IMPROVING AEROMAGNETIC CALIBRATION USING ARTIFICIAL NEURAL NETWORKS

THESIS

Mitchel Hezel

AFIT-ENG-MS-20-M-027

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENG-MS-20-M-027

IMPROVING AEROMAGNETIC CALIBRATION USING ARTIFICIAL NEURAL NETWORKS

THESIS

Presented to the Faculty

Department of Electrical and Computer Engineering

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Electrical Engineering

Mitchel Hezel, B.S. Physics

March 2020

AFIT-ENG-MS-20-M-027

IMPROVING AEROMAGNETIC CALIBRATION USING ARTIFICIAL NEURAL NETWORKS

THESIS

Mitchel Hezel, B.S. Physics

Committee Membership:

Maj Aaron Canciani, Ph.D
Chair

Capt Joseph Curro, Ph.D
Member

Dr. Robert Leishman, Ph.D
Member

AFIT-ENG-MS-20-M-027

Abstract

The Global Positioning System (GPS) has proven itself to be the single most accurate positioning system available, and no navigation suite is found without a GPS receiver. Even basic GPS receivers found in most smartphones can easily provide high quality positioning information at any time. Even with its superb performance, GPS is prone to jamming and spoofing, and many platforms requiring accurate positioning information are in dire need of other navigation solutions to compensate in the event of an outage, be the cause hostile or natural. Indeed, there has been a large push to achieve an alternative navigation capability which performs nearly as well as GPS. One navigation method which has shown promise to increase navigation performance of aircraft utilizes magnetic anomalies[1] - local variations in the Earth's crust - to discern position. One significant drawback to this approach is the magnetic disturbance generated by the aircraft itself, which must be accounted for and eliminated. Current calibration procedures involve placing the magnetometer on a long stinger far from the aircraft body to minimize interference with the magnetic anomaly signal. While some aircraft permit the addition of stingers, many do not. No calibration procedure exists which satisfies potential location restraints of the magnetometer and the calibration problem for these less ideal aircraft, especially potentially magnetically noisy platforms such as an F16. Current linear models which attempt to correct mild disturbance fields on more ideal aircraft exist. We propose that a more sophisticated model is necessary to make magnetic navigation platform agnostic. Specifically, we show that a deep learning approach and the utilization of more inputs than the current *de facto* calibration procedure - known in the literature as Tolles-Lawson - can achieve a 90% reduction in platform based magnetic disturbance signals.

Table of Contents

Page

List of Figures

List of Tables

IMPROVING AEROMAGNETIC CALIBRATION USING ARTIFICIAL NEURAL NETWORKS

I. Introduction

1.1 Problem Background

There is a great and urgent desire to create dependable navigation alternatives to GPS. GPS provides unmatched accuracy for an unlimited amount of time, but is prone to jamming and spoofing, wherein the GPS signal processed by GPS receivers is either made unavailable or made to suggest a location other than that of the user. This vulnerability in GPS is its largest pitfall, and one whose GPS-equivalent solution has evaded engineers and researchers alike. In fact, completely replacing GPS is probably unrealistic, and the effort has been to compliment GPS and provide solutions while GPS is unavailable and until the signal is recovered. Many GPS-denied navigation solutions exist, each with their own pros and cons. Inertial Measurement Units (IMUs) are found in nearly all aircraft, ships, many land based vehicles, and even in our personal phones. An IMU is a collection of accelerometers and gyroscopes which measure translational and rotational forces, respectively. In order to receive position estimates from an accelerometer, a double integration must be performed. Any errors or drifts present in the accelerometer or gyroscope are then propagated through time quadratically. This renders even the best IMUs ineffective for navigation over long periods of time. For this reason, IMUs are not typically used as a sole means of navigation for extended periods of time.

Other alternative navigation solutions include visual odometry, wherein a feature detection algorithm determines useful image features to track from frame to frame. The translation of the tracked features between images is then translated to a relative position, with error accruing proportional to the duration of its use. Visual odometry is closely related to landmark marching. In landmark matching, images that are taken by the vehicle are cross-referenced to a database of landmarks whose GPS position information is well known. Once a landmark is recognized, the vehicle's positioning is known with absolute accuracy. Landmark matching relies on other navigation solutions (such as an IMU) until the next update is acquired. Of course, there are drawbacks to visual forms of navigation. Visual odometry works under the assumption that the image features being tracked are sufficiently illuminated, unobscured, and remain visible and otherwise identical from image to image. Abrupt changes in lighting or weather can cause degraded performance. Highly accelerated movements which result in a rapid loss of many features over a short time will also negatively impact performance. Landmark navigation suffers from the same environmental drawbacks as visual odometry. In addition, the removal or alteration of a landmark must be accounted for each time it occurs, necessitating periodic database updates. Finally, both forms of navigation fail when there are no useful features or landmarks to identify, and operate poorly over areas where features stagnate, such as over water.

Still other forms of alternative navigation exist. Sensors which measure celestial signals - specifically sextants - have been used for hundreds of years. More recently, startrackers have shown promise to navigate when a clear view of the night sky is present. The position of stars in the night sky is known to a very high accuracy, and the recognition of certain constellations or other celestial objects can yield positioning information. Of course, this solution is not without its drawbacks. Chiefly, most star-

trackers only work when there is an unobstructed view of the stars - when it's dark. This represents an obvious, significant weakness to celestial navigation. Needless to say, having another navigation solution which is more robust and can provide positioning information in environments and circumstances where other methods cannot would be very beneficial to the GPS-denied navigation problem.

1.1.1 Magnetic Navigation

Recently [1] a new means of airborne GPS-denied navigation has been proposed and validated. Navigation via magnetic anomalies is a process by which magnetic signals created in the earth's crust - magnetic anomalies - are used to navigate. As an aircraft flies over a region whose magnetic anomaly characteristics are known and mapped, the readings recorded by the magnetometer are referenced to the anomaly values on the map. A nonlinear state estimator, like an extended Kalman filter or particle filter, then determines the likely position of the vehicle given a series of these measurements, resulting in an absolute position estimate of the vehicle. Importantly, even though any single magnetic anomaly measurement may fail to yield a unique position solution - there could be many areas on the map with the same anomaly value - a series of measurements along a given trajectory could very well be unique. The advantages of magnetic anomaly navigation are apparent: magnetic measurements can be taken and are available regardless of weather, time of day, location, and other environmental factors. Canciani also determined that magnetic anomaly maps are viable over long periods of time, potentially many years and even geological timescales. This implies that, once an area has been mapped, it won't need remapping unless significant geological activity occurs in the area. Additionally, this form of navigation is virtually unjammable. Due to how the magnetic field decays with

distance, the amount of energy needed to disable a magnetic navigation capability from a ground station is unrealistically large. An adversary looking to impair an aircraft's navigation hardware would be unlikely to target magnetometers.

Of course, like the other navigation methods, magnetic navigation has its drawbacks, primarily two: map availability and platform effects. There exist a very small number of high quality magnetic anomaly maps, and their widespread creation is costly and time consuming. Additionally, the platform performing the navigation will impart its own magnetically corrupting signature to the measurements of interest, resulting in potentially severe degradation of the anomaly signal. The resolution of these two problems would permit wide scale use of magnetic anomaly navigation, and would provide a navigation technique which synergizes well with existing alternative navigation solutions.

1.2 Research Objectives and Outcomes

This work furthered the utility of the research performed by Canciani by improving platform calibration procedures and showed that the traditional platform compensation model is often insufficient to reduce platform disturbance fields to levels that are conducive to magnetic anomaly navigation. In particular, we show that the inclusion of other sensors and inputs, as well as more complicated machine learning models, can more accurately reduce platform effects than has been shown in the literature. By utilizing control surface inputs and accelerometer and gyroscope measurements, this research succeeded in reducing the error of the state-of-the-art compensation procedure by 90%, from nearly 48nT RMSE down to less than 5nT RMSE on a real UAV. Research was conducted at the Air Force Institute of Technology (AFIT).

1.3 Document Overview

Chapter 2 focuses on introducing magnetic anomaly navigation as well as the platform calibration problem. We explain why magnetic navigation constitutes a promising alternative to traditional navigation procedures and the relevant research that has been performed so far. We also discuss different machine learning models and why it is believed that a more sophisticated machine learning approach can aid in modelling complicated disturbance field generation by the platform. Chapter 3 outlines the experiment performed and provides a detailed look at the sensors and aircraft that were used to facilitate the experiment. Chapter 4 discusses the data that was collected by the Air Force Institute of Technology's UAV and the different machine learning models that were ultimately used to drastically improve compensation performance. Chapter 5 concludes the thesis, providing a summary of the results, and discusses future work to be done which would likely further the calibration problem solution.

II. Background and Literature Review

2.1 Overview

This section begins with a detailed description of the earth's magnetic field, magnetic anomalies, and the construction of the maps used for magnetic anomaly navigation, as well as a brief summary of different navigation techniques and their shortcomings. It continues by illustrating why navigation via magnetic anomalies constitutes a promising alternative to other GPS denied navigation solutions, and showcases early and current aeromagnetic compensation techniques and the limitations thereof. Finally, we discuss machine learning methods and neural networks and discuss why it is believed that deep learning models can improve current aeromagnetic compensation efforts; one of the two primary roadblocks to fieldable magnetic navigation.

2.2 The Earth's Magnetic Field

2.2.1 The Main Earth Field

The bulk of Earth's magnetic field is generated deep within the earth. Escaping heat from the core and the rotation of the earth move the liquid outer core, which contains conductive, molten iron, and creates electric currents inside the outer core[13]. As a result, a strong magnetic field is produced, and its value at the surface of the earth has a typical range of 20-70μT (20,000-70,000nT). This is the field that magnetometers will mainly measure. To a first approximation, Earth's magnetic field is a dipole. Figure 1 shows a computer simulation of the earth's magnetic field, highly reminiscent of that of a magnetic dipole, similar to if a bar magnet were placed at the center of the earth.

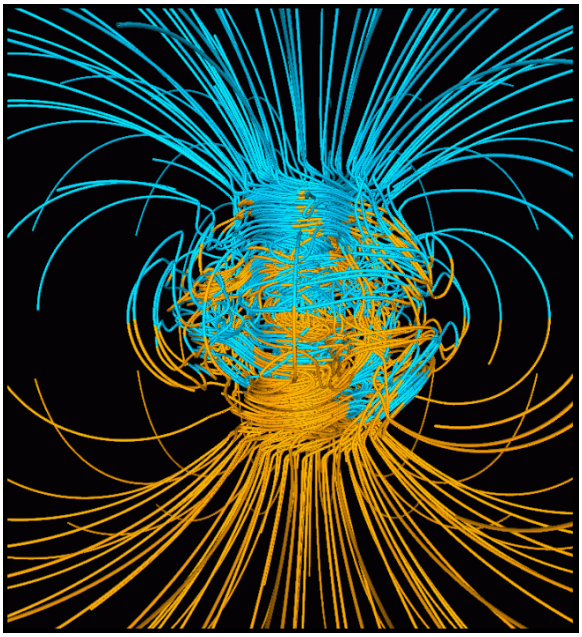

Figure 1: A computer simulation shows the Earth's magnetic field lines. [2]

.

There are multiple models used to estimate the magnetic field due to the core at any location on Earth. The World Magnetic Model (WMM) is one commonly chosen, utilizing a degree and order 12 spherical harmonic model to predict the core field at a given latitude, longitude, radius, and time [14]. In addition, because of the complicated and time-dependent physics governing the motion of the liquid outer core, the model incorporates a linear model for modelling *secular* variation in the field; changes in the field that occur over a time span of a couple of years. The physics of the field change are nonlinear, however, and so the entire WMM must be updated every 5 years. The current model, implemented in 2015, is valid through 2019 [14].

The WMM models the main core field, which makes up the vast majority of the overall fields strength. Unfortunately for navigation, the WMM predicts that the wavelengths of the main earth field are quite long; the shortest are around 3000km

[14]. The frequency of this signal is therefore far too low to be of use from a navigation standpoint.

2.2.2 The Crustal Field

While the main earth field supplies the bulk of the magnetic signal measured by a magnetometer, it is not the only source. The earth's crust - the lithosphere - contains many magnetic and magnetically susceptible materials that either naturally create their own magnetic field or do so because they are induced by the earth's core field. Often times these sources come from large deposits of ferrous or otherwise magnetic material such as magnetite. Figure 2 shows a map of magnetic strengths due to these crustal sources.

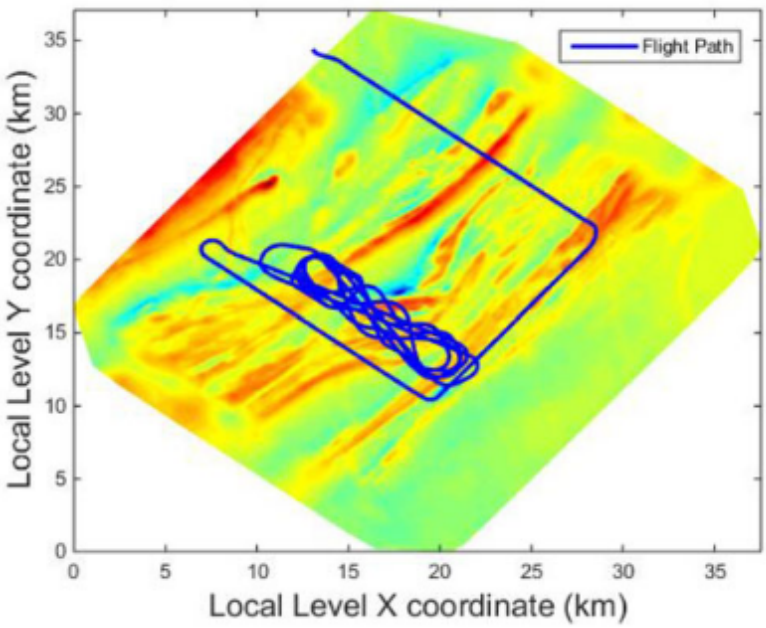

Figure 2: A magnetic anomaly map over Lousia, Virginia [1]

.

As stated earlier, the main core field generates a magnetic strength of tens of

thousands of nano-Tesla. Typical ranges for crustal field sources are on the order of hundreds of nano-Tesla. The frequency content of crustal field sources is much higher than that of the main field. It is therefore it is a primary candidate for use in navigation.

It is important to understand how magnetic sources from the crust are measured. For many years crustal magnetic maps and crustal field sources have been used as a means to search for mineral deposits and oil because they enable quick exploration of shallow crustal substructure. Typically for airborne surveys, a highly accurate scalar magnetometer is placed on a stinger far from the surveying aircraft. This has the effect of minimizing magnetic noise coming from the aircraft - a phenomenon which this thesis is chiefly concerned with. Next, a survey flight pattern is flown over the area of interest at a predetermined altitude and line spacing, and scalar magnetic data is collected. The height of the survey and the spacing between back-and-forth flight lines determine the resolution of the map. Once other variations - such as diurnal variations - are accounted for and removed, the scalar magnetic reading is subtracted from the WMM model's prediction of the magnetic field in the area to yield what is known as the *magnetic anomaly*. That is, for any point along the survey flight,

$$\delta b \approx B_{core} - B_T \tag{1}$$

Note that short term fluctuations such as space-related effects are eliminated by consulting with a nearby reference station. Also important to note is that (1) is an approximation, since magnetic fields are vector fields. What is really being measured during an anomaly survey is the projection along the main core field made by the anomaly. Consider figure 3.

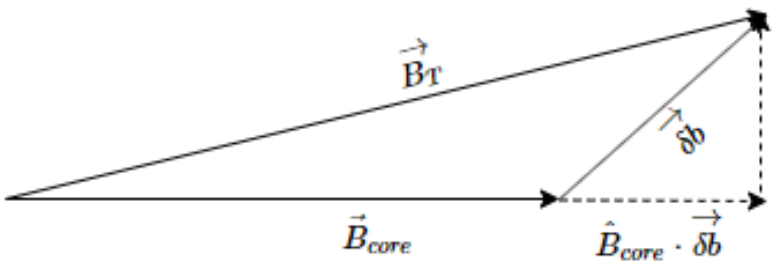

Figure 3: A measurement of the magnetic field, decomposed into core and anomaly field contributions. Not to scale.

.

Because the aircraft collecting the magnetic measurements is doing so with a scalar magnetometer, it is difficult to capture the entire vector contribution to the core field. Instead, what is measured is the projection onto the core field by the anomaly field vector. Consider the following:

$$\begin{aligned} |\vec{B}_T| &= |\vec{B}_{core} + \vec{\delta b}| \\ &= (\vec{B}_{core} \cdot \vec{B}_{core} + 2\vec{B}_{core} \cdot \vec{\delta b} + \vec{\delta b} \cdot \vec{\delta b})^{1/2} \\ &\approx (B_{core}^2 + 2\vec{B}_{core} \cdot \vec{\delta b})^{1/2} \\ &= B_{core}(1 + \frac{2\vec{B}_{core} \cdot \vec{\delta b}}{B_{core}^2})^{1/2} \\ &\approx B_{core}(1 + \frac{\vec{B}_{core} \cdot \vec{\delta b}}{B_{core}^2}) \\ &= B_{core} + \hat{B}_{core} \cdot \vec{\delta b} \end{aligned} \tag{2}$$

Equation (2) indicates that, for magnetic anomalies, what we measure is the deviation from the core field along the direction of the core field. As a quick example, suppose the angle between $\vec{B}_{core}$ and $\vec{\delta b}$ is 45°, $B_{core} = 50,000nT$, $\delta b = 100nT$. The difference between the approximated total field strength and the actual field strength is $\vec{B}_T - (B_{core} + \hat{B}_{core} \cdot \vec{\delta b}) = 0.05nT$. This is an extremely small difference, and

validates the approximation. Most of the time, the approximation is made better because the angle between the anomaly and core field vectors is typically small.

2.2.3 Ionospheric Field

As discussed earlier, the WMM must be updated every 5 years to account for the Earth's changing core. There are other contributions to the overall magnetic field that vary much quicker than this. During the day, sunlight striking the ionosphere ionizes air molecules, thereby creating a conductive medium. As the sun heats part of the earth, the resulting winds push the ionosphere through the magnetic field of the earth. The so-called "Solar-Quiet (Sq) currents" - named for their ease of observability during times of low solar activity - generated from this produce magnetic fields of their own, which can be detected by base-stations armed with sensitive magnetometers. See figures 4 and 5.

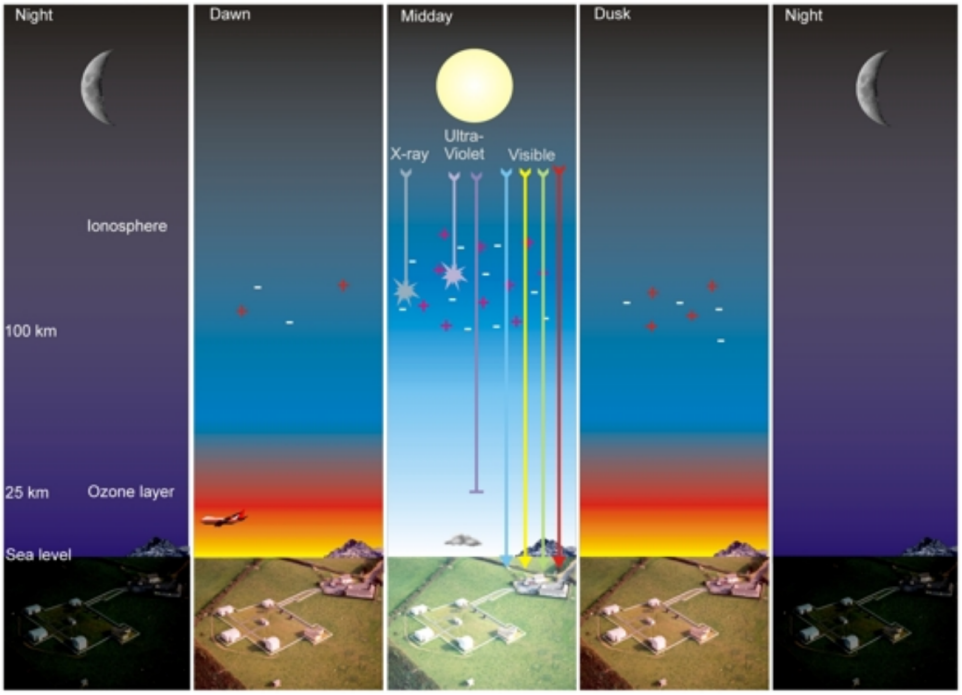

Figure 4: An illustration of solar radiation ionizing the air, leading to the Sq currents.[3]

.

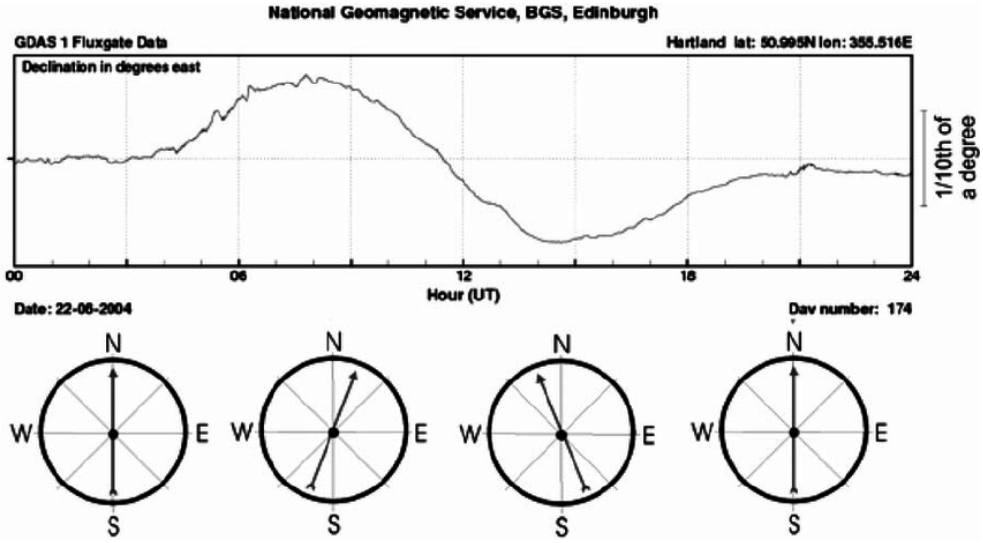

Figure 5: Magnetic variation due to Sq currents taken at a base station in Hartland, UK. The compass deflection is exaggerated, but represents a real phenomenon. [3]

.

The period of Sq induced magnetic fields is 24-hours and their strengths are typically on the order of tens of nano-Tesla. Reference to a nearby ground station can aid the elimination of Sq-induced error sources corrupting a magnetic reading, if one exists.

2.3 Aeromagnetic Interference

To adequately perform aeromagnetic surveys or, more recently, aeromagnetic navigation, it is important to remove the effects of the aircraft. Most aircraft contain magnetic and magnetically susceptible materials. This means that they will have a non-negligible effect on nearby magnetometers. For example, an aircraft engine will cause significant swings in the magnetometer reading if the magnetometer is within close proximity. Additionally, if the sensor is housed in a metal pod the readings can be near useless. These swings can be on the order of thousands of nano-Tesla. This is catastrophic for any survey where magnetometer accuracy is required. As a preventative measure, magnetometers used for airborne magnetic surveys have either been placed on a "stinger" far from the hull of the aircraft or slung below the body. Figure 6 shows a typical example of a magnetometer on a stinger. Since magnetic field strength falls off following an inverse distance cubed law, this decades-old technique is still the go-to method for reducing aircraft-induced magnetic interference. However, not all aircraft come equipped with stingers, and not all aircraft will accept them. Additionally, there may be a desire to house the magnetometer anywhere on board the aircraft, for a multitude of reasons. Military aircraft which don't have stingers may be difficult to augment with one, and many airborne vehicles are unsuitable for magnetometer towing, such as an F16. Moreover, if magnetometers are to be used in conjunction with other sensors and space is limited, it might be very handy to have the ability to place the magnetometer anywhere on-board.

Figure 6: Fixed wing aircraft with stinger seen extending beyond the rudder. [4]

.

2.3.1 Traditional Calibration and Tolles-Lawson

At any time during flight, an on-board scalar magnetometer will read a magnetic signal given by the following

$$B_{sensed}(\boldsymbol{x}, \boldsymbol{\theta}, t) = B_E(\boldsymbol{x}, t) + B_{dist}(\boldsymbol{x}, \boldsymbol{\theta}, t, \epsilon) \tag{3}$$

where $\boldsymbol{x} = \begin{bmatrix} x & y & z \end{bmatrix}^T$ denotes the spatial positioning of the magnetometer on Earth and $\boldsymbol{\theta} = \begin{bmatrix} \theta & \phi & \psi \end{bmatrix}^T$ denotes the orientation of the aircraft relative to the geomagnetic field vector. Each term in $\boldsymbol{\theta}$ denotes the angle between the geomagnetic vector and one of the aircraft's primary axes.

There are 3 important components to (3). B_{sensed} is the value the scalar magnetometer is outputting, and the value we care about when performing magnetic navigation and magnetic anomaly map making. B_E is the geomagnetic field - the field produced by the earth. Ideally, B_{sensed} will consist of only this part, and B_{dist} will be nearly 0.

B_{dist} consists of all corrupting sources. This biggest component of B_{dist} comes from the platform effects. However, this term actually encompasses all nuisance sources of magnetic disturbance, like ionospheric effects and even lightning strikes.

The backbone of aeromagnetic compensation for platform effects was established by Tolles and Lawson [15]. The Tolles-Lawson equations, given in their modern form in [16], assert that the disturbance field is given by

$$B_{dist} = B_{perm} + B_{ind} + B_{eddy} \tag{4}$$

B_{perm} is the portion of the disturbance field created by objects in the aircraft with permanent magnetization. A permanent magnet, for example, creates a field on board the aircraft which impairs the magnetometer's ability to measure the Earth's magnetic field accurately. This field is static relative to the platform, but will interact with the magnetic field differently depending on the orientation of the aircraft. B_{ind} is created from magnetically susceptible materials which produce an induced magnetic field when placed inside the earth's main field. Unlike permanent sources, this field is not static relative to the platform, and will change depending on the orientation of the aircraft. Induced effects are not unlike the ionospheric fields produced during the movement of the newly ionized medium in the earth's field. The aircraft engine, often magnetically susceptible, will typically create a large, induced field. Finally, B_{eddy} models the effects of eddy currents; undesired currents created in the bulk of a conductor similar in nature to induced currents.

The different components of (4) are defined [17] as:

$$B_{perm} = a_1 \cos\theta + a_2 \cos\phi + a_3 \cos\psi$$

$$B_{ind} = B_t(a_4 \cos^2\theta + a_5 \cos\theta\cos\phi + a_6 \cos\theta\cos\psi + a_7 \cos^2\phi + a_8 \cos\phi\cos\psi + a_9 \cos^2\psi)$$

$$\begin{aligned} B_{eddy} = B_t(&a_{10}\cos\theta(\cos\theta)' + a_{11}\cos\theta(\cos\phi)' + a_{12}\cos\theta(\cos\psi)' \\ &+ a_{13}\cos\phi(\cos\theta)' + a_{14}\cos\phi(\cos\phi)' + a_{15}\cos\phi(\cos\psi)' \\ &+ a_{16}\cos\psi(\cos\theta)' + a_{17}\cos\psi(\cos\phi)' + a_{18}\cos\psi(\cos\psi)') \end{aligned}$$

or, more compactly,

$$B_{dist} = \mathbf{A}\mathbf{u} + \mathbf{u}^T\mathbf{B}\mathbf{u} + (\mathbf{u}^T)'\mathbf{C}\mathbf{u} \tag{5}$$

where $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ represent the coefficient matrices to be determined using linear regression, and $\mathbf{u} = \begin{bmatrix} cos\theta & cos\phi & cos\psi \end{bmatrix}^T$ is a vector of the *direction cosines* which describe the orientation of the aircraft in the earth's field. These terms are determined using a 3-axis fluxgate magnetometer. Namely,

$$\begin{aligned} \cos\theta &= F_{x'}/B_t \\ \cos\phi &= F_{y'}/B_t \\ \cos\psi &= F_{z'}/B_t \end{aligned} \tag{6}$$

where $F_{i'}$ is the magnetic flux along the i-th aircraft axis, as in figure 7.

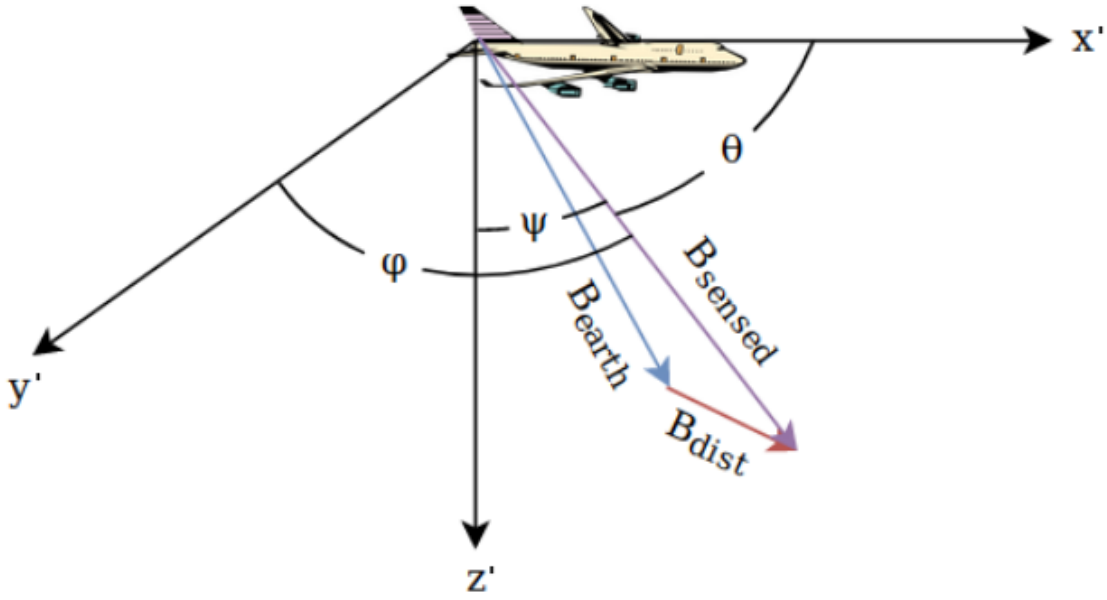

Figure 7: Aircraft reference frame used during aeromagnetic compensation

The reason why we use a 3-axis vector magnetometer can be subtle, but is important. In magnetic navigation, the signal we care about is the scalar magnetic value. Scalar values measure intensity, and not direction. However, we know that the platform effects are often highly correlated with the aircraft's orientation with respect to the geomagnetic vector. The sensor which measures magnetic direction and the orientation of the aircraft within the earth field is the vector magnetometer, since it gives information about the strength of the magnetic field along each of the aircraft axes. It has therefore always played a key role in platform magnetic calibration.

To determine the Tolles-Lawson coefficients which will ultimately be used to compensate for platform induced effects, a calibration flight is performed. During a calibration flight, the magnetometer is flown away from man-made and other sources of magnetic disturbance to isolate the aircraft's effect on the magnetometer. Additionally, the aircraft is flown in an area where the geomagnetic field varies as little as

possible. This further helps to ensure any variation is due solely to the aircraft. To do this, multiple steps must be taken. First, calibration will preferably be done in an area which has already been magnetically mapped. This allows for the selection of a calibration area where the geomagnetic gradient is at or near 0. A low gradient indicates a flat, constant field, and flying through that field should ideally not cause any variation in the scalar magnetometer. Next, the aircraft is flown high - potentially several kilometers above ground level - to ensure that man-made effects have been neutralized due to the $1/r^3$ effect. Finally, the aircraft is flown in a tight area. This has the effect of ensuring that the magnetic gradient of the area, if not 0, is at least linear over the calibration flight. The simpler the underlying geomagnetic field, the easier it is to model it during calibration. Planar geomagnetic variations are simpler to model than those which are not planar. Once the calibration area has been determined, the calibration flight is performed. Calibration flights typically include flying in four orthogonal directions and performing small-angle banking maneuvers along each heading, typically $5 - 10^\circ$ [16]. This consists of actuating the ailerons, elevator and rudder to induce roll, pitch and yaw maneuvers, respectively. These maneuvers are very important: they allow us to "inject" data into the scalar magnetometer that lie in a very restricted frequency range. By utilizing a bandpass filter, we are able to match the Tolles-Lawson coefficients to magnetic variation which lie in that frequency range, extracted by the bandpass filter. This allows us to ultimately understand how only the platform affects the disturbance field. (7) shows the justification for this approach, assuming A is the matrix of the direction cosines, y are scalar magnetic readings, and x represents the Tolles-Lawson coefficients to be determined.

$$
\begin{aligned}
y_{earth} + y_{aircraft} &= Ax \\
bpf(y_{earth} + y_{aircraft}) &= bpf(Ax) \\
bpf(y_{earth}) + bpf(y_{aircraft}) &= bpf(Ax) \\
bpf(y_{aircraft}) &= bpf(Ax)
\end{aligned}
\tag{7}
$$

where the last line follows because of the behavior of the geomagnetic field in the area where calibration is performed. Once the 18 coefficients have been determined this way, they can be used for real-time compensation of the magnetic signal any time. This assumes that the magnetic structure of the aircraft and the location of the magnetometer within haven't changed. Figure 8 shows an example calibration flight pattern.

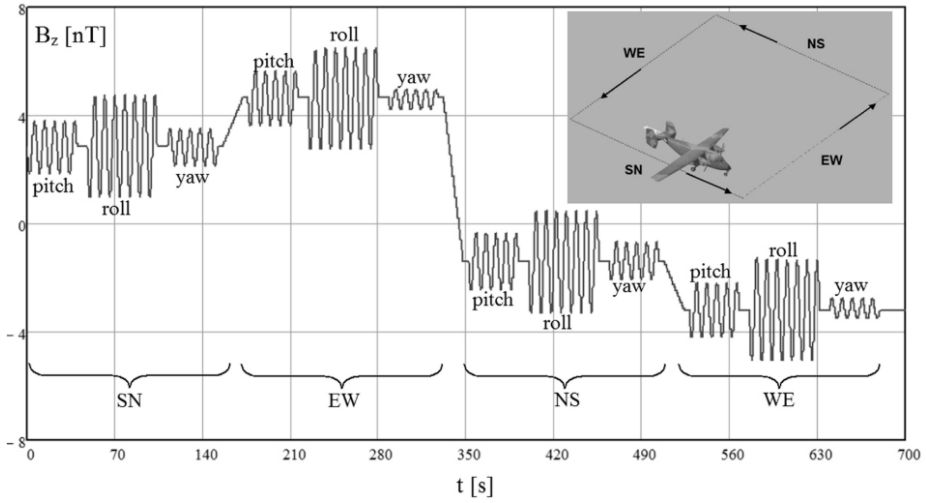

Figure 8: Typical calibration flight pattern with clear heading and banking effects disturbing the magnetometer reading [5]

Tolles-Lawson calibration attempts to remove the variations seen in figure 8, since a calibrated scalar magnetometer should read a signal whose changes are geomagnetic in nature. Once the magnetometer has been calibrated, a "figure-of-merit" (FOM) is flown to demonstrate the effectiveness of the procedure. This consists of flying in a

clover-leaf pattern over the same spot in all four cardinal directions. In addition, the peak-to-peak difference between the maneuver readings is summed for each type of maneuver over each heading. That is, the difference in magnetic reading for $+5^\circ$ and -5° roll is added to that of pitch and yaw for each heading and summed. Successful calibrations can result in figures-of-merit of less than 1 nano-Tesla. Figure 9 shows a typical figure-of-merit flight pattern.

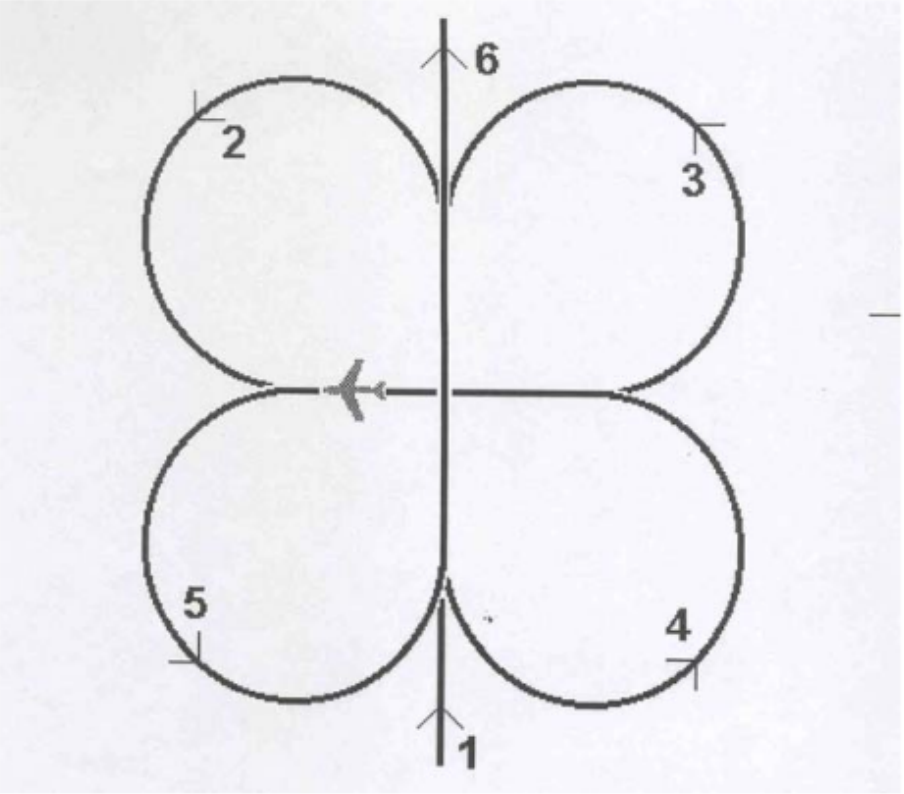

Figure 9: A FOM flight pattern [4]

Currently, for aeromagnetic surveying and, recently, magnetic navigation, the magnetometer is placed as far from the engine and hull of the aircraft as possible. This is because the engine produces a large induced effect and the running engine itself can produce unwanted magnetic noise. Additionally, the control surfaces on the aircraft can be sources of magnetic disturbance. Many modern aircraft are made of a non-magnetic metal like aluminum, but there are many potentially magnetic materials

that go into making a control surface move - servos are one example. The Tolles-Lawson equation in (5) is not equipped to deal with such effects because its only input is the vector magnetometer, and any potential non-linearity will go unnoticed. It is unwise to expect that the vector magnetometer could account for disturbance field sources that aren't completely explained by orientation within the earth's magnetic field. It is proposed that machine learning can beat Tolles-Lawson's shortcomings in these aspects.

2.4 Machine Learning and Neural Networks

Machine learning is a process by which relationships between inputs and outputs can be determined through automated inference and pattern recognition when no clear instructions are given. Given a set of inputs $\boldsymbol{X}$ and corresponding truth outputs $\boldsymbol{Y}$, supervised machine learning aims to find a function which adequately maps $\boldsymbol{X}$ to $\boldsymbol{Y}$. That is, it aims to find f such that $f(\boldsymbol{X}) = \hat{\boldsymbol{Y}} \approx \boldsymbol{Y}$ for all output samples $y \in \boldsymbol{Y}$, where $\boldsymbol{X}, \boldsymbol{Y}$ are input and truth output from the training set; a set of data that the computer knows the correct outputs for. In theory there are many such functions, and the trick is training the computer to find a function that will perform well on data it has not seen. To do this, it ideally learns meaningful relationships on its own. This is different from classical programming, where inputs and rules are fed in and outputs are returned. Here, we seek the model.

There are many examples of machine learning algorithms. Linear and logistic regression are two very simple algorithms for regression and classification. Other examples include k-nearest neighbors, k-means clustering, and decision trees and random forests. One machine learning paradigm which has experienced waves of excitement is that of Artificial Neural Networks (ANNs). In their most basic form, feed-forward

ANNs are neural networks which transform data by passing it through layers of "neurons" and activation functions. Data is taken from one layer, acted upon by weights, passed through a potentially non-linear activation function, and finally sent to the next layer, where the process can be repeated. Bias terms are also typically added before being passed to the activation function. This handling of the data allows neural networks to approximate any continuous function [6]. This is an incredibly useful property of neural networks, but designing the neural network with the proper architecture can often be very difficult. Figure 10 shows an example of an ANN.

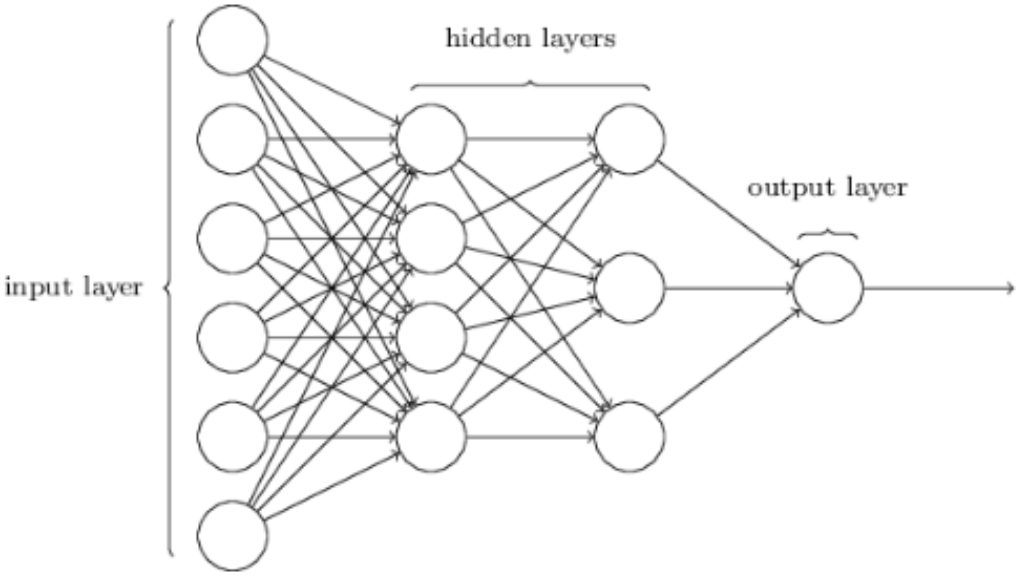

Figure 10: A 3-layer feed-forward ANN [6]

The basic principle behind most ANNs involves reducing a cost function J using a technique known as *gradient descent*. This involves updating the parameters or weights of the network in a way that improves the network's output on the current training sample. The cost function is often taken as the average of all of the *losses* incurred over one *epoch* - one full pass of all M training samples. The loss function simply measures how well the network prediction $\hat{y}$ matches the true training output y on the current training example:

$$L_i = L(y_i, \hat{y}_i) = L(y_i, \boldsymbol{w})$$

$$J = \frac{1}{M} \sum_{i=1}^{M} L_i$$

Next, gradient descent determines how the weights should be updated so as to perform better on the next epoch:

$$w^k := w^k - \eta \frac{\partial J}{\partial w^k} \tag{8}$$

for the k^{th} weight. η is the learning rate which determines how much the parameters adjust after feedback from the training samples. (8) is written most generally as

$$\boldsymbol{w} := \boldsymbol{w} - \eta \nabla J \tag{9}$$

Passing the input to the network and generating an output $\hat{y}$ is known as *forward propagation*. The updating of the weights by calculation of the relevant derivatives is known as *back propagation*. The final component of neural networks that will be discussed is the activation function. The ANN shown in figure 10 depicts the weights with arrows. Each node in the input layer is connected to every node in the next layer, and so there are 24 weights to tune in the first layer along with 4 bias terms. If this was all the neural network consisted of, it wouldn't be able to produce non-linear transformations, since everything up to this point has consisted of linear operations. Adding more layers does not expand the ANN's ability to produce non-linear functions, nor does layer widening. This is the role of the activation function. Activation functions introduce non-linearities when data is passed through each node. Several activation functions are shown in figure 11.

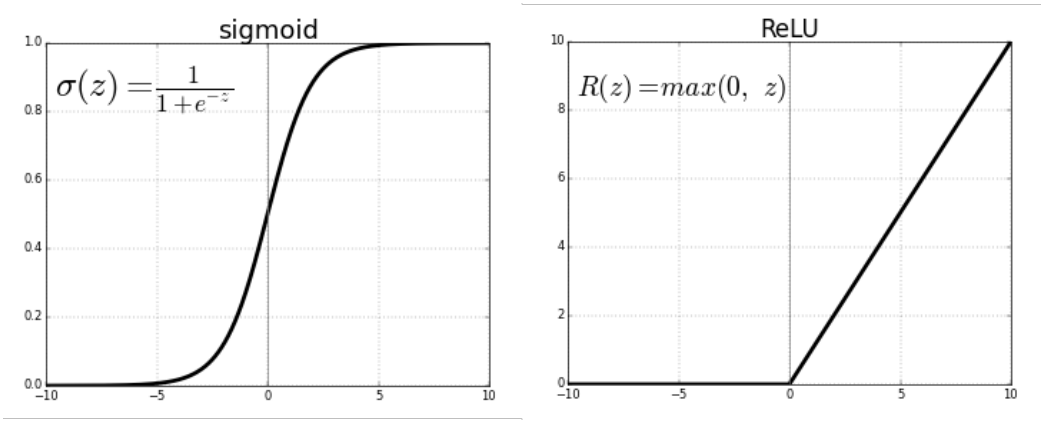

Figure 11: Two commonly used nonlinear activation functions [7]

The application of nonlinear activation functions permits ANNs to learn nearly any relationship between input and output, if a meaningful one exists. The rectified linear unit (ReLU) is a relatively recent addition to the list of popular activation functions, and is the most popular one for deep learning currently.

2.4.1 Recurrent Neural Networks

The ANN shown in figure 10 is a densely connected neural network (DNN). All nodes in a layer are connected to all nodes in the previous layer. This architecture is great at finding relationships between input and output samples that do not correlate in time. For example, a task that requires classifying objects in an image would be better handled with a DNN than a task that requires predicting the next word in a sentence. This is because sequential information is very important in temporal signals. DNNs have no built in memory and so are not able to remember temporal relationships.

Recurrent Neural Networks (RNN) are networks which use outputs at time t and feed them as inputs to time $t + \Delta t$. Many signals whose properties are autocorrelated

or correlate with another signal are best handled by some kind of recurrent neural network. A basic recurrent neural network is shown in figure 12.

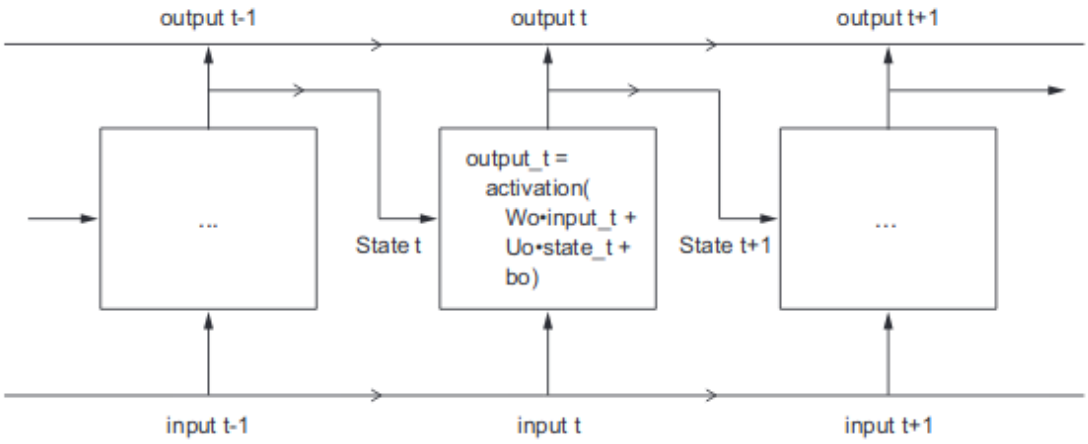

Figure 12: A basic RNN architecture [8]

Figure 12 shows the addition of a hidden state propagating the output of the previous state to the activation of the next state. This allows the passage of sequential information, which is vitally important for understanding relationships between time-dependent inputs and outputs. There are many applications where information from the previous timestep is vital in knowing the next output. RNNs like that of figure 12 aren't typically used in complex time-series prediction, because they tend to suffer from the vanishing-gradient problem, which is encountered when training ANNs via back propagation. Because of the chain-rule, the calculation of the weight adjustment in (9) can result in many multiplications of derivatives, especially in deeper networks. The sigmoid activation shown in figure 11 has derivatives whose values are nearly zero at the ends of the function, and this can lead to a vanishingly small gradient by the time the update reaches the weight. Small weight updates result in very small learning rates and extremely long computation times. Other activation functions can help (ReLU doesn't suffer from small gradients), but other RNN architectures can also be beneficial. Probably the most commonly used RNN architecture is the

Long-Short Term Memory (LSTM) RNN. The architecture of LSTMs allows them to remember important sequences arbitrarily many timesteps apart by introducing a transport track that is immune from the effects of the vanishing gradient problem. Figure 13 shows an LSTM architecture.

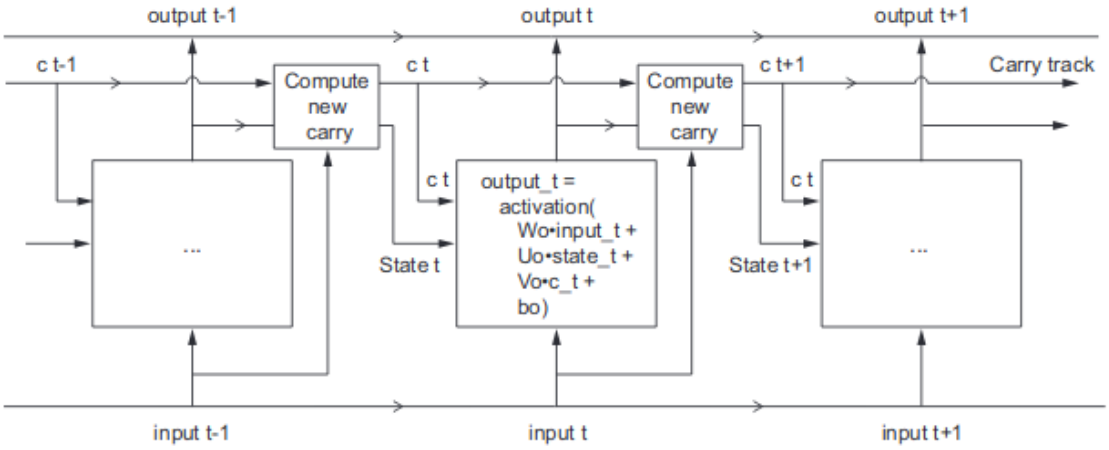

Figure 13: An LSTM architecture [8]

The improvement to the RNN architecture shown in figure 12 is that now the LSTM can learn relationships over many timesteps by determining when the data from the transport track is useful in determining output.

2.4.2 1-Dimensional Convolutional Neural Networks (1DCNNs)

Convolutional neural networks have found great success in machine learning tasks which involve images. Dense networks aren't as useful for identifying an object in an image, for example, because there isn't a functionality built in DNNs that allow for translation invariance. A machine learning model designed to look for cats should be able to find one regardless of its position in the image. Convolutional Neural Networks (CNNs) are excellent at object classification and image recognition. They operate by using a series of *convolutional kernels* and *pooling* layers. When used with

images, convolutional kernels consist of a grid of weights which pass over the image in a lawn-mower pattern. When pixel intensity values match the weights of the kernel, the kernel activates maximally. Pooling is also performed to retain the most salient features and reduce computational complexity and time. Pooling layers consist of an operation which downsamples the output from the convolutional layer, by applying some operation over a grid of specified size. Pooling layers might perform *max pooling*, where the maximum value over the grid is taken, *average pooling*, and others. Since it is a downsampling procedure, it also reduces the number of parameters to be tuned, thus fighting overfitting. Figure 14 illustrates a 2DCNN in action.

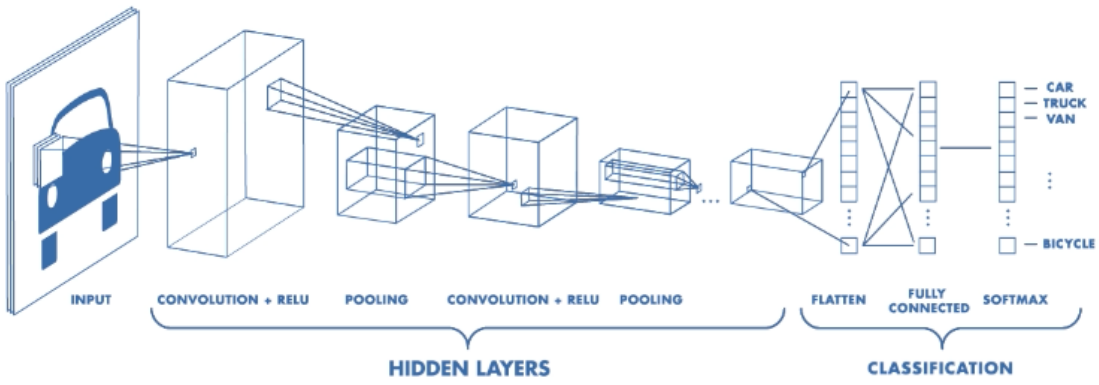

Figure 14: A 2DCNN architecture [9]

1-Dimensional convolutional neural networks work in a similar fashion to 2DCNNs. They are mostly used for time-series data, and the learning process consists of fine-tuning 1-dimensional kernels to activate maximally on salient features. Because the kernels are 1-dimensional, they are far less computationally expensive than 2DCNNs. Pooling can also be performed during training. Figure 15 illustrates the working principle of a 1DCNN.

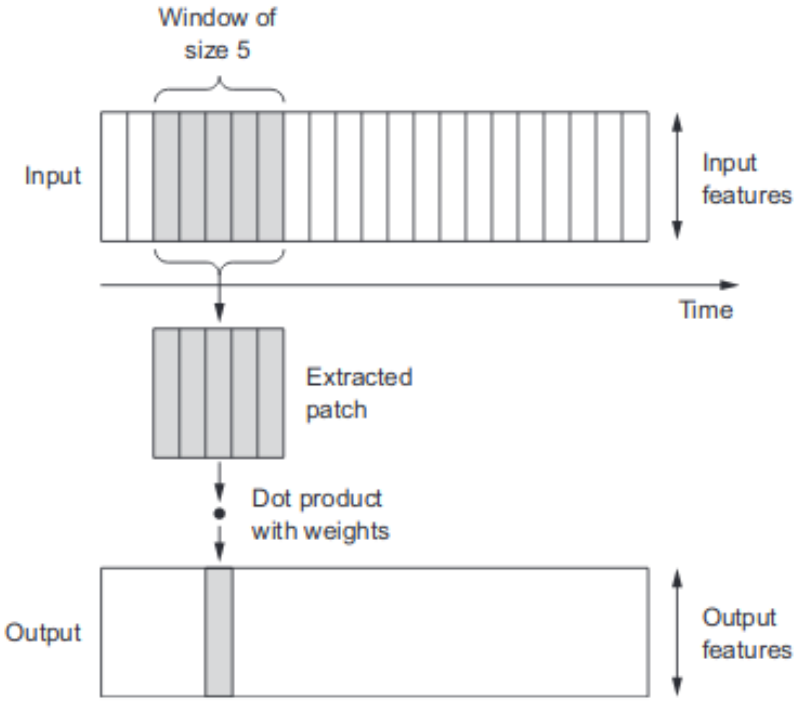

Figure 15: Input features being passed through a 1DCNN layer [8]

1DCNNs, like their 2-D counterpart, are able to extract temporal feature patches and scan entire input sequences for similar features. This makes them potentially very powerful for time-series prediction. More complicated neural network architectures might combined LSTMs and 1DCNNs for improved performance.

2.5 Literature Review

Because navigation via magnetic anomalies provides an essentially jam-proof method of navigation, is completely passive, and is available globally and regardless of time of day and weather conditions, it represents an appealing method of navigation. It therefore is no surprise that it is an active area of research. The most significant contribution to the area came from Canciani [1] in 2016. Canciani proposed it was possible to significantly constrain inertial errors over long-range flights using only

magnetic anomaly measurements. First, Canciani showed that very accurate magnetic anomaly maps constructed and taken over the same area 3 years apart had remained virtually unchanged, with a disagreement standard deviation of 1.55nT after temporal and Tolles-Lawson corrections were implemented. This demonstration was paramount for proving the viability of magnetic navigation. Since map availability and construction is one of magnetic navigation's most significant limitations, showing that these maps are reliable on the order of years was a crucial achievement. These accurate maps were then used during a 1-hour flight over Louisa, Virginia which combined a navigation-grade INS with a Geometrics 823A optically pumped cesium atomic magnetometer. This flight represented the best-case scenario, since the platform and magnetometer position allowed optimal magnetometer calibration and the map quality and flight altitude were nearly ideal. The unaided INS yielded an error of 344m DRMS over the flight, whereas the full MagNav filter solution yielded an error of only 13.1m [1]. This result constitutes the most impressive aeromagnetic navigation solution to date, and proves the concept.

A second, cross-country flight from Louisa, VA to Iowa was flown and utilized the North American Magnetic Anomaly Database (NAMAD). The NAMAD is shown in figure 16

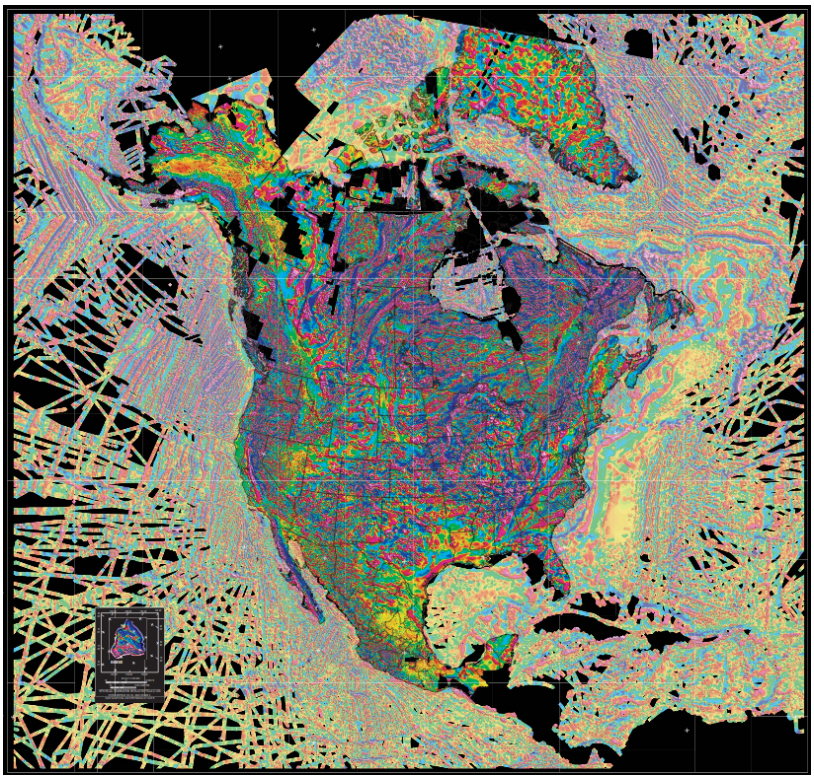

Figure 16: North American Magnetic Anomaly Database (NAMAD) [10]

The NAMAD is a lower-quality, continental scale magnetic anomaly map which was stitched together from many flights and whose GPS-position accuracy is somewhat dubious; certainly of much lower quality than the one collected over Louisa, VA. Because the performance of magnetic anomaly navigation is directly tied to the accuracy of the anomaly measurements and the map quality, it is expected that this flight would be somewhat less impressive. Indeed, the navigation-grade INS received little assistance with MagNav. However, when the INS was corrupted to have quality closer to that of a tactical version, the MagNav solution constrained the error from 34.3km to 3.2km DRMS. When using a simulated map of similar quality to that over Louisa, MagNav reduced the error incurred by the navigation grade INS from 1.5km to 160m DRMS over the same flight. Even with just NAMAD, these results show that a crude anomaly map is helpful for navigation purposes, and that map quality

is a hurdle for magnetic anomaly navigation.

Canciani's results show that MagNav is a worthy addition to traditional, GPS-denied navigation solutions with impressive advantages over other methods. However, it also has significant limitations in its applicability.

2.5.1 Current Limitations of MagNav

Canciani proved that MagNav is capable of giving highly accurate positioning solutions when two criteria were met. Firstly, the availability of a high quality map is necessary for high quality navigation. This is not a simple task, and making many high quality maps is time consuming and costly. However, the map making problem, at least from a technical standpoint, is solved. Making high quality maps has a known solution. Secondly, Canciani showed MagNav under ideal disturbance platform conditions. The calibration and map making procedure were performed by Sander Geophysics, a professional geo-survey company with optimized aircraft. MagNav has not been shown to work on aircraft which do not permit stingers and whose disturbance field is more complicated than what is traditionally handled by the Tolles-Lawson equations and a vector magnetometer. Platform calibration beyond what is handled with a stinger and the Tolles-Lawson equations is an open problem, and one of the last technically and scientifically challenging problems preventing widespread use of MagNav. Obviously, the ultimate goal for fieldable MagNav is a solution or calibration procedure which can be used on any aircraft and in any location on the aircraft, but currently this is not the case. Additionally, many platforms do not permit the use of stingers. An F16, for example, would not accept a stinger. Some platforms allow for the "towing" of a magnetometer either below or behind. Obviously for an F16 this wouldn't work. The current research focus for magnetic navigation is deter-

mining a way to make it platform agnostic, and showing that there are more sensors and relationships to explore and utilize toward this effort is the primary purpose of this thesis.

2.5.2 Improvements to Tolles-Lawson and Machine-Learning Aided Sensor Calibration

Several important improvements to the original Tolles-Lawson equations have been made. Recall that the equation in (5) requires only the vector magnetometer to perform calibration. This means that the calibration procedure is highly dependent on the accuracy of that sensor. Vector magnetometers are much less accurate than their scalar counterparts at measuring scalar values, and reasons for this include bias, scale factor and misalignment errors. Han et. al. combated this problem by introducing a λ term, which captures the relationship between the magnitude of the noisy vector measurement $\boldsymbol{B_v}$ and the scalar measurement B_s [18]. Namely, $\lambda = B_s/||\boldsymbol{B}_v||$. They also introduce a τ_s term which handles errors incurred by the Gaussian wideband noise and the null-shift error. The "modified" disturbance field equation they present is

$$B_{dist,mod} = \lambda \mathbf{A}\mathbf{u} + \lambda^2 \mathbf{u}^T \mathbf{B}\mathbf{u} + \lambda^2 (\mathbf{u}^T)' \mathbf{C}\mathbf{u} + \tau_s \tag{10}$$

They additionally discuss the addition of parameters η_0, η_1, η_2 for modeling a linearly sloped magnetic field at lower altitudes as opposed to a constant field at higher altitudes. This makes the modified Tolles-Lawson model a 21-term model.

Jiang, Guo, and Zhang [19] showed that the 3-axis fluxgate magnetometer outputs can be mapped to the scalar magnetometer outputs during aeromagnetic data collection using both a linear regression and neural network approach. Their work showed

that the neural network can achieve a mean error of 18nT from the scalar magnetometer. This is a promising result because 2 sensors are traditionally required to perform magnetic navigation. There is a desire to reduce the necessary hardware to just a vector magnetometer, since a vector magnetometer provides more information than an equally accurate scalar magnetometer. This would help to alleviate size, weight and power (SWaP) concerns.

Chen et. al. designed a compensation method which employs ridge regression and a second scalar magnetometer to compensate both sensors mounted on a boom on a helicopter [20]. Since the two magnetometers are measuring the same geomagnetic and disturbance fields, there is a spatial correlation between the two. They designed a transfer function between the two signals detected by the respective magnetometers, and used ridge regression on Tolles-Lawson to calibrate both sensors simultaneously. A 31% decrease in disturbance field signal was achieved. This implies that an ensemble of scalar magnetometers (and probably also fluxgate magnetometers) can result in improved compensation.

Williams' neural network contribution represents one of the most important contributions to aeromagnetic compensation using alternative methods. In his 1993 paper [11], Williams proposed that a neural model of the total measured field can be calculated by assuming it is well-modeled as a linear combination of the local, diurnal, and interference effects. He built a 2 layer neural network (not including the input layer) which was composed of 3 heads. Each head modeled each component of the assumed measurement listed previously. Figure 17 shows the architecture implemented by Williams.

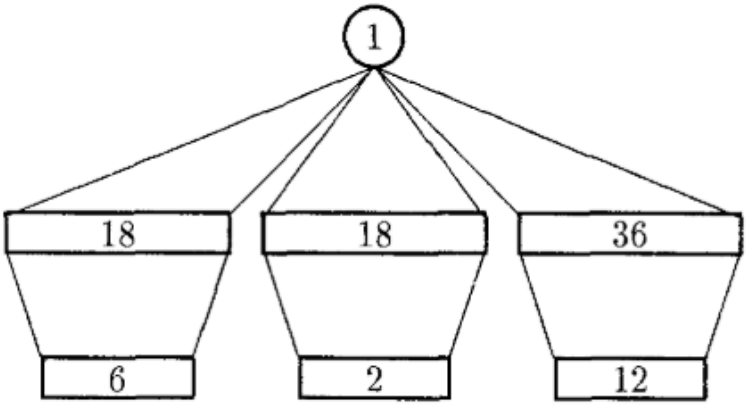

Figure 17: Neural network architecture implemented by Williams. From left to right: the local, diurnal, and interference heads [11]

The inputs to the local field head consisted of 6 position measurements and position cross terms. The diurnal head was given time t and t^2. The interference head was given the sine and cosine of roll, pitch, and heading values and their derivatives. Each head's individual predictions were summed with a linear activation function, and this value was assumed to be the total magnetic field reading. Williams successfully showed that the three heads successfully calculated their assigned components on the training data, and the performance was similar to that of normal calibration. This shows that, at least from a flexibility standpoint, neural networks are well conditioned to handle more inputs than traditional Tolles-Lawson and are capable of more detailed modelling. Even more impressive is the fact that this research was conducted in the early 90s. Much more advanced architectures have been synthesized since, and Williams' neural network contained no time-sensitive layers, like LSTMs.

Sensor calibration is an extremely common and necessary practice for many sensors, and its requirement isn't limited to magnetic sensors. Advanced machine learning techniques have been shown to yield improvement over more simple models for sensor

calibration. For example, in their 2017 report [21], Yamamoto et. al. demonstrated the calibration of a low cost and quality air temperature sensor using a higher quality "AMeDAS" sensor. They considered both a linear regressive model and a 2-layer feed-forward ANN. Their results showed that the ANN was able to better predict the AMeDAS temperature outputs than the low quality sensor when both AMeDAS and the low quality sensor were in roughly the same area, and outperformed the linear regression approach in each case.

It is clear that much work has been done in the area of aeromagnetic calibration, but most of it has concerned itself with small absolute improvements to the Tolles-Lawson model assuming optimal or near optimal magnetometer location. Additionally, there are no examples of experiments, in this background review or in the broader literature, which tax Tolles-Lawson in any meaningful way. Moreover, it is not clear that any substantive research exists which identifies and remedies significant corrupting sources not readily handled by the vector magnetometer. To illustrate the point, [20] showed a 31% decrease in disturbance field signal, but the calibrated disturbance field strength was 0.065nT after the application of the Tolles-Lawson equations and with the use of a stinger-equipped helicopter. Even before Tolles-Lawson compensation, their level of disturbance was low enough for effective magnetic navigation anyway. Papers and reports which deal with aeromagnetic calibration for a magnetometer inside the fuselage of the aircraft and subject to significant disturbance are unknown to the author. However, it is clear that neural networks and machine learning constitute powerful tools for sensor calibration, and so it is believed that they can be of great utility in this area.

III. Methodology

This section focuses on the experimental design used to demonstrate improved compensation performance over traditional Tolles-Lawson approaches using machine learning and supplementary sensors. First, we discuss the aircraft used for the experiment and the expected noises that will corrupt the magnetometer. We proceed by showing the hardware that was used to provide supplementary information to enhance calibration. We then describe the calibration procedure which took place and justify the different phases of

3.1 Magnetometers and Supplementary Sensors

Magnetic fields are vector fields. Because of this, at any point where the magnetic field is being sampled, it has both a direction and magnitude. A scalar magnetometer is able to measure the magnitude of the magnetic field vector only. In theory, if a perfect scalar magnetometer is rotated perfectly about any axis and is without the influence of a disturbance field, it will maintain a constant reading. These sensors are also the sensors that are used to measure the magnetic anomaly during flight, and the sensor whose readings must be calibrated to ensure aircraft induced noise is minimized. Vector magnetometers measure the components of the magnetic field along each of their 3 axes, and the signal measured by each channel will change if the orientation of the vector magnetometer is altered. Vector magnetometers can also measure the magnitude of the magnetic vector they are measuring by taking the root-sum-square of the channel outputs. It is therefore natural to wonder why vector magnetometers aren't the only sensor used for magnetic navigation if they are capable of supplying scalar information as well. Vector magnetometers are subject to many more error sources and are therefore significantly less accurate than scalar sen-

sors at measuring magnetic intensity. For example, misalignments in the 3 axes can result in significant errors. Therefore, scalar magnetometers are the sensor of choice for magnetic navigation. Of course, platform calibration requires a sense of direction with respect to the geomagnetic vector, and so vector magnetometer readings have traditionally been the only necessary input to Tolles-Lawson.

3.1.1 Micro-Fabricated Atomic Magnetometer (MFAM)

The scalar magnetometer chosen for this experiment was the Geometrics micro-fabricated atomic magnetometer development kit (MFAM). The MFAM consists of 2 laser pumped cesium atomic magnetometer scalar sensor heads attached to the main board via flex cables. The MFAM also comes equipped with a Bosch BMI160 gyroscope and accelerometer and an HMC5983 vector magnetometer. All sensors excluding the scalar magnetometers are MEMS grade. The MFAM development kit is shown in figure 18.

Figure 18: MFAM development kit

The MFAM offers very high sensitivity - $1pT/\sqrt{Hz}$ - and is therefore very useful for aeromagnetic navigation, where accurate measurements of small magnetic changes

are necessary. 2 separate magnetometer sensing heads are attached by long cable circuits, enabling placement far from the native metal shielding. Each head offers the same measurement accuracy and sensitivity, and are both affected by one polar dead zone. A dead zone results in a dubious reading (or none at all) and is a result of the total magnetic vector being too aligned with either sensor's optical axis. If the geomagnetic vector (the primary contributor to the sensed magnetic vector) is aligned to within $\pm 35^{\circ}$ of the optical axis, the sensor enters the dead zone. See Figure 19.

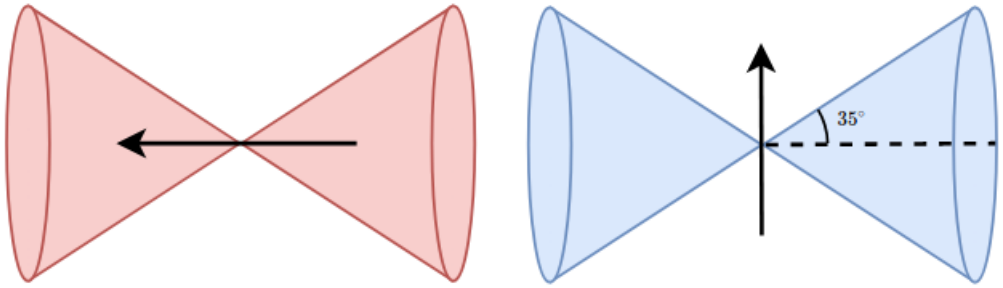

Figure 19: Polar dead zones for the MFAM magnetometer. Left: No signal generated due to magnetic field intersecting dead zone. Right: No dead zone intersection.

Two sensors heads allows for strategic relative orientation of the sensors. Placing both sensors heads next to and parallel to each other can yield heading error cancellation. Additionally, placing the sensors orthogonally to each other can eliminate signal dropout due to dead zones. If the magnetic vector falls into the dead zone of one sensor, it will be outside of the dead zone of the other, given proper relative orientation. Figure 20 shows an ideal example of the proper orientation to eliminate simultaneous dead zone dropout of the sensors.

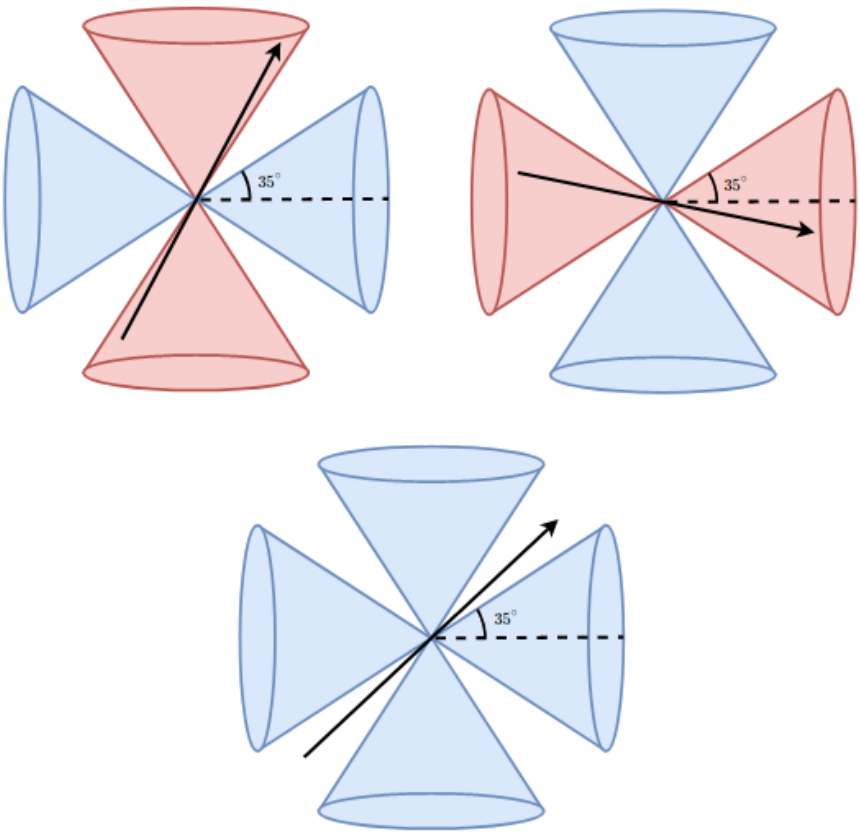

Figure 20: An example of orthogonal sensor placement to eliminate simultaneous dead zone dropout. Top left: dead zone incurred in vertical sensor. Top Right: dead zone incurred in horizontal sensor. Bottom: No dead zone intersection.

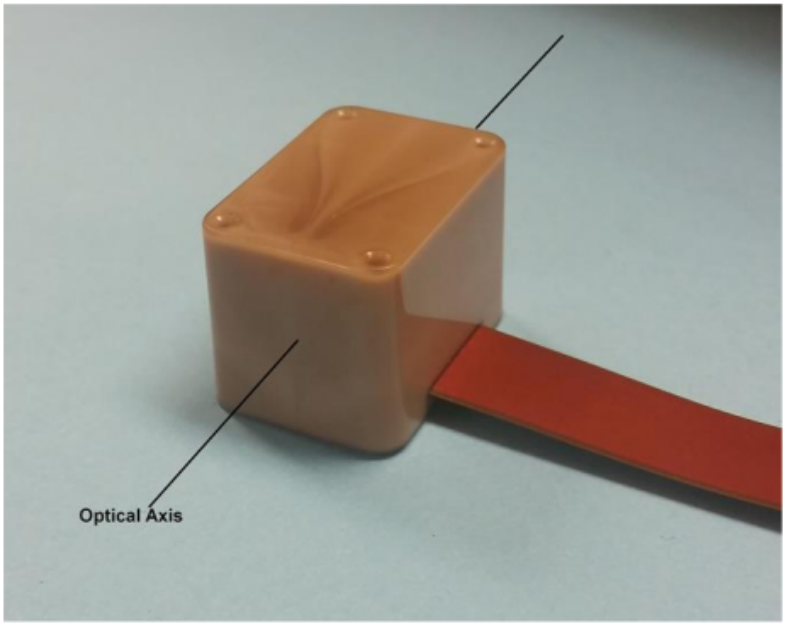

Figure 21: MFAM sensor head optical axis [12]

Prolonged loss of magnetic signal can be detrimental to magnetic navigation. A magnetometer setup like that shown in figure 20 will ensure that a magnetic signal is always captured by at least one sensor.

3.1.2 Additional Hardware and Aircraft

The MFAM contains two highly accurate scalar atomic magnetometers, and one MEMS grade accelerometer, gyroscope and vector magnetometer. Additionally, more accurate sensors offer higher quality inputs to the chosen machine learning models. For this reason, the VN100 IMU and HMR2300 vector magnetometer were also included on-board the aircraft. All sensors except those housed within the MFAM module have access to GPS time via an on-board precision time protocol (PTP) server controlled by a Linux computer. The MFAM contains its own GPS receiver which allows for microsecond level accuracy of all housed sensors.

The aircraft used for this experiment was an Aeroworks 40% scale carbon cub. This aircraft contains a 2-piston, 120cc, 2-stroke gas engine and is composed of a lightweight wooden frame and a thin, protective plastic coating. Figure 22 shows the aircraft in question.

Figure 22: 40% scale Aeroworks Carbon Cub

The Cub also contained a PixHawk 2 autopilot. This allows for autonomous flight patterns via set waypoints. Additionally, the PixHawk 2 supports remote controlled flight. The ability to control the aircraft using custom dynamics has obvious importance for magnetic platform calibration, which necessitates the traditional Tolles-Lawson calibration flight as well as more advanced disturbance field signal creation for potential machine learning model development. Finally, the PixHawk flight controller records the voltage pulse-width modulation (PWM) required to induce and maintain control surface deflections for roll, pitch, and yaw dynamics. It also records throttle percentage and flap inputs.

The sensors and hardware on-board the aircraft are summarized below:

- 2 laser pumped cesium scalar atomic magnetometers (MFAM)
- 1 BMI160 MEMS accelerometer/gyroscope (MFAM)
- 1 HMC5983 MEMS vector magnetometer (MFAM)
- 1 AdaFruit GPS receiver (MFAM)
- 1 VN100 IMU
- 1 HMR2300 vector magnetometer
- 1 PixHawk 2 flight controller and RC receiver
- 1 Piksi global navigation satellite system (GNSS) module

A summary of the specifications for all sensors is shown in Table 1:

Table 1: Sensor Payload Specs

Sensor	Sensitivity	Noise Density	Bandwidth	Sampling Rate
Scalar Mags	$1pT/\sqrt{Hz}$	$5pT/\sqrt{Hz}$	$400Hz$	$1kHz$
BMI160 (A)	$60\mu g$	$180\mu g/\sqrt{Hz}$	$80Hz$	$200Hz$
BMI160 (G)	$0.06°/s$	$0.007°/\sqrt{Hz}$	$74.6Hz$	$200Hz$
HMC5983	$125nT$	-	-	$100Hz$
VN100 (A)	-	$140\mu g/\sqrt{Hz}$	256Hz	$400Hz$
VN100 (G)	-	$\frac{0.0035°/s}{\sqrt{Hz}}$	$256Hz$	$400Hz$
HMR2300	-	-	$85Hz$	$100Hz$

A summary of aircraft specifications for the Aeroworks 40% scale Carbon Cub is shown in Table 2:

Table 2: Aeroworks Aircraft Specs

Specs for Aeroworks Scale Carbon Cub	
Engine	120cc gas
RPM	1300 - 6900
Wing Span	4.27m
Fuselage Length	2.79m
Weight	17.69kg

3.1.3 Expected Magnetic Noise Sources from a Non-Ideal Aircraft

A magnetometer located very near the fuselage of, say, an F16 will very likely experience magnetic disturbance from a wide range of sources. The following list summarizes some of the sources that might have a detrimental impact on the scalar magnetometer in this and many other types of platform.

- Permanent magnetic sources
- Induced magnetic sources
- Eddy currents
- Control surfaces
- High frequency magnetic noise from the running engine
- Periods signals, such as lights
- Electrical impulses
- Fuel pumps

- Temperature Effects
- DC shifts from electrical engagement/disengagement

Permanent, induced, and eddy current sources are already modelled by Tolles-Lawson. Control surfaces will have an effect on a nearby magnetometer if either they, the servos that go into moving them, or any attached metal objects are magnetic or magnetically susceptible. The engine, aside from the induced field handled by Tolles-Lawson, can create high frequency noise as well, either in the form of vibration in the aircraft or from the rapid oscillations of the pistons themselves. Electrical impulses will create momentary magnetic pulses. Any magnetic action as a result of fuel pumps in the aircraft may have an effect on the magnetometer. Finally, variations due to temperature effects and DC shifts due to electrical equipment must also be taken into consideration. It is important to remember that this isn't even a comprehensive list of all possible corrupting sources on a less-ideal aircraft. It is therefore very important to have sensors and other technologies on the aircraft which are able to account for these new sources of magnetic disturbance. Clearly Tolles-Lawson will be unable to account for any of these sources - except of course the first 3 - since it only considers the vector magnetometer.

3.1.4 MFAM Noise Analysis

All sensors are subject to noise. There could be many sources for sensor noise, such as noise generated within the electronics, imperfect electronic operation, heating induced noise, and others. Regardless of its origin, noise is almost never desirable. Because the scalar and vector magnetometers are the most important sensors for magnetic navigation, an understanding of their noise sources is paramount.

Electronic power grids operate at 60Hz. Often called "power line noise", it is an ever-present, sinusoidal signature that can be found in raw electronic and magnetic sensor data streams while in the presence of energized, man-made sources on the ground. These noises are byproducts of the alternating current power sources that power various devices. Magnetometers are particularly susceptible to power grid noise, and so measures to eliminate that noise must be made. Additionally, noise from other electronics is also a concern. Because the magnetic sensors on the MFAM are connected by long ribbon cables, there is some flexibility for decreasing the noise from the MFAM board and other sensors, if any is present.

To examine the noise corrupting the magnetometers, the MFAM was taken to a location in WPAFB where it was expected that much of the electronic noise would be minimized - far from buildings, power lines, and other energized devices. The sensor heads were placed orthogonally to each other, and a power spectrum was made for each sensor head when they were as far as possible from the development kit, seen in Figure 23.

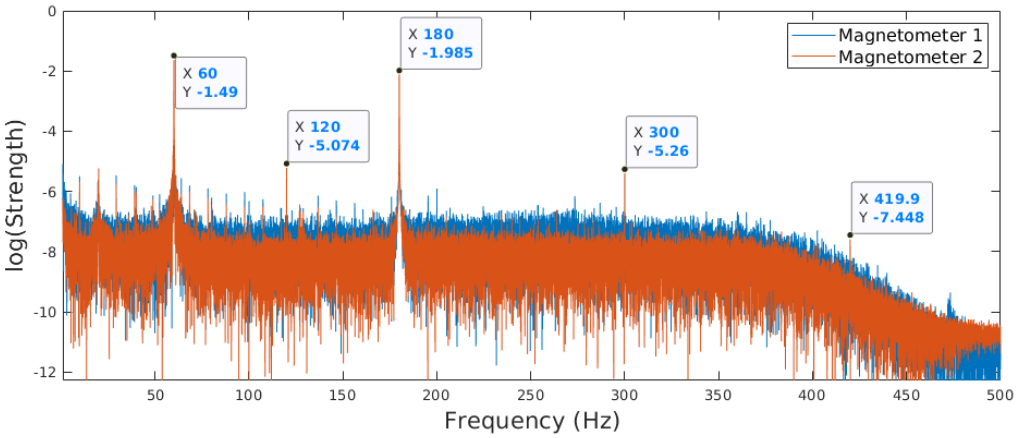

Figure 23: Power Spectrum of MFAM sensor

There is a very strong 60Hz component and other harmonics present in the signal

when the sensor head is as far as possible from the development kit, even when far from any obvious AC power sources. This implies that it will be hard to escape power grid noise for any location on the ground. This leaves two options: the power grid harmonics can be filtered out, leaving only those frequencies that are part of unrelated signals, or the aircraft can escape much of the noise by going airborne. Of course, it is likely that filtering will still be necessary to completely remove power grid noise.

3.2 Experimental Design

Tolles-Lawson is only able to account for disturbance field sources that are directly accounted for by the vector magnetometer, as shown in equation (5). However, there are many potential sources of noise that remain unaccounted for. An on-board magnetometer ought to be indifferent to its location within the aircraft as it relates to measuring the geomagnetic vector, as well as any happenings on or off-board which impact that measurement in the absence of a compensation scheme. For example, if the magnetometer is placed near a large metal control surface (or some magnetic material on the control surface) or a servo which controls the deflection of any magnetic component, Tolles-Lawson will be unable to account for the resulting disturbance. As another example, the oscillations of magnetic engine pistons are also not taken into consideration. In extreme cases, a highly magnetic platform may evoke serious, high-frequency periodic magnetic noise simply due to the shaking of the magnetic fuselage near the engine. Other examples, like electrical impulses resulting from pilot-initiated commands could manifest themselves as magnetic impulses in the magnetometer reading. Even rarer phenomena like nearby lightning strikes present challenges to a reliable compensation system in inclement weather. Many of these noise sources can be tracked by supplementary sensors whose readings can be used as inputs to a machine learning model. Additionally, some of these noise sources lend

themselves to some kind of filtering, such as low-pass filtering, and some do not. Even so, a compensation scheme which can eliminate these sources without the need for digital filtering may be preferable in some cases.

The purpose of this experiment was to develop a machine learning model that can out perform Tolles-Lawson. Specifically, Tolles-Lawson is known to be very effective when the disturbance field noise is relatively minor and caused by magnetic material on-board and by variations in relative orientation of the aircraft in the earth field, like in figure 7. Since the only inputs to Tolles-Lawson are vector magnetometer inputs, sources that are independent of aircraft attitude cannot be modelled. This is problematic in the case of on-board, dynamic disturbance sources. This experiment will evaluate the possibility of modeling control surface deflections as well as engine induced noise. Tolles-Lawson is incapable of modeling either, and a comparison between Tolles-Lawson and the learned approach will be shown.

3.2.1 Control Surface Disturbance Analysis

The frame and control surfaces of the Aeroworks aircraft shown in figure 22 are composed of balsa wood. They therefore have very little magnetic signature of their own, even when close to the magnetometer. Figure 24 shows the result of the MFAM placed near the tail of the aircraft while the rudder was actuated, after the signal was low-pass filtered to eliminate power grid noise. Figure 25 illustrates the setup.

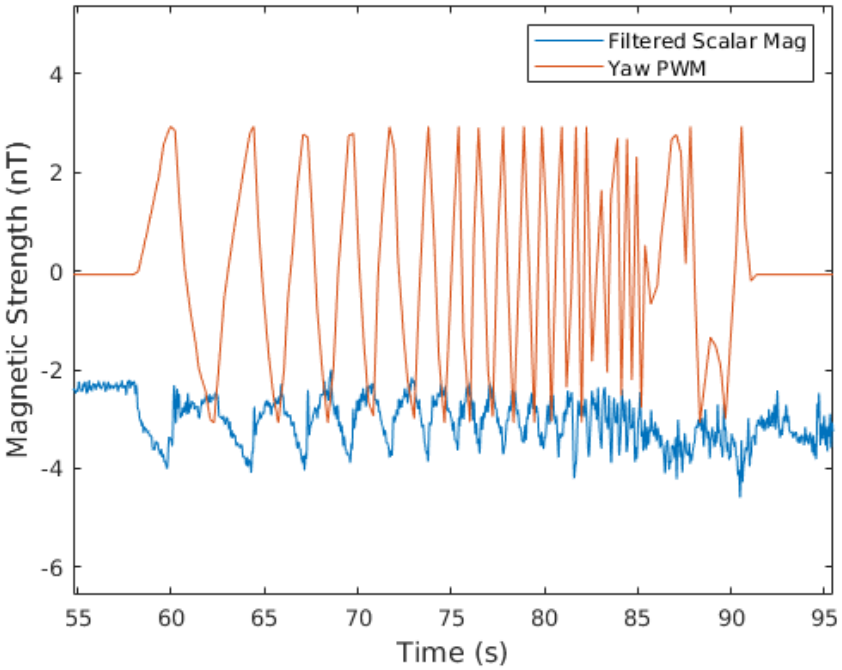

Figure 24: Magnetic interference due to accelerating rudder deflections

Figure 25: MFAM near the tail of the Aeroworks aircraft

Figure 24 shows a clear relationship between yaw PWM and magnetic signal generation, but the peak-to-peak signal is on the order of only a few nano-Tesla. Similar experiments were performed for other control surfaces by placing the magnetometer closer to them, with similar disturbance strengths. To boost the signal coming from the control surfaces, neodymium magnets were attached to each control surface. Neodymium magnets are permanent magnets that provide a strong magnetic field. This is desirable, because even though some aircraft, even more complex military UAVs, are indeed made of a non-magnetic material like wood, many are not. There is good reason to believe that control surfaces from many platforms would create disturbance fields that are disastrous from a magnetic navigation point of view, at least when tackled only with Tolles-Lawson. By placing neodymium magnets on all control surfaces, it was possible to create strong disturbance fields from each surface, even though detecting every disturbance field from each control surface might not happen on an actual flight. More realistically, the magnetometer may sense only a couple of nearby control surfaces. But presumably if the compensation scheme can model disturbances coming from every control surface, doing so on another aircraft where only a few are involved should be trivial. The magnets used were 1/2"x1/8". Figure 26 shows the type of magnet used in this experiment.

Figure 26: 6 Neodymium magnets

A series of tests were performed to determine the proper number of magnets to add to each control surface to bring the resulting disturbance field to a realistic strength. With the MFAM sensors placed in the middle of the fuselage, it was determined that 6 magnets attached to every control surface except the flaps yielded significant, but not overwhelming, control surface disturbances. Since the flaps were the closest control surface to the magnetometers, only 3 neodymium magnets were added, for a total of 36 magnets. Because the magnets were so strong, they stuck to each other through the width of the control surfaces. Even still, tape was used to keep them in place.

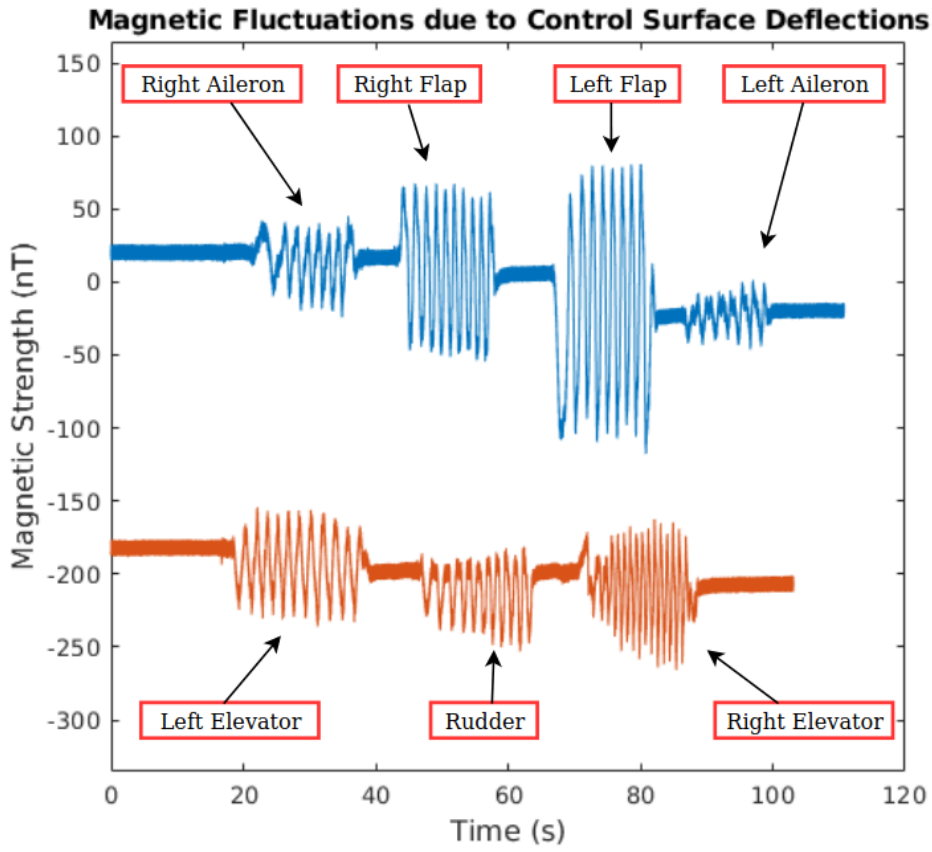

Figure 27: Magnetic disturbance from neodymium magnets on control surfaces

Figure 27 shows the deflection of these control surfaces. The peak-to-peak amplitudes for similar surfaces are dissimilar because the actuation was done by hand. Additionally, the disturbance breaks through the noise floor. The smallest disturbance - from the left aileron - is still around 50nT peak-to-peak, while the flaps generate hundreds of nano-Tesla. The roughly 200nT offset between the front and back control surface data is for ease of display. These noise strengths are much more reminiscent of a field which is a) more plausible on an actual aircraft with the magnetometer installed inside and b) detrimental to airborne navigation via magnetic anomalies.

This experiment also concerned itself with attempting to calibrate magnetic noise

coming from the engine. There are multiple sources of noise from a running engine. The most common relates to the induced field. As the aircraft rotates within the main earth field, potentially very large DC shifts can occur from the large metal engine, depending on the distance to the magnetometer. Tolles-Lawson attempts to correct for this, and can do so well when the induced noise isn't too large. However, Tolles-Lawson cannot alleviate the periodic noise coming from the running engine. Specifically, either the engine shaking the magnetic material of the aircraft within close proximity, or the actual pistons moving up and down at high frequency.

In order to evaluate the possibility of creating an improved platform compensation model, it was necessary to choose a location to perform the experiment which was conducive to detecting magnetic changes coming solely from the aircraft. Any location close to buildings, power lines or any other energized infrastructure would elicit a harmful 60Hz signal. The runway on WPAFB was determined to be a prime location for the tests, as it was far from any signs of electrical power. The nearest buildings were hundreds of meters away. Figure 28 shows the location of the ground test used to gather data for the machine learning model, circled in red.

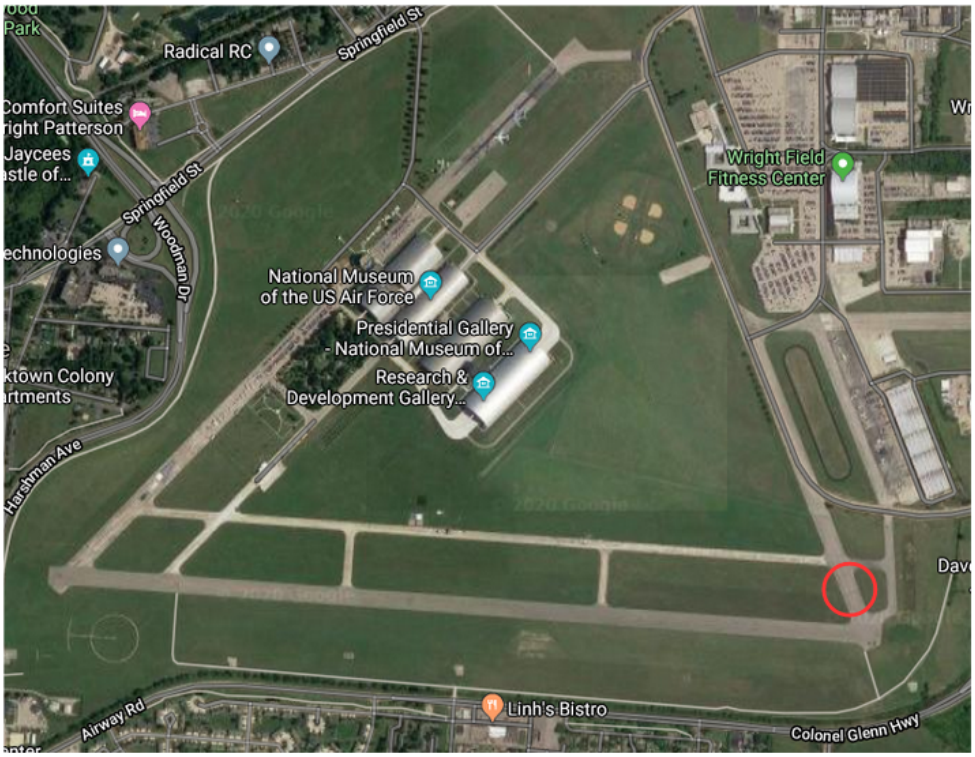

Figure 28: Location of the ground test in WPAFB

The test consisted of 4 phases along 4 headings: southwest, southeast, northwest, and northeast. Each heading lasted 20 minutes and each phase lasted 5 minutes. Each heading consisted of the following phases, in order:

1. Regular flight: simulated take off, normal flight, landing
2. Stationary
3. Realistic control surface maneuvers (engine off)
4. Tolles-Lawson maneuvers by hand

Importantly, this was a ground test. The aircraft did not at any time become airborne. Again, each of the four phases were performed at different headings. This allows for the investigation of orientation on the permanent and induced fields. Each

phase along each of the headings had a very particular purpose. Phase 1 simulated real flight. Everything about phase 1 was reminiscent of what would occur with the air. The engine was running and the aircraft control surfaces were actuated as they would be during regular flight. Phase 2 allows us to capture heading induced effects. Because the DC level of the magnetic field will change as a function of heading, incorporating a large amount of information at different headings allows us to gain substantial information of heading-related effects on the disturbance field. The third phase was identical to the first one, except the engine was not running. This allows for a clean collect of just the control surfaces. The last interval at each heading consisted of roll, pitch, and yaw movements by hand to simulate the traditional Tolles-Lawson compensation procedure, wherein the platform undergoes about 10 °of these movements. The control surfaces were also moving during this time. The phase served two purposed. Firstly, it served as a baseline for the machine learning model to beat. The importance of an alternative model will be determined based on how well Tolles-Lawson performs. Secondly, it allowed for the collection of more orientation related and control surface data. Finally, the last phase of the last heading, southwest, also included 2 clockwise and counterclockwise turns, as a means to feed continuous heading information into the machine learning algorithm.

Figure 29: Roll, pitch, and yaw maneuvers being implemented by hand, similar to a calibration flight

3.3 Conclusion

This section detailed the MFAM - the magnetometer used to detect changes in magnetic fields - and the other supplementary sensors in the MFAM. This included a detailed description of what magnetometers measure and the dead zones affecting them. We also discussed the other sensors on board the aircraft that were used as input features into a deep learning model to surpass Tolles-Lawson. Additionally, we discussed the primary noise source affecting most magnetometers on the ground: power grid noise, and the steps taken to reduce that noise during the test. We also showed the steps taken to bolster the magnetic signature from each control surface in order to make them discernible through the noise floor provide more realistic magnetic interference - accomplished through the use of neodymium magnets. Lastly, the performed experiment was detailed and the location of the experiment was shown.

IV. Results and Analysis

4.1 Introduction

This chapter will discuss the procedure that was used to ultimately beat out Tolles-Lawson by considering additional inputs, such as control surface maneuvers. We will begin by showing the data that was collected from the experiment and the different procedures taken to eliminate the disturbance field created from multiple sources. This will include a comparison of different machine learning approaches as well as a comparison to the state of the art compensation procedure, Tolles-Lawson.

Figure 30 shows the magnetic field generated during all 4 phases at each heading.

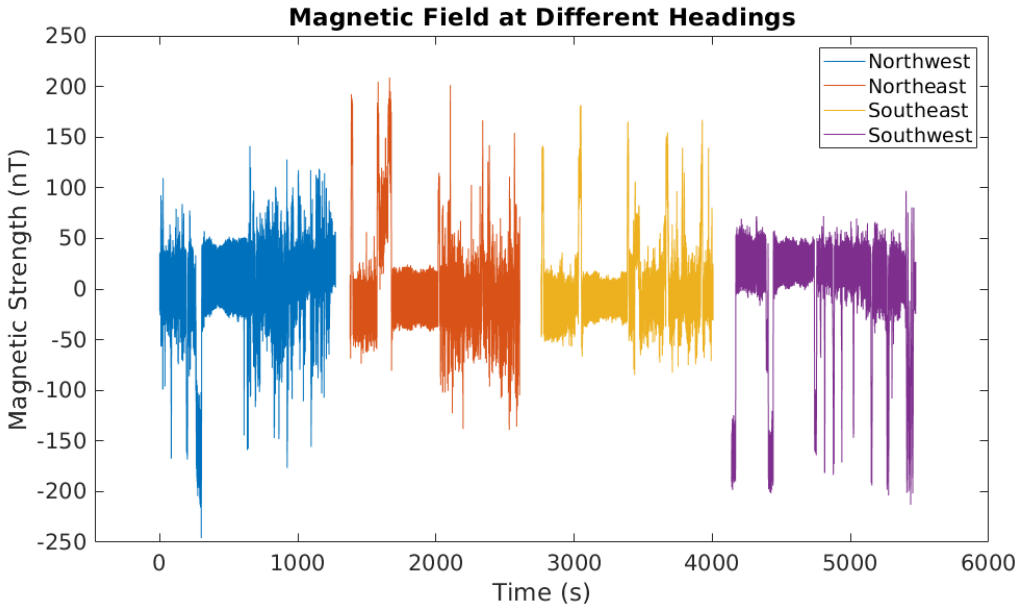

Figure 30: Magnetic data to be calibrated

The data shown was collected by only one of the two scalar magnetometers. The other magnetometer fell into a dead zone during all but the southeast heading, illustrating the importance of having 2 sensor heads. As long as one sensor remains functional,

aircraft compensation can be performed.

Still present in the data shown in Figure 30 was a strong 60Hz signal. Because neither control surface maneuvers nor magnetic anomalies operate at this frequency, a 55Hz low-pass filter was used to clean the data. Figure 31 shows the filtered data. Because the MFAM collects data at a very fast 1kHz, the data shown in Figure 31 was decimated to 10Hz to reduce unnecessary computational overhead during training.

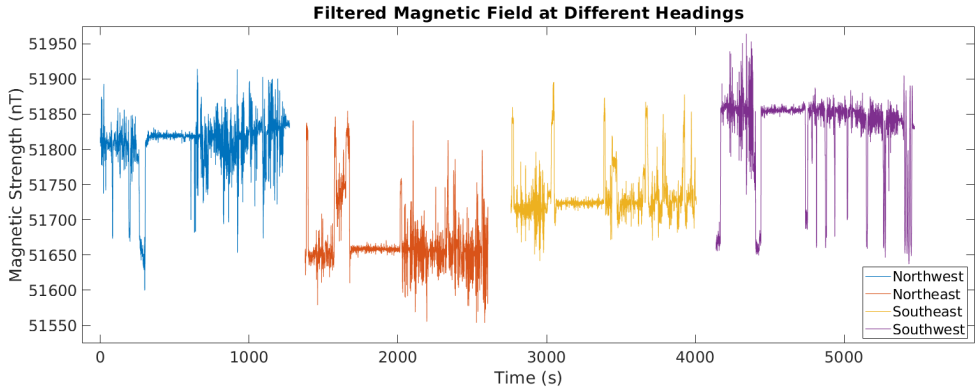

Figure 31: Filtered magnetic data

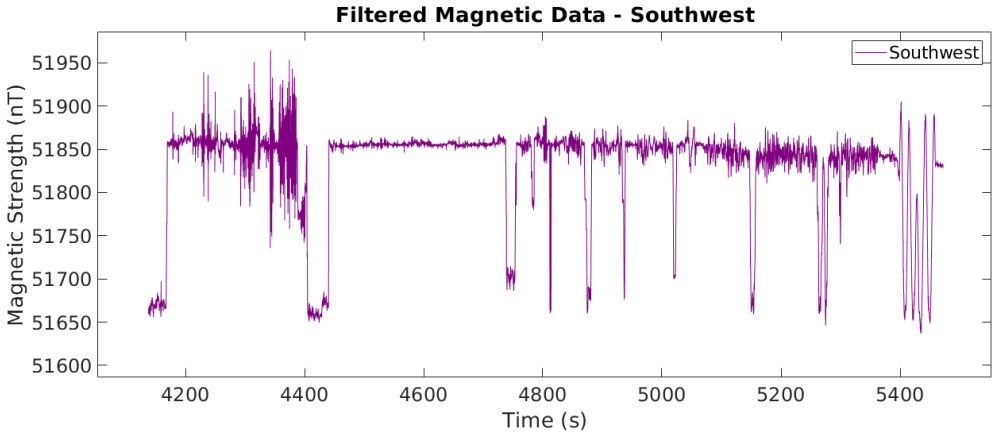

Figure 32: Filtered magnetic data - Southwest heading

Figure 32 shows the different phases of the southwest heading in closer detail. The 4 phases are clearly seen: phase 1 ends at about 4450 sec, phase 2 ends around 4750 sec, phase 3 ends around 5100 sec, and phase 4 ends around 5500 sec. Also clearly visible is the sinusoidal effect resulting from the clockwise and counterclockwise turns at the end. This behavior demonstrates the permanent and induced effects of the platform disturbance field in action. As the aircraft rotates in the main earth field, the permanent and induced magnetic fields on the aircraft interact differently with the geomagnetic vector, which is stationary. This results in the sinusoidal behavior. Also clearly seen are the effects the control surfaces have. The noise resulting from control surface actuation can span hundreds of nano-Tesla.

The primary interest for this test concerned the modelling of control surface inputs and mapping them to disturbance field outputs. Because the aircraft is stationary and the resultant data was filtered, any deviation from a constant value at any heading is the result of platform interference or instrument noise.

To evaluate the need for a more advanced machine learning model, Tolles-Lawson was implemented over phases 3 and 4 at every heading, since the movements performed during these phases are reminiscent of what would happen in the air. To reiterate, Tolles-Lawson is an 18 coefficient linear model which utilizes only vector magnetometer readings as inputs, specifically, the direction cosines. In least-squares linear regression, the aim is to create a coefficient vector $\vec{\beta}$ which minimizes $\sum_{i=1}^{M}(\vec{\beta}x_i - y_i)^2$ over all M training inputs $\mathbf{x}$ and observations $\mathbf{y}$. However, Tolles-Lawson is known to suffer from multicollinearity under many circumstances, wherein some of the predictors are highly correlated with other predictors. To combat this, a regularized form of linear regression known as ridge regression was implemented. Ridge regression is

similar to linear regression, except now $\vec{\beta}$ is the coefficient vector determined by

$$\vec{\beta}_{RR} = argmin \sum_{i=1}^{M} (x_i \vec{\beta} - y_i)^2 + \alpha ||\vec{\beta}||^2 \tag{11}$$

the closed form solution of which is found as

$$\vec{\beta}_{RR} = (\mathbf{X}^T \mathbf{X} + \alpha \mathbf{I})^{-1} \mathbf{X}^T \mathbf{Y} \tag{12}$$

Ridge regression accomplishes several tasks that normal linear regression does not. Equation 12 shows the addition of the ridge parameter α along the diagonal of the previous term. This helps ensure that all columns are linearly independent and that the inverse is well defined. Secondly, the penalty term shown at the end of equation 11 serves to force all non-essential weights toward 0, thus pruning the model. Only those weights which sufficiently predict the label y are allowed to survive.

In order to pick the optimal ridge parameter α, the data at each heading was split into three sections. Recall figure 31 shows the filtered magnetic data, and figure 32 shows the detail of the southwestern heading. First, only phases 3 and 4 of each heading were considered for analysis here. Because phase 4 consisted of Tolles-Lawson maneuvers as well as control surface movements, the entirety of it was chosen as the training set - the set the machine learning model will learn from. The first half of phase 3 was chosen as the test set - the set the machine learning model will ultimately be tested on to see how well it has learned to predict the disturbance field response. The ridge parameter α is a hyperparameter which can be tuned to any value. Optimality of this value was determined using the validation set, chosen to be the second half of the phase 3 data. For each value of alpha the model was retrained using the training set. The chosen alpha will be the one which performs best on the validation set. Once the

optimal value for alpha was picked, the model was sent to the test set. Note here that "optimal" refers only to that alpha which performed best on the validation set over all considered alphas. It is impossible to know a-priori which alpha is best because a) there are infinitely many alphas and b) the validation set does not concern itself with all possible input-output ranges; the chosen alpha could've gotten "lucky".

Figure 33 shows the train-validation-test split on the control surface only (phases 3 and 4) portions of the southwestern heading. 70% of the data was reserved for testing. The other 30% was split in half for validation and testing. Red: test - 15%, Blue: validation - 15%, Green: training - 70%. The other headings were split similarly.

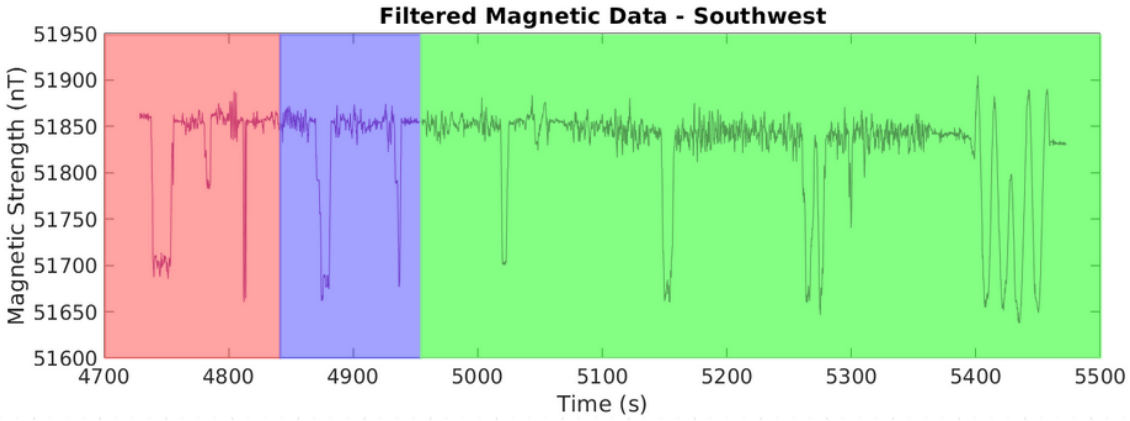

Figure 33: Training/validation/test split for the southwest heading

The ridge regression for Tolles-Lawson was implemented in Python using the machine learning package scikit-learn. 200 different alpha parameters were chosen for analysis and separated linearly in log-space from -8 to 2. Again, for each ridge parameter alpha the model was trained on the entire training set from each heading and evaluated on the validation set. The model's performance on the validation set as a function of alpha is shown in Figure 34.

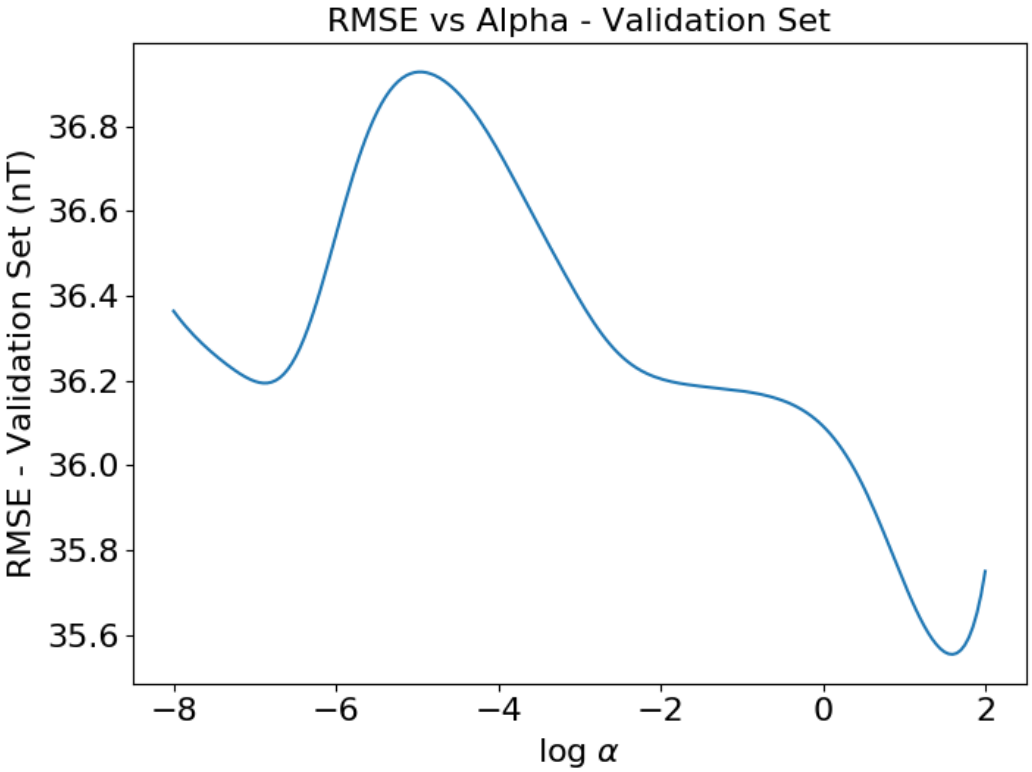

Figure 34: RMSE vs ridge parameter α

The ideal α is $\log_{10}(\alpha_{opt}) = 1.598$, however the actual performance doesn't vary significantly until alpha becomes very large. Figure 35 shows how the coefficients change as the ridge parameter alpha changes.

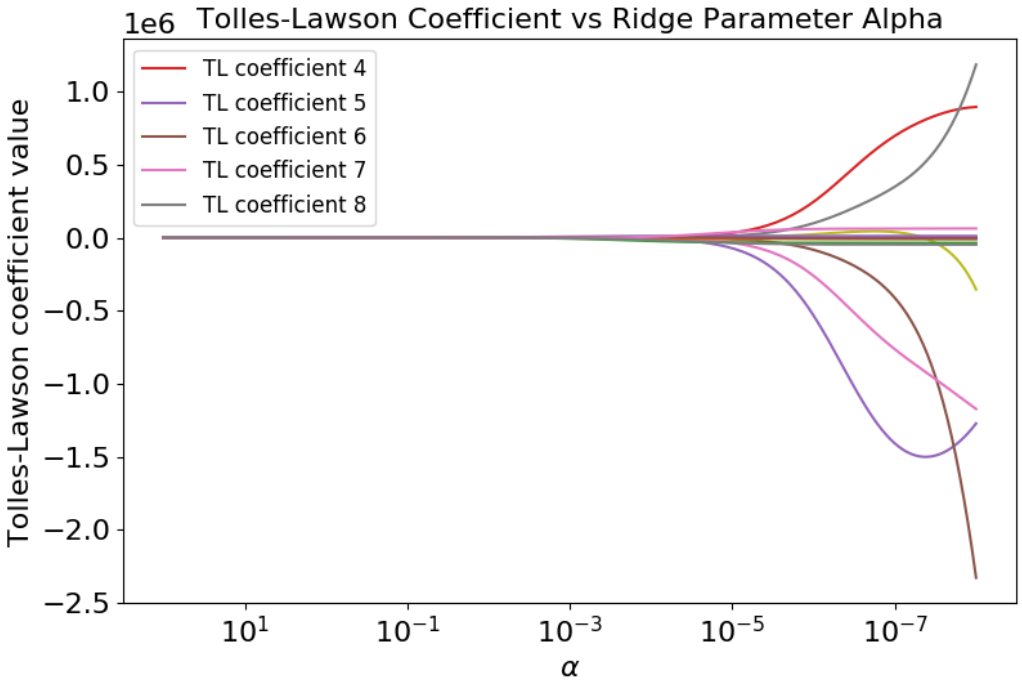

Figure 35: Tolles-Lawson coefficients vs ridge parameter alpha

The tendency for the coefficients to go to zero as alpha gets very large is clearly seen.

Figure 36 shows the final Tolles-Lawson coefficients for $\alpha = \alpha_{opt}$

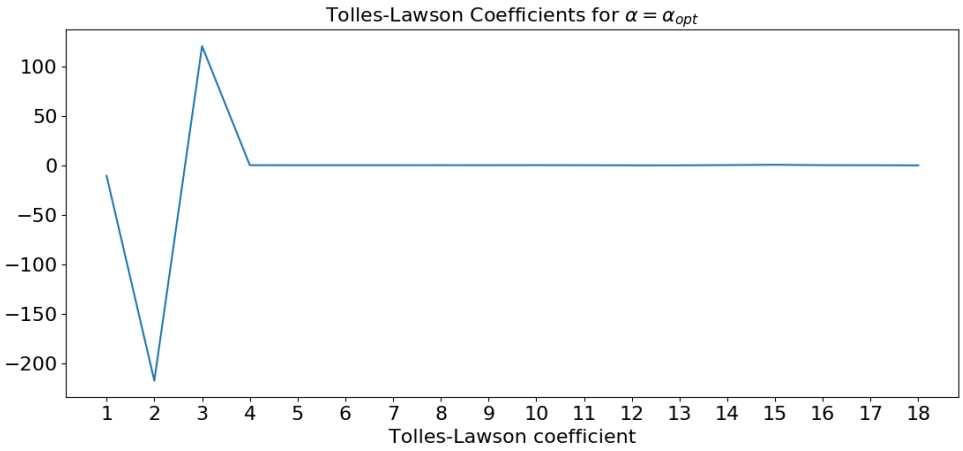

Figure 36: Tolles-Lawson coefficients for $\alpha = \alpha_{opt}$

Once the optimal ridge parameter was found, the model was retrained on the training

and validation data. That is, the training and validation data became the new set of training data for the optimized ridge regression model, where $\alpha = \alpha_{opt}$. Figures 37 and 38 show the performance of Tolles-Lawson on the test sets for each heading.

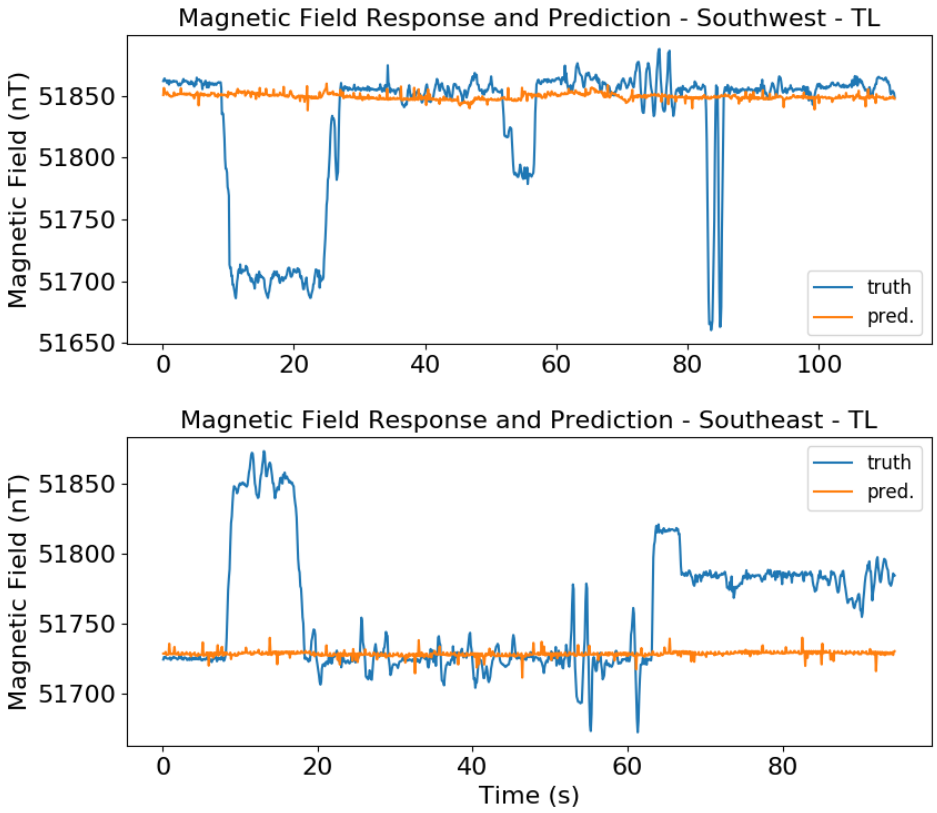

Figure 37: Tolles-Lawson predictions on the southern headings

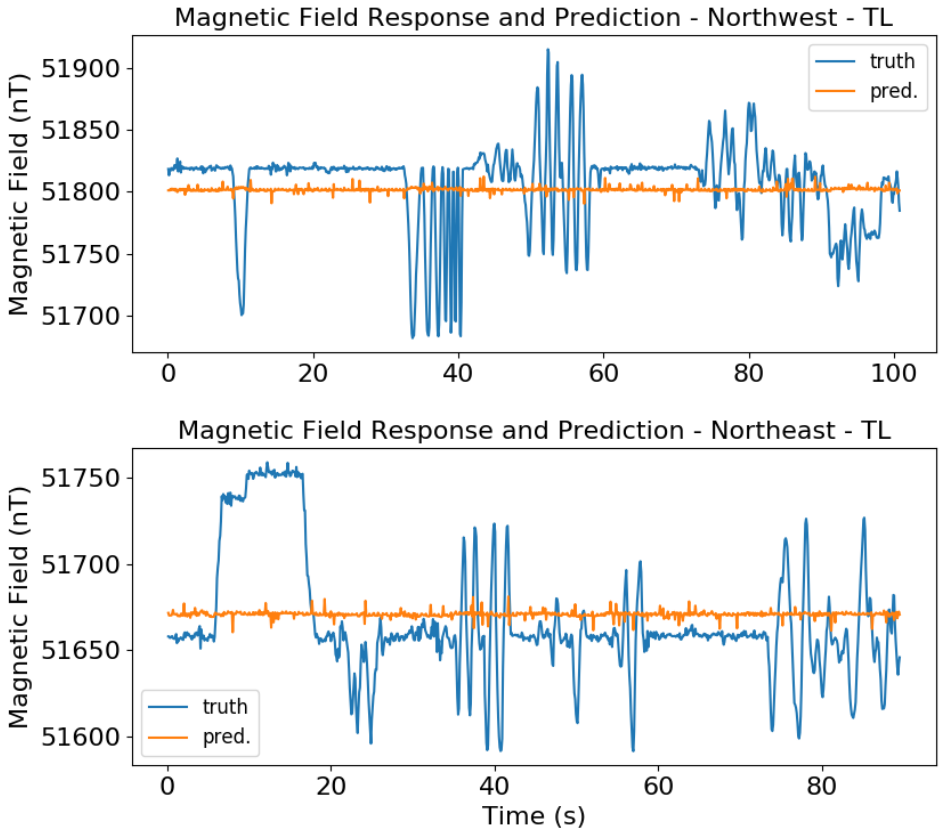

Figure 38: Tolles-Lawson predictions on the northern headings

The RMSE over all headings for Tolles-Lawson was 47.04nT.

4.2 Model Linearity vs Nonlinearity

Tolles-Lawson on its own was bound to fail because the disturbance field contributions from the control surfaces were clearly not capture-able entirely by the vector magnetometer. Since Tolles-Lawson only considers the vector magnetometer as input, and therefore only measures orientation changes, it would be unreasonable to expect that it could model control surface related disturbances. Not only that, but during phase 3 the aircraft was completely stationary, so it was unlikely that the vector mag-

netometer would be useful at all for this phase, because any changes in the sensed magnetic field would've been entirely due to sources that aren't related to orientation change. That being said, the PixHawk flight controller did record voltage PWMs of the control surfaces during the entire experiment, and it would be much more reasonable to expect these PWMs to be appropriate predictors of the disturbance field. To reiterate what was stated in chapter 3, the voltage PWMs measure how much voltage is required to keep a control surface at a particular deflection, and is reported by the PixHawk autopilot as an integer value. In total, 23 inputs were captured: 6 inputs from the 2 vector magnetometers, 12 inputs from the accelerometers and gyros, and 5 PWM values from each control surface and the throttle.

Figure 39 shows the correlation coefficient for each control surface at each heading, showing how predictable the process is linearly.

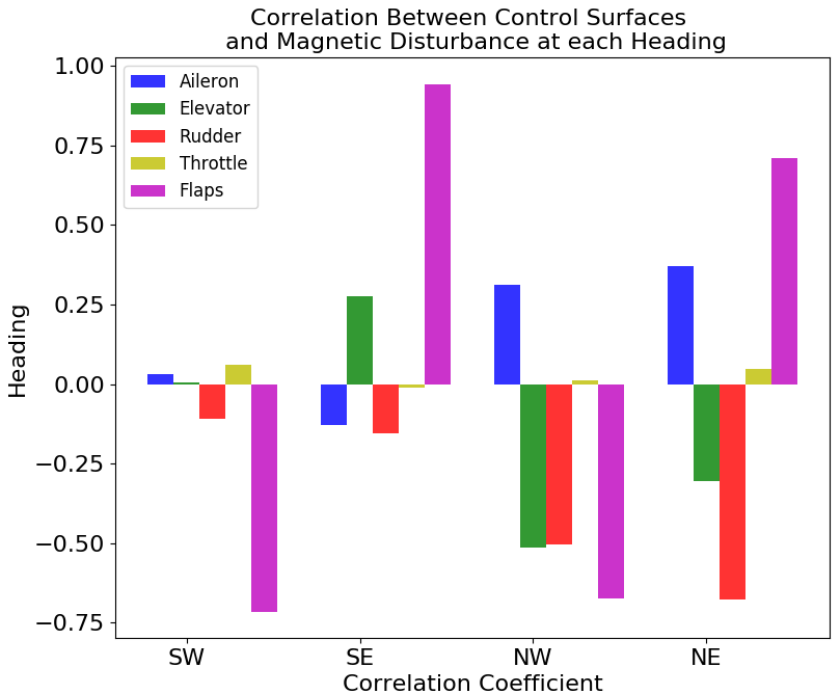

Figure 39: Control surface correlations with disturbance field at each heading

Figure 39 shows that along any given heading, the mapping between control surface input and disturbance field output is linear. However, the coefficients themselves change as a function of heading. Three of the four headings clearly show a magnetic response to an actuated elevator, but the southwestern heading indicates indifference. A different linear model - known as the augmented Tolles-Lawson model (ATL) - was trained to see how well the control surfaces model the disturbance field outputs. Specifically, in addition to the accelerometer, gyroscope, and PixHawk inputs mentioned earlier, only the first 3 Tolles-Lawson terms were included in the AGL model. These are the permanent field terms. Indeed, figure 36 shows that the only coefficients which are very useful at predicting the disturbance field are the first three, This makes sense, because the aircraft harbors many strong permanent magnets. So instead of the typical A_{TL} matrix, whose 18 columns consisted of the typical Tolles-Lawson

terms, the new A matrix, A_{ATL}, was defined as

$$A_{ATL} = \begin{bmatrix} \vdots & \vdots & \vdots & \vdots \\ \mathbf{Perm}_{Nx6} & \mathbf{Acc}_{Nx6} & \mathbf{Gyro}_{Nx6} & \mathbf{PWM}_{Nx5} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{13}$$

where $\mathbf{Perm}_{Nx6}$ is an Nx6 matrix of permanent Tolles-Lawson terms from both magnetometers, $\mathbf{Acc}_{Nx6}$ is an Nx6 matrix consisting of reading from both accelerometers, $\mathbf{Gyro}_{Nx6}$ is an Nx6 matrix consisting of readings from both gyroscopes, and $\mathbf{PWM}_{Nx5}$ is an Nx5 matrix of the voltage PWMs which measure the value of the throttle and the deflection of the control surfaces. N denotes the number of training samples. A_{ATL} is then used to find the augmented Tolles Lawson solution, as in (14)

$$\vec{\beta}_{ATL} = (\mathbf{A}_{ATL}^T \mathbf{A}_{ATL})^{-1} \mathbf{A}_{ATL}^T \mathbf{Y} \tag{14}$$

Because there were no hyperparameters over which to sweep, the training and validation sets from each heading were used as the entire training set again. Once the linear regressive model was trained and the ATL coefficients were found, they were used to predict the outputs from the test set. Figures 40 and 41 show the ATL predictions.

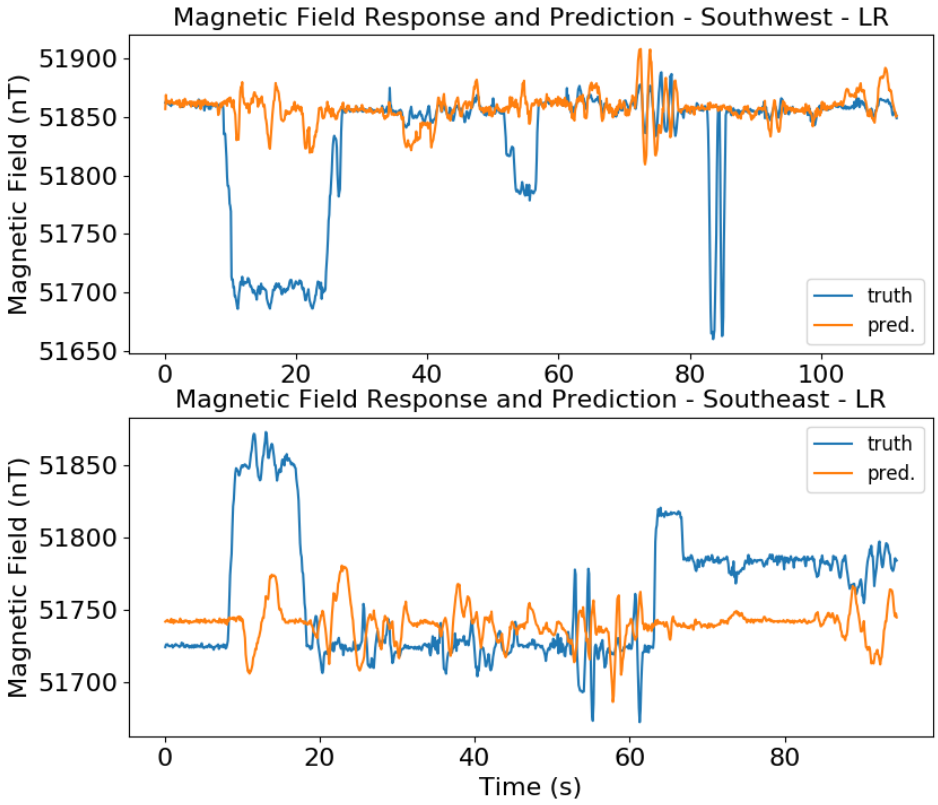

Figure 40: ATL model regression predictions on the southern headings

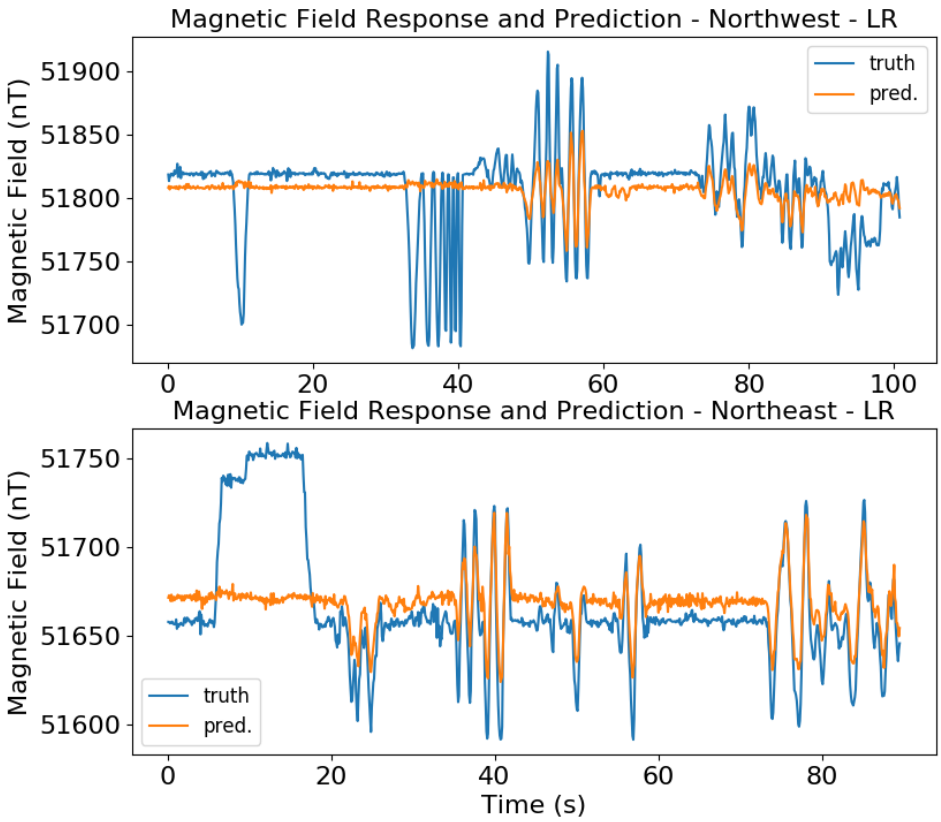

Figure 41: ATL model regression predictions on the northern headings

The ATL model is able to account for some of the disturbance better than Tolles-Lawson can, but still struggles with the overall DC level of the disturbance. Additionally, it is unable to model flap-induced disturbance at all, which according to figure 39 is the most correlated over any given heading. This indicates that the relationship is nonlinear and must be modeled accordingly. The ATL model did beat out Tolles-Lawson by showing some ability to predict control surface disturbance, but not by much. The RMSE incurred by the ATL model was 44.98nT.

4.3 Neural Network Approach

Neural networks are state-of-the-art nonlinear estimators, and are often used for modeling complicated relationships that aren't obviously modeled otherwise. As mentioned in chapter 2, they are capable of modeling nearly any function if endowed with enough flexibility. They can be difficult to engineer because of the vast number of hyperparameters with which they can be tuned, such as batch-size, training duration, learning rate, layer count, neuron count, momentum, etc. Tolles-Lawson and the ATL approach incorporating all inputs were both linear models, and were incapable of modeling the input-output relationship satisfactorily.

The neural network type chosen to analyze the control surface contribution to the disturbance field was a dense network, like that seen in Figure 10. In this model, all nodes - colloquially known as neurons - in one layer are connected to every node in the previous layer. Activation functions enable the neural network to model nonlinearity. Without them, non-linear relationships would be impossible to learn. The final layer produces the output that is checked against the known output during training. The analysis was performed using the Python deep learning library Keras.

As in the ridge regression approach, the validation set was used to pick the best model from several differing sets of hyperparameters. The hyperparameters which were tuned included batch-size - how many training samples the neural network sees before the neuron weights are updated, first layer activation function, number of epochs - the number of times the network passes through the entire set of training data, number of hidden layers in the network, and the number of neurons in each layer. The batch-size count was set at 4, 8, 16, or 32. The first layer activation function was set at either ReLU or hyperbolic tangent function. The number of epochs was

set at 100, 200 or 300. Finally, 10 different combinations of neuron count per layer and number of layers were analyzed, and are detailed in Table 3.

Table 3: Different NN Model Types

Type	H. L. 1 Neurons	H. L. 2 Neurons	H. L. 3 Neurons
1	2		
2	4		
3	8		
4	16		
5	4	2	
6	8	4	
7	16	8	
8	16	8	4
9	32	16	8
10	32	16	4

4 of the neural network types consisted of only 1 hidden layer, 3 had 2 hidden layers, and the remainder had 3. These 5 hyperparameters yielded 240 different neural networks to train and validate. Training was performed on a Lenovo P52 Thinkpad running Ubuntu 18.04.2. This computer operates with an i7 intel processor with 12 cores at 4.3GHz and 48GB DDR4 RAM. It also contains an NVIDIA Quadro P3200 GPU. Training took approximately 22 hours. The training set consisted of 19,263 samples, the validation set consisted of 4126 samples, and the test set consisted of 4125 samples.

Figure 42 shows how the validation RMSE varied as the different hyperparameters were changed.

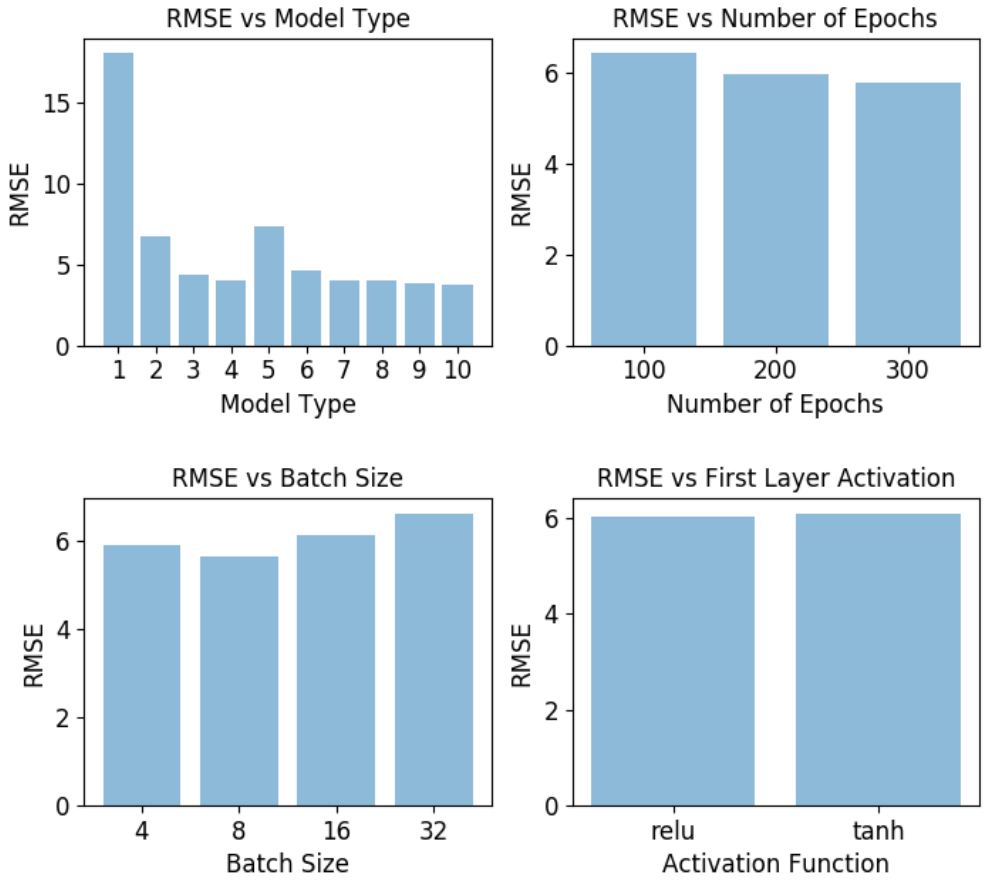

Figure 42: Validation set RMSE incurred by different hyperparameters

After all 240 neural networks were trained, the one which incurred the lowest validation set RMSE was model 9, with a ReLU activation function in the first layer, trained for 200 epochs with 4 training samples per batch. This neural network had an overall validation set RMSE of 3.4nT and utilized 1,441 trainable weights and biases. This neural network is shown in figure 43.

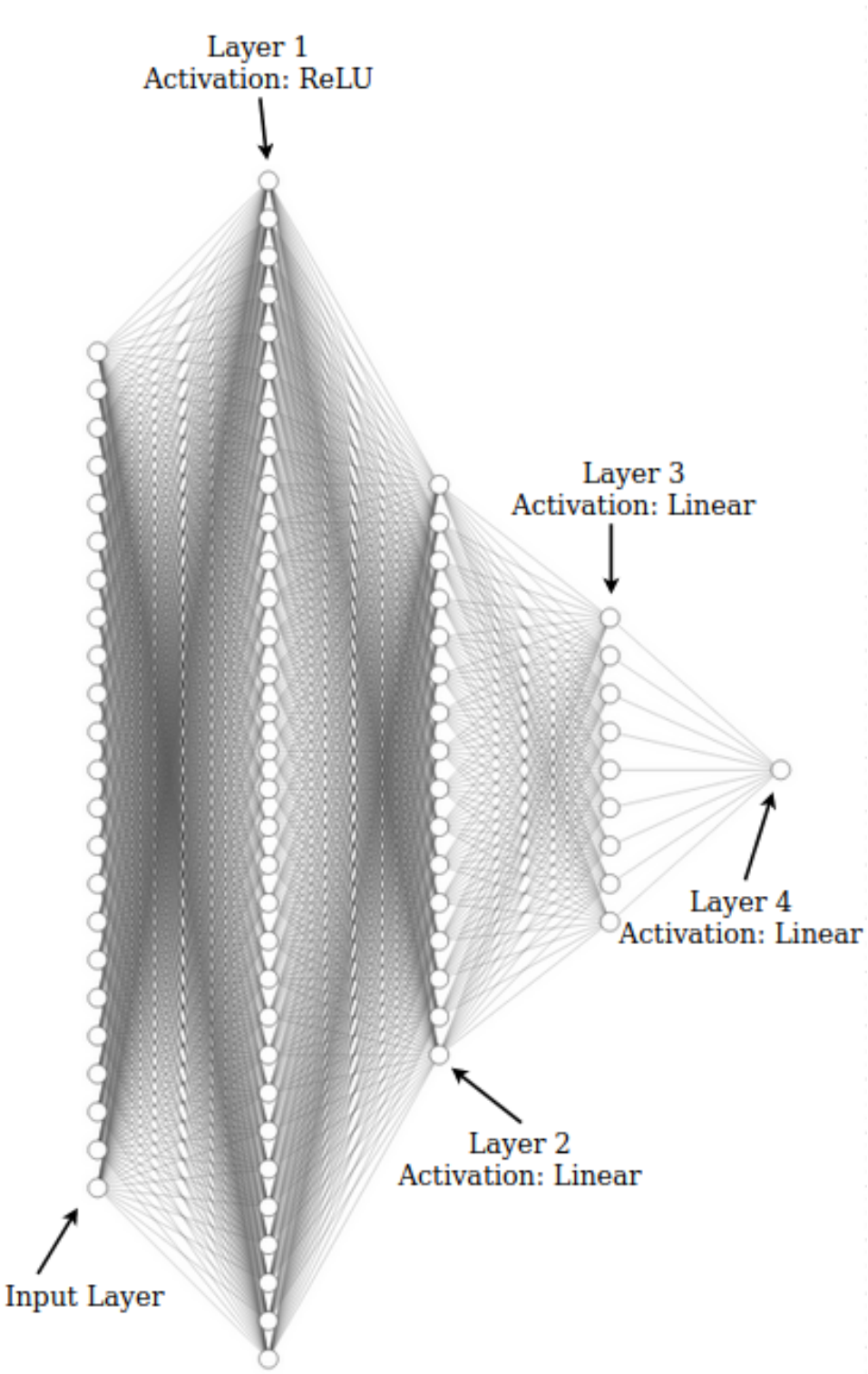

Figure 43: Final architecture of the dense neural network

This model was retrained on the combined training and validation set. Figures 44 and 45 show the neural networks predictions on the test set.

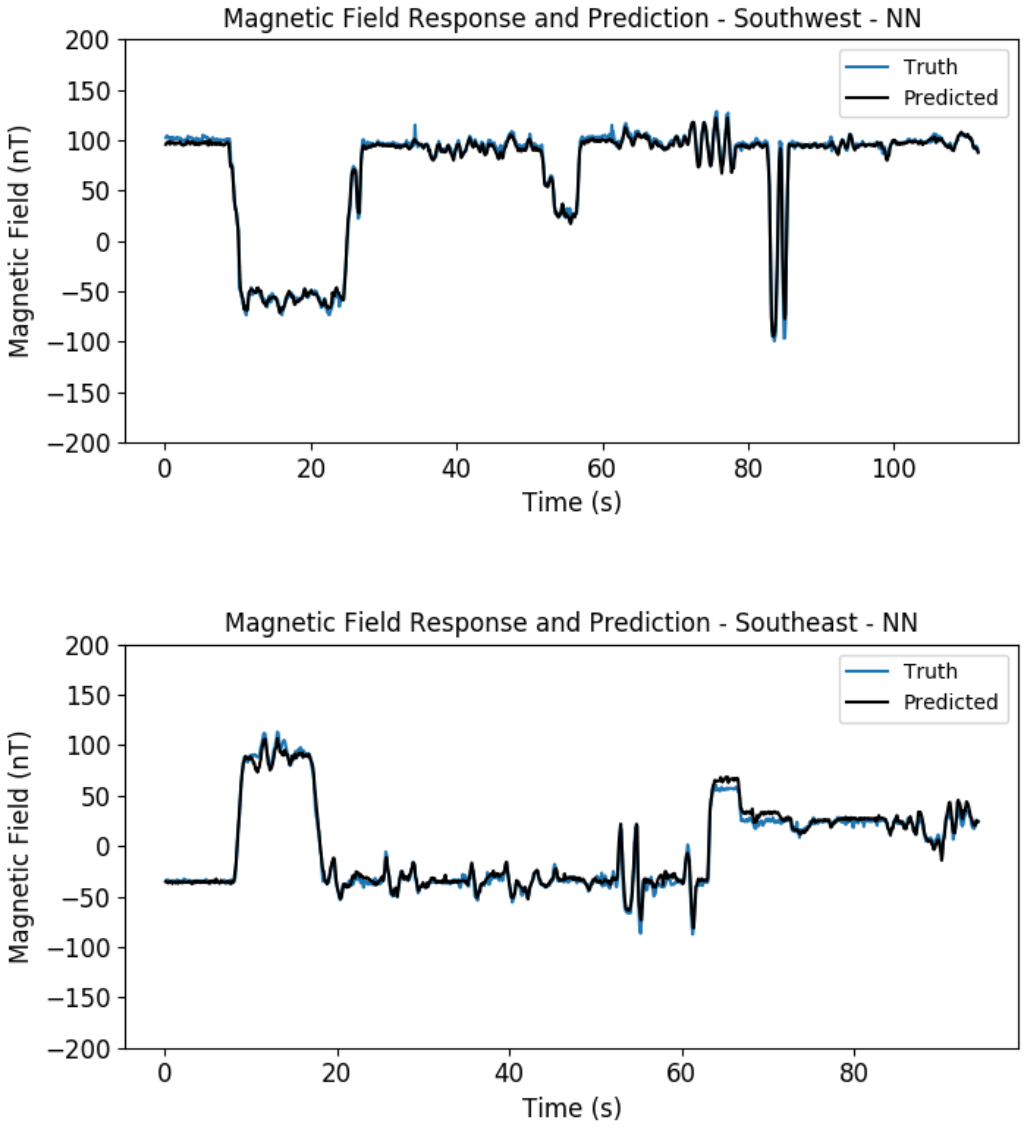

Figure 44: Dense neural network predictions on the southern headings

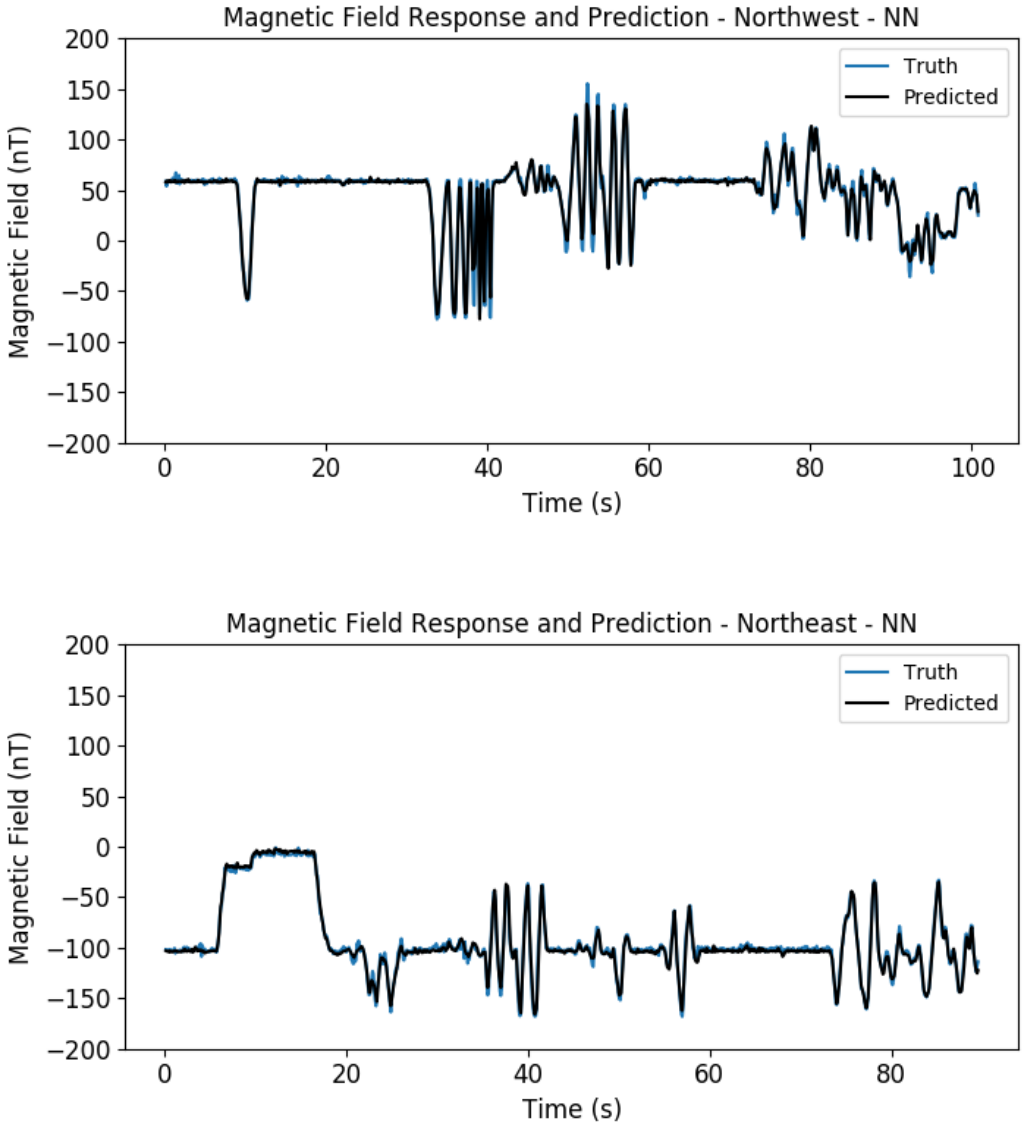

Figure 45: Dense neural network predictions on the northern headings

The dense neural network incurred a test set RMSE of 4.79nT and an MAE of 2.97nT. This is a significant improvement over the linear models, and proves that the control surface disturbance field problem greatly benefits from a more flexible, non linear

model. Figures 46 and 49 show detail of figures 44 and 45.

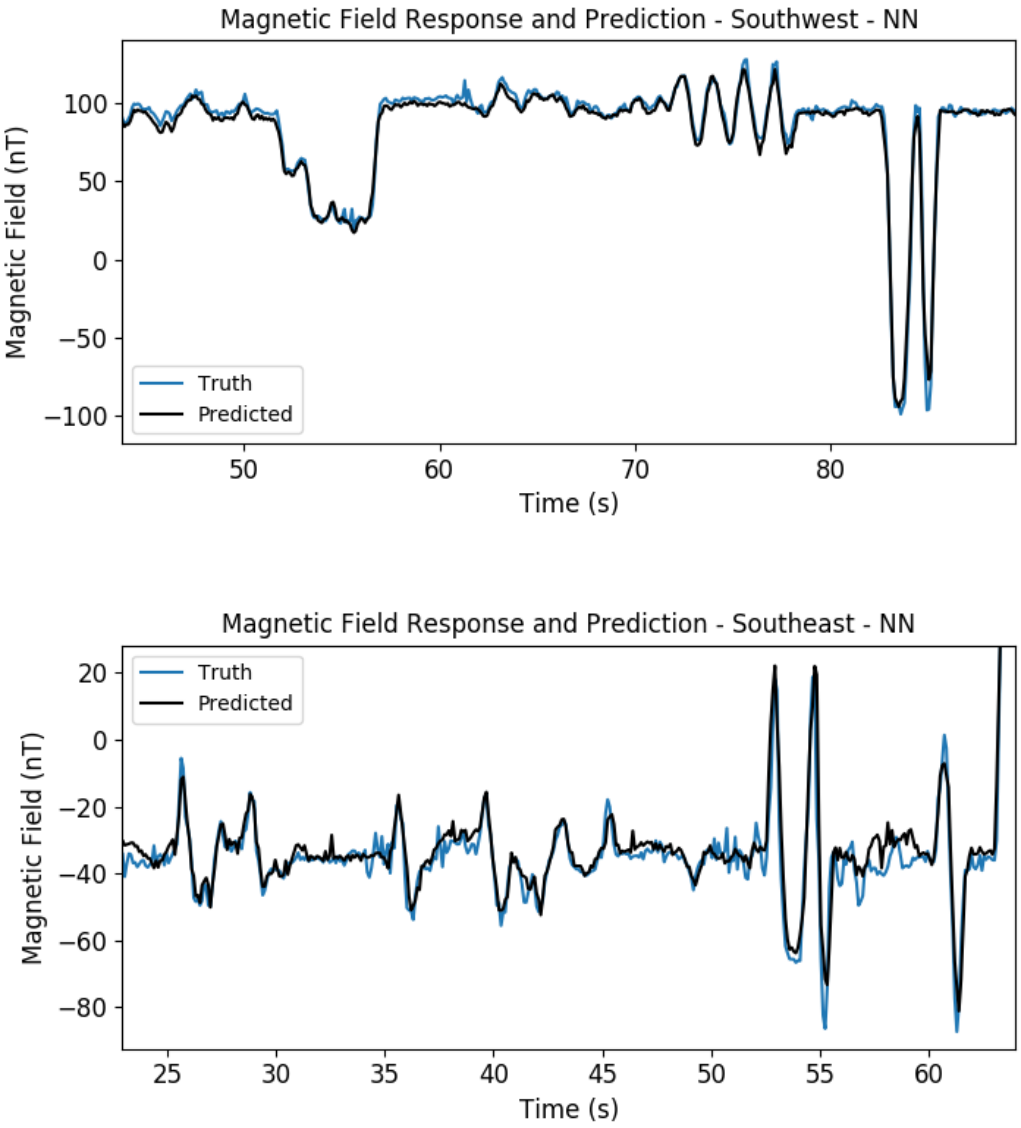

Figure 46: Detail of DNN performance on the southern headings

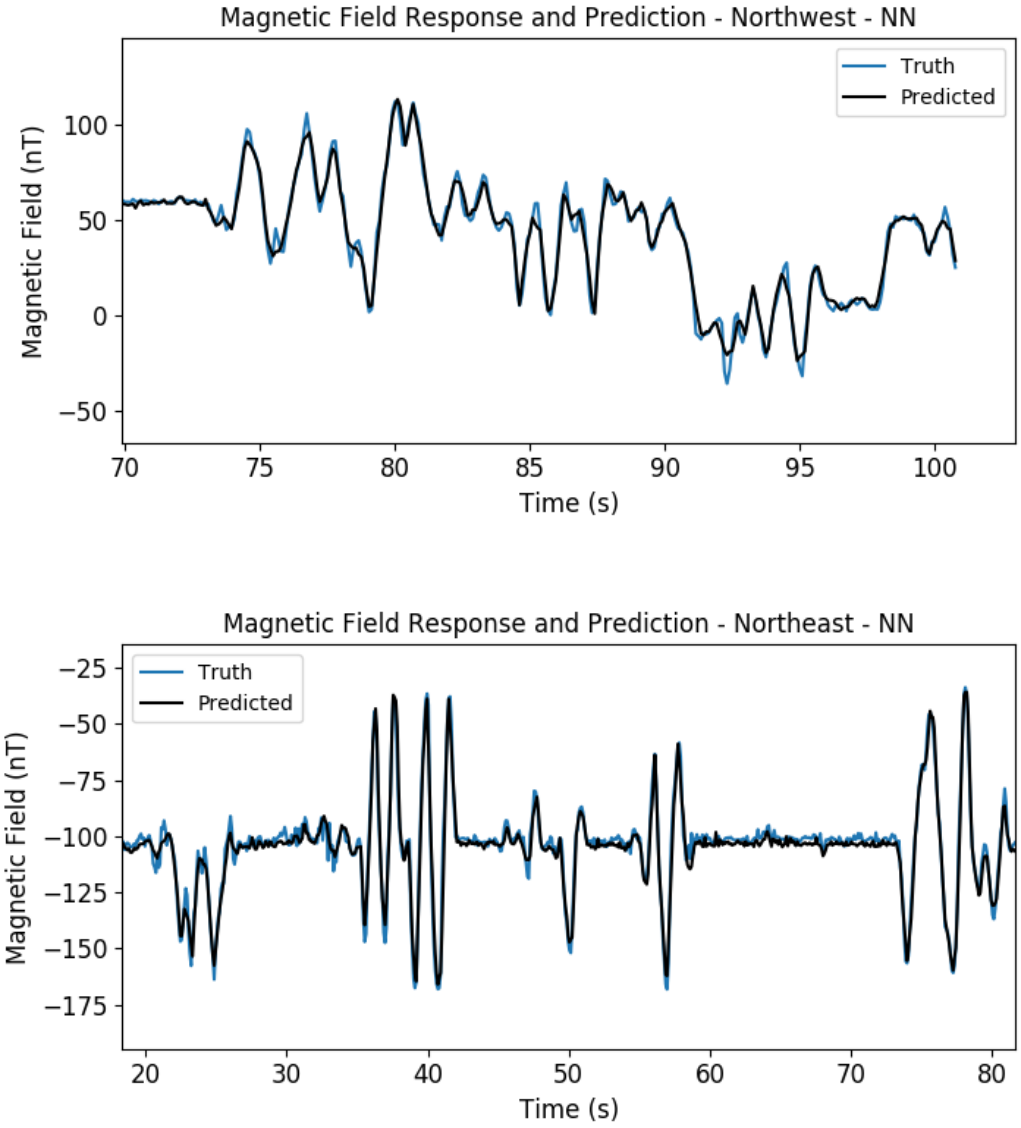

Figure 47: Detail of DNN performance on the northern headings

The dense neural network seemed to very slightly underpredict the mean of the test data on the southwestern heading. More training data might suffice in fixing that. Additionally, greater accuracy could likely have been attained by taking additional

measures such as sweeping over more hyperparameters and trying even more flexible models. Figure 42 shows that the RMSE on the validation set kept decreasing as the model was given extra flexibility. Table 4 details the performance of the different machine learning algorithms.

Table 4: RMSE of Different ML Models

	TL	ATL	DNN
RMSE	47.04nT	44.98nT	4.79nT
% dec. over TL	—	4.38%	89.43%

4.3.1 Other Sources of Magnetic Disturbance

In chapter 4 we discussed many potential sources of the magnetic disturbance field on the aircraft. One of those, the aircraft engine, will very often contribute a large proportion of the overall disturbance field in the form of an induced field. As the engine moves in 3-dimensional space, its orientation relative to the geomagnetic vector produces a varying disturbance field. However, a running engine close to the magnetometer (or the shaking fuselage or control surfaces due to the running engine) can also corrupt the magnetic signal with high frequency noise. Phase 1 of each heading consisted of actual flight with the running engine, and figure 32 shows the effect the running engine has on the magnetic signal. At its worst, the running engine and control surfaces together generate around 200nT of noise. Since the spatial frequency of the crustal anomalies translate ultimately to a temporal frequency on the order of single Hertz in the scalar magnetometer data stream [1], this high frequency noise would usually be eliminated by the application of a low-pass filter. However, no filtering scheme is perfect, and filters will often introduce artifacts into the data which don't exactly honor the data at the desired frequency. If a sufficiently sophisticated,

time-sensitive deep learning model, like an LSTM or 1DCNN, was able to remove this high frequency noise, it might be preferred over filtering. The sensors on-board the cub included accelerometers and gyroscopes, and it might seem reasonable that these sensors would be the sensors of choice for removing high frequency engine noise. Presumably if either moving components of the aircraft or the shaking of magnetic components in the fuselage by the engine - or both - are contributing to the high frequency magnetic disturbance in phase one of each heading, like that shown in figure 32, one might expect there to be a correlation with sensors that measure that shaking process - accelerometers. Indeed, spectrograms of the scalar magnetometer and accelerometer channels show a clear relationship in the frequency domain. Figures 48 and 49 show the spectrograms for phase 1 of the southwest flight.

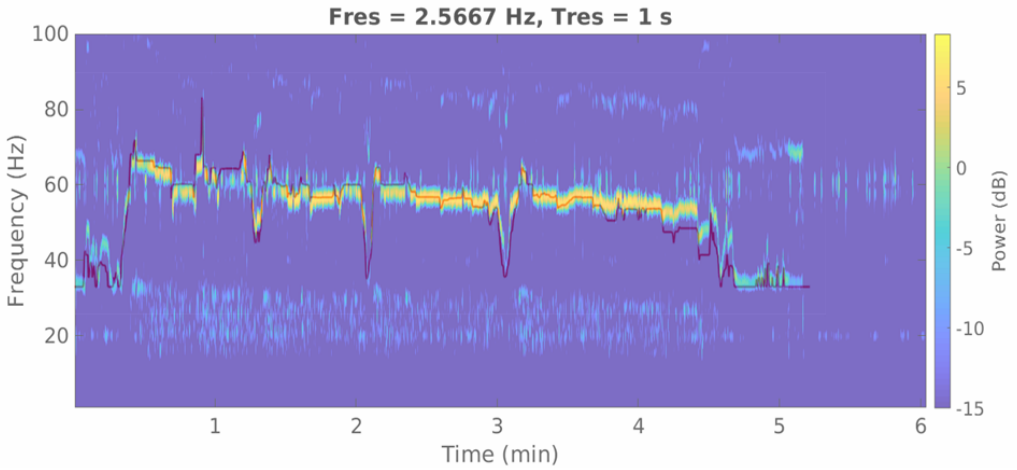

Figure 48: Magnetometer spectrogram of phase 1 along the southwest heading with throttle overlain

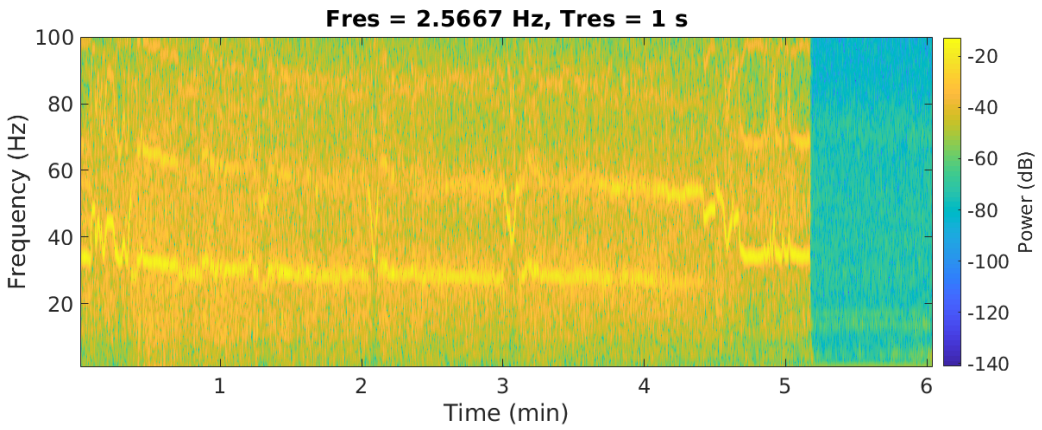

Figure 49: Y-accelerometer spectrogram of phase 1 along the southwest heading

Figure 48 also shows the throttle overlain on top of the magnetic spectrogram in a red color. To make the frequency band of interest more obvious, a minimum decibel threshold was set. Intriguingly, the spectrogram frequency waveform in both the accelerometer and scalar magnetometer signals are identical. This is the case along all 4 headings during phase 1. This indicates that there are very powerful relationships a neural network could learn from in the frequency domain. Prediction of engine noise using accelerometer and gyroscope data was briefly examined during this research using both LSTMs and 1DCNNs. Overall, the neural networks seemed to learn the frequency patterns, but struggled to learn the amplitude. Nevertheless, an exhaustive and in-depth research push was not done in this domain, and deserves further study.

4.4 Analysis Conclusion

This chapter discussed the different machine learning models that were used to perform advanced disturbance field compensation. The traditional compensation method, Tolles-Lawson, takes into account only the direction cosines and has no knowledge of other potential corrupting sources. Tolles-Lawson is typically used to compensate small induced field disturbances when the magnetometer is far from mag-

netized material, and is not conditioned to model sources not well captured by a vector magnetometer. We have shown that control surface disturbances are not well modeled by Tolles-Lawson, nor are they well accounted for by a linear regressive approach which incorporates linear terms related to the control surface PWM. A sufficiently advanced non-linear learning technique is required to model control surface interference to a high degree of accuracy. Figure 39 indicates that along any given heading, control surface action may be linear, but when coupled with a variable heading becomes nonlinear. We showed the variation in performance of many different dense neural networks with varying hyperparameters and determined the best one using the validation set. The test set RMSEs for Tolles-Lawson and the linear regressive model with control surface PWM inputs were 47.04nT and 44.98nT. The dense neural network approach was able to reduce the test set RMSE to 4.79nT; a 90% decrease in disturbance field strength. In addition, we showed an interesting relationship between the accelerometer spectrogram data and magnetometer spectrogram data, as well as throttle, during phase 1 of the 4 different headings. These relationships may prove to be powerful input features to a sufficiently advanced deep learning algorithm for use in engine noise compensation.

V. Conclusions

This research presented a method for incorporating more inputs and sensor readings than is typically accounted for by the traditional Tolles-Lawson model. We showed that a more advanced, non-linear machine learning model is capable of reducing control surface noise to single digit nano-tesla levels, and presented a means to make magnetic anomaly navigation more feasible on less-ideal aircraft, such as those without stingers. Using control surface inputs, vector magnetometer readings, and other sensors, we showed an improvement over Tolles-Lawson by nearly 90% using a fully connected dense neural network and a gridded hyperparameter search.

5.1 Future Work

5.1.1 Advanced Airborne Calibration

There is much left to be done in the airborne platform calibration area. While this research did show a large improvement in disturbance field noise, it did not do so with airborne data. Platform calibration is fundamentally simpler on the ground because any deviation from zero magnetic variation is a form of disturbance. In an actual flight, the calibration routine must distinguish between desirable geomagnetic change and fluctuations due to the undesired disturbance field, so it is not quite as clear what the "truth labels" the machine learning algorithm should learn from are. Additionally, the ground calibration was susceptible to 60Hz power grid noise and, as discussed in section 2, potential time-varying effects due to space-weather. The assumption that the only variation in the magnetic signal while calibrating on the ground must be solely due to platform effects is somewhat exaggerated, although clearly mostly held true given the performance of the neural network. A likely next step would be to perform calibration using real flight data over an area whose magnetic field behavior

is well known with all of these considerations taken into account.

5.1.2 Alternative Machine Learning Models

This thesis concerned itself only with neural networks for nonlinear learning. Neural networks, especially those that take into account temporal variations like LSTMs, require a large amount of computational overhead and training data and can take a very long time to train. This research searched a fairly tame number of hyperparameters using a decently powerful computer, and training took nearly an entire day. A more in-depth hyperparameter search might span multiple days. Other, more lightweight nonlinear modelling might suffice in performing as well, or better than, neural networks, both in performance and training time. Random forest regression would be a good first start.

5.1.3 Other Platforms and Inputs

Aeromagnetic navigation has been shown to work very effectively when the aircraft is magnetically optimized. Figure 6 shows the type of aircraft that is typically used for magnetic anomaly cartography, and the class of aircraft used by Canciani in [1]. It has not been shown to work when the disturbance field provides a substantial proportion of the overall sensed magnetic field, nor when the disturbance field is a byproduct of orientation within the geomagnetic field. Many aircraft which would benefit from magnetic navigation are considered difficult platforms, because they do not readily enable the placement of the scalar magnetometer in a location which is magnetically quiet. More flight data from other aircraft would be beneficial in determining whether advanced machine learning models are sufficient at tackling the calibration problem.

Finally, we presented a striking relationship between accelerometer and magnetome-

ter data in the frequency domain during the portion of the test where the engine was running, and a clear correlation between both those and the throttle data recorded by the PixHawk 2 flight controller software was discussed. It is believed that a sufficiently sophisticated deep learning algorithm can learn how to predict high frequency magnetic engine noise by using inertial sensors and throttle data.

Bibliography

1. Aaron Canciani. *Absolute Positioning Using the Earth's Magnetic Anomaly Field*. PhD thesis, 2016.

2. Gary A Glatzmaier and Paul H Roberts. A three-dimensional self-consistent computer simulation of a geomagnetic field reversal. *Nature*, 377(6546):203–209, 1995.

3. British Geological Survey. The earth's magnetic field: An overview.

4. Colin Reeves. Aeromagnetic Surveys: Principles, Practice and Interpretation. 2009.

5. Miroslaw Woloszyn. Analysis of aircraft magnetic interference. *International Journal of Applied Electromagnetics and Mechanics*, 39(1-4):129–136, 2012.

6. Michael Nielsen. Neural Networks and Deep Learning. *Artificial Intelligence*, pages 389–411, 2018.

7. Sagar Sharma. Activation functions in neural networks. `https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6`, September 2017. [Accessed: Nov-2019].

8. F. Chollet. *Deep Learning with Python*. Manning Publications Company, 2017.

9. Chandra Chatterjee. Basics of the classic cnn. `https://towardsdatascience.com/basics-of-the-classic-cnn-a3dce1225add`, July 2019. [Accessed: Dec-2019].

10. Viki et. al. Bankey. Digital data grids for the magnetic anomaly map of north america. `https://pubs.usgs.gov/of/2002/ofr-02-414/`, 2002. [Accessed: Nov-2019].

11. Peter M. Williams. Aeromagnetic compensation using neural networks. *Neural Computing & Applications*, 1993.

12. Geometrics. MFAM Development Kit User Guide. 2017.

13. Arnaud Chulliat, Susan Macmillan, Patrick Alken, Ciaran Beggan, Manoj Nair, Brian Hamilton, Adam Woods, Victoria Ridley, Stefan Maus, and Alan Thomson. The us/uk world magnetic model for 2015-2020. 2015.

14. Stefan Maus et. al. The US / UK World Magnetic Model for 2010-2015. *NOAA Technical Report NESDIS/NGDC.*, 2010.

15. W.E. Tolles and J.D. Lawson. Magnetic Compensation of MAD Equipped Aircraft, Airborne Instruments Lab Inc., Mineola, NY. 1950.

16. Paul Leliak. Identification and Evaluation of Magnetic-Field Sources of Magnetic Airborne Detector Equipped Aircraft. *IRE Transactions on Aerospace and Navigational Electronics*, ANE-8(3):95–105, 1961.

17. Mioara Mandea and Korte Monika. *Geomagnetic Observations and Models*. Springer Science and Business Media, 2010.

18. Qi Han, Zhenjia Dou, Xiaojun Tong, Xiang Peng, and Hong Guo. A Modified Tolles-Lawson Model Robust to the Errors of the Three-Axis Strapdown Magnetometer. *IEEE Geoscience and Remote Sensing Letters*, 2017.

19. Lei Jiang, Ziqi Guo, and Baogang Zhang. Scalar Calibration of Aeromagnetic Data Using BPANN and LS Algorithms Based on Fixed-Wing UAV Platform. 64(7):1968–1976, 2015.

20. L. Chen, P. Wu, Z. Wanhua, Y. Feng, and G. Fang. A Novel Strategy for Improving the Aeromagnetic Compensation Performance of Helicopters. *Sensors*, 2018.

21. Kyosuke Yamamoto, Takashi Togami, Norio Yamaguchi, and Seishi Ninomiya. Machine learning-based calibration of low-cost air temperature sensors using environmental data. *Sensors (Switzerland)*, 17(6):1–16, 2017.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
03/26/2020	Master's Thesis	Sept 2018 — Mar 2020

4. TITLE AND SUBTITLE

Improving Aeromagnetic Calibration Using Artificial Neural Networks

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Mitchel C. Hezel

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENG-MS-20-M-027

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

The Charles Stark Draper Laboratory
555 Technology Square, Cambridge, MA 02139
617-258-1000

10. SPONSOR/MONITOR'S ACRONYM(S)

Draper Labs

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A:
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Magnetic anomaly navigation has been shown to be highly successful but only under idealized scenarios. In particular, it has proven itself a viable alternative to GPS only on optimized, geosurvey aircraft. No body of research exists which attempts to address the platform effect problem outside the scope of the de facto calibration equations, known in the literature as Tolles-Lawson. We show an alternative calibration technique which achieves a 90% decrease in aircraft disturbance field strength over the Tolles-Lawson equations using deep learning and artificial neural networks.

15. SUBJECT TERMS

Aeromagnetic calibration, Tolles-Lawson, Neural networks, Deep learning, Magnetic anomaly navigation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Maj Aaron J. Canciani, AFIT/ENG
U	U	U	UU	99	19b. TELEPHONE NUMBER *(include area code)* (937) 255-3636 x4618 aaron.canciani@afit.edu