter observability of the aircraft magnetic field, potentially en-route to the primary mission.

There may be a more optimal manner in which a pilot should fly for online compensation and thus navigation performance. For online applications, the aim should be to have as few calibration maneuvers as possible to avoid unnecessary pre-mission flight requirements. The goal here is to determine if some maneuvers are more important than others for online compensation.

Table 5.7: Aircraft maneuver sets for filter observability. Quick maneuvers, similar to the typical calibration box shown in Figure 3-4, were flown. Different maneuver types, counts, and amplitudes *per leg* were evaluated.

Number	Maneuvers	Amplitude [deg]
1	pitch ×2 box	2.5
2	pitch ×2 box	5
3	pitch ×2 box	10
4	pitch ×6 box	2.5
5	pitch ×6 box	5
6	pitch ×6 box	10
7	pitch, roll ×2 box	2.5, 5
8	pitch, roll ×2 box	5, 10
9	pitch, roll ×2 box	10, 20
10	pitch, roll ×6 box	2.5, 5
11	pitch, roll ×6 box	5, 10
12	pitch, roll ×6 box	10, 20
13	pitch, roll, yaw ×2 box	2.5, 5, 2.5
14	pitch, roll, yaw ×2 box	5, 10, 5
15	pitch, roll, yaw ×2 box	10, 20, 10
16	pitch, roll, yaw ×6 box	2.5, 5, 2.5
17	pitch, roll, yaw ×6 box	5, 10, 5
18	pitch, roll, yaw ×6 box	10, 20, 10

Within the 2021 SGL dataset are a variety of aircraft maneuver sets specifically chosen for filter observability, which are shown in Table 5.7. This is a similar idea to the sinusoidal aircraft maneuvers proposed by Leliak [45] for observability while performing a calibration flight, as shown in Figures 3-4 and 3-7. These maneuvers have the potential to be used for better determining the Tolles-Lawson coefficients (or neural network weights) during online compensation. All maneuvers were performed twice – once at a high altitude over a region with a small magnetic gradient and once at a low altitude over a region with a large magnetic gradient, as shown in Figure 5-7. Overall, the goal was to provide a set of maneuvers that can be individually analyzed in respect to compensation and navigation performance.

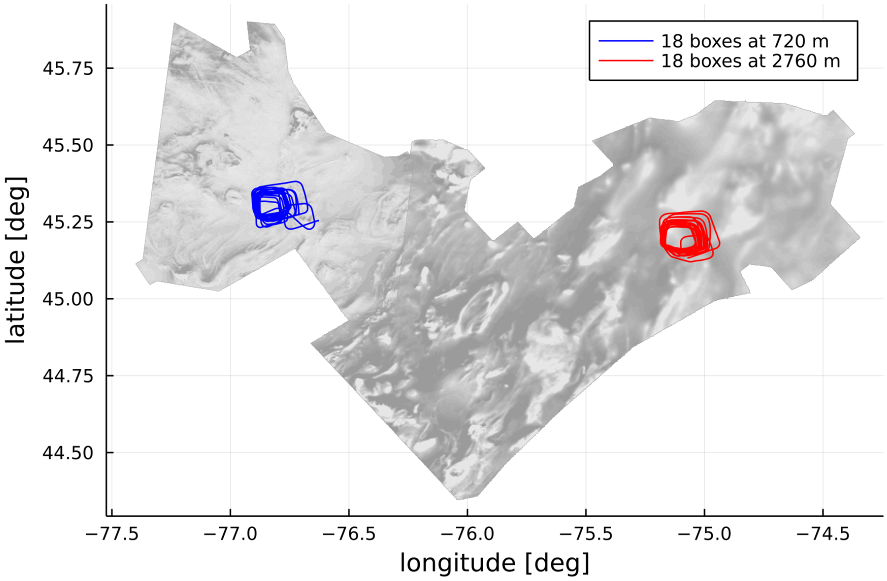


Figure 5-7: 18 boxes with aircraft maneuvers. Each box was completed with different aircraft maneuvers, as listed in Table 5.7.

The flight plan included a nominally straight and level flight portion between each maneuver in order to isolate the effect of the maneuver on filter observability. These nominally straight and level flight portions between maneuvers ended up being short in duration, meaning the overall flight length for each set of maneuvers only lasted approximately 10 min. This was generally not enough time for significant INS drift, so in many cases less than 50 m DMRS position error could be achieved per maneuver set without any magnetic measurement aiding. Additionally, before each set of maneuvers, a controlled, repeatable change in the aircraft magnetic field (e.g. moving a magnetic object within the cabin) was supposed to be completed, but this did not occur. Together these limit the analysis that can be performed here.

Furthermore, the magnitudes of the magnetic errors in the provided data limit which scalar magnetometers can be examined. This is shown in Table 5.2 and holds true for other flight lines. Mags 1 and 5 are very clean signals at approximately 10 nT of post-compensation error each. Mag 3 is completely unusable with over 1000 nT of post-compensation error. Mag 2 has over 100 nT of post-compensation error, which may be reasonable to examine, except Mag 2 also contains signal dropouts in over half of the maneuver sets listed in Table 5.7. Thus, this leaves Mag 4 to examine more closely, which typically has around 20 nT of post-compensation error. As before, Flux A is used.

The first thing to consider is treating each maneuver set in Table 5.7 as an actual calibration flight line. Tolles-Lawson model coefficients can then be determined and applied to flight lines for evaluation. The results of this are listed in Table 5.8. Flight lines 2005.36, 2005.18, and 2005.38 were all evaluated, but only 2005.36 is shown, as the trends are very similar for each flight line. The first thing to note is that overall the compensation performance is fairly consistent for each set of maneuvers. For reference, the uncompensated error is 37 nT, and the compensated error is 22 nT using calibration flight line 2001.22, as listed in Table 5.2. In all cases, there is benefit in performing compensation. However, in particular it appears that large roll and/or yaw amplitudes should be avoided. The standard set of calibration maneuvers, #17, performs well, which may be expected. Since the high altitude maneuvers were flown in a region with less of a magnetic gradient, those results are likely more reliable. Additionally, it may be desirable for an operational aircraft to do calibration maneuvers at higher altitude, where it would be flying anyways. In either case, a single pair of pitch and roll maneuvers (#7-9) shows promise for good calibration with limited flight requirements. These maneuvers took approximately 7 min to perform at 120 kts (62 m/s). Note that these maneuver sets were not flown on exactly the same path at exactly the same altitude, so there is error and uncertainty introduced by these discrepancies.

Table 5.8: Magnetic errors on flight line 2005.36 using aircraft maneuver sets. "Low" and "High" refer to using the low or high altitude sets of maneuvers to create Tolles-Lawson coefficients, respectively. Errors are compared to the SGL-compensated tail stinger magnetometer.

Number	Maneuvers	Amplitude [deg]	Low [nT]	High [nT]
1	pitch ×2 box	2.5	19	22
2	pitch ×2 box	5	24	20
3	pitch ×2 box	10	17	24
4	pitch ×6 box	2.5	27	24
5	pitch ×6 box	5	32	18
6	pitch ×6 box	10	20	22
7	pitch, roll ×2 box	2.5, 5	24	11
8	pitch, roll ×2 box	5, 10	19	16
9	pitch, roll ×2 box	10, 20	21	18
10	pitch, roll ×6 box	2.5, 5	18	23
11	pitch, roll ×6 box	5, 10	22	23
12	pitch, roll ×6 box	10, 20	27	23
13	pitch, roll, yaw ×2 box	2.5, 5, 2.5	24	19
14	pitch, roll, yaw ×2 box	5, 10, 5	20	19
15	pitch, roll, yaw ×2 box	10, 20, 10	18	21
16	pitch, roll, yaw ×6 box	2.5, 5, 2.5	18	17
17	pitch, roll, yaw ×6 box	5, 10, 5	16	20
18	pitch, roll, yaw ×6 box	10, 20, 10	26	26

The aircraft maneuver sets are now looked at in respect to navigation performance. Only the low altitude maneuvers are examined, since the high altitude maneuvers occurred at nominally 2760 m. This altitude requires too much upward continuation for the Eastern Ontario magnetic anomaly map to be reliable, as explained at the end of section 2.3.3. The navigation results using Tolles-Lawson with an EKF and online Tolles-Lawson are compared in Table 5.9. As stated earlier, the flight paths were nominally the same, but not exactly the same. The INS-only performance ("No Mag" in Table 5.9) varies between 11 and 85 m for exactly the same navigation filter parameters. It appears to vary randomly, as in there is no correlation with the types or amplitudes of the maneuvers. The calibrated stinger does have fairly consistent performance, with navigation errors between 21 and 33 m for all except 2 maneuver sets. For most, but not all, cases, the stinger magnetometer measurement improves navigation accuracy. Navigation performance likely could have been improved if the filter parameters were hand-tuned for each maneuver set, but that was viewed as "cheating" and instead consistent values were used, specifically: $R = 5^2\,\mathrm{nT}^2$ measurement noise covariance, $\sigma_{\mathrm{FOGM}} = 10\,\mathrm{nT}$ first-order Gauss-Markov (FOGM) catch-all bias, and $\tau_{\mathrm{FOGM}} = 50\,\mathrm{s}$ FOGM catch-all time constant.

Now looking at the columns of real interest, "T-L EKF" for Tolles-Lawson with an EKF and "Online T-L" for online Tolles-Lawson, the navigation errors are higher. For about half of the maneuver sets, the navigation performance is improved with the Mag 4 magnetometer measurements. However, the INS would continue to drift after completing each maneuver set, while the navigation error may remain constant or even decrease due to aiding from the magnetic measurements. Looking closer at the values, online Tolles-Lawson does better than offline Tolles-Lawson overall, particularly for large amplitude maneuvers. However, it does not appear that any specific types or amplitudes of the maneuvers result in clearly the best

Table 5.9: Navigation performance using aircraft maneuver sets. Mag 4 and Flux A are used here. DRMS navigation errors [m] are shown. “T-L” refers to the Tolles-Lawson model. “p,” “r,” and “y” refer to pitch, roll, and yaw, respectively.

Number	Maneuvers	Amplitude [deg]	T-L Mag [nT]	T-L EKF	Online T-L	Stinger	No Mag
1	p ×2 box	2.5	17	45	43	55	85
2	p ×2 box	5	15	64	53	25	38
3	p ×2 box	10	14	50	37	22	11
4	p ×6 box	2.5	15	40	31	45	39
5	p ×6 box	5	13	48	47	24	62
6	p ×6 box	10	13	73	38	21	39
7	p, r ×2 box	2.5, 5	14	43	27	27	16
8	p, r ×2 box	5, 10	11	30	21	24	36
9	p, r ×2 box	10, 20	12	111	101	24	44
10	p, r ×6 box	2.5, 5	12	34	43	27	44
11	p, r ×6 box	5, 10	16	56	29	30	23
12	p, r ×6 box	10, 20	18	55	36	27	21
13	p, r, y ×2 box	2.5, 5, 2.5	17	37	29	28	43
14	p, r, y ×2 box	5, 10, 5	19	60	50	23	58
15	p, r, y ×2 box	10, 20, 10	21	95	52	21	44
16	p, r, y ×6 box	2.5, 5, 2.5	15	56	32	25	12
17	p, r, y ×6 box	5, 10, 5	18	48	56	33	54
18	p, r, y ×6 box	10, 20, 10	17	32	36	33	57

navigation performance, though with the exception of the extreme outlier (#9), the pitch and roll maneuvers (#7-12) arguably perform the best overall. As with the tail stinger, consistent filter parameters were used. For Tolles-Lawson with an EKF: $R = 40^2\,\mathrm{nT}^2$ measurement noise covariance, $\sigma_{\mathrm{FOGM}} = 30\,\mathrm{nT}$ FOGM catch-all bias, and $\tau_{\mathrm{FOGM}} = 50\,\mathrm{s}$ FOGM catch-all time constant. For online Tolles-Lawson: $R = 40^2\,\mathrm{nT}^2$ measurement noise covariance, $\sigma_{\mathrm{FOGM}} = 50\,\mathrm{nT}$ FOGM catch-all bias, and $\tau_{\mathrm{FOGM}} = 30\,\mathrm{s}$ FOGM catch-all time constant. These were set based on the autocorrelation of the magnetic signal errors as well as trial and error with multiple flight lines in the 2021 SGL dataset.

Though these sets of maneuvers do provide some insights, they could have been carried out differently for better usability of the flight data. There are several recommendations for a future repeat of this type of flight plan. First, it would be useful to have more usable magnetometer data. Only Mag 4 was used for this analysis, as Mags 1 and 5 had very low corruption, while Mags 2 and 3 had very high corruption and/or dropouts. The placement of the magnetometers in the cabin was purposely carried over from the previous 2020 SGL data collection, during which the magnetometer placements were selected by SGL. It is recommended to do more analysis prior to the first set of maneuvers to achieve magnetometers with approximately the following compensated errors *without dropouts*: 5 nT (Mag 1), 25 nT (Mag 4), 50 nT, 100-150 nT (Mag 2), 200-300 nT. Several test flights may be necessary to achieve these magnetometer corruption levels, but this would be a worthwhile cost, especially if more documentation on how the corruption and dropouts change with magnetometer placement is provided.

Second, either a larger and/or higher magnetic anomaly map should be acquired, or the high altitude sets of maneuvers should be lowered from 2760 m to 2000 m at most. This would allow for navigation performance to be evaluated with the higher altitude maneuver sets. Next, to both better separate the maneuver sets and to allow for more INS drift, a longer nominally straight and level flight portion should be completed between each set of maneuvers. This could be an additional box on the other half of a "boxy figure 8." This would increase the overall flight length for each set of maneuvers to approximately 20 min. There should also be a clear, repeatable change in the aircraft magnetic field (e.g. moving a magnetic object within the cabin). This change should be recorded by the flight crew (e.g. SGL pilot comments) and obvious in all of the in-cabin scalar magnetometers with a goal of 10-50% changes in the aircraft magnetic field (compensation magnitude). A good time during the flight to make the repeatable change would be at the "cross" of the "boxy figure 8."

Finally, it may be desirable to reduce the number of maneuvers sets to evaluate. This is due to the additional flight time for the "boxy figure 8" and a goal to complete all sets of maneuvers in one day to reduce uncertainty in the flight data. In this case, it is recommended to exclude some or all of the pitch, roll, and yaw maneuver sets (#13-18), with #16-18 being the first choice to exclude. The calibration box itself already includes yaw at each corner of the box, and to a lesser extent during roll, as shown in Figure 3-7. The results here show that pitch and roll alone may be adequate, which would reduce the overall pre-mission flight requirements.

5.9 Covariance-Adaptive Filtering

An additional end-to-end model that may improve navigation performance is covariance-adaptive filtering. Here, the noise covariance used by the navigation filter is updated in real-time, rather than using a static value. In this section, first covariance-adaptive Kalman filters are discussed, followed by recurrent neural networks and covariance-adaptive neural filters.

5.9.1 Covariance-Adaptive Kalman Filters

As indicated by (5.1)–(5.5), Kalman filters require knowledge of the process noise covariance matrix $\boldsymbol{Q}$ and the measurement noise covariance matrix $\boldsymbol{R}$. Mehra demonstrated that both covariance matrices can be estimated using covariance-matching techniques [118, 119]. The idea is to make the actual residuals consistent with the theoretical covariances. The covariance matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$ are adapted as measurements are provided. An innovation-based adaptive technique uses

$$\hat{\boldsymbol{R}}_k = \hat{\boldsymbol{C}}_{v_k} - \boldsymbol{H}_k \boldsymbol{P}_{k|k-1} \boldsymbol{H}_k^T \tag{5.46}$$

and

$$\hat{\boldsymbol{Q}}_k = \boldsymbol{K}_k \hat{\boldsymbol{C}}_{v_k} \boldsymbol{K}_k^T \tag{5.47}$$

with

$$\hat{\boldsymbol{C}}_{v_k} = \frac{1}{N} \sum_{j=j_0}^{k} \boldsymbol{v_j} \boldsymbol{v_j}^T \tag{5.48}$$

where $\hat{\boldsymbol{C}}_{v_k}$ is the innovation measurement noise covariance matrix at epoch k, which uses a moving window of length N starting at $j_0 = k - N + 1$ [120]. The innovation sequence at epoch k is

$$\boldsymbol{v_k} = \boldsymbol{z_k} - \hat{\boldsymbol{z}}_{k|k-1} \tag{5.49}$$

where the predicted measurement $\hat{\boldsymbol{z}}_{k|k-1}$ is given by (5.10) for an EKF or is equal to $\boldsymbol{H}_k \hat{\boldsymbol{x}}_{k|k-1}$ for a normal Kalman filter. A similar set of equations can be created for a residual-based adaptive technique, which uses the residual sequence instead of the innovation sequence. For this approach, $\boldsymbol{P}_k$, $\hat{\boldsymbol{z}}_k$, and $\hat{\boldsymbol{x}}_k$ are used instead of $\boldsymbol{P}_{k|k-1}$, $\hat{\boldsymbol{z}}_{k|k-1}$, and $\hat{\boldsymbol{x}}_{k|k-1}$

Notice that the adaptive approach depends on a chosen window length, as shown in (5.48). The optimal window length is not obvious, and may vary in time. For this reason, neural network aided adaptive EKF approaches have been introduced [121, 122, 112]. Here a neural network, described further in the next section, is trained offline using $\hat{\boldsymbol{C}}_{v_k}$ as input data to predict $\boldsymbol{Q}$ and/or $\boldsymbol{R}$. Variations of this idea use additional measurements as training data [123, 124]. It may also be possible to use marginalized particle filter (MPF) if differentiable resampling is used [125].

5.9.2 Recurrent Neural Networks

A recurrent neural network (RNN) is an extension of a neural network that has internal feedback loops within the layers, as shown in Figure 5-8. For clarity, standard neural networks are often referred to as feedforward neural networks, since they do not have feedback loops. Recurrent neural networks, as well as feedforward neural networks, were previously used in the field of navigation by Ellis for pedestrian navigation [126].

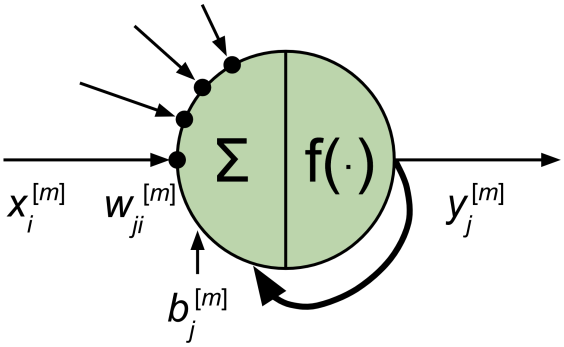


Figure 5-8: Individual node within a recurrent neural network.

The recurrent neural network can be "unrolled," as shown in Figure 5-9 to more clearly illustrate what is happening. The hidden state from one time step is fed back into the same, or potentially different, hidden layer at a later time step.

Long short-term memory (LSTM), shown in Figure 5-10, is a more complex type of recurrent neural network that uses gating for an increased ability to store information over long time intervals [127]. A detailed explanation can be found in [128].

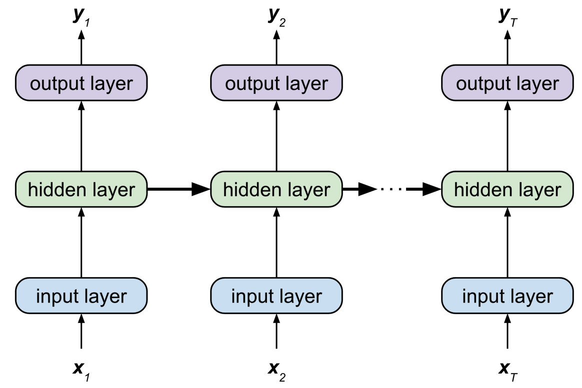


Figure 5-9: Recurrent neural network.

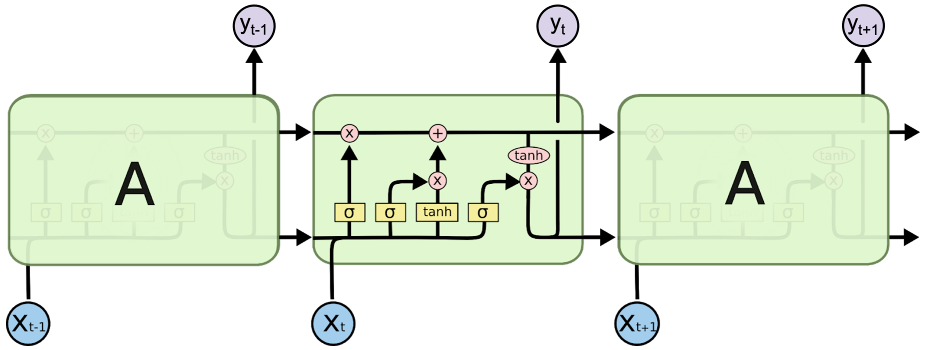


Figure 5-10: Long short-term memory. Modified from [128].

5.9.3 Covariance-Adaptive Neural Filters

As discussed in section 5.9.1, another approach for filtering with uncertain noise covariances is using an adaptive filter. Though the process and measurement noise covariance matrices can generally be well estimated for magnetic navigation using tail stinger measurements, when highly corrupted data is used the covariances are less static. For example, when the aircraft control surfaces move during a maneuver, the measurement noise covariance $\boldsymbol{R}$ can increase substantially compared to straight and level flight. Preliminary results using a measurement noise covariance-matching technique have shown marginally increased navigation performance depending on the window length and measurement noise covariance bounds. Both of these were hand-tuned, which is clearly not ideal for a real-time application.

For magnetic navigation, a machine learning aided adaptive filter may have some benefit. For this approach, a neural network is trained offline using auxiliary data, as shown in Figure 5-11. More specifically, an LSTM-type recurrent neural network, as described in section 5.9.2, with an EKF is used to learn how the measurement noise covariance should vary with the measurements for better navigation performance. The loss function in this case directly uses position (latitude and longitude) rather than a magnetic field value.

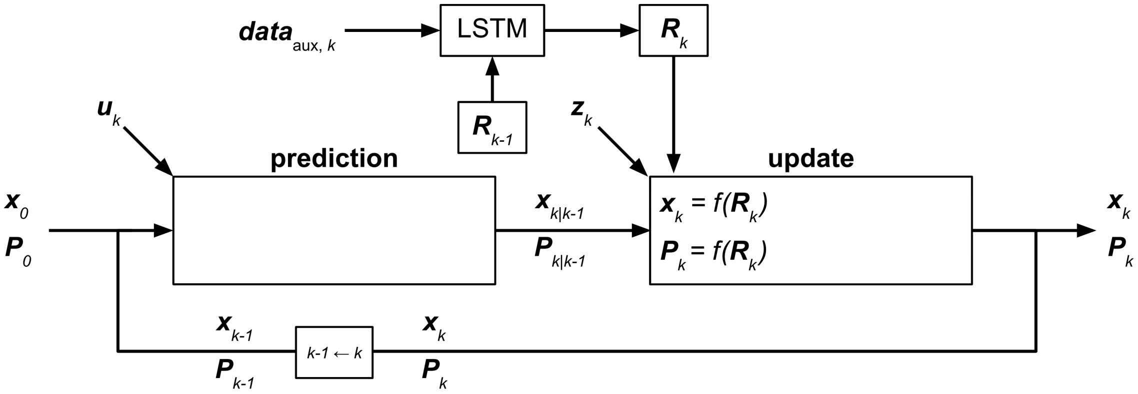


Figure 5-11: Measurement noise covariance-adaptive neural EKF.

As far as implementation, this model would be difficult to implement in any programming language other than Julia. The use of the Flux machine learning library and reverse-mode algorithmic differentiation (AD) with Zygote allows for straightforward backpropagation and differentiable programming [129, 76]. It was not necessary, but custom adjoints can be implemented as well [130].

A similar idea for the process noise is shown in Figure 5-12, where specifically the temporal variation filter state noise covariance and/or time constant are learned. However, this approach is not currently implemented in MagNav.jl or evaluated in this work and is only presented for completion.

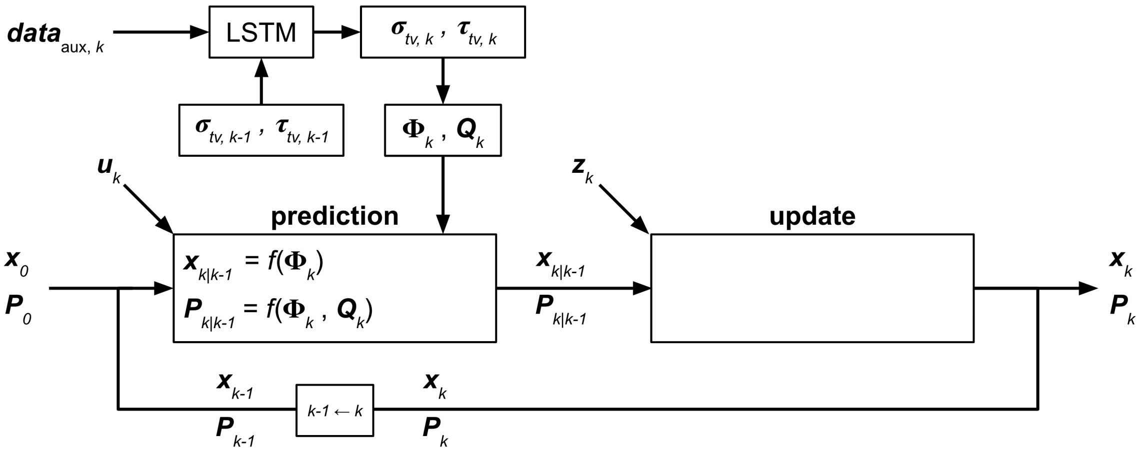


Figure 5-12: Process noise covariance-adaptive neural EKF.

5.9.4 End-to-End Strategy Performance Comparison

To compare these end-to-end approaches, flight line 1007.06 is examined. As mentioned in section 2.2.3, this flight line was both held out and is publicly available. The end-to-end approaches are generally newer and less developed than the pre-processing approaches, and further improvement is likely possible. Specifically, Mags 4 and 5 (independently) and Flux A are used here without any additional non-magnetometer features (as in small feature set 2, described at the end of section 4.9). This allows a fair comparison between the online Tolles-Lawson and online neural network-based aeromagnetic compensation models.

For the (linear) covariance-adaptive EKF, the measurement noise covariance was allowed to vary within a lower and upper bound. This was found to result in better performance than allowing any measurement noise covariance to be set by the algorithm. In particular, the measurement noise covariance was bounded between $30^2\,\mathrm{nT}^2$ and $120^2\,\mathrm{nT}^2$ for Mag 4 and $5^2\,\mathrm{nT}^2$ and $30^2\,\mathrm{nT}^2$ for Mag 5, which was based on their levels of corruption. The scalar magnetometer measurement in this case is Mag 4 or 5 compensated using model 2c.

For the covariance-adaptive neural EKF, due to computational constraints, the LSTM-based model was trained on only flight line 1003.08. This flight line was selected because it contained a significant amount of in-flight "events" (i.e. purposely caused magnetic disturbances), and it was flown in the same region at the same altitude (Renfrew at 400 m HAE) as flight line 1007.06. This means better performance is unlikely to be found using a different flight line for training. Testing performance did not improve if more than 5 epochs were used for the cases that were examined. Various feature sets were looked at, including small feature set 2 and different model 2c output combinations (i.e. neural network and/or Tolles-Lawson portions). Multiple feature sets had similar performance, the simplest being the (map to model 2c prediction) measurement error, which is reported in Table 5.10.

Table 5.10: End-to-end navigation performance comparison on flight line 1007.06. Mag 4, Flux A, and small feature set 2 (only magnetometer-related features) are used here. DRMS navigation errors [m] are shown. "T-L" refers to the Tolles-Lawson model. Model 2c with an EKF is shown for reference. Online Tolles-Lawson does not use the model 2c compensated values, while the others do.

Scalar Mag	Model 2c (EKF)	Cov-Adapt (linear)	Cov-Adapt (neural)	Online Model 2c	Online T-L	Stinger	No Mag
4	51	48	51	37	51	19	300
5	21	22	21	18	21	19	300

As can be seen in Table 5.10 there is little to no benefit to the measurement noise covariance-adaptive EKF models. Both require tuning to have even marginal performance benefits. Specifically, the bounds for the (linear) covariance-adaptive EKF must be set, otherwise the filter often underestimates the measurement noise covariance for this nonlinear problem. For the covariance-adaptive neural EKF, the LSTM sequence length must be set. A sequence length of 500 was used, though 10, 50, and 100 were also examined and had similar or worse results. Note that the lack of performance improvement here assumes that an appropriate measurement noise covariance is set for the baseline, which in this case was model 2c (pre-processing aeromagnetic compensation) with an EKF (navigation algorithm) using a measurement noise covariance of $60^2\,\mathrm{nT}^2$. If an unsuitable measurement noise covariance is set and/or it is completely unknown, both of the covariance-adaptive approaches could provide some benefit. For example, with Mag 4 (and Flux A), if the baseline uses an improper measurement noise covariance of $5^2\,\mathrm{nT}^2$, the navigation error is 124 m. The (linear) covariance-adaptive EKF still achieves 48 m (with the previously discussed bounds) and the covariance-adaptive neural EKF now achieves 86 m. This may have some future use, however in this work the approximate corruption magnitudes are known from the training data, and thus the benefit is minimal. Online Tolles-Lawson performs exactly the same in this case (by coincidence) as the baseline, model 2c with an EKF. However, the online version of model 2c is clearly able to improve navigation, which is further explored in Chapter 6.

Chapter 6

Conclusion

This chapter first compares the navigation performance of the various approaches presented in this work. This is followed by a summary of the thesis contributions and limitations. Finally, suggested future work is described.

6.1 Navigation Performance

This section describes the navigation performance metric then compares the navigation performance of the various approaches on the same flight lines.

6.1.1 Navigation Performance Metric

For navigation, position is the target rather than a compensated magnetic signal. In this work, the distance root mean square (DRMS) metric is used, which is valid for all approaches. The DRMS is calculated as

$$\text{DRMS} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(x_{\text{GPS,k}} - x_{\text{filter,k}})^2 + (y_{\text{GPS,k}} - y_{\text{filter,k}})^2} \tag{6.1}$$

where x and y are Universal Transverse Mercator (UTM) coordinates for epochs k =1:N. UTM x and y coordinates correspond to longitude and latitude in the LLA coordinate system, respectively. Here, GPS position data is considered truth and filter output data is an estimate. Relating performance between two models is simply the difference or fraction of two DRMS values.

6.1.2 Navigation Performance Comparison

This work presented 5 linear and 5 nonlinear pre-processing approaches, and 4 end-to-end approaches for aeromagnetic compensation. The best performing of each group, along with the state-of-the-art model, are compared. The overall system block diagram for the neural network-based approaches is shown in Figure 6-1. Here, the neural network weights (and biases) are determined from during training, and either static or updated online.

Presented in Table 6.1 is compensation performance for the classical Tolles-Lawson, PLSR-based, and neural network-based (model 2c) aeromagnetic compensation models. Each row in Table 6.1 is a flight line with unobserved (held out) data, with the last two rows being flight 1007, which was not used at all during training. These flight lines were selected among the held out testing data as each is over 1 hr in length, which resembles an operational scenario. The first calibration box of flight line 1006.04 was used for classical Tolles-Lawson, while the standard training dataset, described in section 2.2.2, was used for PLSR and model 2c. PLSR and model 2c used the small feature set 1 with only current, voltage, and 2 direction

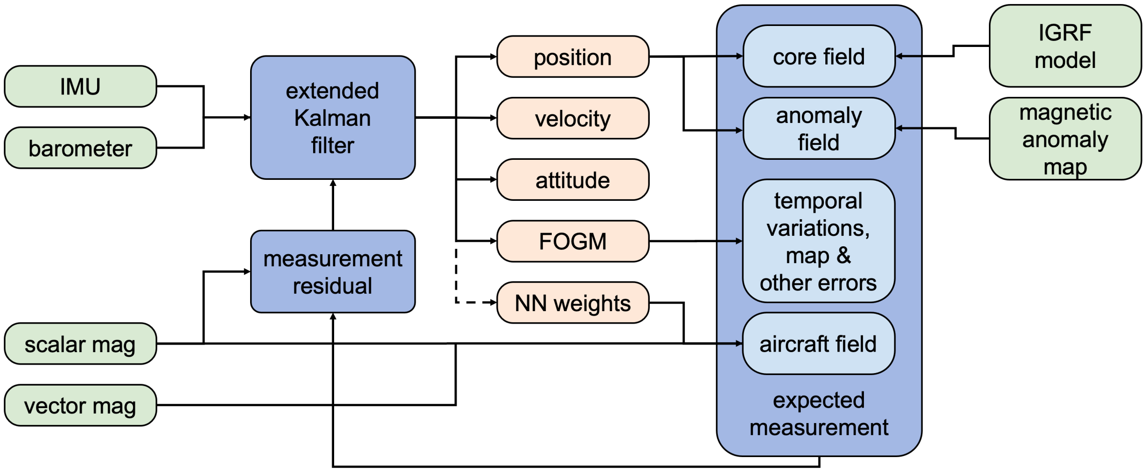


Figure 6-1: MagNav system block diagram using neural network-based aeromagnetic compensation and an extended Kalman filter.

cosine matrix features (in addition to the magnetometer terms), as explained at the end of section 4.9.

The magnetic signal error, which compares the model output to the target value (compensated stinger), is around 100 nT for classical Tolles-Lawson, 60 nT for PLSR, and 40 nT for model 2c. These correspond to signal-to-noise ratios, defined in section 3.8, of approximately 2.2, 3.7, and 5.8, respectively. Due to these high levels of corruption and nonlinearities, the measurement noise covariance was set to $120^2\,\mathrm{nT}^2$, though other values were evaluated. For reference, the stinger measurement noise covariance was $10^2\,\mathrm{nT}^2$. As expected, the corresponding navigation performance is best for model 2c (among the three pre-processing models), as can be seen in Table 6.2. Further navigation performance can be achieved by using online neural network-based aeromagnetic compensation with model 2c. In this case, the neural network weights (and biases) of model 2c are used as initial states in the online EKF, and then they are updated during navigation. Note that the Tolles-Lawson coefficients used in model 2c were first initialized with classical Tolles-Lawson, then updated during the initial model 2c training. During online learning they are now static, similar to model 2b. Online neural network-based aeromagnetic compensation can result in significant navigation accuracy improvements, up to $\sim$ 2x lower navigation error. This performance, which again uses a *noisy, in-cabin* scalar magnetometer, approaches the performance of the Tolles-Lawson-compensated *tail stinger* magnetometer. For reference, the INS-only (no magnetometer) performance is also shown in Table 6.2. This can be determined by simply setting the measurement noise covariance to infinity.

Table 6.1: Compensation performance comparison with Mag 3, Flux A, and small feature set 1. Only current, voltage, and 2 direction cosine matrix features are used here (in addition to the magnetometer terms), as explained at the end of section 4.9. "T-L" refers to the Tolles-Lawson model.

Flight	Flight Line	T-L [nT]	PLSR [nT]	Model 2c [nT]
1003	1003.10	83	60	34
1007	1007.04	113	69	40
1007	1007.06	102	48	44

Table 6.2: Navigation performance comparison with Mag 3, Flux A, and small feature set 1. Only current, voltage, and 2 direction cosine matrix features are used here (in addition to the magnetometer terms), as explained at the end of section 4.9. DRMS navigation errors are shown. "T-L" refers to the Tolles-Lawson model.

Flight	Flight Line	T-L [m]	PLSR [m]	Model 2c [m]	Online 2c [m]	Stinger [m]	No Mag [m]
1003	1003.10	271	296	122	83	32	350
1007	1007.04	88	88	68	25	25	88
1007	1007.06	71	58	48	25	19	300

As a reference point, previous work by the author achieved 3.5 nT and 6.3 nT compensation errors and 42 m and 23 m navigation errors on flight lines 1003.10 and 1007.04, respectively [131]. Clearly, the compensation performance here is far worse, but the work in [131] used all available in-cabin magnetometers, i.e. Mags 2, 3, 4, and 5 and Flux A, C, and D. This would be a challenge to implement in an operational aircraft. Despite worse compensation, navigation accuracies are still less than 100 m when using the online approach, which would provide some benefit in an operational scenario.

Table 6.3: Compensation performance comparison with Mag 3, Flux A, and small feature set 2. Only magnetometer-related features are used here. "T-L" refers to the Tolles-Lawson model.

Flight	Flight Line	T-L [nT]	PLSR [nT]	Model 2c [nT]
1003	1003.10	83	70	55
1007	1007.04	113	72	58
1007	1007.06	102	52	55

Table 6.4: Navigation performance comparison with Mag 3, Flux A, and small feature set 2. Only magnetometer-related features are used here. DRMS navigation errors are shown. "T-L" refers to the Tolles-Lawson model.

Flight	Flight Line	T-L [m]	PLSR [m]	Model 2c [m]	Online 2c [m]	Stinger [m]	No Mag [m]
1003	1003.10	247	187	120	122	32	350
1007	1007.04	88	88	70	34	25	88
1007	1007.06	71	57	53	32	19	300

Presented in Table 6.3 is compensation performance with small feature set 2. This feature set only contains magnetometer terms, as explained at the end of section 4.9. Again the first calibration box of flight line 1006.04 was used for classical Tolles-Lawson, while the standard training dataset, described in section 2.2.2, was used for

PLSR and model 2c. The magnetic signal error, which compares the model output to the target value (compensated stinger), is now around 100 nT for classical Tolles-Lawson, 70 nT for PLSR, and 55 nT for model 2c. These correspond to signal-to-noise ratios, defined in section 3.8, of approximately 2.2, 3.5, and 4.0, respectively. Similar to Table 6.1, the corresponding navigation performance is best for model 2c (among the three pre-processing models), as can be seen in Table 6.4. Further navigation performance can again be achieved by using online neural network-based aeromagnetic compensation with model 2c. However, the overall performance here is worse than with the current, voltage, and 2 direction cosine matrix terms included.

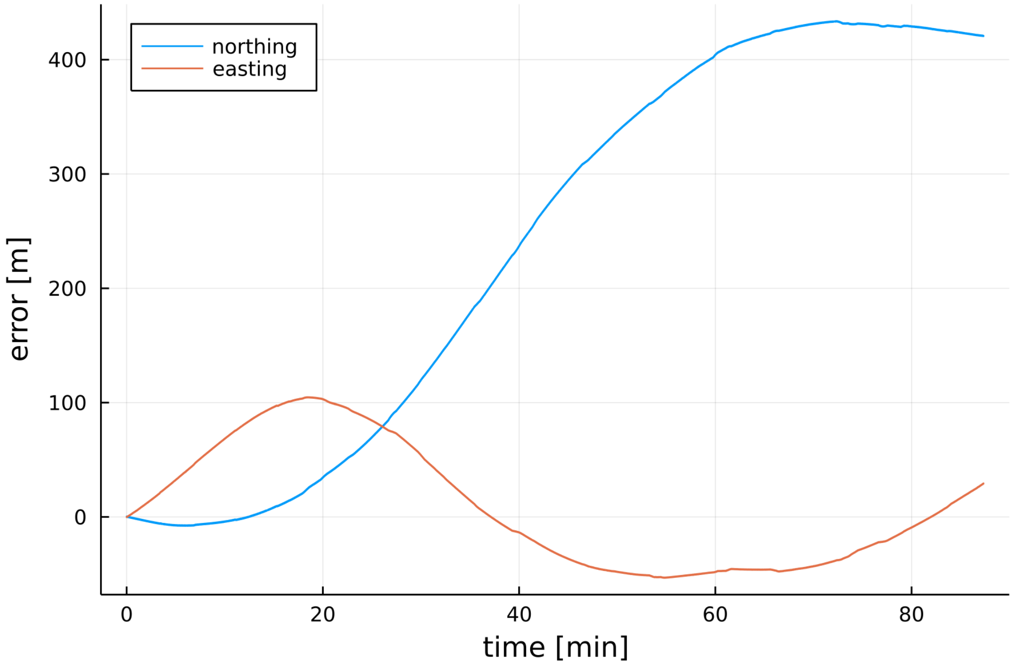


Figure 6-2: Flight line 1007.06 INS drift.

Looking closer at the performance of flight line 1007.06, first notice the INS drift shown in Figure 6-2. This is the navigation performance without any magnetic measurements. Now looking at Figures 6-3 and 6-4, the navigation performance greatly improves when compensated magnetic measurements are used. The online neural network-based aeromagnetic compensation (model 2c) clearly has the best performance, as it avoids the "spikey" behavior seen with the other pre-processing approaches. This would be good for an operational scenario, where the model does not abruptly give extremely poor position estimates. Once again, only a *noisy, in-cabin* scalar magnetometer is used here.

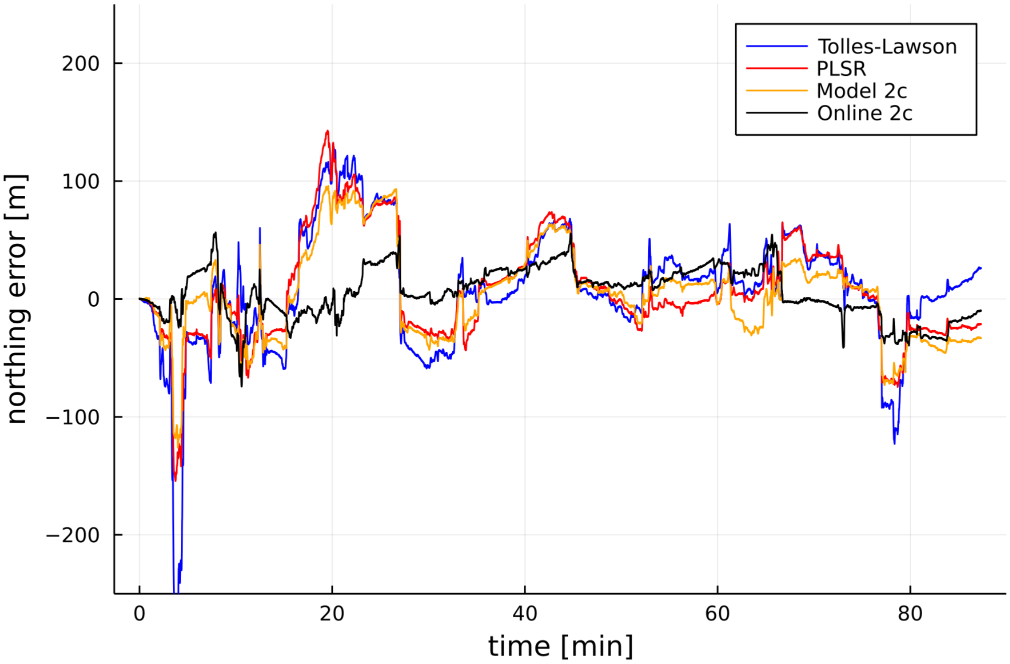


Figure 6-3: Flight line 1007.06 northing error comparison with Mag 3.

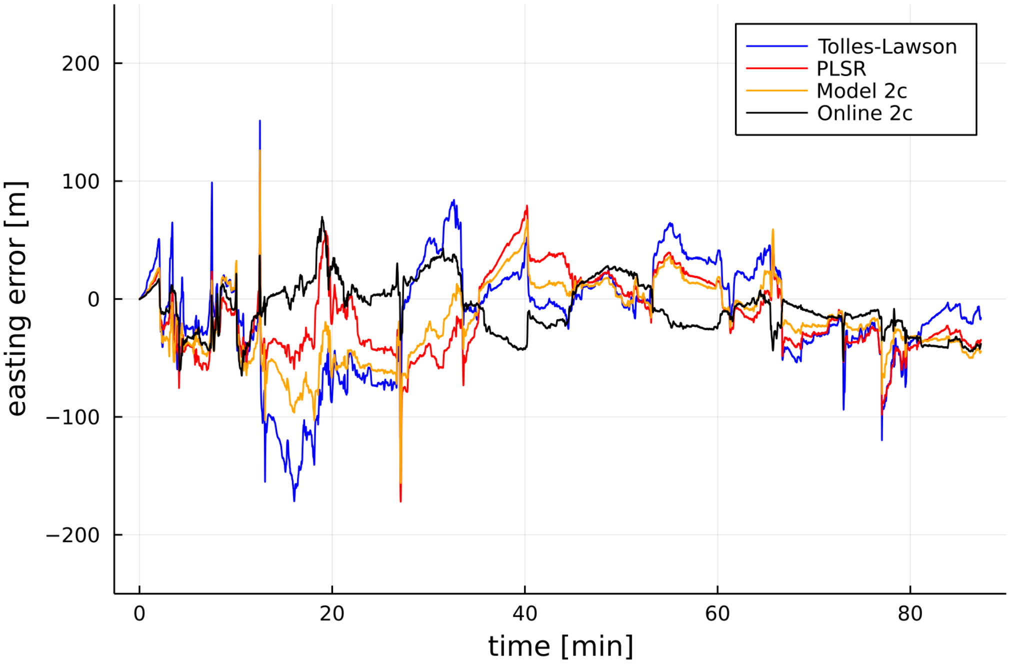


Figure 6-4: Flight line 1007.06 easting error comparison with Mag 3.

However, it is worth noting that the benefit of online neural network-based aeromagnetic compensation (model 2c) diminishes if a scalar magnetometer with less corruption is used. Table 6.5 shows the compensation performance as before, but with Mag 5 instead of Mag 3. As can be seen in Table 2.3 in section 2.2.1, Mag 5 has nearly an order of magnitude less corruption. In this case, the classical Tolles-Lawson model performs fairly well, and PLSR or model 2c perform comparably. As shown in Table 6.6, navigation performance is fairly consistent when using classical Tolles-Lawson, PLSR, or model 2c for typical pre-processing. The online version of model 2c does still perform better than any of these, but not significantly. As shown in Figures 6-5 and 6-6, the navigation errors for each of the different models are very similar and often overlap.

Table 6.5: Compensation performance comparison with Mag 5, Flux A, and small feature set 2. Only magnetometer-related features are used here. “T-L” refers to the Tolles-Lawson model.

Flight	Flight Line	T-L [nT]	PLSR [nT]	Model 2c [nT]
1003	1003.10	13	12	15
1007	1007.04	26	12	9.5
1007	1007.06	19	8.9	9.4

Table 6.6: Navigation performance comparison with Mag 5, Flux A, and small feature set 2. Only magnetometer-related features are used here. DRMS navigation errors are shown. “T-L” refers to the Tolles-Lawson model.

Flight	Flight Line	T-L [m]	PLSR [m]	Model 2c [m]	Online 2c [m]	Stinger [m]	No Mag [m]
1003	1003.10	58	70	58	57	32	350
1007	1007.04	33	23	23	21	25	88
1007	1007.06	20	21	21	18	19	300

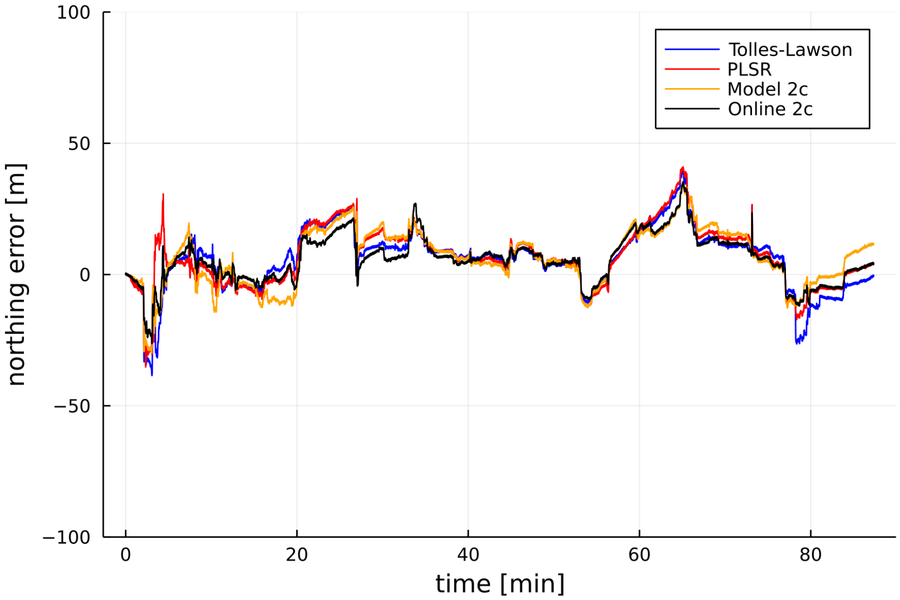


Figure 6-5: Flight line 1007.06 northing error comparison with Mag 5.

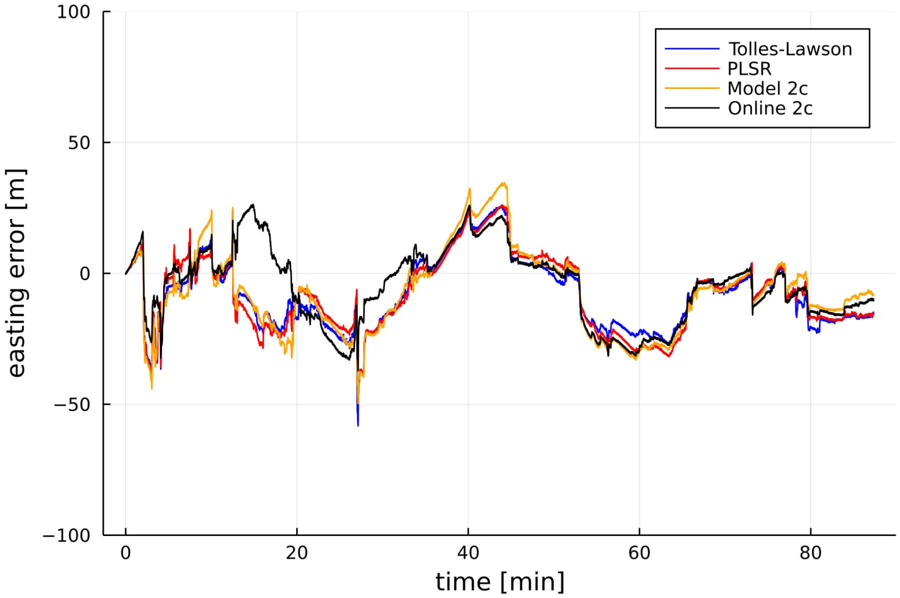


Figure 6-6: Flight line 1007.06 easting error comparison with Mag 5.

6.2 Summary of Findings and Limitations

Airborne magnetic anomaly navigation has the potential to serve as a valuable backup navigation system to GPS. However, dealing with the corruption in the magnetic measurements is a long-standing and difficult problem. This work explored many approaches to aeromagnetic compensation with magnetic field measurements that contained various levels of corruption in an effort to resolve this problem. These models have been shown to improve on the state-of-the-art aeromagnetic compensation model, which resulted in improved navigation performance. More specifically, the findings and limitations of this work by chapter are as follows.

In Chapter 2, the open-source MagNav.jl Julia package was described. This software package provides a full suite of open-source aeromagnetic compensation and airborne magnetic anomaly navigation tools. No publicly available software package of this type current exists at the time of writing. The package was designed to handle new flight datasets and new magnetic anomaly maps. MagNav.jl may be valuable to both other MagNav researchers and for implementing compensation models and navigation algorithms in future real-time demonstrations. However, the package has only been used by a small team of people thus far. It may suffer from a usability standpoint, though each function is documented and some examples are provided. All of the models described next are available within MagNav.jl.

In Chapter 3, the classical Tolles-Lawson aeromagnetic compensation model was described in detail. Two variations of the Tolles-Lawson model were then detailed. Finally, two additional linear aeromagnetic compensation models were introduced, which were shown to outperform the classical model when a large flight dataset is available. An unresolved research question would be what is a “large” flight dataset, though this was somewhat evaluated in section 5.7. In this work, 11 hr and 52 min of flight data was typically used to fit or train models, however, this is likely not necessary to improve on the classical model.

In Chapter 4, five neural network-based approaches were introduced. Extensive neural network design and feature selection evaluations were performed to determine what form the models should take. It was determined that no more than 16 nodes in a single hidden layer are necessary, and as few as 2 nodes are adequate when only magnetometer terms are used. Some features, other than magnetometers, did improve performance. In particular, including some current sensors and terms from the direction cosine matrix in the feature set consistently reduced the compensation error. Model 2c was marginally the best performing neural network-based aeromagnetic compensation. This model contains both a linear and nonlinear portion, allowing for greater reliability and interpretability. The performance of small feature sets was evaluated, which showed consistent improvement over linear aeromagnetic compensation models. Though model 2c was deemed the best model, performance was often close between the various neural network-based approaches. Deciding which model to incorporate in an operational scenario could depend on additional factors, such as implementation complexity and runtime, which were not explored in this work.

In Chapter 5, multiple end-to-end approaches were explored. These models are generally less developed that the previously discussed models, but navigation performance could still be improved compared to the baseline EKF. The measurement noise covariance-adaptive filtering approaches were shown to have marginal or no benefit in typical use cases. The online Tolles-Lawson model was able to decrease the navigation error on nearly all of the maneuver sets that were examined. Finally, the online neural network-based model was able to further reduce the navigation error. However, though navigation performance was shown to improve on multiple flight lines, it has not been proven to work on additional aircraft types.

With these findings and limitations explained, it may be beneficial to describe the overall recommended procedure for using these results with different flight data or in an operational setting. First, some sort of initial training data should be collected, which should include at least one calibration flight segment. If tail stinger measurements are not possible, the flight should occur over an area with an available magnetic anomaly map. During the flight data collection, it is desirable to replicate the aircraft magnetic field for future testing (or missions) as close as possible. As shown in section 5.7, even an aircraft of the same type can have a dramatically different magnetic field. A new configuration results in worse performance when using a previously developed model.

Next, evaluate the magnetometer measurements, both with and without classical Tolles-Lawson aeromagnetic compensation. This gives a sense of how corrupted the magnetic signal is and how well the state-of-the-art model performs on it, i.e. this is a measure of the nonlinearities present in the data. Using the determined Tolles-Lawson coefficients, train and then evaluate model 2c on split portions of the flight data. Additional sensor data may be included here, such as current measurements. The value of additional data can be determined using the feature selection approaches described in section 4.8 and/or by looking at the data itself, as in Table 4.8.

Finally, evaluate the navigation performance using an EKF with both the Tolles-Lawson and model 2c compensated data, as well as the online versions of both models. From a risk and to a lesser degree, computational, standpoint it may be better to implement the simplest model that still performs adequately. However, in general the online Tolles-Lawson model may offer navigation improvement on less corrupted data, while the online neural network-based model offer navigation improvement on more corrupted data.

6.3 Recommendations for Future Work

Airborne magnetic anomaly navigation has many complexities, and it has only been seriously investigated in the past decade or so. Much more research needs to be done in order to use MagNav in an operational situation. First and foremost, MagNav is completely reliant on having a high-quality magnetic anomaly map. The navigation results presented in this thesis was only possible due to having the maps near Ottawa, Ontario, Canada. However, given the small size of these maps and the limits imposed by upward continuation, flights over approximately 2000 m were unable to be evaluated in respect to navigation performance. Additionally, in many areas of the world no high-resolution magnetic anomaly maps exist at all. Further flight surveys and methods to collect vast amounts of magnetic anomaly data are recommended. This may include using commercial or military flights. These flights would not collect data in the conventional manner with a tail stinger and gridded flight path, but instead collect unsystematic magnetic anomaly data. There may be a way to use this data to better downward continue wide-coverage, large-cell, high-altitude maps, such as NAMAD [9], to typical flight altitudes.

Next, this research used flight data from a Cessna 208B Grand Caravan. Magnetometers were placed in the cabin so that the magnetic signal was corrupted to various degrees based on the placement location. However, in an operational scenario, the magnetometers may be placed on the wing. It is recommended to research how the magnetic corruption differs between these placement locations, as well as how well the models presented here work on other aircraft. Furthermore, the flight data used here was collected at fairly low speeds and altitudes, and it would be valuable to evaluate performance at higher speeds and altitudes. The relationship between post-compensation magnetometer errors and map values, and the effect on navigation performance, should be investigated. This would be suitable for simulation using MagNav.jl.

Online aeromagnetic compensation shows promise based on the results presented here and by Canciani [20]. To tune the compensation states in the navigation filter, specific aircraft maneuvers may be beneficial, as explored in section 5.8. However, the available flight data had limitations, and the recommended modifications should be used during a future flight data collection.

For the neural network-based approaches, there are several potential improvements that should be explored. As shown in Figure 5-6, there is potential to use transfer learning for improved performance. For this purpose, a two or three layer neural network may be more appropriate, allowing for some layers to remain static during retraining. For less than 100 features, layers with 2, 4, 8, or 16 nodes are recommended to be evaluated. Additionally, though Swish was found to work best for the pre-processing approach, for the online (end-to-end) approach, a different activation function may perform better and this should be investigated. Model 2d was implemented much later than the other neural network-based models, and thus should be experimented with further, especially examining the additive corrections applied to the various Tolles-Lawson coefficients. Finally, the relative scaling between the neural network and Tolles-Lawson portions of models 2b and 2c, both within the loss function and the resulting compensation, should be further investigated to understand when the neural network is most “active” and the linear model is insufficient.

The models presented here vary from only using a single scalar and vector magnetometer pair up to using many magnetometers and auxiliary data. Reliance on more sensors introduces more risk, which is why this work investigated small feature sets. Due to different maneuvers performed by an aircraft, signal dropouts can occur. More research should be conducted to evaluate performance and to find mitigation solutions to this potential issue. This may include having redundant magnetometers, so multiple models could be used, e.g. one with both magnetometers, one with only the first magnetometer, and one with only the second magnetometer. Additionally, the models available in MagNav.jl are not immediately capable of being used in real-time. Further work must be conducted to allow for a real-time implementation and evaluation of MagNav.

Appendix A

SGL Flight Data Fields

In addition to scalar and vector magnetometer measurements, various auxiliary sensor data from the SGL flights is included in the datasets. Below is a description of each available data field. (WGS-84) indicates elevation above the WGS-84 ellipsoid.

Table A.1: SGL 2020 flight data fields.

Field	**Units**	**Description**
tie_line	-	line number
flight	-	flight number
year	-	year
doy	-	day of year
tt	s	fiducial seconds past midnight UTC
utmX	m	x-coordinate, WGS-84 UTM zone 18N
utmY	m	y-coordinate, WGS-84 UTM zone 18N
utmZ	m	z-coordinate, GPS altitude (WGS-84)
msl	m	z-coordinate, GPS altitude above EGM2008 Geoid
lat	deg	latitude, WGS-84
lon	deg	longitude, WGS-84
baro	m	barometric altimeter
radar	m	filtered radar altimeter
topo	m	radar topography (WGS-84)
dem	m	digital elevation model from SRTM (WGS-84)
drape	m	planned survey drape (WGS-84)
ins_pitch	deg	INS-computed aircraft pitch
ins_roll	deg	INS-computed aircraft roll
ins_yaw	deg	INS-computed aircraft yaw

diurnal	nT	measured diurnal
mag_1_c	nT	Mag 1: compensated magnetic field
mag_1_lag	nT	Mag 1: lag-corrected magnetic field
mag_1_dc	nT	Mag 1: diurnal-corrected magnetic field
mag_1_igrf	nT	Mag 1: IGRF & diurnal-corrected magnetic field
mag_1_uc	nT	Mag 1: uncompensated magnetic field
mag_2_uc	nT	Mag 2: uncompensated magnetic field
mag_3_uc	nT	Mag 3: uncompensated magnetic field
mag_4_uc	nT	Mag 4: uncompensated magnetic field
mag_5_uc	nT	Mag 5: uncompensated magnetic field
mag_6_uc	nT	Mag 6: uncompensated magnetic field
flux_a_x	nT	Flux A: fluxgate x-axis
flux_a_y	nT	Flux A: fluxgate y-axis
flux_a_z	nT	Flux A: fluxgate z-axis
flux_a_t	nT	Flux A: fluxgate total
flux_b_x	nT	Flux B: fluxgate x-axis
flux_b_y	nT	Flux B: fluxgate y-axis
flux_b_z	nT	Flux B: fluxgate z-axis
flux_b_t	nT	Flux B: fluxgate total
flux_c_x	nT	Flux C: fluxgate x-axis
flux_c_y	nT	Flux C: fluxgate y-axis
flux_c_z	nT	Flux C: fluxgate z-axis
flux_c_t	nT	Flux C: fluxgate total
flux_d_x	nT	Flux D: fluxgate x-axis
flux_d_y	nT	Flux D: fluxgate y-axis
flux_d_z	nT	Flux D: fluxgate z-axis
flux_d_t	nT	Flux D: fluxgate total
ogs_mag	nT	OGS survey diurnal-corrected, levelled, magnetic field
ogs_alt	m	OGS survey, GPS altitude (WGS-84)
ins_acc_x	m/s^2	INS x-acceleration
ins_acc_y	m/s^2	INS y-acceleration
ins_acc_z	m/s^2	INS z-acceleration
ins_wander	rad	INS-computed wander angle (ccw from north)
ins_lat	rad	INS-computed latitude
ins_lon	rad	INS-computed longitude
ins_alt	m	INS-computed altitude (WGS-84)
ins_vn	m/s	INS-computed north velocity
ins_vw	m/s	INS-computed west velocity

ins_vu	m/s	INS-computed vertical (up) velocity
pitch_rt	deg/s	avionics-computed pitch rate
roll_rt	deg/s	avionics-computed roll rate
yaw_rt	deg/s	avionics-computed yaw rate
lon_acc	g	avionics-computed longitudinal (forward) acceleration
lat_acc	g	avionics-computed lateral (starboard) acceleration
alt_acc	g	avionics-computed normal (vertical) acceleration
true_as	m/s	avionics-computed true airspeed
pitot_p	kPa	avionics-computed pitot pressure
static_p	kPa	avionics-computed static pressure
total_p	kPa	avionics-computed total pressure
cur_com_1	A	current sensor: aircraft radio 1
cur_ac_hi	A	current sensor: air conditioner fan high
cur_ac_lo	A	current sensor: air conditioner fan low
cur_tank	A	current sensor: cabin fuel pump
cur_flap	A	current sensor: flap motor
cur_strb	A	current sensor: strobe lights
cur_srvo_o	A	current sensor: INS outer servo
cur_srvo_m	A	current sensor: INS middle servo
cur_srvo_i	A	current sensor: INS inner servo
cur_heat	A	current sensor: INS heater
cur_acpwr	A	current sensor: aircraft po wer
cur_outpwr	A	current sensor: system output power
cur_bat_1	A	current sensor: battery 1
cur_bat_2	A	current sensor: battery 2
vol_acpwr	V	voltage sensor: aircraft power
vol_outpwr	V	voltage sensor: system output power
vol_bat_1	V	voltage sensor: battery 1
vol_bat_2	V	voltage sensor: battery 2
vol_res_p	V	voltage sensor: resolver board (+)
vol_res_n	V	voltage sensor: resolver board (-)
vol_back_p	V	voltage sensor: backplane (+)
vol_back_n	V	voltage sensor: backplane (-)
vol_gyro_1	V	voltage sensor: gyroscope 1
vol_gyro_2	V	voltage sensor: gyroscope 2
vol_acc_p	V	voltage sensor: INS accelerometers (+)
vol_acc_n	V	voltage sensor: INS accelerometers (-)
vol_block	V	voltage sensor: block

vol_back	V	voltage sensor: backplane
vol_servo	V	voltage sensor: servos
vol_cabt	V	voltage sensor: cabinet
vol_fan	V	voltage sensor: air conditioner fan

Table A.2: SGL 2021 flight data fields.

Field	**Units**	**Description**
tie_line	-	line number
flight	-	flight number
year	-	year
doy	-	day of year
tt	s	fiducial seconds past midnight UTC
utmX	m	x-coordinate, WGS-84 UTM zone 18N
utmY	m	y-coordinate, WGS-84 UTM zone 18N
utmZ	m	z-coordinate, GPS altitude (WGS-84)
msl	m	z-coordinate, GPS altitude above EGM2008 Geoid
lat	deg	latitude, WGS-84
lon	deg	longitude, WGS-84
baro	m	barometric altimeter
ins_pitch	deg	INS-computed aircraft pitch
ins_roll	deg	INS-computed aircraft roll
ins_yaw	deg	INS-computed aircraft yaw
diurnal	nT	measured diurnal
mag_1_c	nT	Mag 1: compensated magnetic field
mag_1_lag	nT	Mag 1: lag-corrected magnetic field
mag_1_dc	nT	Mag 1: diurnal-corrected magnetic field
mag_1_igrf	nT	Mag 1: IGRF & diurnal-corrected magnetic field
mag_1_uc	nT	Mag 1: uncompensated magnetic field
mag_2_uc	nT	Mag 2: uncompensated magnetic field
mag_3_uc	nT	Mag 3: uncompensated magnetic field
mag_4_uc	nT	Mag 4: uncompensated magnetic field
mag_5_uc	nT	Mag 5: uncompensated magnetic field
flux_a_x	nT	Flux A: fluxgate x-axis
flux_a_y	nT	Flux A: fluxgate y-axis
flux_a_z	nT	Flux A: fluxgate z-axis
flux_a_t	nT	Flux A: fluxgate total
flux_b_x	nT	Flux B: fluxgate x-axis

flux_b_y	nT	Flux B: fluxgate y-axis
flux_b_z	nT	Flux B: fluxgate z-axis
flux_b_t	nT	Flux B: fluxgate total
flux_c_x	nT	Flux C: fluxgate x-axis
flux_c_y	nT	Flux C: fluxgate y-axis
flux_c_z	nT	Flux C: fluxgate z-axis
flux_c_t	nT	Flux C: fluxgate total
flux_d_x	nT	Flux D: fluxgate x-axis
flux_d_y	nT	Flux D: fluxgate y-axis
flux_d_z	nT	Flux D: fluxgate z-axis
flux_d_t	nT	Flux D: fluxgate total
ins_acc_x	m/s^2	INS x-acceleration
ins_acc_y	m/s^2	INS y-acceleration
ins_acc_z	m/s^2	INS z-acceleration
ins_wander	rad	INS-computed wander angle (ccw from north)
ins_lat	rad	INS-computed latitude
ins_lon	rad	INS-computed longitude
ins_alt	m	INS-computed altitude (WGS-84)
ins_vn	m/s	INS-computed north velocity
ins_vw	m/s	INS-computed west velocity
ins_vu	m/s	INS-computed vertical (up) velocity
cur_com_1	A	current sensor: aircraft radio 1
cur_ac_hi	A	current sensor: air conditioner fan high
cur_ac_lo	A	current sensor: air conditioner fan low
cur_tank	A	current sensor: cabin fuel pump
cur_flap	A	current sensor: flap motor
cur_strb	A	current sensor: strobe lights
vol_block	V	voltage sensor: block
vol_back	V	voltage sensor: backplane
vol_cabt	V	voltage sensor: cabinet
vol_fan	V	voltage sensor: air conditioner fan

Appendix B

Full Feature Selection Results

Table B.1: Feature selection for neural network-based aeromagnetic compensation, model 1. "Drop" is drop-column feature importance, "Perm" is permutation feature importance, "SGL" is sparse group Lasso, and "Shap" is for Shapley values, as described in sections 4.8.1–4.8.4. "Train" and "Test" are the standard training and testing datasets described in sections 2.2.2 and 2.2.3, respectively. Sorted by the average of the four right columns.

Feature	Drop Train	Drop Test	Perm Test	SGL Train	Shap Train	Shap Test
TL_A_flux_d_Z	4	79	2	1	3	1
TL_A_flux_c_Z	94	26	1	2	1	2
TL_A_flux_d_XZ	45	22	3	3	2	3
mag_5_uc	36	60	4	5	5	5
TL_A_flux_c_XZ	33	11	5	6	4	4
TL_A_flux_a_YY	40	97	9	4	16	6
TL_A_flux_d_ZZ	14	52	7	7	7	8
TL_A_flux_a_XX	108	18	6	11	9	7
mag_4_uc	34	15	8	8	6	9
mag_4_5_uc	53	53	10	9	8	11
TL_A_flux_c_ZZ	57	81	12	10	26	10
mag_3_uc	52	89	11	18	10	12
TL_A_flux_a_YZ	82	33	14	15	11	13
TL_A_flux_c_X	10	106	17	14	29	14
TL_A_flux_d_XX	29	10	15	13	20	17
dcm_1	63	27	16	22	13	16

TL_A_flux_a_XY	35	78	23	20	23	20
mag_3_5_uc	9	94	21	24	18	19
TL_A_flux_d_X	38	107	24	16	34	24
TL_A_flux_c_XY	44	48	22	23	19	22
TL_A_flux_c_XX	77	8	29	12	32	26
TL_A_flux_c_YZ	69	32	13	44	12	15
TL_A_flux_d_YZ	72	40	19	32	17	21
TL_A_flux_d_YY	47	9	20	29	14	25
TL_A_flux_d_XY	32	59	31	19	22	27
mag_3_4_uc	91	17	25	34	21	23
dcm_4	70	31	33	30	25	28
TL_A_flux_a_ZZ	19	95	36	42	24	33
TL_A_flux_a_X	21	104	47	21	55	43
dcm_2	103	63	42	40	49	32
TL_A_flux_d_Y	49	96	26	60	30	29
dcm_8	96	36	44	41	39	40
TL_A_flux_a_XZ	42	29	27	72	27	30
dcm_3	15	80	45	50	59	36
TL_A_flux_a_YX_dot	41	47	37	71	36	31
cur_heat	43	25	61	31	45	48
TL_A_flux_c_Y	105	37	18	107	15	18
mag_2_uc	18	83	48	54	37	44
vol_acpwr	90	54	52	39	54	55
dcm_9	39	69	55	53	87	41
dcm_5	75	23	65	25	47	60
cur_outpwr	50	72	69	28	53	54
mag_2_3_uc	92	87	59	27	46	67
TL_A_flux_c_XZ_dot	25	82	39	75	52	42
mag_2_4_uc	55	14	67	17	38	72
vol_outpwr	93	91	53	38	48	66
dcm_7	2	76	63	37	43	57
TL_A_flux_a_XX_dot	28	103	30	100	40	34
TL_A_flux_c_ZZ_dot	24	71	43	70	50	53
TL_A_flux_a_YY_dot	98	28	28	101	42	38
mag_2_5_uc	68	90	49	73	31	46
TL_A_flux_c_YY	20	98	35	99	28	37
TL_A_flux_a_Y	48	88	32	103	33	39
TL_A_flux_a_Z	5	92	34	106	35	35

TL_A_flux_d_ZZ_dot	65	67	41	84	63	50
TL_A_flux_c_XX_dot	87	108	40	79	41	58
TL_A_flux_d_XZ_dot	8	50	54	66	51	59
TL_A_flux_c_ZY_dot	13	84	50	86	68	45
vol_srvo	83	73	79	36	57	69
TL_A_flux_a_ZZ_dot	62	30	46	93	70	47
TL_A_flux_d_XX_dot	12	58	38	98	67	52
TL_A_flux_a_YZ_dot	31	70	56	76	74	56
dcm_6	86	41	66	61	62	63
vol_bat_1	64	64	80	35	61	77
TL_A_flux_c_XY_dot	46	85	51	95	65	49
TL_A_flux_d_ZY_dot	95	56	64	81	81	51
baro	11	2	90	26	44	81
cur_bat_2	67	6	83	48	75	71
TL_A_flux_a_XY_dot	76	74	62	83	66	62
TL_A_flux_c_YX_dot	59	5	60	90	60	61
ins_acc_y	22	3	84	58	86	76
cur_bat_1	84	24	81	52	91	86
TL_A_flux_d_YZ_dot	66	19	70	77	83	73
ins_acc_x	37	62	86	55	95	80
cur_flap	81	7	92	33	98	96
vol_bat_2	89	65	85	49	64	91
TL_A_flux_c_ZX_dot	27	102	58	104	69	64
TL_A_flux_d_XY_dot	99	99	77	85	79	65
cur_tank	23	35	87	51	89	89
vol_acc_n	100	20	93	43	92	92
TL_A_flux_d_YY_dot	54	86	57	97	56	75
vol_cabt	17	75	82	64	85	83
vol_back_p	101	46	91	45	84	93
TL_A_flux_a_XZ_dot	97	43	75	82	82	74
vol_back_n	30	77	95	56	80	85
TL_A_flux_c_YY_dot	74	16	68	92	73	79
vol_block	71	42	89	62	96	88
cur_acpwr	78	21	76	94	58	70
vol_acc_p	16	105	100	46	78	95
TL_A_flux_a_ZX_dot	61	93	73	88	71	82
TL_A_flux_d_ZX_dot	3	100	74	87	94	84
cur_com_1	1	38	88	65	90	94

cur_ac_hi	7	49	94	47	88	106
TL_A_flux_a_ZY_dot	106	44	72	108	77	68
TL_A_flux_c_YZ_dot	104	61	71	102	72	78
TL_A_flux_d_YX_dot	6	101	78	89	76	90
vol_res_p	102	66	98	63	104	100
vol_gyro_2	26	51	107	57	101	97
vol_back	88	39	96	67	97	99
vol_fan	58	13	99	69	102	98
vol_gyro_1	51	1	108	74	93	87
cur_ac_lo	60	57	109	59	99	101
cur_strb	80	4	97	68	108	108
cur_srvo_m	107	45	101	78	107	102
vol_res_n	85	12	106	80	100	103
cur_srvo_i	79	34	104	91	105	104
ins_acc_z	73	55	105	96	103	107
cur_srvo_o	56	68	102	105	106	105

Table B.2: Feature selection for neural network-based aeromagnetic compensation, model 2c. “Drop” is drop-column feature importance, “Perm” is permutation feature importance, “SGL” is sparse group Lasso, and “Shap” is for Shapley values, as described in sections 4.8.1–4.8.4. “Train” and “Test” are the standard training and testing datasets described in sections 2.2.2 and 2.2.3, respectively. Sorted by the average of the four right columns.

Feature	Drop Train	Drop Test	Perm Test	SGL Train	Shap Train	Shap Test
TL_A_flux_c_Z	12	49	1	2	1	1
mag_5_uc	35	53	3	1	6	4
TL_A_flux_c_XZ	1	1	2	5	2	2
TL_A_flux_d_Z	6	6	4	6	3	3
TL_A_flux_d_XZ	101	89	6	7	5	6
mag_4_uc	16	78	5	8	4	7
TL_A_flux_d_ZZ	2	11	7	10	8	5
mag_4_5_uc	45	62	8	9	7	10
TL_A_flux_c_ZZ	17	33	9	11	12	8
TL_A_flux_a_YY	34	85	16	3	30	11
TL_A_flux_a_YZ	68	90	11	15	9	9
TL_A_flux_c_XX	102	24	13	4	11	19

TL_A_flux_c_X	14	16	12	14	19	12
TL_A_flux_d_YZ	72	67	10	21	10	13
TL_A_flux_a_XY	3	2	15	17	14	14
TL_A_flux_a_XX	60	48	17	18	20	17
TL_A_flux_a_XZ	57	87	14	28	13	15
dcm_1	49	84	19	22	15	16
TL_A_flux_c_XY	74	3	18	24	17	18
TL_A_flux_c_YZ	36	96	20	20	16	20
TL_A_flux_d_XY	77	99	22	16	21	23
TL_A_flux_d_XX	24	22	26	12	31	27
TL_A_flux_d_Y	27	10	21	27	26	21
TL_A_flux_c_YY	28	45	23	25	18	24
TL_A_flux_a_X	99	70	29	19	25	25
TL_A_flux_d_X	71	71	27	23	32	29
dcm_4	48	43	30	26	24	26
mag_3_uc	15	21	31	36	23	28
TL_A_flux_c_Y	11	95	32	33	22	30
mag_3_5_uc	89	72	37	31	28	34
TL_A_flux_a_ZZ	78	75	36	34	34	43
TL_A_flux_a_Y	20	36	33	50	36	33
dcm_5	86	38	39	49	37	32
TL_A_flux_a_Z	62	55	41	42	33	44
TL_A_flux_d_YY	82	41	46	41	39	41
TL_A_flux_c_ZZ_dot	10	18	25	73	38	31
mag_3_4_uc	97	31	38	55	42	36
dcm_2	105	108	28	89	27	22
TL_A_flux_c_XX_dot	7	44	24	83	45	35
vol_acpwr	13	29	50	43	44	49
vol_outpwr	33	105	51	39	48	54
TL_A_flux_a_YX_dot	42	9	40	69	50	38
dcm_8	69	74	48	48	52	56
dcm_7	9	7	65	37	40	52
TL_A_flux_c_XZ_dot	88	82	43	64	51	48
TL_A_flux_a_YY_dot	25	8	34	84	46	40
mag_2_5_uc	91	83	44	76	29	39
dcm_3	31	37	59	47	69	65
TL_A_flux_d_XZ_dot	61	63	57	65	55	50
TL_A_flux_a_XX_dot	64	101	35	105	41	37

cur_outpwr	43	47	71	40	63	66
TL_A_flux_a_YZ_dot	38	28	54	77	65	53
mag_2_3_uc	23	73	56	74	49	55
TL_A_flux_a_ZX_dot	83	103	45	96	54	46
vol_bat_1	39	61	90	13	82	88
TL_A_flux_c_YZ_dot	47	46	42	108	43	42
vol_bat_2	55	26	82	29	64	81
mag_2_4_uc	96	107	66	66	47	61
mag_2_uc	37	12	49	101	35	45
TL_A_flux_d_XY_dot	67	17	70	78	67	51
baro	5	5	87	30	53	83
TL_A_flux_c_YY_dot	106	35	47	107	58	47
ins_acc_x	32	30	86	56	93	59
TL_A_flux_d_ZZ_dot	90	97	53	88	56	62
cur_heat	70	68	83	35	71	86
dcm_6	30	57	76	59	77	73
vol_srvo	103	102	75	52	62	82
cur_ac_lo	58	92	79	61	76	70
dcm_9	51	104	74	60	83	76
cur_acpwr	54	40	69	86	57	57
TL_A_flux_c_YX_dot	56	42	55	91	59	67
TL_A_flux_a_ZZ_dot	22	86	58	93	68	64
TL_A_flux_d_ZX_dot	29	32	68	79	88	69
vol_back_p	40	50	91	32	85	93
TL_A_flux_d_XX_dot	18	27	52	95	60	72
cur_bat_2	108	59	84	62	94	75
TL_A_flux_c_ZY_dot	8	34	61	104	81	58
cur_bat_1	80	98	81	57	89	85
vol_back_n	59	56	95	38	87	90
cur_ac_hi	81	19	88	44	92	92
TL_A_flux_a_XZ_dot	41	91	62	106	66	60
TL_A_flux_c_ZX_dot	50	51	67	99	72	63
ins_acc_y	87	80	80	58	79	94
cur_tank	107	88	109	46	91	77
vol_gyro_2	44	81	101	53	95	80
TL_A_flux_a_ZY_dot	85	23	60	102	73	74
vol_cabt	65	100	85	67	90	84
TL_A_flux_d_ZY_dot	52	65	72	97	84	68

vol_acc_n	84	25	89	51	80	97
vol_acc_p	73	52	94	45	74	98
TL_A_flux_c_XY_dot	94	79	64	103	61	71
TL_A_flux_d_YY_dot	79	60	63	100	70	78
TL_A_flux_d_YX_dot	26	15	73	94	78	79
TL_A_flux_d_YZ_dot	19	58	77	82	75	87
cur_com_1	4	13	92	63	96	99
cur_flap	93	4	96	54	105	105
vol_res_p	21	20	93	68	97	96
TL_A_flux_a_XY_dot	98	94	78	92	86	89
vol_fan	63	54	97	75	98	91
vol_back	75	39	98	71	99	95
vol_gyro_1	53	76	99	72	100	101
vol_block	92	106	102	80	104	104
cur_strb	46	64	108	70	108	108
cur_srvo_m	100	93	106	81	103	100
cur_srvo_i	66	77	100	85	106	103
cur_srvo_o	76	14	107	87	102	102
ins_acc_z	95	66	104	90	107	107
vol_res_n	104	69	105	98	101	106

Appendix C

Down-Selected Feature Set

Table C.1: Down-Selected Feature Set.

dcm_1	TL_A_flux_a_X
dcm_2	TL_A_flux_a_Y
dcm_3	TL_A_flux_a_Z
dcm_4	TL_A_flux_a_XX
dcm_5	TL_A_flux_a_XY
dcm_6	TL_A_flux_a_XZ
dcm_7	TL_A_flux_a_YY
dcm_8	TL_A_flux_a_YZ
dcm_9	TL_A_flux_a_ZZ
baro	TL_A_flux_a_X_dot
cur_com_1	TL_A_flux_a_Y_dot
cur_ac_hi	TL_A_flux_a_Z_dot
cur_ac_lo	TL_A_flux_c_X
cur_flap	TL_A_flux_c_Y
cur_strb	TL_A_flux_c_Z
cur_heat	TL_A_flux_c_XX
cur_acpwr	TL_A_flux_c_XY
cur_outpwr	TL_A_flux_c_XZ
vol_bat_1	TL_A_flux_c_YY
igrf	TL_A_flux_c_YZ
ins_lat	TL_A_flux_c_ZZ
ins_lon	TL_A_flux_c_X_dot
ins_alt	TL_A_flux_c_Y_dot

mag_2_uc	TL_A_flux_c_Z_dot
mag_3_uc	TL_A_flux_d_X
mag_4_uc	TL_A_flux_d_Y
mag_5_uc	TL_A_flux_d_Z
mag_2_uc_dot	TL_A_flux_d_XX
mag_2_uc_dot4	TL_A_flux_d_XY
mag_3_uc_dot	TL_A_flux_d_XZ
mag_3_uc_dot4	TL_A_flux_d_YY
mag_4_uc_dot	TL_A_flux_d_YZ
mag_4_uc_dot4	TL_A_flux_d_ZZ
mag_5_uc_dot	TL_A_flux_d_X_dot
mag_5_uc_dot4	TL_A_flux_d_Y_dot
mag_2_3_uc	TL_A_flux_d_Z_dot
mag_2_4_uc	
mag_2_5_uc	
mag_3_4_uc	
mag_3_5_uc	
mag_4_5_uc	

References

[1] P. Misra and P. Enge, *Global Positioning System: Signals, Measurements, and Performance.* Ganga-Jamuna Press, 2006.

[2] A. J. Canciani and J. F. Raquet, “Airborne Magnetic Anomaly Navigation,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 1, pp. 67–80, 2017. [Online]. Available: https://doi.org/10.1109/TAES.2017.2649238

[3] United States Air Force, “U.S. Air Force 2030 Science and Technology Strategy,” 2019.

[4] M. Mandea and M. Korte, Eds., *Geomagnetic Observations and Models.* Dordrecht, Netherlands: Springer Netherlands, 2011. [Online]. Available: https://doi.org/10.1007/978-90-481-9858-0

[5] NASA Earth Observatory, “Measuring Earth’s Magnetism,” 2014. [Online]. Available: https://earthobservatory.nasa.gov/images/84266/measuring-earths-magnetism

[6] P. Alken, E. Thébault, C. D. Beggan, H. Amit, J. Aubert, J. Baerenzung, T. N. Bondar, W. J. Brown, S. Califf, A. Chambodut, A. Chulliat, G. A. Cox, C. C. Finlay, A. Fournier, N. Gillet, A. Grayver, M. D. Hammer, M. Holschneider, L. Huder, G. Hulot, T. Jager, C. Kloss, M. Korte, W. Kuang, A. Kuvshinov, B. Langlais, J.-M. Léger, V. Lesur, P. W. Livermore, F. J. Lowes, S. Macmillan, W. Magnes, M. Mandea, S. Marsal, J. Matzka, M. C. Metman, T. Minami, A. Morschhauser, J. E. Mound, M. Nair, S. Nakano, N. Olsen, F. J. Pavón-Carrasco, V. G. Petrov, G. Ropp, M. Rother, T. J. Sabaka, S. Sanchez, D. Saturnino, N. R. Schnepf, X. Shen, C. Stolle, A. Tangborn, L. Tøffner-Clausen, H. Toh, J. M. Torta, J. Varner, F. Vervelidou, P. Vigneron, I. Wardinski, J. Wicht, A. Woods, Y. Yang, Z. Zeren, and B. Zhou, “International Geomagnetic Reference Field: the

thirteenth generation," *Earth, Planets and Space*, vol. 73, no. 1, p. 49, 2021. [Online]. Available: https://doi.org/10.1186/s40623-020-01288-x

[7] T. J. Sabaka, N. Olsen, and R. A. Langel, "A comprehensive model of the quiet-time, near-Earth magnetic field: phase 3," *Geophysical Journal International*, vol. 151, no. 1, pp. 32–68, oct 2002. [Online]. Available: https://doi.org/10.1046/j.1365-246X.2002.01774.x

[8] R. A. Langel and W. J. Hinze, *The Magnetic Field of the Earth's Lithosphere: The Satellite Perspective.* Cambridge University Press, 1998.

[9] U.S. Geological Survey, "Magnetic anomaly maps and data for North America," 2021. [Online]. Available: https://mrdata.usgs.gov/magnetic/

[10] The European Space Agency, "Magnetic field sources," 2012. [Online]. Available: https://www.esa.int/

[11] A. J. Canciani and J. F. Raquet, "Absolute Positioning Using the Earth's Magnetic Anomaly Field," *Navigation*, vol. 63, no. 2, pp. 111–126, 2016. [Online]. Available: https://doi.org/10.1002/navi.138

[12] K. S. Ariffin, "Geophysical Surveying Using Magnetics Methods," Universiti Kebangsaan Malaysia, Tech. Rep., 2004. [Online]. Available: http://www.ukm.my/rahim/G-MagneticsMethods.pdf

[13] A. J. Canciani, "Absolute Positioning Using the Earth's Magnetic Anomaly Field," Doctoral dissertation, Air Force Institute of Technology, 2016. [Online]. Available: https://scholar.afit.edu/etd/251/

[14] F. Goldenberg, "Geomagnetic Navigation beyond the Magnetic Compass," in *IEEE/ION Position, Location, And Navigation Symposium.* Coronado, CA: IEEE, 2006, pp. 684–694. [Online]. Available: https://doi.org/10.1109/PLANS.2006.1650662

[15] Geometrics, "GEOMETRICS' MODEL G-823A CESIUM MAGNETOMETER," pp. 1–2, 2019. [Online]. Available: http://www.rmsinst.com/products/magnetometers/G823A.pdf

[16] G. Noriega, "Aeromagnetic Compensation in Gradiometry—Performance, Model Stability, and Robustness," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 1, pp. 117–121, jan 2015. [Online]. Available: https://doi.org/10.1109/LGRS.2014.2328436

[17] C. Reeves, *Aeromagnetic Surveys: Principles, Practice & Interpretation.* Geosoft Incorporated, 2005.

[18] M. C. Hezel, "Improving Aeromagnetic Calibration Using Artificial Neural Networks," Master's thesis, Air Force Institute of Technology, 2020. [Online]. Available: https://scholar.afit.edu/etd/3589/

[19] K. A. Emery, "Modeling Aircraft Disturbance Fields for Magnetic Navigation Using Dense ANNs and the Novel MANNTL Architecture," Master's thesis, Air Force Institute of Technology, 2021. [Online]. Available: https://scholar.afit.edu/etd/4894/

[20] A. J. Canciani, "Magnetic Navigation on an F-16 Aircraft using Online Calibration," *IEEE Transactions on Aerospace and Electronic Systems*, pp. 1–15, 2021. [Online]. Available: https://doi.org/10.1109/TAES.2021.3101567

[21] USAF-MIT Artificial Intelligence Accelerator, "MagNav.jl: airborne Magnetic anomaly Navigation," 2022. [Online]. Available: https://github.com/MIT-AI-Accelerator/MagNav.jl

[22] MIT Lincoln Laboratory, "Diamond Magnetometer," 2022. [Online]. Available: https://www.ll.mit.edu/r-d/projects/diamond-magnetometer-0

[23] Y. Ma, "SafeTestsets," 2022. [Online]. Available: https://github.com/YingboMa/SafeTestsets.jl

[24] Sander Geophysics Ltd., "High Resolution Magnetic Gradiometer Surveys: Fixed-Wing or Helicopter," Ottawa, CA, pp. 1–2, 2020. [Online]. Available: http://www.sgl.com/MagGradient.html

[25] A. R. Gnadt, J. Belarge, A. J. Canciani, L. Conger, J. A. Curro, A. Edelman, P. Morales, M. F. O'Keeffe, J. Taylor, and C. Rackauckas, "Signal Enhancement for Magnetic Navigation Challenge Problem," *arXiv*, pp. 1–21, 2020. [Online]. Available: http://arxiv.org/abs/2007.12158

[26] A. R. Gnadt, J. Belarge, A. Canciani, L. Conger, J. Curro, A. Edelman, P. Morales, M. F. O'Keeffe, J. Taylor, and C. Rackauckas, "Signal Enhancement for Magnetic Navigation Challenge Problem," 2020. [Online]. Available: https://doi.org/10.5281/zenodo.4271803

[27] A. R. Gnadt, C. Rackauckas, A. Edelman, J. Taylor, G. Carl, and A. Wollaber, "Signal Enhancement for Magnetic Navigation Challenge Problem," 2022. [Online]. Available: https://magnav.mit.edu/

[28] J. Dyment, V. Lesur, M. Hamoudi, Y. Choi, E. Thebault, M. Catalan, WDMAM Task Force, WDMAM Evaluators, and WDMAM Data Providers, "World Digital Magnetic Anomaly Map version 2.0," 2015. [Online]. Available: http://www.wdmam.org

[29] W. J. Hinze, R. R. B. von Frese, and A. H. Saad, *Gravity and Magnetic Exploration: Principles, Practices, and Applications.* Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2013.

[30] A. J. Canciani and J. F. Raquet, "Magnetic Anomaly Navigation Accuracy with Respect to Map Quality and Altitude," in *Proceedings of the 2016 International Technical Meeting of The Institute of Navigation*, Monterey, CA, jan 2016, pp. 110–116. [Online]. Available: https://doi.org/10.33012/2016.13415

[31] W. M. Telford, L. P. Geldart, and R. E. Sheriff, *Applied Geophysics*, 2nd ed. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 1990.

[32] R. J. Blakely, *Potential Theory in Gravity and Magnetic Applications.* Cambridge University Press, 2009. [Online]. Available: https://doi.org/10.1017/CBO9780511549816

[33] L. E. Cordell, J. D. Phillips, and R. H. Godson, "US Geological Survey potential-field geophysical software version 2.0," Geological Survey (US), Tech. Rep., 1992. [Online]. Available: https://pubs.er.usgs.gov/publication/ofr9218https://pubs.er.usgs.gov/publication/ofr97725https://pubs.usgs.gov/of/1997/0725/report.pdf

[34] J. W. Cooley, P. A. W. Lewis, and P. D. Welch, "Historical Notes on the Fast Fourier Transform," *Proceedings of the IEEE*, vol. 55, no. 10, pp. 1675–1677, 1967. [Online]. Available: https://doi.org/10.1109/PROC.1967.5959

[35] J. D. Phillips, "Potential-field continuation: Past practice vs. modern methods," in *SEG Technical Program Expanded Abstracts 1996.* Society of Exploration Geophysicists, jan 1996, pp. 1411–1414. [Online]. Available: https://doi.org/10.1190/1.1826376

[36] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems.* Winston, 1977.

[37] R. Pašteka, R. Karcol, D. Kušnirák, and A. Mojzeš, "REGCONT: A Matlab based program for stable downward continuation of geophysical potential fields using Tikhonov regularization," *Computers & Geosciences*, vol. 49, pp. 278–289, 2012. [Online]. Available: https://doi.org/10.1016/j.cageo.2012.06.010

[38] X. Zeng, D. Liu, X. Li, D. Chen, and C. Niu, "An improved regularized downward continuation of potential field data," *Journal of Applied Geophysics*, vol. 106, pp. 114–118, 2014. [Online]. Available: https://doi.org/10.1016/j.jappgeo.2014.04.015

[39] A. B. Reid, "Aeromagnetic survey design," *GEOPHYSICS*, vol. 45, no. 5, pp. 973–976, 1980. [Online]. Available: https://doi.org/10.1190/1.1441102

[40] Ontario Geological Survey, "Ontario Airborne Geophysical Surveys, Magnetic and Gamma-Ray Spectrometric, Grid and Profile Data (ASCII and Geosoft® Formats) and Vector Data, Renfrew Area," Ontario Geological Survey, Tech. Rep., 2014. [Online]. Available: http://www.geologyontario.mndm.gov.on.ca/mndmfiles/pub/data/records/GDS1074.html

[41] ——, "Ontario Airborne Geophysical Surveys, Magnetic Data, Grid and Profile Data (ASCII and Geosoft® Formats) and Vector Data, Eastern Ontario Area," Ontario Geological Survey, Tech. Rep., 2014. [Online]. Available: http://www.geologyontario.mndm.gov.on.ca/mndmfiles/pub/data/records/GDS1075.html

[42] W. E. Tolles and J. D. Lawson, "Magnetic compensation of MAD equipped aircraft," *Report 201-1*, 1950.

[43] W. E. Tolles, "Compensation of Aircraft Magnetic Fields," US Patent 2,692,970, pp. 1–8, 1954. [Online]. Available: https://patents.google.com/patent/US2692970A/en

[44] ——, "Magnetic Field Compensation System," US Patent 2,706,801, pp. 1–5, 1955. [Online]. Available: https://patents.google.com/patent/US2706801A/en

[45] P. Leliak, "Identification and Evaluation of Magnetic-Field Sources of Magnetic Airborne Detector Equipped Aircraft," *IRE Transactions on Aeronautical and Navigational Electronics*, vol. ANE-8, no. 3, pp. 95–105, 1961. [Online]. Available: https://doi.org/10.1109/TANE3.1961.4201799

[46] L. E. Tuck, “Characterization and compensation of magnetic interference resulting from unmanned aircraft systems,” Doctoral dissertation, Carleton University, 2019. [Online]. Available: https://curve.carleton.ca/d2b99a48-2c1d-4c7c-b9ec-a411ad390020

[47] L. E. Tuck, C. Samson, J. Laliberté, and M. Cunningham, “Magnetic interference mapping of four types of unmanned aircraft systems intended for aeromagnetic surveying,” *Geoscientific Instrumentation, Methods and Data Systems*, vol. 10, no. 1, pp. 101–112, 2021. [Online]. Available: https://doi.org/10.5194/gi-10-101-2021

[48] R. Versteeg, M. McKay, M. Anderson, R. Johnson, B. Selfridge, and J. Bennett, “Feasibility study for an Autonomous UAV-Magnetometer system,” Idaho National Laboratory, Idaho Falls, ID, Tech. Rep., 2007. [Online]. Available: http://www.serdp-estcp.org/content/download/8195/100898/file/mm-1509-fr.pdf

[49] A. Malehmir, L. Dynesius, K. Paulusson, A. Paulusson, H. Johansson, M. Bastani, M. Wedmark, and P. Marsden, “The potential of rotary-wing UAV-based magnetic surveys for mineral exploration: A case study from central Sweden,” *The Leading Edge*, vol. 36, no. 7, pp. 552–557, 2017. [Online]. Available: https://doi.org/10.1190/tle36070552.1

[50] S. H. Bickel, “Small Signal Compensation of Magnetic Fields Resulting from Aircraft Maneuvers,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-15, no. 4, pp. 518–525, 1979. [Online]. Available: https://doi.org/10.1109/TAES.1979.308736

[51] Q. Han, Z. Dou, X. Tong, X. Peng, and H. Guo, “A Modified Tolles–Lawson Model Robust to the Errors of the Three-Axis Strapdown Magnetometer,” *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 3, pp. 334–338, 2017. [Online]. Available: https://doi.org/10.1109/LGRS.2016.2640188

[52] G. Noriega, “Performance measures in aeromagnetic compensation,” *The Leading Edge*, vol. 30, no. 10, pp. 1122–1127, oct 2011. [Online]. Available: https://doi.org/10.1190/1.3657070

[53] M. Drela, *Flight Vehicle Aerodynamics.* MIT Press, 2013.

[54] S. H. Bickel, “Error Analysis of an Algorithm for Magnetic Compensation of Aircraft,” *IEEE Transactions on Aerospace and Electronic Systems*,

vol. AES-15, no. 5, pp. 620–626, 1979. [Online]. Available: https://doi.org/10.1109/TAES.1979.308850

[55] B. W. Leach, "Automatic aeromagnetic compensation," *Technical Report LTR-FR-69*, 1979.

[56] J. B. Nelson, "Analysis of the Sander Geophysics Flt1002 and Flt1003 data," Aeromagnetic Solutions Incorporated, Gloucester, ON, Canada, Tech. Rep., 2020.

[57] B. Gu, Q. Li, and H. Liu, "Aeromagnetic Compensation Based on Truncated Singular Value Decomposition With an Improved Parameter-choice Algorithm," in *6th International Congress on Image and Signal Processing*. IEEE, 2013, pp. 1545–1551. [Online]. Available: https://doi.org/10.1109/CISP.2013.6743921

[58] P. Wu, Q. Zhang, L. Chen, W. Zhu, and G. Fang, "Aeromagnetic compensation algorithm based on principal component analysis," *Journal of Sensors*, pp. 1–7, 2018. [Online]. Available: https://doi.org/10.1155/2018/5798287

[59] L. Chen, P. Wu, W. Zhu, Y. Feng, and G. Fang, "A Novel Strategy for Improving the Aeromagnetic Compensation Performance of Helicopters," *Sensors*, vol. 18, no. 6, p. 1846, 2018. [Online]. Available: https://doi.org/10.3390/s18061846

[60] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301–320, 2005. [Online]. Available: https://doi.org/10.1111/j.1467-9868.2005.00503.x

[61] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996. [Online]. Available: https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

[62] S. Kornblith and J. Dunn, "GLMNet," 2022. [Online]. Available: https://github.com/JuliaStats/GLMNet.jl

[63] T. Lienart, "MLJLinearModels," 2022. [Online]. Available: https://github.com/JuliaAI/MLJLinearModels.jl

[64] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization Paths for Generalized Linear Models via Coordinate Descent," *Journal of Statistical Software*, vol. 33, no. 1, pp. 1–22, 2010. [Online]. Available: https://doi.org/10.18637/jss.v033.i01

[65] B.-H. Mevik and R. Wehrens, "Introduction to the pls Package," The Comprehensive R Archive Network, Tech. Rep., 2021. [Online]. Available: https://cran.r-project.org/web/packages/pls/vignettes/pls-manual.pdf

[66] P. M. Williams, "Aeromagnetic Compensation using Neural Networks," *Neural Computing & Applications*, vol. 1, no. 3, pp. 207–214, 1993. [Online]. Available: https://doi.org/10.1007/BF01414949

[67] M. Ma, D. Cheng, S. Chalup, and Z. Zhou, "Uncertainty Estimation in the Neural Model for Aeromagnetic Compensation," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 12, pp. 1942–1946, 2018. [Online]. Available: https://doi.org/10.1109/LGRS.2018.2864239

[68] P. Yu, X. Zhao, and J. Jiao, "Aeromagnetic Data Preprocessing Method Based on Deep Learning," in *AGU Fall Meeting Abstracts*, 2019, pp. NS13B–0664. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2019AGUFMNS13B0664Y/exportcitation

[69] X. Xu, L. Huang, X. Liu, and G. Fang, "DeepMAD: Deep Learning for Magnetic Anomaly Detection and Denoising," *IEEE Access*, vol. 8, pp. 121 257–121 266, 2020. [Online]. Available: https://doi.org/10.1109/ACCESS.2020.3006795

[70] K. Hornik, M. Stinchcombe, and H. White, "Multilayer Feedforward Networks are Universal Approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989. [Online]. Available: https://doi.org/10.1016/0893-6080(89)90020-8

[71] J. A. Farrell and M. M. Polycarpou, *Adaptive Approximation Based Control: Unifying Neural, Fuzzy and Traditional Adaptive Approximation Approaches*. John Wiley & Sons, 2006, vol. 48. [Online]. Available: https://www.wiley.com/en-us/Adaptive+Approximation+Based+Control{%}3A+Unifying+Neural{%}2C+Fuzzy+and+Traditional+Adaptive+Approximation+Approaches-p-9780471727880

[72] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for Activation Functions," *arXiv*, pp. 1–13, 2017. [Online]. Available: http://arxiv.org/abs/1710.05941

[73] G. Noriega and A. Marszalkowski, “Adaptive techniques and other recent developments in aeromagnetic compensation,” *First Break*, vol. 35, no. 9, pp. 31–38, sep 2017. [Online]. Available: https://doi.org/10.3997/1365-2397.2017018

[74] C. Du, H. Wang, H. Wang, M. Xia, X. Peng, Q. Han, P. Zou, and H. Guo, “Extended aeromagnetic compensation modelling including non-manoeuvring interferences,” *IET Science, Measurement & Technology*, vol. 13, no. 7, pp. 1033–1039, sep 2019. [Online]. Available: https://doi.org/10.1049/iet-smt.2018.5654

[75] M. J. Innes, E. Saba, K. Fischer, D. Gandhi, M. C. Rudilosso, N. M. Joy, T. Karmali, A. Pal, and V. B. Shah, “Fashionable Modelling with Flux,” *CoRR*, vol. abs/1811.0, pp. 1–7, 2018. [Online]. Available: https://arxiv.org/abs/1811.01457

[76] M. J. Innes, “Flux: Elegant machine learning with Julia,” *Journal of Open Source Software*, vol. 3, no. 25, p. 602, 2018. [Online]. Available: https://doi.org/10.21105/joss.00602

[77] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, “Julia: A Fresh Approach to Numerical Computing,” *SIAM Review*, vol. 59, no. 1, pp. 65–98, jan 2017. [Online]. Available: https://doi.org/10.1137/141000671

[78] D. P. Kingma and J. L. Ba, “Adam: A Method for Stochastic Optimization,” in *3rd International Conference on Learning Representations, ICLR 2015*, 2014, pp. 1–15. [Online]. Available: https://arxiv.org/abs/1412.6980

[79] B. B. Chu, K. L. Keys, C. A. German, H. Zhou, J. J. Zhou, E. M. Sobel, J. S. Sinsheimer, and K. Lange, “Iterative hard thresholding in genome-wide association studies: Generalized linear models, prior weights, and double sparsity,” *GigaScience*, vol. 9, no. 6, pp. 1–13, 2020. [Online]. Available: https://doi.org/10.1093/gigascience/giaa044

[80] B. B. Chu, S. Ko, J. J. Zhou, H. Zhou, J. S. Sinsheimer, and K. Lange, “Multivariate Genomewide Association Analysis with IHT,” *bioRxiv*, pp. 1–24, 2021. [Online]. Available: https://doi.org/10.1101/2021.08.04.455145

[81] B. Zhang, Z. Guo, and Y. Qiao, “A simplified aeromagnetic compensation model for low magnetism UAV platform,” in *2011 IEEE International*

Geoscience and Remote Sensing Symposium. IEEE, jul 2011, pp. 3401–3404. [Online]. Available: http://doi.org/10.1109/IGARSS.2011.6049950

[82] D. R. S. Saputro and P. Widyaningsih, "Limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method for the parameter estimation on geographically weighted ordinal logistic regression model (GWOLR)," in *AIP Conference Proceedings*, vol. 1868, 2017, p. 040009. [Online]. Available: https://doi.org/10.1063/1.4995124

[83] J. Feng and N. Simon, "Sparse-Input Neural Networks for High-dimensional Nonparametric Regression and Classification," *arXiv*, pp. 1–37, 2017. [Online]. Available: https://arxiv.org/abs/1711.07592

[84] C. Molnar, "Interpretable Machine Learning, 5.9 Shapley Values," 2021. [Online]. Available: https://christophm.github.io/interpretable-ml-book/shapley.html

[85] N. Redell and 4S Analytics & Modelling Ltd., "ShapML," 2021. [Online]. Available: https://github.com/nredell/ShapML.jl

[86] D. J. Teskey, R. Barlow, P. J. Hood, D. Lefebvre, N. Paterson, M. Reford, and D. Watson, "Guide to aeromagnetic specifications and contracts," Geological Survey of Canada, Tech. Rep., 1991. [Online]. Available: https://geoscan.nrcan.gc.ca/starweb/geoscan/servlet.starweb?path=geoscan/fulle.web{&}search1=R=131845

[87] J. A. Curro, "Navigation with Artificial Neural Networks," Doctoral dissertation, Air Force Institute of Technology, 2018. [Online]. Available: https://scholar.afit.edu/etd/1948/

[88] E. Oja, "Principal Components, Minor Components, and Linear Neural Networks," *Neural Networks*, vol. 5, no. 6, pp. 927–935, 1992. [Online]. Available: https://doi.org/10.1016/S0893-6080(05)80089-9

[89] M. Kok, J. D. Hol, and T. B. Schön, "Using Inertial Sensors for Position and Orientation Estimation," *Foundations and Trends in Signal Processing*, vol. 11, no. 1-2, pp. 1–153, 2017. [Online]. Available: https://doi.org/10.1561/2000000094

[90] C. Tyren, "Magnetic terrain navigation," in *Proceedings of the 1987 5th International Symposium on Unmanned Untethered Submersible Technology.* IEEE, 1987, pp. 245–256.

[91] F. Jie, C. Guoxiong, and L. Tianyou, "Kalman Filter Underwater Passive Geomagnetic Navigation Technology Research Considering the Impact of Diurnal Variation," in *2009 2nd International Congress on Image and Signal Processing.* IEEE, oct 2009, pp. 1–4. [Online]. Available: https://doi.org/10.1109/CISP.2009.5304917

[92] G. Shorshi and I. Bar-Itzhack, "Satellite autonomous navigation and orbit determination using magnetometers," in *Proceedings of the 31st IEEE Conference on Decision and Control*, IEEE. IEEE, 1992, pp. 542–548. [Online]. Available: https://doi.org/10.1109/CDC.1992.371675

[93] M. L. Psiaki, L. Huang, and S. M. Fox, "Ground tests of magnetometer-based autonomous navigation (MAGNAV) for low-earth-orbiting spacecraft," *Journal of Guidance, Control, and Dynamics*, vol. 16, no. 1, pp. 206–214, jan 1993. [Online]. Available: https://doi.org/10.2514/3.1144

[94] M. L. Psiaki, "Autonomous Low-Earth-Orbit Determination from Magnetometer and Sun Sensor Data," *Journal of Guidance, Control, and Dynamics*, vol. 22, no. 2, pp. 296–304, mar 1999. [Online]. Available: https://doi.org/10.2514/2.4378

[95] J. A. Shockley, "Ground Vehicle Navigation Using Magnetic Field Variation," Doctoral dissertation, Air Force Institute of Technology, 2012. [Online]. Available: https://scholar.afit.edu/etd/1156/

[96] J. A. Shockley and J. F. Raquet, "Navigation of Ground Vehicles Using Magnetic Field Variations," *NAVIGATION: Journal of the Institute of Navigation*, vol. 61, no. 4, pp. 237–252, dec 2014. [Online]. Available: https://doi.org/10.1002/navi.70

[97] A. J. Canciani and C. J. Brennan, "An Analysis of the Benefits and Difficulties of Aerial Magnetic Vector Navigation," *IEEE Transactions on Aerospace and Electronic Systems*, pp. 1–15, 2020. [Online]. Available: https://doi.org/10.1109/taes.2020.2987475

[98] B. Li, T. Gallagher, A. G. Dempster, and C. Rizos, "How feasible is the use of magnetic field alone for indoor positioning?" in *2012 International Conference on Indoor Positioning and Indoor Navigation.* IEEE, nov 2012, pp. 1–9. [Online]. Available: https://doi.org/10.1109/IPIN.2012.6418880

[99] J. Haverinen and A. Kemppainen, “A global self-localization technique utilizing local anomalies of the ambient magnetic field,” in *2009 IEEE International Conference on Robotics and Automation.* Kobe, Japan: IEEE, may 2009, pp. 3142–3147. [Online]. Available: https://doi.org/10.1109/ROBOT.2009.5152885

[100] H.-S. Kim, W. Seo, and K.-R. Baek, “Indoor Positioning System Using Magnetic Field Map Navigation and an Encoder System,” *Sensors*, vol. 17, no. 3, p. 651, mar 2017. [Online]. Available: https://doi.org/10.3390/s17030651

[101] W. F. Storms, “Magnetic Field Aided Indoor Navigation,” Master’s thesis, Air Force Institute of Technology, 2009. [Online]. Available: https://scholar.afit.edu/etd/2567/

[102] W. F. Storms, J. A. Shockley, and J. F. Raquet, “Magnetic Field Navigation in an Indoor Environment,” in *2010 Ubiquitous Positioning Indoor Navigation and Location Based Service.* Kirkkonummi, Finland: IEEE, oct 2010, pp. 1–10. [Online]. Available: https://doi.org/10.1109/UPINLBS.2010.5653681

[103] J. M. Wilson, R. J. Kline-Schoder, M. A. Kenton, P. H. Sorensen, and O. H. Clavier, “Passive Navigation Using Local Magnetic Field Variations,” in *Proceedings of the 2006 National Technical Meeting of the Institute of Navigation*, Monterey, CA, 2006, pp. 770–779. [Online]. Available: https://www.ion.org/publications/abstract.cfm?articleID=6581

[104] A. Doucet, N. Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice.* New York, NY: Springer New York, 2001. [Online]. Available: https://doi.org/10.1007/978-1-4757-3437-9

[105] T. B. Schön, F. Gustafsson, and P.-J. Nordlund, “Marginalized Particle Filters for Mixed Linear/Nonlinear State-Space Models,” *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2279–2289, 2005. [Online]. Available: https://doi.org/10.1109/TSP.2005.849151

[106] T. N. Lee, “Aerial Simultaneous Localization and Mapping Using Earth’s Magnetic Anomaly Field,” Master’s thesis, Air Force Institute of Technology, 2019. [Online]. Available: https://scholar.afit.edu/etd/2268/

[107] T. N. Lee and A. J. Canciani, “MagSLAM: Aerial simultaneous localization and mapping using Earth’s magnetic anomaly field,” *NAVIGATION*, vol. 67, no. 1, pp. 95–107, mar 2020. [Online]. Available: https://doi.org/10.1002/navi.352

[108] A. V. Oppenheim and G. C. Verghese, *Signals, Systems and Inference.* Pearson, 2015.

[109] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960. [Online]. Available: https://doi.org/10.1115/1.3662552

[110] S. Haykin, Ed., *Kalman Filtering and Neural Networks.* New York, USA: John Wiley & Sons, Inc., 2001. [Online]. Available: https://doi.org/10.1002/0471221546

[111] A. Gelb, *Applied Optimal Estimation.* MIT Press, 1974.

[112] D.-J. Jwo and H.-C. Huang, "Neural Network Aided Adaptive Extended Kalman Filtering Approach for DGPS Positioning," *Journal of Navigation*, vol. 57, pp. 449–463, 2004. [Online]. Available: https://doi.org/10.1017/S0373463304002814

[113] D. Titterton and J. Weston, *Strapdown Inertial Navigation Technology*, 2nd ed. London, UK: The Institute of Engineering and Technology and The American Institute of Aeronautics, 2004. [Online]. Available: http://digital-library.theiet.org/docserver/fulltext/books/ra/pbra017e/PBRA017E.pdf?expires=1515449011{&}id=id{&}accname=ietid004563{&}checksum=D9C89688A47C097CDA87C8862A59461D

[114] Z. Dou, K. Ren, Q. Han, and X. Niu, "A Novel Real-Time Aeromagnetic Compensation Method Based on RLSQ," in *Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing.* IEEE, 2014, pp. 243–246. [Online]. Available: https://doi.org/10.1109/IIH-MSP.2014.67

[115] T. Beravs, S. Beguš, J. Podobnik, and M. Munih, "Magnetometer Calibration Using Kalman Filter Covariance Matrix for Online Estimation of Magnetic Field Orientation," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 8, pp. 2013–2020, aug 2014. [Online]. Available: https://doi.org/10.1109/TIM.2014.2302240

[116] B. Siebler, S. Sand, and U. D. Hanebeck, "Localization with Magnetic Field Distortions and Simultaneous Magnetometer Calibration," *IEEE Sensors Journal*, vol. 21, no. 3, pp. 3388–3397, 2020. [Online]. Available: https://doi.org/10.1109/JSEN.2020.3024073

[117] Flux ML, "Transfer Learning with Flux," 2020. [Online]. Available: https://fluxml.ai/tutorials/2020/10/18/transfer-learning.html

[118] R. K. Mehra, "On the Identification of Variances and Adaptive Kalman Filtering," *IEEE Transactions on Automatic Control*, vol. 15, no. 2, pp. 175–184, apr 1970. [Online]. Available: https://doi.org/10.1109/TAC.1970.1099422

[119] ——, "Approaches to Adaptive Filtering," *IEEE Transactions on Automatic Control*, vol. 17, no. 5, pp. 693–698, 1972. [Online]. Available: https://doi.org/10.1109/TAC.1972.1100100

[120] A. H. Mohamed and K. P. Schwarz, "Adaptive Kalman Filtering for INS/GPS," *Journal of Geodesy*, vol. 73, no. 4, pp. 193–203, 1999.

[121] S. C. Stubberud, R. N. Lobbia, and M. Owen, "An Adaptive Extended Kalman Filter Using Artificial Neural Networks," in *Proceedings of 34th IEEE Conference on Decision and Control.* New Orleans, LA: IEEE, 1995, pp. 1852–1856. [Online]. Available: https://doi.org/10.1109/CDC.1995.480611

[122] S. C. Stubberud and K. A. Kramer, "System Identification Using the Neural-Extended Kalman Filter For Control Modification," in *2006 IEEE International Joint Conference on Neural Network Proceedings.* IEEE, 2006, pp. 4449–4455. [Online]. Available: https://doi.org/10.1109/IJCNN.2006.247047

[123] J. J. Wang, W. Ding, and J. Wang, "Improving Adaptive Kalman Filter in GPS/SDINS Integration with Neural Network," in *20th International Technical Meeting of the Satellite Division of The Institute of Navigation*, 2007, pp. 571–578. [Online]. Available: https://www.ion.org/publications/abstract.cfm?articleID=7555

[124] J.-K. Lee and C. Jekeli, "Neural Network Aided Adaptive Filtering and Smoothing for an Integrated INS/GPS Unexploded Ordnance Geolocation System," *Journal of Navigation*, vol. 63, no. 2, pp. 251–267, 2010. [Online]. Available: https://doi.org/10.1017/S0373463309990397

[125] M. Zhu, K. Murphy, and R. Jonschkowski, "Towards Differentiable Resampling," *arXiv*, pp. 1–8, apr 2020. [Online]. Available: http://arxiv.org/abs/2004.11938

[126] D. J. Ellis, "Pedestrian Navigation using Artificial Neural Networks and Classical Filtering Techniques," Master's thesis, Air Force Institute of Technology, 2020. [Online]. Available: https://scholar.afit.edu/etd/3618/

[127] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997. [Online]. Available: https://doi.org/10.1162/neco.1997.9.8.1735

[128] C. Olah, “Understanding LSTM Networks,” 2015. [Online]. Available: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

[129] M. J. Innes, “Don’t Unroll Adjoint: Differentiating SSA-Form Programs,” *arXiv*, pp. 1–9, 2018. [Online]. Available: http://arxiv.org/abs/1810.07951

[130] Flux ML, “Custom Adjoints,” 2022. [Online]. Available: https://fluxml.ai/Zygote.jl/stable/adjoints/

[131] A. R. Gnadt, “Machine Learning-Enhanced Magnetic Calibration for Airborne Magnetic Anomaly Navigation,” in *AIAA SCITECH 2022 Forum.* AIAA, 2022, pp. 1–16. [Online]. Available: https://doi.org/10.2514/6.2022-1760