Magnetic Field Mapping for Indoor Aerial Navigation

by

Prince E. Kuevor

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Robotics)
in the University of Michigan
2023

Doctoral Committee:

Professor Ella M. Atkins, Co-Chair,
Associate Professor James W. Cutler, Co-Chair
Assistant Professor Maani J. Ghaffari
Professor Mark B. Moldwin

Prince E. Kuevor

kuevpr@umich.edu

ORCID iD: 0000-0003-0750-0895

DEDICATION

This work is dedicated to the intellectual and empathetic growth I've realized.
To the confidence I've found, at last, near the end of this journey and learned to harness with the help of my community.
To the promise of service to those nearby and their smiling faces as I now wave goodbye.

Through life, move in a way that allows others to grow and learn in your wake.
For the ripples you leave may last long after you.

ACKNOWLEDGMENTS

I would first like to thank my advisors Ella Atkins and Jamie Cutler. They both supported me through a number of difficulties through my PhD and I am incredibly grateful for their kindness. Thanks to Maani Ghaffari for introducing me to mapping techniques and Mark Moldwin for providing his expertise on magnetism.

Thanks to my labmates in both the Michigan eXploration Lab (MXL) and the Autonomous Aerospace Systems lab (A2sys). I specifically thank Matthew Romano, Jeremy Castagno, Prashin Sharma, Akshay Mathur, Mark Nail, Nathanael England, Justin Schachter, Matthew Szczerba, Srinagesh Sharma, Tran Anh Nguyen, Paul Flanigen, and Brandon Apodaca for their support in fabricating my test vehicle, tuning the flight controller, assisting with flight tests, developing/soldering PCBs for the magnetometers, and helping me convey my technical findings through easily-accessible visuals. Thanks to Dae Yong Lee who was a PhD student when I started working in MXL as a sophomore in undergrad. My work with you planted the seeds of what became this dissertation. Special thanks to Mark Moldwin, Alex Hoffmann, and Brady Strabel from the University of Michigan Magnetometer Laboratory for help with configuring the PNI RM3100 magnetometer.

I have had the great fortune of really getting to be a part of the campus community here at the University of Michigan. Throughout my time here, I've helped lead efforts and student organizations like Black Students in Aerospace, Graduates Out in STEM, Black Students in Robotics, the Aerospace DEI Committee, and Operation Graduation. These experiences have taught me a lot in terms of what it means to serve a community, how to empathize and connect with people, and how to organize a group of people towards a common goal. Thank you to everyone I have worked with on these projects for truly giving me a holistic college experience. Special thanks to the Pizza Calzone volleyball crew for keeping me sane with some spicy serves and chaotic rallies.

The work in this dissertation relied heavily on experimental data that required support and facilities from many staff and faculty members on campus. Specifically, I would like to thank Peter Gaskell, Denise Edmund, Kimberly Johnson, Damen Provost, Tanner Schudlich, Michael Lee, and Christopher Chartier for their support throughout various aspects of my PhD.

Finally, none of this would be possible without the immeasurable support from my family. Many thanks to my parents (Golda Kumedzro and Aubrey Kuevor) and sisters (Princess, Christiana, and Mawuli Kuevor). Specifically, thanks to my mom and two younger sisters. It is not lost on me how much you three have done to let me be here and I cannot thank you enough.

TABLE OF CONTENTS

Chapter

LIST OF FIGURES

FIGURE

LIST OF TABLES

TABLE

LIST OF APPENDICES

ABSTRACT

Earth's magnetic field is a ubiquitous signal commonly used to orient one's heading relative to North. Typical magnetic field navigation techniques assume the ambient field points Northward and remains constant in a local area. Both assumptions break down inside buildings due to distortions caused by ferrous materials in modern structural components. In this dissertation, we show how to make maps of magnetic field distortions to improve our ability to orient and traverse through indoor spaces. Because the magnetic field is ubiquitous on Earth and magnetometers are present on most modern smartphones and IMUs (inertial measurement units), our methods can be applied nearly anywhere on Earth.

This dissertation presents three major contributions towards the use of magnetic fields for indoor navigation. First, we show how to use a small UAV (unmanned aerial vehicle) to measure the ambient field throughout a workspace and create three-dimensional magnetic field maps. We leverage a machine learning tool called Gaussian process regression (GPR) as the backbone of our magnetic field maps to interpolate the field at unobserved locations. This first contribution is grounded in practical tradeoffs between the convenience of using an autonomous UAV for indoor mapping and the magnetic disturbances created by the UAV's electronics. Here, we present methods to reduce the UAV-induced magnetic disturbances, a new technique to create magnetic field maps, and a set of best practices creating and utilizing indoor magnetic field maps.

Second, we use a multiplicative extended Kalman filter (MEKF) with our GPR-based maps to estimate the attitude (orientation) of a UAV. Here, we introduce the concept of spatial variation which describes how much the magnetic field changes locally. Essentially, outdoor environments have low spatial variation while indoor spaces typically have higher spatial variation in the magnetic field. Results show that our magnetic maps yield a two-fold improvement of attitude estimates indoors (where there is high spatial variation), but are unnecessary for outdoor environments where a constant-field assumption is appropriate.

Finally, we use a particle filter to estimate the position of a UAV using indoor magnetic fields. This last innovation is important because GPS (a staple for position estimation outdoors) is not available inside buildings. Instead, we use our GPR-based magnetic field maps to track the UAV's motion through our mapped space. Our results give three-dimensional position estimates of a UAV within 0.2m for six of our eight test cases. In addition, we show how the amount of magnetic field's

spatial gradient correlates with our position estimation accuracy.

With some improvements, our methods can be used to transform the way people navigate through buildings. Imagine an indoor route planning application that guides someone to their terminal at an airport, to a book at their local library, or to their office at a new job. Because magnetic fields are everywhere on Earth, we can apply our mapping and navigation techniques to any building on the planet. In addition, the presence of magnetometers in modern smartphones gives everyone the ability to benefit from the invisible field all around them.

To make this vision a reality, our methods need to be less sensitive to changes in the magnetic field. Although this dissertation does not investigate robust navigation or time-varying magnetic field mapping, it does present important foundations on the practice of creating magnetic field maps and their value in enabling indoor navigation.

CHAPTER 1

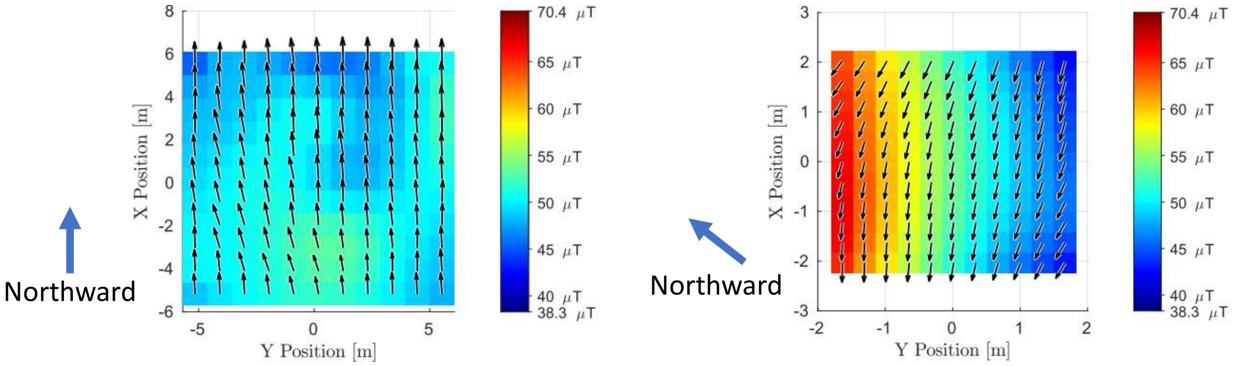
Introduction

1.1 Motivation

Sensor fusion is a technique that leverages information from different *sensing modalities* to estimate the position and attitude state of a vehicle [1]. Vision and LIDAR are common *exteroceptive* sensing modalities that provide information about the vehicle's surroundings while gyroscopes, accelerometers, and magnetometers give *proprioceptive* information about the vehicle's motion. A key goal of sensor fusion is for one sensing modality to enable accurate navigation in the event another modality fails (e.g., a sudden change in ambient light when using a vision-based system). The goal of this dissertation is to improve the use of magnetometers as a sensing modality for navigation.

Magnetometers provide point-wise measurements of the magnetic field created by Earth, nearby ferromagnetic objects, and electric currents. They are often present in inertial measurement units (IMUs) that consist of a gyroscope, accelerometer, and a magnetometer to provide proprioceptive data. Today, IMUs with magnetometers are ubiquitous, appearing on cell phones and low-cost robotics sensing kits. Despite the prevalence of magnetometers in modern electronics, their use case is often restricted to estimating heading relative to magnetic north while outdoors [2, 3] or assisting in three degree of freedom (3DOF) attitude estimation for spacecraft in Earth's orbit [4]. They find less use inside or near buildings where ferromagnetic materials distort the local magnetic field making it unintuitive and noisy.

This dissertation investigates how magnetometers can be used to improve indoor position and attitude estimates. Typically, we use GPS and cell towers to estimate our position on Earth [5], however both become unavailable or less reliable when near or inside a building. Alternatives arise to estimate position in GPS-denied environments like visual fiducial markers [6], feature-based tracking [7], point-cloud-based spatial mapping [8], and signal mapping of Wi-Fi router signals [9] or magnetic fields [10]. However, many of these require additional hardware to be placed on the vehicle (e.g., camera, LiDAR, or Wi-Fi receiver) or in the environment the vehicle will move through (e.g., fiducial markers). We propose the use of magnetic fields to improve navigation



(**a**) (**b**)

Figure 1.1: Motivation for local magnetic field maps. Outdoor magnetic fields (**a**) generally point northwards, have low spatial variation, and typically agree with WMM. Indoor fields (**b**) can have high spatial variation that is not captured by WMM or other global magnetic maps.

in GPS-denied environments. Since there is an ambient magnetic field everywhere on Earth and magnetometers are commonplace (via IMUs), use of this sensing modality requires no additional hardware on or off the vehicle. However, as with all sensing modalities, there are tradeoffs.

Most uses of magnetometers assume that the magnetic field is generally pointing Northward and remains spatially constant throughout the workspace. These assumptions tend to hold well outdoors and away from buildings but are inaccurate for the magnetic field inside modern buildings (Figure 1.1). Figure 1.1a shows an outdoor magnetic field map where the planar magnetic field (black arrows) generally point North. Figure 1.1b is the same visualization method for an indoor workspace showing that materials in the building create high spatial variation and distort the planar direction of the magnetic field. This means many classical methods of navigating using the magnetic field (which assume a constant, Northward field) are not applicable in or near buildings. Spacecraft applications like Ref. [4] relax these assumptions by using global models of the Earth's magnetic field like the World Magnetic Model (WMM) [11]. However, WMM estimates do not account for distortions in Earth's magnetic field due to human-made structures. As an example, WMM returns a constant magnetic field vector for all locations in Figure 1.1b despite the depicted change in vector strength from $\sim 40 \mu T$ to $\sim 70 \mu T$.

In this dissertation, we show how to improve the usability of magnetometers in locations where global maps like WMM are inaccurate. We do this by creating *local* magnetic field maps and showing how these local maps can improve IMU-based attitude estimates in regions where the WMM is inaccurate. Additionally, the distortions caused by human-made structures create spatial variations in the magnetic field that, when accurately mapped, can be utilized to improve position estimates of a vehicle. We expect our methods to be of most use in places where GPS is not available, WMM is not accurate, and there is sufficient spatial variation in the ambient magnetic field to localize position. Though we typically refer to such locations as “indoor” spaces, we acknowledge that not all indoor locations have these properties while some outdoor ones might.

The emphasis of our work is not that magnetometers and magnetic fields should replace other sensing modalities for GPS-denied navigation. Instead, we aim to highlight that magnetometers already exist on most sensing platforms yet find little use for most navigation-based applications. This dissertation demonstrates the potential and value of magnetic field measurements in GPS-denied navigation to improve sensor fusion methods.

1.2 Problem Statement

In this dissertation, we aim to address the following challenges:

1. Indoor magnetic fields are complex and difficult to model without full knowledge of ferrous materials and their location. How can we create magnetic field maps of indoor spaces and ensure they accurately represent the ambient magnetic field? In this dissertation, we propose gathering observations with an unmanned aerial vehicle (UAV) and interpolating them with Gaussian Process Regression (GPR) to make a GPR-based magnetic field map.

2. Magnetometers are used to assist attitude estimates outdoors but find little use for indoor and GPS-denied navigation. The inaccuracies of the World Magnetic Model (WMM) in indoor environments make it difficult to extract value from magnetometers inside. In this work, we investigate how local magnetic field maps might improve both position and attitude estimates of a flight vehicle.

3. Buildings create distortions in the magnetic field that are often treated as a disturbance. However, these distortions create spatial variation in the field that can improve our ability to localize position. This dissertation presents novel metrics and tools that show the relationship between magnetic field gradients and position localization accuracy.

1.3 Related Works

This dissertation aims to address problems related to creating accurate indoor magnetic field maps, using magnetometers to estimate attitude indoors, and comparing magnetic field spatial gradient to position localization accuracy. In this section, we present related works on each topic.

1.3.1 Indoor Magnetic Field Mapping

A magnetic field map must have an input dimension p and output dimension m, which we will denote as a $p \to m$ map. The choice of input dimension p often depends on the agent making the map. For example, $p = 2$ for wheeled robots [12, 13, 14, 15] and most pedestrian localization while $p = 3$ for UAVs [16, 17], multi-floor pedestrian localization [18], or specially-outfitted ground robots [19, 20]. The output m is set by how the map will be utilized. Works like [21, 22] use $m = 1$ for the orientation or magnitude of the magnetic field vector, while other works like [23, 19, 18] use all three vector components of the magnetic field with $m = 3$.

A number of works using indoor magnetic field maps are focused on pedestrian localization. Such works often utilize foot-mounted, or calf-mounted sensor suites [24, 25, 26] that can leverage zero-velocity updates to help with localization. Alternatively, many such works utilize the ubiquity of smartphones [27] and even consider constructing maps by crowdsourcing magnetic field measurements from many users [28]. To our knowledge, all pedestrian-localization-based magnetic field maps are $2 \to m$ (or "$2.5 \to m$" maps [26]), which assume a constant (or near-constant) altitude.

Ground vehicles are another common platform for indoor magnetic field mapping [29, 20, 30, 24]. These platforms benefit from wheel encoder odometry to assist their state estimates. Most of such works also create $2 \to m$ maps since the platform is confined to the ground. However, some researchers outfit their wheeled robots with a magnetometer on a vertical actuator [19] or vertically-spaced magnetometers [20] that let them create $3 \to m$ maps. The latter is a recent work by Hanley et al. showing that indoor magnetic fields can be quite sensitive to changes in altitude [20] and emphasizing some issues with $2 \to m$ planar magnetic field maps.

UAVs are used more regularly for *outdoor* magnetic field surveys than for *indoor* mapping. A recent review paper by Zhang et al. [31] presents works on outdoor magnetic field surveillance and also addresses methods of characterizing and suppressing UAV-induced magnetic noise. From the papers in their review, it is clear that the size and cost of the outdoor survey vehicles and sensors tend to be much larger. For example, [32] mounted a Geometrics G823A cesium magnetometer on a gas-powered helicopter and measured 800nT of magnetic variation caused by components on the vehicle when it is not powered. This reduced to 80nT and 40nT when the magnetometer was attached to a boom of length 0.5m and 1.2m, respectively. Additionally, they consider vehicles that cost \$2K - \$45K USD and weigh 7lbs - 51lbs. The larger vehicle size of outdoor magnetic surveys allows for interventions like suspending a magnetometer at the end of a 4.5m-long cable [33]. Thus, the vehicle size and project budget of outdoor surveying techniques allow for solutions to vehicle-induced magnetic noise (and magnetometers with better sensing capabilities) that are not applicable to (or not used on) indoor platforms. Zheng et al. cover other UAV noise characterization and mitigation efforts for outdoor magnetic surveys in Section 3 of ref. [31].

By contrast, some works use UAVs with magnetometers indoors, but do not create a map of the magnetic field for their implementation. Brzozowski et al. describe methods to *support* the use of indoor magnetic fields on UAVs in refs. [16, 17], but mostly present methods of gathering and visualizing magnetic field observations. They stop short of interpolating their magnetic field observations to create a queryable map or perform any state estimation. Furthermore, they do not actually use a flight vehicle (real or simulated) to gather or validate their magnetic field measurements. The authors in ref. [34] perform 3D position localization of a UAV near a two-wire power line setup in their lab space, but do not attempt to map the magnetic field near the wires in the process. Li et al. estimate the 6DOF pose of a UAV indoors using magnetometers (among other sensors), but only use the magnetometer for heading estimates and achieve this without a magnetic field map [35]. Zahran et al. leverage hall effect sensors (a type of magnetometer) in a clever way to estimate the velocity of a quadrotor and improve the dead reckoning of such flight vehicles (again, without magnetic field maps) [36].

There are works that use UAVs to map signals like ultra-wideband (UWB) radio [37] and visual light communication (VLC) [38]. Some works often use the term “fingerprinting” to describe the association of unique features in a reference signal to specific locations in a workspace. In this dissertation, we describe the same process as mapping.

To our knowledge, however, there is only one other paper that uses a UAV to make or leverage *indoor* magnetic field *maps*. Lipovsky et al. [39] uses observations from a UAV and cubic spline interpolation to create $3 \rightarrow 3$ maps of the DC, 50Hz, and 150Hz magnetic fields in a [5×5×2.5]m room. Ref. [39] mounts their magnetometer 40cm away from the drone based on lessons learned from their previous investigation on drone-induced magnetic noise [40].

In Chapter 2, we use a UAV to create $3 \rightarrow 3$ indoor magnetic field maps. A flight vehicle allows for quick and repeatable mapping and validation of an indoor volume but can also impede accurate measurements of the ambient field with magnetic noise generated by the motors and ESCs (electronic speed controllers). The focus of Chapter 2 is on identifying the type of magnetic noise our UAV creates, reducing the amount of magnetic noise, and finally incorporating any remaining measurable variations into the magnetic field map.

1.3.2 Attitude Estimation

Attempts to estimate attitude indoors tend towards detecting and rejecting perturbations in Earth’s magnetic field [41, 42, 43]. However, the methods from these works do not easily extend to full 3DOF roll, pitch, and yaw estimation and require multiple magnetometers. References [44, 45, 46, 47, 48] and others mitigate the effect of disturbances by adjusting the credibility of the magnetometer’s or accelerometer’s measurements when distortions in either are detected. Such works propose different methods of detecting disturbances, but generally assume that disturbances are brief and the magnetic field will return back to some expected value. Reference [49] relaxes this assumption by requiring only that the norm of the magnetic field remains constant over small time periods. However, Reference [49] does not provide 3DOF attitude estimation metrics of their indoor tests to understand how well this method works for indoor navigation. Reference [48] compares their disturbance rejection method to that of [49] and other works with brief, yet large, magnetic field disturbances.

In Chapter 3, we perform 3DOF attitude estimation using a gyroscope, an accelerometer, and a single magnetometer despite magnetic field disturbances from buildings and human-made structures. Using magnetic field maps, we leverage the spatial variation of the field to assist in 3DOF attitude estimation rather than rejecting these features as disturbances. Such capabilities are important for full pose estimation using magnetometers indoors since spatial variation in the magnetic field is desirable for position localization. Although we draw a distinction in our method from the disturbance rejection paradigm, our mapping-based methods can still be paired with the disturbance rejection methods if a magnetic field map is locally inconsistent in a region near doors, elevators, or other mobile ferromagnetic objects. Time-varying magnetic field mapping is addressed in [10], but not in this dissertation.

1.3.3 Position Localization

The position estimation accuracy of works that estimate position with the magnetic field range by a couple orders of magnitude. A direct comparison across all related works is difficult due to differences in the size of their workspaces and the varying sensors and algorithms used to compute ground-truth pose estimates.

Recent results from Almeida et al. in [30] give an average of 0.07m of 2D position error on a ground robot across three different test cases in an 8m×10m room. Their work has a 2D LIDAR sensor for ground-truth, uses a particle filter to estimate position, and presents estimation accuracy after their particle filter (PF) has converged (i.e., after the PF first achieves error within 0.1m). Similarly in 2013 Frassl, Robertson, et al. obtain 2D position estimates of 0.064m for their wheeled robot and 0.08m for a pedestrian walking through a 6m×10m room in [24]. They use the same room in a SLAM-based approach and obtain 0.01m to 0.2m of average error for foot-mounted pedestrian localization on the last 80% of their four test trials [25]. For both works, the authors use Vicon motion capture for ground-truth, a particle filter for position estimates, and initialize their PF with ground-truth position and attitude like in our work.

We suggest Refs. [30, 24, 25] provide the most direct and fair comparisons to our work in Chapter 4 given the size of the respective workspaces and their method of computing position estimation error throughout the trajectory. This is despite differences in the respective vehicles (i.e., wheeled robots and pedestrians vs UAV). Finally, we mention [50] by Lee et al. which achieves 0.1m of minimum mean squared error (MMSE) across four trials using a SLAM-based magnetic field mapping method on a wheeled robot. Though [50] also tests in rooms of similar size (e.g., 5m×4m), they compute their position estimation error at ∼10 known ground-truth locations instead of throughout their trajectories.

In Chapter 4, we use an accelerometer, altimeter, and magnetometer to estimate the position of a UAV in our indoor workspace. These position estimates are compared against errors in our magnetic map to investigate the cause of our position estimation errors. Here, we use an altimeter to demonstrate how another sensing modality can improve position estimates when the given magnetic field map is erroneous.

The core focus of Chapter 4 is to relate the gradients of the magnetic field to localization accuracy. In a workspace with a constant magnetic field in all locations, it would be impossible to localize position using the magnetic field alone. As such, we see the gradient of the magnetic field as a source of information and aim to analyze the relationship between the signal's spatial gradient and our ability to localize. To our knowledge, there are no other works that quantitatively compare the spatial gradient of an indoor magnetic field to the accuracy of their position localization. A closely-related set of works, however, utilize gradients of a signal as their fingerprints and obtain less spatial ambiguities in their maps [51], better motion tracking through a workspace [52], or better agreement in fingerprints across different devices [53]

1.4 Research Approach

This dissertation focuses on experimental analysis and validation of magnetic-field-based navigation techniques applied to small, low-cost unmanned aerial vehicles (UAVs). Some works like Ref. [20] have shown that the magnetic field inside buildings can vary significantly with changes in altitude. As such, we use a UAV to repeatably gather observations at various altitudes since it can be time consuming to do so by hand [19].

The approach and ability to navigate using magnetic fields depends on how much the magnetic field varies in a target area. If the magnetic field is constant in the region of interest, then there are well-known tools to incorporate magnetometer measurements to assist with attitude estimation. However, in this constant-field case, it is impossible to distinguish one location from another simply by measuring the magnetic field. Basically, it becomes difficult to track the vehicle's motion if the measured magnetic field is the same whether we remain stationary or move around. On the other hand, a local magnetic field that varies significantly within a workspace can yield rich, unique signatures that can improve position localization. However, indoor magnetic fields with such spatial variation are not accurately mapped by WMM making it more difficult to use magnetic fields for attitude estimation without local maps.

Since the approach to magnetic-field-based navigation is so dependent on the *spatial variation* of the magnetic field within a workspace, it is important to understand how much spatial variation one would experience in real environments. Simulations would enable a careful analysis of how valuable magnetic fields are for improving navigation and how that value relates to spatial variation in the field. For this, it can be tempting to simulate arbitrarily complex or trivially simple magnetic fields that show promising results. However, if these simulated fields would never be seen anywhere on Earth, then the simulation fails to demonstrate the value of magnetic fields in practice.

To our knowledge, there are not any rules or best practices on how to model, or even characterize, indoor magnetic fields. As such, this dissertation relies on gathering measurements of the magnetic field in various locations to better understand how much spatial variation exists in real-world magnetic fields and how to leverage the information stored in magnetic fields for indoor navigation. We believe our findings and data uncover targeted questions that can now be addressed more readily with modeling and simulation.

Anyone interested in developing best practices for modeling indoor magnetic fields may benefit from the following references that include open-source data on indoor magnetic fields [54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66] with reviews like [55] covering more details of such databases.

1.5 Contributions and Innovations

Contributions are best described as “things I worked hard on”. The contributions of this thesis are:

- Taking attitude estimation methods traditionally used for spacecraft in low Earth orbit and applying them to UAVs for indoor navigation.
- Several contributions to the `rc_pilot_a2sys` autopilot software developed by myself and other labmates in the A2sys (Autonomous Aerospace Systems) lab at University of Michigan.
- A comprehensive experimental analysis using data from 130+ test flights.
- Analytical tools to determine if position localization errors are caused by low spatial variation in the magnetic field, error in the GPR map, or magnetic ambiguities in nearby areas.
- An investigation showing that higher magnetic field spatial variation correlates with more accurate position estimates.
- An investigation showing magnetic field maps can improve 3DOF (degree of freedom) *attitude* estimates in regions of high magnetic field spatial variation but are less valuable in regions of low spatial variation.

Innovations are novel methods and tools. The innovations of this thesis are:

- GPR (Gaussian Process Regression)-based “compromise maps” that learn the time-varying magnetic biases from a UAV (unmanned aerial vehicle) over several flights without significantly increasing prediction time.

- A visualization method for 3D magnetic fields that show planar direction and downward-vector component of the magnetic field at various altitudes.
- A new magnetic gradient localization metric meant to predict the accuracy of position estimates using magnetic fields.
- A new "consistency metric" meant to indicate when magnetometer measurements should be used or rejected.

1.6 Thesis Outline

The outline of this dissertation is summarized in Figure 1.2.

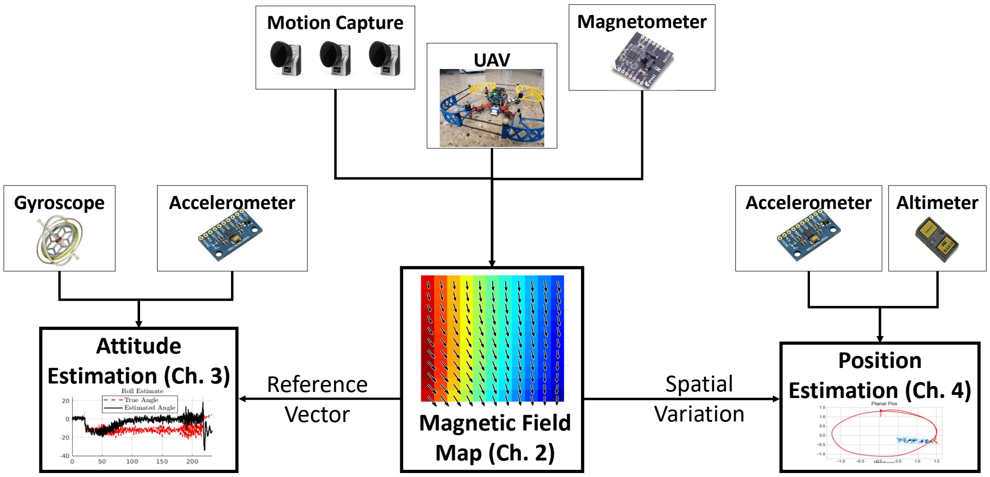


Figure 1.2: Dissertation outline.

Chapter 2 explains how to create 3D magnetic field maps using observations from a UAV and Gaussian Process Regression (GPR) for interpolation. The UAV helps by allowing us to quickly, and repeatably, gather observations of the magnetic field at various altitudes. However, it also creates magnetic noise that makes it difficult to map the ambient magnetic field. We discuss tools and procedures on how to identify, reduce, and incorporate UAV-induced magnetic noise into the maps.

In Chapter 3, we demonstrate the use of local magnetic field maps for attitude estimation using a Multiplicative Extended Kalman Filter. Here, we find that local maps are not valuable in outdoor locations (really regions accurately represented by WMM). However, when navigating indoors, local magnetic field maps become necessary for IMU-based attitude estimation since the magnetic field is significantly different from WMM estimates. This chapter is largely similar to Ref. [23] with some updates on lessons learned since then.

Chapter 4 investigates how spatial variation in the magnetic field related to position localization estimates from a particle filter. Here, we introduce metrics for describing magnetic spatial variation to show that position estimates often improve as the magnetic field gradient increases.

Finally, Chapter 5 summarizes our major findings and proposes ideas for future work.

CHAPTER 2

Enhanced Magnetic Field Mapping

2.1 Introduction

This chapter, outlined in Figure 2.1, aims to augment indoor sensor suites with magnetic field data by demonstrating how to create and validate magnetic field maps of indoor spaces. The key idea is to enable the creation of local indoor magnetic field maps that can then be leveraged to improve position and attitude estimates. The value of indoor magnetic field maps is demonstrated in works that use them to estimate the position of pedestrians and robotic vehicles inside buildings and to estimate the attitude of a drone indoors. As shown by ref. [19], gathering observations for a magnetic field map by hand is time-consuming and does not scale to mapping increasingly large spaces. As such, we propose using an unmanned aerial vehicle (UAV) to gather observations in the target volume of the workspace, using Gaussian Process Regression (GPR) to interpolate between these observations, then finally leveraging the map in future experiments within the mapped volume.

The contributions of this chapter are as follows.

1. The magnetic noise induced by the motors and ESCs on a UAV can be reduced by removing high-frequency commands from the flight controller to the motors.
2. Our UAV produces magnetic biases that occasionally change in magnitude and direction. We were unable to identify the underlying cause but found that distancing the magnetometer from the electronics, a commonly used intervention to reduce the measured magnetic field noise, decreased the measured variation of these biases.
3. Our "compromise map" can be trained on large datasets to prevent overfitting to a single flight test and learn the flight-by-flight variations of the drone without incurring high computation costs when predicting the magnetic field.
4. We find that our compromise map has similar accuracy if the location of training observations is no further than 0.55m apart. This agrees with a similar, qualitative study by Akai and Ozaki in ref. [19].

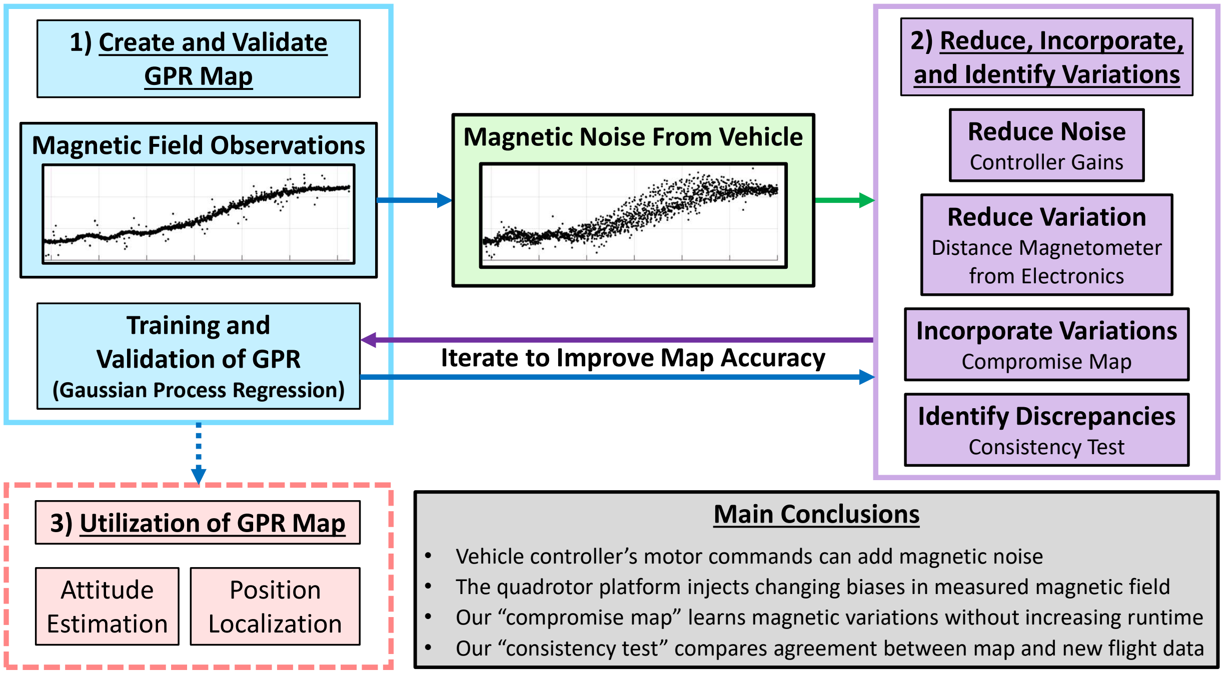


Figure 2.1: Methods to reduce, incorporate, and identify vehicle-induced magnetic field noise. Our suggested interventions can be progressively added until the user's GPR map is sufficiently accurate for the application.

5. We introduce a consistency test to indicate if a given GPR map has good agreement with new measurements from a subsequent flight test. This consistency test was important for contextualizing our position localization accuracy in Chapter 4.

6. In a comparison of two methods of mapping the norm (magnitude) of the magnetic field, we find (empirically) that creating a specialized map on the norm of the field is equivalent to taking the norm of estimates from a vector-valued magnetic field map.

The remainder of this chapter is structured as follows. Section 2.2 presents related works, and Section 2.3 introduces our mathematical notation for Gaussian process regression and our compromise map. Section 2.4 introduces our UAV and testing methodology for this chapter. Finally, Section 2.5 presents our results, while conclusions and future work appear in Section 2.6.

2.2 Related Works

A magnetic field map must have an input dimension p and output dimension m, which we will denote as a $p \rightarrow m$ map. The choice of input dimension p often depends on the agent making the map. For example, $p = 2$ for wheeled robots [12, 13, 14, 15] and most pedestrian localization while $p = 3$ for UAVs [16, 17], multi-floor pedestrian localization [18], or specially-outfitted ground robots [19, 20]. The output m is set by how the map will be utilized. Works like [21, 22] use $m = 1$

for the orientation or magnitude of the magnetic field vector, while other works like [23, 19, 18] use all three vector components of the magnetic field with $m = 3$.

A number of works using indoor magnetic field maps are focused on pedestrian localization. Such works often utilize foot-mounted, or calf-mounted sensor suites [24, 25, 26] that can leverage zero-velocity updates to help with localization. Alternatively, many such works utilize the ubiquity of smartphones [27] and even consider constructing maps by crowdsourcing magnetic field measurements from many users [28]. To our knowledge, all pedestrian-localization-based magnetic field maps are $2 \to m$ (or "$2.5 \to m$" maps [26]), which assume a constant (or near-constant) altitude.

Ground vehicles are another common platform for indoor magnetic field mapping [29, 20, 30, 24]. These platforms benefit from wheel encoder odometry to assist their state estimates. Most of such works also create $2 \to m$ maps since the platform is confined to the ground. However, some researchers outfit their wheeled robots with a magnetometer on a vertical actuator [19] or vertically-spaced magnetometers [20] that let them create $3 \to m$ maps. The latter is a recent work by Hanley et al. showing that indoor magnetic fields can be quite sensitive to changes in altitude [20] and emphasizing some issues with $2 \to m$ planar magnetic field maps.

UAVs are used more regularly for *outdoor* magnetic field surveys than for *indoor* mapping. A recent review paper by Zhang et al. [31] presents works on outdoor magnetic field surveillance and also addresses methods of characterizing and suppressing UAV-induced magnetic noise. From the papers in their review, it is clear that the size and cost of the outdoor survey vehicles and sensors tend to be much larger. For example, [32] mounted a Geometrics G823A cesium magnetometer on a gas-powered helicopter and measured 800nT of magnetic variation caused by components on the vehicle when it is not powered. This reduced to 80nT and 40nT when the magnetometer was attached to a boom of length 0.5m and 1.2m, respectively. Additionally, they consider vehicles that cost \$2K - \$45K USD and weigh 7lbs - 51lbs. The larger vehicle size of outdoor magnetic surveys allows for interventions like suspending a magnetometer at the end of a 4.5m-long cable [33]. Thus, the vehicle size and project budget of outdoor surveying techniques allow for solutions to vehicle-induced magnetic noise (and magnetometers with better sensing capabilities) that are not applicable to (or not used on) indoor platforms. Zheng et al. cover other UAV noise characterization and mitigation efforts for outdoor magnetic surveys in Section 3 of ref. [31].

By contrast, some works use UAVs with magnetometers indoors, but do not create a map of the magnetic field for their implementation. Brzozowski et al. describe methods to *support* the use of indoor magnetic fields on UAVs in refs. [16, 17], but mostly present methods of gathering and visualizing magnetic field observations. They stop short of interpolating their magnetic field observations to create a queryable map or perform any state estimation. Furthermore, they do not actually use a flight vehicle (real or simulated) to gather or validate their magnetic field measurements. The authors in ref. [34] perform 3D position localization of a UAV near a two-wire power line setup in their lab space, but do not attempt to map the magnetic field near the wires in the process. Li et al. estimate the 6DOF pose of a UAV indoors using magnetometers (among other sensors), but only use the magnetometer for heading estimates and achieve this without a magnetic field map [35]. Zahran et al. leverage hall effect sensors (a type of magnetometer) in a clever way to estimate the velocity of a quadrotor and improve the dead reckoning of such flight vehicles (again, without magnetic field maps) [36].

There are works that use UAVs to map signals like ultra-wideband (UWB) radio [37] and visual light communication (VLC) [38]. Some works often use the term "fingerprinting" to describe the association of unique features in a reference signal to specific locations in a workspace. In this dissertation, we describe the same process as mapping.

To our knowledge, however, there are only two works that use a UAV (drone, multirotor, flight vehicle, etc...) to make or leverage *indoor* magnetic field *maps*. One is Chapter 3 of this dissertation which uses Gaussian process regression (GPR) to make $3 \to 3$ magnetic field maps of the DC magnetic field in an indoor workspace of dimensions [4×3×2.25]m. The other work, by Lipovsky et al. [39] uses observations from a UAV and cubic spline interpolation to create $3 \to 3$ maps of the DC, 50Hz, and 150Hz magnetic fields in a [5×5×2.5]m room. Ref. [39] mounts their magnetometer 40cm away from the drone based on lessons learned from their previous investigation on drone-induced magnetic noise [40].

In this chapter, we use a UAV to create $3 \to 3$ indoor magnetic field maps. A flight vehicle allows for quick and repeatable mapping and validation of an indoor volume, but can also impede accurate measurements of the ambient field with magnetic noise generated by the motors and ESCs (electronic speed controllers). The focus of this chapter is on identifying the type of magnetic noise our UAV creates, reducing the amount of magnetic noise, and finally incorporating any remaining measurable variations into the magnetic field map.

Additionally, we study how the spatial density of magnetic observations affect the accuracy of a magnetic field map (similar to a study done in [19]) and empirically demonstrate the equivalence of two maps that estimate the norm of the ambient magnetic field. Finally, we introduce a consistency check to reason about when a GPR-based magnetic field map is appropriate to use for state estimation.

The interventions and studies done in this chapter are lessons we learned while investigating the 3D position localization of a UAV using a low-cost IMU and magnetometer (Chapter 4). It is important to emphasize that some of our interventions may not be needed for other applications of indoor magnetic field maps (e.g., they were not applied for our study on attitude estimation using indoor magnetic field maps in Chapter 3).

2.3 Mathematical Preliminaries and Methodology

This section introduces our use of Gaussian Process Regression (GPR) to interpolate a set of magnetic field observations in our workspace. Special notation is used to distinguish a set of n_2 observations used to train hyperparameters and a separate set of n_1 observations used to perform inference. Additionally, we introduce performance metrics used later in our analysis of our GPR-based magnetic field maps.

We first introduce some notation. A single measurement of the magnetic field at an unspecified location is $\tilde{\boldsymbol{y}} \in \mathbb{R}^3$ with the x, y, and z components of this measurement denoted as $\tilde{y}_x$, $\tilde{y}_y$, $\tilde{y}_z \in \mathbb{R}$ respectively. In general, a subscript of $_x$, $_y$, or $_z$ denotes that respective component of the magnetic field while an overhead tilde $\tilde{}$ denotes a *measured* value. $\tilde{\boldsymbol{Y}}^n \in \mathbb{R}^{3\times n}$ is a set of n magnetic field measurements while $\tilde{\boldsymbol{Y}}_z^n \in \mathbb{R}^n$ denotes the z component of each magnetic field measurement in the set. Similarly, $\boldsymbol{X}^n \in \mathbb{R}^{3\times n}$ is a collection of n spatial locations in our workspace. The *predicted* or *estimated* magnetic field $\hat{\boldsymbol{m}}$ at some location $\boldsymbol{r} \in \mathbb{R}^3$ is denoted as $\hat{\boldsymbol{m}}(\boldsymbol{r}) \in \mathbb{R}^3$ and the x component of this prediction is $\hat{m}_x(\boldsymbol{r}) \in \mathbb{R}$. Thus, a collection of n magnetic field estimates will be $\hat{\boldsymbol{M}}^n(\boldsymbol{X}^n) \in \mathbb{R}^{3\times n}$. Similarly, $\hat{\boldsymbol{M}}^n$ is a set of predictions of the magnetic field where the location of these predictions is arbitrary and $\hat{\boldsymbol{M}}_y^n \in \mathbb{R}^n$ gives just the predictions from $\mathcal{GP}_y$. Generally, an overhead hat $\hat{}$ denotes a prediction/estimate while a superscript integer n denotes the number of measurements/predictions/locations of the respective matrix.

2.3.1 Gaussian Process Regression

Gaussian Process Regression (GPR) is a machine learning tool that can be used to estimate a signal given a set of noisy measurements. Here, we use GPR as the backbone of our magnetic field map leveraging methods from Rasmussen et al. in [67]. The goal is to have a magnetic field map of the flight workspace which will provide an estimated magnetic field vector at any location in the working volume.

Throughout this dissertation, we use GPR to create three separate $3 \rightarrow 1$ maps each responsible for estimating the x, y, and z components of the magnetic field vector in our workspace. To create the GPR-based map, we first gather n observations of the magnetic field throughout the workspace. Observation sets give the three components of the *measured* magnetic field $\tilde{\boldsymbol{Y}}_x^n$, $\tilde{\boldsymbol{Y}}_y^n$, $\tilde{\boldsymbol{Y}}_z^n \in \mathbb{R}^n$ at each 3D position in our design matrix $\boldsymbol{X}^n \in \mathbb{R}^{3\times n}$. Together, these quantities define the training sets $D_x = (\boldsymbol{X}, \tilde{\boldsymbol{Y}}_x)$, $D_y = (\boldsymbol{X}, \tilde{\boldsymbol{Y}}_y)$, and $D_z = (\boldsymbol{X}, \tilde{\boldsymbol{Y}}_z)$ for the x, y, and z GPRs respectively.

In Ref. [67], Rasmussen and Williams define a Gaussian process as a distribution over functions written as

$$f(\boldsymbol{r}) = \mathcal{N}\left(\bar{f}, \mathbb{V}(f)\right) \sim \mathcal{GP}\left(\boldsymbol{0}, k(\boldsymbol{r}, \boldsymbol{r}')\right) \tag{2.1}$$

where $\boldsymbol{0}$ is the zero-mean function, k is the covariance function (or kernel) and $\boldsymbol{r}$, $\boldsymbol{r}' \in \mathbb{R}^3$ are 3D positions. This dissertation uses the squared exponential covariance function

$$k(\boldsymbol{r}, \boldsymbol{r}') = \sigma_f^2 \exp\left(-\frac{1}{2l^2}(\boldsymbol{r} - \boldsymbol{r}')^\top(\boldsymbol{r} - \boldsymbol{r}')\right) + \sigma_n^2 \tag{2.2}$$

with σ_f as the signal variance, l as the length scale, and σ_n as the sensor noise variance which define our set of *hyperparameters* $\boldsymbol{\Theta} = \{\sigma_f, l, \sigma_n\}$ for this kernel. In essence, the squared exponential kernel is a measure of similarity between two 3D locations $\boldsymbol{r}$ and $\boldsymbol{r}'$ (scaled by the hyperparameters) and will serve as a weighting term in the inference of the magnetic field at unobserved locations.

Recall that we perform regression using three Gaussian processes $\mathcal{GP}_x$, $\mathcal{GP}_y$, and $\mathcal{GP}_z$ to represent the full magnetic field anywhere in the workspace. As such, we need three sets of hyperparameters ($\boldsymbol{\Theta}_x$, $\boldsymbol{\Theta}_y$, $\boldsymbol{\Theta}_z$) each computed using the Gaussian Processes for Machine Learning `gpml-matlab` [1] toolbox by minimizing the negative log marginal likelihood of the respective training sets D_x, D_y, and D_z. We define hyperparameters optimized over the observations in a training set D_x as $\boldsymbol{\Theta}_x^*(D_x)$.

Now, to estimate the magnetic field at some location $\boldsymbol{r}^*$, we need a training set D and a set of hyperparameters $\boldsymbol{\Theta}$ corresponding to a selected kernel k. A squared exponential kernel k using optimal hyperparameters $\boldsymbol{\Theta}^*(D)$ is denoted as $k_{\boldsymbol{\Theta}^*(D)}(\boldsymbol{r}, \boldsymbol{r}^*)$ where the subscript on the kernel k specifies which set of hyperparameters are used and $\boldsymbol{r} \in \boldsymbol{X}$ is a location in the training set.

This is the extent of our notation, but an example is helpful. Say we want to estimate the z component of the magnetic field $\hat{m}_z(\boldsymbol{r}^*)$ at some location $\boldsymbol{r}^*$. For this, $\mathcal{GP}_z$ requires a kernel k to compare the target location $\boldsymbol{r}^*$ against locations $\boldsymbol{X}^{n_1}$ in its training set $D_z^{n_1} = (\boldsymbol{X}^{n_1}, \tilde{\boldsymbol{Y}}_z^{n_1})$. Additionally, its hyperparameters $\boldsymbol{\Theta}^*(D_z^{n_2})$ are optimized over a separate set of observations $D_z^{n_2} = (\boldsymbol{X}^{n_2}, \tilde{\boldsymbol{Y}}_z^{n_2})$. Mathematically, we express this as

$$\begin{aligned} \hat{m}_z(\boldsymbol{r}^*) &= GPR_z(\boldsymbol{r}^*) \\ &= \mathbb{E}\left[\mathcal{GP}_z\left(\boldsymbol{0},\ K_{\boldsymbol{\Theta}_z^*(D_z^{n_2})}(\boldsymbol{X}^{n_1}, \boldsymbol{r}^*)\right)\right] \end{aligned} \tag{2.3}$$

where a matrix $K(\boldsymbol{X}, \boldsymbol{r}^*)$ has scalar elements $k\left(\boldsymbol{X}^{\{i\}}, \boldsymbol{r}^*\right)$ with each $\boldsymbol{X}^{\{i\}} \in \mathbb{R}^3$ a column of $\boldsymbol{X} \in \mathbb{R}^{3\times n}$ while $\hat{m}_z(\boldsymbol{r}^*) \in \mathbb{R}$. Finally, $GPR_z(\boldsymbol{r}^*)$ is a shorthand for the predicted magnetic field value of the z-component Gaussian process $\mathcal{GP}_z$ at $\boldsymbol{r}^*$.

The reason for the verbosity in Equation 2.3 is to allow an observation set for optimizing hyperparameters $D_z^{n_2} = (\boldsymbol{X}^{n_2}, \tilde{\boldsymbol{Y}}_z^{n_2})$ and a separate "inference set" $D_z^{n_1} = (\boldsymbol{X}^{n_1}, \tilde{\boldsymbol{Y}}_z^{n_1})$ for computing the similarity weights when predicting the magnetic field at some location $\boldsymbol{r}^*$. This is important because hyperparameter optimization, typically done offline, scales with $\mathcal{O}\left({n_2}^3\right)$ while inference scales with

---

[1] `gpml-matlab` was created by Carl Edward Rasmussen and Hannes Nickisch `http://www.gaussianprocess.org/gpml/code/matlab/doc/` accessed on September 2021

$\mathcal{O}\left({n_1}^2\right)$. With this formulation, our map can encode information from a large set of flights offline, then use a subset of these observations during the actual inference to reduce computational load. We leverage this idea in Section 2.4.4 to create and query our magnetic field map.

2.3.2 Magnetometer Calibration

To calibrate our vehicle's magnetometer, we use the model and iterative least squares solver from Ref. [68] along with the two-step calibration procedure from Ref. [69]. The magnetometer model from [68] is repeated here

$$\tilde{B}_x = \theta_a B_x + \theta_{x_0} + \eta_x \tag{2.4}$$

$$\tilde{B}_y = \theta_b(\tilde{B}_y \cos(\theta_\rho) + B_x \sin(\theta_\rho)) + \theta_{y_0} + \eta_y \tag{2.5}$$

$$\begin{aligned}\tilde{B}_z = \theta_c(&B_x \sin(\theta_\lambda) + B_y \sin(\theta_\phi)\cos(\theta_\lambda) \\ &+ B_z \cos(\theta_\phi)\cos(\theta_\lambda)) + \theta_{z_0} + \eta_z\end{aligned} \tag{2.6}$$

where ($\tilde{B}_x$, $\tilde{B}_y$, $\tilde{B}_z$) are measured magnetic field values, (B_x, B_y, B_z) are the true magnetic field, and the parameters to solve for are bias ($\theta_{x_0}, \theta_{y_0}, \theta_{z_0}$), scale factor ($\theta_a, \theta_b, \theta_c$), and non-orthogonality terms ($\theta_\rho, \theta_\lambda, \theta_\phi$).

Ultimately, we aim to find parameters $\theta = [\theta_a, \theta_b, \theta_c, \theta_{x_0}, \theta_{y_0}, \theta_{z_0}, \theta_\rho, \theta_\lambda, \theta_\phi]^\top$ by minimizing the following cost function via an iterative non-linear least squares solver

$$\begin{aligned}\Delta B &= B_R^2 - B^2 \\ &= B_R^2 - (B_x^2 + B_y^2 + B_z^2) \\ &= B_R^2 - g(\tilde{B}_x, \tilde{B}_y, \tilde{B}_z, \alpha)\end{aligned} \tag{2.7}$$

where B_R is the reference magnetic field strength (taken from the World Magnetic Model) and $g()$ is obtained by solving Equations 2.4 - 2.6 for (B_x, B_y, B_z).

The key difference from Springmann's method [68] is to estimate the nine θ parameters in two separate minimization steps like Wu et al. demonstrate in [69]. We found that this two-step approach gave more consistent results for another magnetometer (not used in this dissertation) on our UAV. As such, we adopted the two-step calibration technique for our primary magnetometer as well.

First, a simplified magnetometer model is created by setting all scaling terms to 1 and non-orthogonalities to 0 leaving just the bias terms ($\theta_{x_0}, \theta_{y_0}, \theta_{z_0}$) in Equations 2.4 - 2.6. The first minimization is done on this simplified model to find an optimal set of bias terms ($\theta^*_{x_0}, \theta^*_{y_0}, \theta^*_{z_0}$) which are in turn used as initial conditions for the second optimization where $\theta_0 = [1, 1, 1, \theta^*_{x_0}, \theta^*_{y_0}, \theta^*_{z_0}, 0, 0, 0]^\top$. The nine optimal parameters from this second optimization are used as the calibration terms for the magnetometer.

2.3.3 Performance Metrics

We use the root mean squared error (RMSE) of each GPR's prediction $\hat{m}_x$ against the corresponding component of the measured field $\tilde{y}_x$ across a validation set. With this, the RMSE for $\mathcal{GP}_x$ over some validation set $D_x^{n_v} = (\boldsymbol{X}^{n_v}, \tilde{\boldsymbol{Y}}_x^{n_v})$ is defined as

$$RMSE_x = \sqrt{\frac{1}{n_v}\sum_{i=1}^{n_v}(\hat{m}_x^{\{i\}} - \tilde{y}_x^{\{i\}})^2} \tag{2.8}$$

where $\hat{\boldsymbol{M}}_x^{n_v} = GPR_x(\boldsymbol{X}^{n_v})$ is the x component of the predicted magnetic vector, $\hat{m}_x^{\{i\}}$ is the ith prediction in $\hat{\boldsymbol{M}}_x^{n_v}$, and $\tilde{y}_x^{\{i\}}$ is the corresponding ith measurement in $\tilde{\boldsymbol{Y}}_x^{n_v}$.

Although we have three GPRs, it is sometimes convenient to summarize the accuracy of their composite estimate as

$$RMSE_{norm} = \sqrt{{RMSE_x}^2 + {RMSE_y}^2 + {RMSE_z}^2}. \tag{2.9}$$

2.4 Experimental Procedure

This section introduces our quadrotor unmanned aerial vehicle (UAV), testing locations, and flight trajectories used to construct and validate our magnetic field maps.

2.4.1 Equipment, Facilities, and Setup

All tests for this chapter were conducted at the Robot Fly Lab in the University of Michigan Ford Motor Company Robotics Building. This indoor flight arena is equipped with eight OptiTrack motion capture cameras that give a working volume of 4m × 3m × 2.25m as shown in Figure 2.2. A ground station computer connected to the OptiTrack system provides ground-truth position and attitude estimates of the vehicle at 120Hz. The communication setup for the ground station, pilot transmitter, and flight vehicle is the same as that depicted in Figure 3.3 of Chapter 3, but with one less BeagleBone Blue on the UAV.

The motion capture pose estimate is streamed to the flight vehicle in real-time, but with a communication latency of about 40ms. For this, the vehicle's onboard estimate of roll and pitch is used for control while the remaining states (position and heading) are taken from the 40ms-delayed motion capture packets.

Time synchronization is important to properly associate each onboard magnetometer measurement with a ground-truth pose when creating and validating our magnetic field maps. During flights,

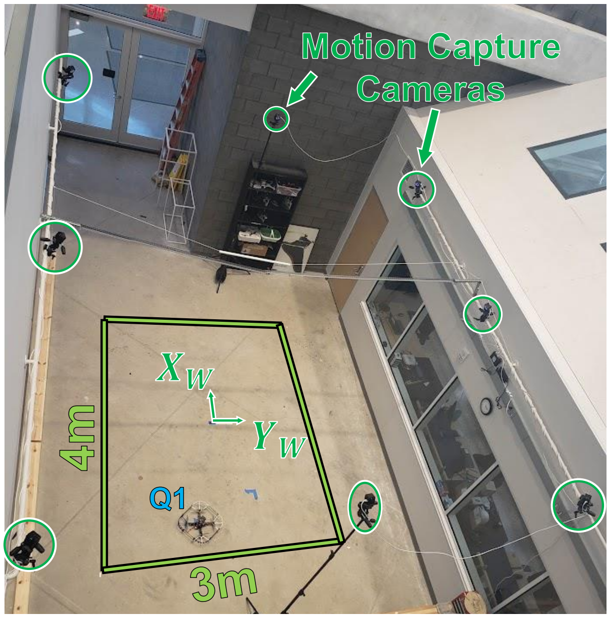


Figure 2.2: Flight vehicle Q1 in the Robot Fly Lab within the Ford Motor Company Robotics Building at the University of Michigan.

our UAV's BeagleBone Blue is synchronized with the ground station laptop and ground station BeagleBone Green using a Linux tool called `chrony` [2] .

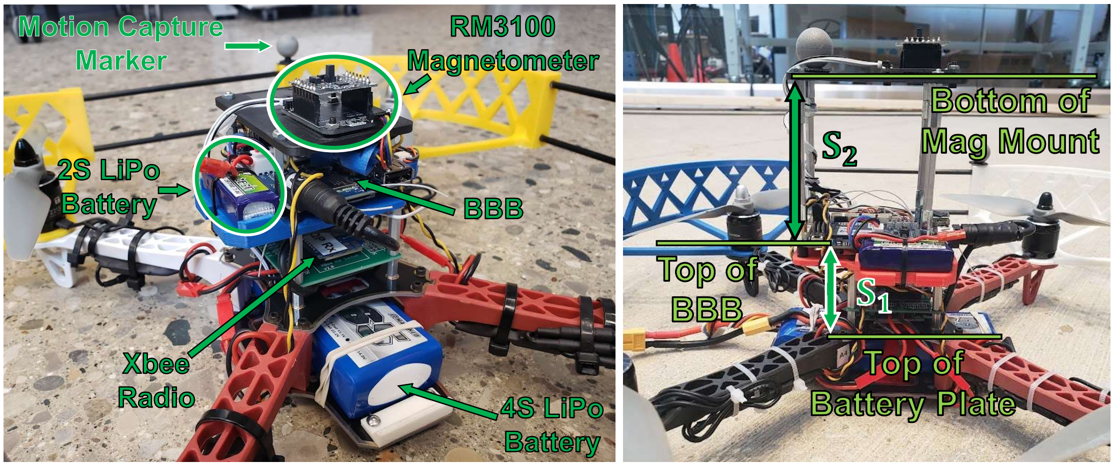
In post-processing, this synchronization allows us to associate the non-delayed motion capture data with the drone's magnetometer readings. Additionally, we use motion capture pose (position and attitude) estimates to rotate the magnetometer data into the world frame and determine the location each magnetic field measurement was collected.

This dissertation uses time-invariant magnetometer calibration combining techniques from Refs. [68] and [69] (see Section 2.3.2). Magnetometer calibration is performed outdoors, just South of the University of Michigan's outdoor netted flight facility, M-Air. This location is far enough away from any buildings that the magnetic field strength is constant over a few meters and should be accurately reflected by the World Magnetic Model (WMM) [11]. For the data gathered on September 1st, 2022 (test series t6), the ambient magnetic field reference term used in our calibration is $B_R = 53.1351\mu T$ as taken from a WMM online calculator for M-Air's location at 42.294431°N, 83.710442°W, and 270m above sea level.

Our flight vehicle "Q1" (Quadrotor one) is shown in Figure 2.3 and has an RM3100 magnetometer that is sampled at 200Hz and an MPU9250 IMU (gyroscope, accelerometer, and magnetometer) sampled at 200Hz. This dissertation made no use of the magnetometer on the MPU9250 relying solely on the RM3100 for all magnetic field measurements. The BeagleBone Blue (BBB) micropro-

---

[2] `chrony` was created by the Red Hat Software company `https://chrony.tuxfamily.org/` accessed on February 2022



(**a**) Main components of Q1. S_2 = 2cm. (**b**) Definition of S_1 and S_2. S_2 = 8cm.

Figure 2.3: Q1: The flight vehicle used for experiments in this chapter.

cessor handles communication with the ground station via Xbee radio, logs all sensor data locally, and commands the four motors to achieve the desired flight trajectory. A Turnigy 4S 2650mAh 20C LiPo battery provides the main power on Q1 while a Turnigy 2S 300mAh LiPo battery keeps the BBB powered on when swapping 4S batteries between flights.

Figure 2.3b shows Q1 with different distances from the magnetometer to the rest of the vehicle's electronics. S_1 is defined as the distance from the top of the battery mounting plate to the top of the BBB. This distance is S_1 = 4.79cm on Q1 and will not change in any experiments through this chapter. S_2 is an adjustable distance defined from the top of the BBB to the bottom of the magnetometer's mounting plate that allows us to investigate how the magnetometer's proximity to other electronics affects the consistency of magnetic field measurements on a quadrotor.

Finally, we placed a stationary RM3100 magnetometer in the corner of the flight lab that gathers data every 10 seconds (0.1Hz). This stationary sensor is on the ground at approximately (-3.5m, +2.5m) in the (X_W, Y_W) frame defined in Figure 2.2 and was used to confirm that the ambient magnetic field remained constant during our experiments.

2.4.2 Flight Profiles

The flight tests conducted for this chapter are all designed around single-altitude scanning patterns that gather observations at a planar slice of the working volume. Figure 2.4 shows the desired trajectory (blue, solid line) has 4m x-axis strides separated by 0.25m y-axis strides. Since all flights begin with the drone at the origin (0m, 0m) of our working volume, the trajectory in Figure 2.4 includes diagonal strides to leave and return to the origin at the start and end of each flight test. Finally, some trajectories gather observations at multiple altitudes in a single flight and include z-axis strides between each planar trajectory. For multi-altitude tests, the drone traverses a

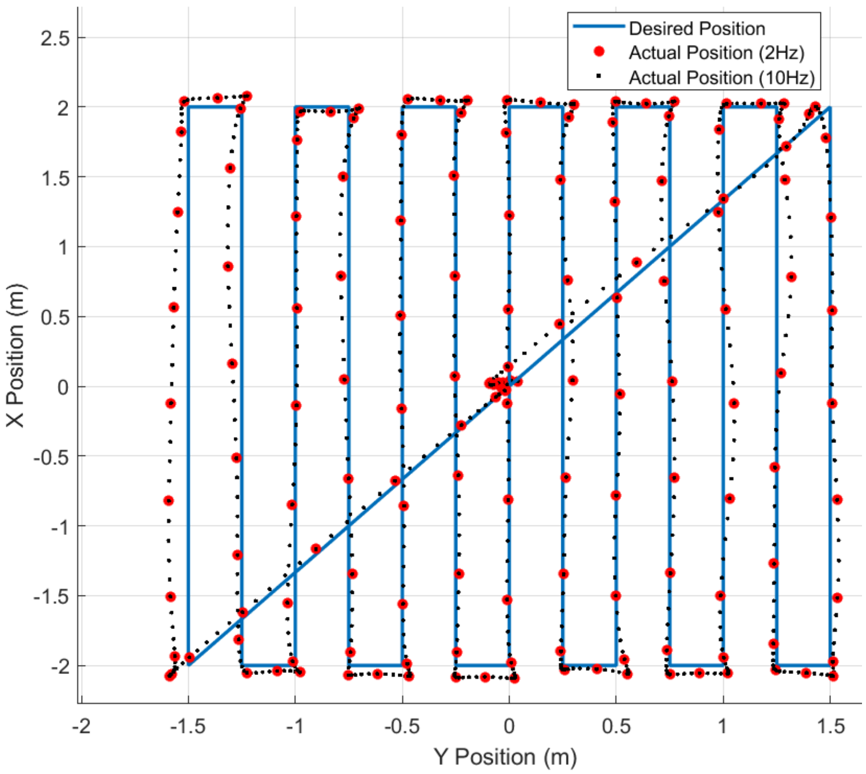


Figure 2.4: Single-altitude lawnmower trajectory used in all trajectories in this chapter.

single diagonal stride from the origin to the corner at the first altitude, and another diagonal stride from the corner to the origin at the final altitude.

All linear strides are created from quintic spline trajectories that are hand-tuned to enforce a 1.9m/s speed limit which is a compromise between the desire for shorter flight tests (that allow for more frequent data collection) and the limitations of our flight controller.

The 0.25m y-axis spacing is selected to enable a study on how the spatial density of observations used to train a GPR-based map affects the accuracy of the said map (Section 2.5.4). A similar study was done in ref. [19] where magnetic field observations in a 3m × 3m × 2.2m volume are hand-gathered at 0.2m increments in all axes. For our study, having 0.2m y-axis and 0.2m z-axis spacing that would scan all altitudes in our working volume created a flight trajectory duration that exceeded the maximum flight time of our vehicles. Thus, a 0.25m minimum separation distance is used in this chapter instead.

2.4.2.1 Flight Profile Nomenclature

Throughout this dissertation, we refer to flights with ID tags like "tY_XX" where *Y* refers to the flight test *series* and *XX* is the two-digit ID of the flight test in that series. This allows the reader to reference our raw experimental data to improve the reproducibility of our results. Appendix A has more details on the matter.

2.4.3 Pre-Processing Magnetic Observations

When creating our hyperparameter (D^{n_2}) and inference sets (D^{n_1}), we first pre-process the magnetic field data gathered by the RM3100.

Figure 2.5 shows data from the UAV's RM3100 for a segment of time when the drone is not moving and the motors are not spinning. The black dots are raw magnetometer measurements that are mostly constant, but frequently have spurious measurements that vary by $\pm 3\mu$T for each axis respectively. We have seen such spurious measurements from three different RM3100 sensors; even one mounted on a platform with no motors or ESCs (electronic speed controllers). Other works using RM3100s have not reported such spurious readings [20, 70]. Thus, we believe this could be a result of our 200Hz sampling rate (600Hz for the whole device to achieve 200Hz per axis) and something with the RM3100's firmware that does not like such a fast sampling rate. Alternatively, this could be caused by some poor voltage regulation on the BeagleBone Blue. We have not tried sampling from other microprocessors to see if the BeagleBone is the cause of this issue.

Nonetheless, all RM3100 data we use throughout this dissertation is first put through a moving median filter with a window size of 5. The resultant observations from the median filter are shown in red in Figure 2.5. Clearly, there are still some outliers in red, but increasing the window size of the moving median filter to remove these outliers delays the signal enough to cause problems for state estimation. Since we want to construct and validate maps with the same pre-processing steps used when performing state estimation, we fix a window size of 5 and work with the remaining outliers.

In addition to a median filter, we downsample the 200Hz of data gathered during the flight test. Since the drone does not move very far in 1/200th of a second, many observations are spatially redundant. As such, we temporally downsample our observations to 2Hz, 4Hz or 10Hz when creating or validating our GPR maps. Figure 2.4 shows the spatial distribution of 2Hz and 10Hz downsampling during a flight test.

Since motion capture data is used to rotate the magnetometer data from the body frame to the world frame, we perform one final check to ensure a timely ground-truth pose estimate. Any temporally downsampled observation with a motion capture pose older than 50ms is considered "stale" and replaced with the nearest (in time) observation that has a more timely motion capture pose. If the "fresh" replacement is already in the downsampled observation set, the redundant observation is removed leaving one instance of it in the final set.

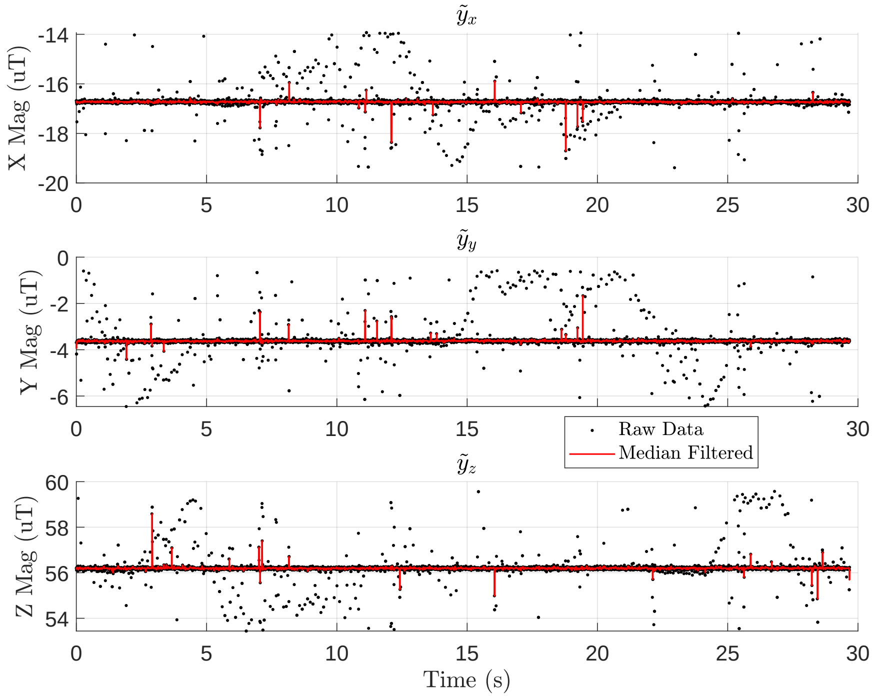


Figure 2.5: Data from a stationary RM3100 placed at the origin of our flight arena and sampled at 200Hz. Spurious measurements are reduced with a moving median filter with a window size of 5.

2.4.4 Creating and Querying the "Compromise" Map

As mentioned in Section 2.3.1, creating a GPR-based map requires each Gaussian process $\{\mathcal{GP}_x, \mathcal{GP}_y, \mathcal{GP}_z\}$ to have a set of hyperparameters optimized over an observation set $\{\boldsymbol{\Theta}_x^*(D_x^{n_2}),$ $\boldsymbol{\Theta}_y^*(D_y^{n_2}),\ \boldsymbol{\Theta}_z^*(D_z^{n_2})\}$ and an inference set $\{D_x^{n_1}, D_y^{n_1}, D_z^{n_1}\}$ to compare the target locations. This section explains how we use flight data to create the hyperparameter observation sets (with n_2 observations) and then use inference from an intermediate magnetic field map to make "compromise" inference sets (with n_1 observations) where $n_1 < n_2$. The goal of the compromise map is to leverage the flight-by-flight variation in magnetic field measurements (Section 2.5.2) by training hyperparameters on n_2 observations from many flights without incurring the computational cost of having a large inference set size n_1.

The "hyperparameter observation set" for the y component Gaussian process $\mathcal{GP}_y$ is $D_y^{n_2} = (\boldsymbol{X}^{n_2}, \tilde{\boldsymbol{Y}}_y^{n_2})$ and similar for $D_x^{n_2}$ and $D_z^{n_2}$. All three hyperparameter observation sets share the same observed locations $\boldsymbol{X}^{n_2} \in \mathbb{R}^{3\times n_2}$ taken from downsampled observations from any number of *training* flights.

These observations are used to compute $\{\boldsymbol{\Theta}_x^*(D_x^{n_2}),\ \boldsymbol{\Theta}_y^*(D_y^{n_2}),\ \boldsymbol{\Theta}_z^*(D_z^{n_2})\}$ by minimizing the negative log marginal likelihood of the observations over the hyperparameters of each respective Gaussian process. The optimal hyperparameters, along with their corresponding observation sets of size n_2, will serve as an "intermediate" magnetic field map.

From here, we query our intermediate map at n_1 user-selected locations to generate n_1 magnetic field observations for the compromise map. The result of this step is a set of n_1 locations around the workspace that make up $\boldsymbol{X}^{n_1} \in \mathbb{R}^{3 \times n_1}$. The magnetic field "measurements" that complete the inference set $D_x^{n_1} = (\boldsymbol{X}^{n_1}, \hat{\boldsymbol{M}}_x^{n_1})$ come from the estimated magnetic field values at each location in $\boldsymbol{X}^{n_1}$ where

$$\hat{\boldsymbol{M}}_x^{n_1} = \mathbb{E}\left[\mathcal{GP}_x\left(\mathbf{0},\ K_{\boldsymbol{\Theta}_x^*(D_x^{n_2})}(\boldsymbol{X}^{n_2}, \boldsymbol{X}^{n_1})\right)\right] \in \mathbb{R}^{n_1} \tag{2.10}$$

and similarly for the magnetic field "measurements" $\hat{\boldsymbol{M}}_y^{n_1}$ and $\hat{\boldsymbol{M}}_z^{n_1}$ for inference sets $D_y^{n_1}$ and $D_z^{n_1}$ respectively. Here, the GPR maps estimate the magnetic field at locations $\boldsymbol{X}^{n_1}$ by computing their similarity to locations $\boldsymbol{X}^{n_2}$. The notation used in Equation 2.10 was introduced in Section 2.3.1.

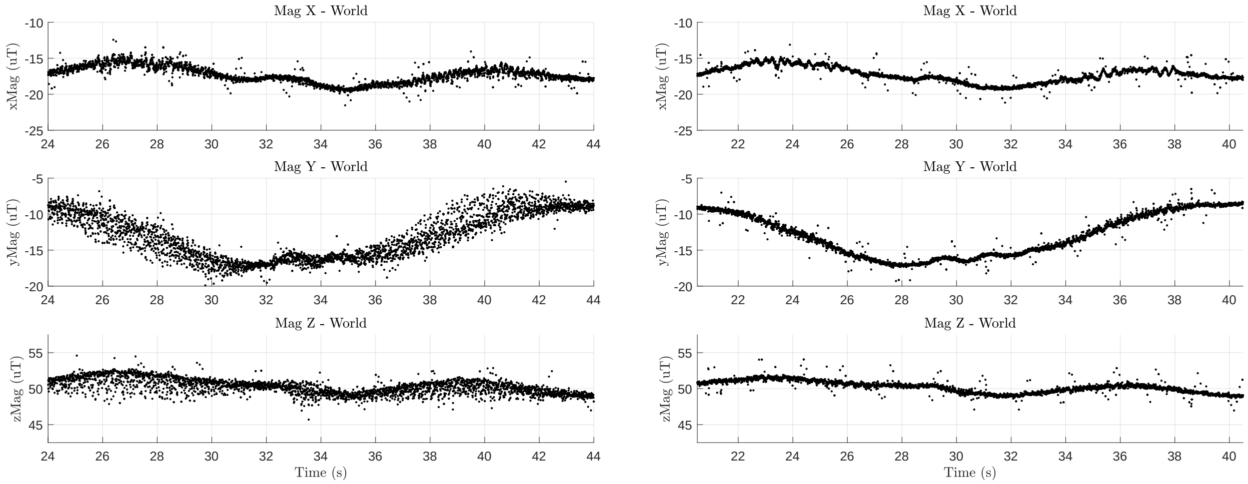
At last, we have what we call our "compromise inference sets" $(D_x^{n_1}, D_y^{n_1}, D_z^{n_1})$ (or just "inference sets") comprised of magnetic field estimates from the intermediate map at n_1 user-selected locations around our workspace.

Our construction of the compromise map gives two benefits. First, it allows our GPR-based maps to avoid overfitting to a single training flight by allowing large hyperparameter observation sets (n_2) that can incorporate observations from several flights of a quadrotor (or potentially even flights from multiple quadrotors for a multi-agent system). Next, it allows us to perform inference in $\mathcal{O}({n_1}^2 n_q)$ rather than $\mathcal{O}({n_2}^2 n_q)$ where n_q is the number of points queried.

2.5 Results and Discussion

This section starts by explaining how the derivative gains from the PID controller of our UAV's autopilot create measurable magnetic noise (Section 2.5.1). Next, Section 2.5.2 introduces the unusual magnetic biases our UAV injects into our measurements and how distancing the magnetometer from the electronics improves the consistency of measurements. We then show that our compromise map (which uses $n_1 = 511$ observations for inference) yields estimates within 0.013μT of our intermediate map (which uses $n_2 = 2001$ observations for inference) in Section 2.5.3.

We follow with an analysis of the spatial density of observation points used to train the GPR map in Section 2.5.4 and find that using observations within 0.55m of one another is sufficient to accurately represent the magnetic field in our flight arena. Section 2.5.5 shows that it is equivalent to either create a specialized map on the norm of the field or simply take the norm of estimates from a vector-valued magnetic field map. Finally, Section 2.5.6 presents our consistency metric which is a tool to identify when a user can rely on the predictions from their GPR-based magnetic field map.



(**a**) Test t5_00: Flight with "noisy" gains. (**b**) Test t5_01: Flight with "quiet" gains.

Figure 2.6: Vehicle "Maggie" flying two instances of the same flight trajectory showing how the flight controller can inadvertently create noise in the measured magnetic field.

2.5.1 Flight Controller Creating Magnetic Field Noise

Our vehicles use the `rc_pilot_a2sys` autopilot originally forked from the open-source `rc_pilot` [3] repository. `rc_pilot_a2sys` uses a four-stage cascaded PID controller as explained in Section IV.B.1 of [71]. Our work started with the gains used in Ref. [71] but adjusted them to reduce noise from the motors and ESCs.

Figure 2.6 shows magnetometer data from a segment of an experiment where the flight vehicle (not Q1, but another vehicle of the same construction) was commanded to a single-altitude scanning pattern (Figure 2.4) at 1.5m altitude. The data in Figure 2.6 was rotated from the vehicle's body frame to the world frame using attitude estimates from motion capture cameras.

Figure 2.6a (t5_00) is the magnetometer data with the gains inherited from ref. [71]. Here, the drone moves along the x axis (4m stride) of the flight space from 24s - 30s, and again from about 36s - 42s. During these time segments, there is a large variance in the measured magnetic field due to commands from the flight controller to the motors and ESCs.

Figure 2.6b (t5_01) shows another instance of the same trajectory, but with new controller gains. The plots are temporally aligned to easily compare corresponding flight segments. At a glance, it is clear that the variance in the magnetic field measurements is much lower in Figure 2.6b than in Figure 2.6a, but not completely gone. This is most evident in the "Mag Y - World" plot of Figure 2.6b where the drone moves along the x axis (4m stride) of the flight space from 21s-26s and again from 34s - 40s.

---

[3] `rc_pilot` was created by James Strawson and Librobotcontrol `https://github.com/StrawsonDesign/rc_pilot`

As explained in Section IV.B.1 of ref. [71], the controller used on this vehicle is a four-stage cascaded PID controller with an outer loop that operates on position error and an inner-most loop controlling angular rate error. The math for the inner loop is shown below in Equation 2.11

$$\begin{aligned}\boldsymbol{\tau}_{\phi} &= K^{P}_{4,\phi}\,(\dot{\phi}_d - \tilde{\omega}_x) \\ &+ K^{I}_{4,\phi}\int_0^t (\dot{\phi}_d - \tilde{\omega}_x)\,ds, \\ &+ K^{D}_{4,\phi}\frac{d}{dt}(\dot{\phi}_d - \tilde{\omega}_x)\end{aligned} \tag{2.11}$$

where $\boldsymbol{\tau}_{\phi}$ is the desired torque for roll (ϕ), $\{K^{P}_{4,\phi}, K^{I}_{4,\phi}, K^{D}_{4,\phi}\}$ are the P, I, and D gains for the fourth-stage controller along the roll axis, $\dot{\phi}_d$ is the desired roll rate, and $\tilde{\omega}_x$ is the measured roll rate from the gyroscope. There are similar fourth-stage PID loops for pitch ($\dot{\theta}$) and yaw ($\dot{\psi}$) rates respectively.

The 'noisy' gains, inherited from [71], used $K^{D}_{4,\phi} = K^{D}_{4,\theta} = K^{D}_{4,\psi} = 0.01$ while the modified 'quiet' gains set all three of these values to 0. This change alone is responsible for the reduction in motor-induced noise in Figure 2.6.

The derivative term of the attitude rate controllers (e.g., Equation 2.11 for roll rate) uses numerical derivatives of gyroscope measurements (Figure 2.7) as part of their computation. As the quadrotor flies, the propellers induce vibrations that make the gyroscope (and accelerometer) measurements rather noisy. Taking numerical derivatives of these noisy gyroscope measurements causes the motors to be commanded with high-frequency inputs that induce measurable electromagnetic noise. We believe this problem is amplified as the total amount of current pulled from the 4S batteries increases. This would explain why the magnetometer variance is larger during the 4m, X-axis stride (with a max commanded velocity of 0.75m/s for this set of tests) than during the 0.25m, Y-axis stride (max commanded velocity of 0.47m/s).

Aside from changing controller gains, other possible solutions include commanding less-aggressive maneuvers during flight tests, moving the magnetometer further from the motors and ESCs (Section 2.5.2), or time-varying magnetometer calibration [68]. In addition, from some preliminary data gathered when trying to understand this problem, we believe that differences in motors and ESCs (even those of the same make and model) may create different magnitudes of measurable magnetic noise.

The remainder of this chapter uses only the 'quiet' controller gains. Note that the analysis for this section used S_2 = 2cm (Figure 2.3b). In the following section, we analyze how the distance of the magnetometer from the motors and ESCs affects the *variation* in the measured magnetic field.

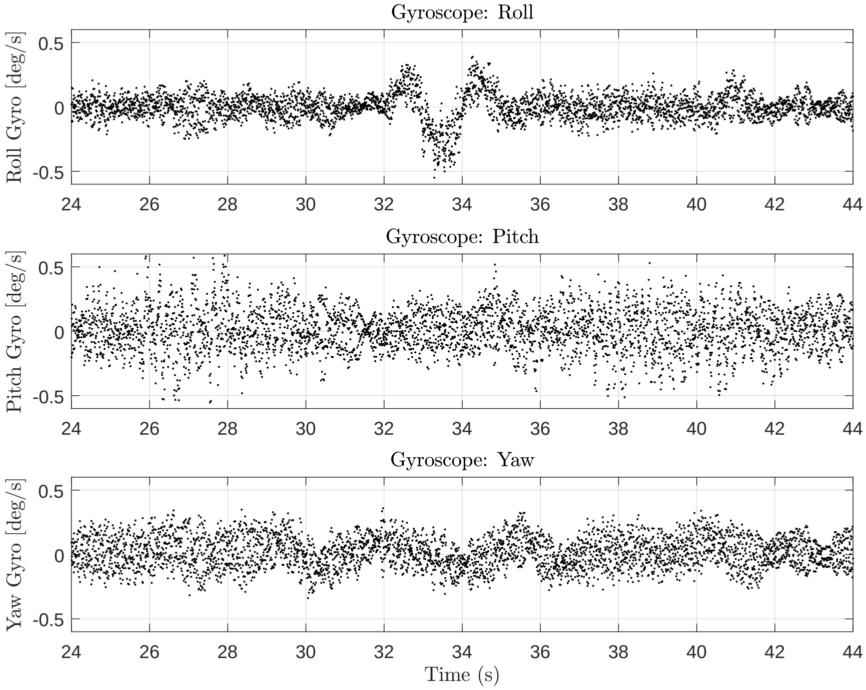


Figure 2.7: Raw gyroscope data from t5_00 . Vibrations from UAV propellers create noisy gyroscope data. Differentiating this noisy data creates the high-frequency magnetic noise seen in Figure 2.6a.

2.5.2 Varying S_2: Distance from Magnetometer to Electronics

Though the flight controller created motor-induced, high-frequency magnetic field noise, there is another measurable magnetic anomaly at play with our quadrotors that causes flight-by-flight *variation* in the measured magnetic field. Here, we distinguish variance (the spread of a distribution of points) from variation (more of an offset). We will reserve variance to describe the spread of points around a probabilistic mean. For this, removing the derivative gains from a particular PID loop of our flight controller reduced the variance of motor-induced magnetic noise (Figure 2.6). To avoid ambiguity, variation will be used to describe the differences in the magnetic field between subsequent flights of the same trajectory. As we will show in this section, the differences in measured magnetic field values are not obviously spread around some average signal.

Recall that S_2 (Figure 2.3b) is the distance from the top of the BeagleBone Blue (BBB) to the bottom of the magnetometer's mounting plate. In this section, we will vary the parameter S_2 to show that increasing the distance of the magnetometer from the other electronics reduces the flight-by-flight variations we see in the measured field. It is easier to explain these anomalies graphically, which we will do in a moment, but first, we explain our testing methodology.

2.5.2.1 Testing Methodology

Q1 was commanded to fly the same single-altitude scanning trajectory (Figure 2.4) at an altitude of 1.5m. Initially, in our investigation of these flight-by-flight variations, we believed that each 4S LiPo battery could have a different amount of “resting current” it provided to the motors to achieve the same flight maneuver as another 4S battery. Though we could not conclusively prove or disprove this hypothesis with our data, it did drive the design of our experiments. For this, each battery started at full charge and Q1 flew the same 1.5m scanning trajectory several times in a row. This allowed us to see if the flight-by-flight variations were due to the battery’s voltage during a flight. After a few repetitions with the first 4S battery, a new (fully charged) 4S battery was used to power Q1 as it flew more repetitions of the same trajectory. By testing different batteries, we could see if the cause of the flight-by-flight anomalies was tied to which battery was powering the motors and ESCs.

Three Turnigy 4S 2650mAh 20C batteries (denoted as #02, #04, and #14) were used for this analysis. The charge/discharge and storage history of these batteries are largely unknown as they have been used by multiple members in the lab since they were purchased in May of 2021. The only thing we could rigorously control for is fully charging each battery to 16.8V just before a flight test.

For this analysis, each flight was downsampled to 10Hz or 2Hz to reduce spatially redundant observations resulting in $\sim$900 or $\sim$180 observations respectively for each flight. Figure 2.4 shows the observation set from one flight with 10Hz and 2Hz downsampling. Although 10Hz clearly shows better coverage of the flight space, the RMSE values for this analysis changed by at most 0.03μT on the S_2=8cm dataset when training on 10Hz vs 2Hz downsampling sets. Thus, we sometimes use 10Hz downsampling to illustrate specific points but typically use 2Hz downsampling for faster training and validation of the magnetic field maps.

With this gathered data, we can train a GPR-based map of the magnetic field at a 1.5m single-altitude slice of the working volume and compare the predictions of that single-altitude map to the measurements gathered from the other repetitions of the same trajectory. Since we assume the ambient magnetic field is not changing throughout our experiments, we expect the magnetic field measurements to be nearly identical between each flight.

2.5.2.2 Flight-by-flight variations

By training the map on a single repetition and validating on another, we get data that looks like Figure 2.8 which has two types of plots: one with red and blue lines with blue shading and another type of plot in grayscale.

For the first type (e.g., Figure 2.8a), each actual magnetometer measurement (in the world frame) from the *validation* dataset is shown as a red cross. The blue line is the GPR’s predicted mean at the location the drone was during that timestamp in the validation flight. The blue shading around the solid blue line depicts the two standard deviations (2σ) of the GPR's uncertainty with its prediction.

The second type of plot (e.g., Figure 2.8b) depicts a black dot as the error between the red cross and blue line while the GPR's 2σ uncertainty is the gray shading. Here, we plot the absolute value of the error since the sign of the error does not indicate anything of interest. The percentage in the grayscale plots depicts how often the black dots (GPR error) are within the gray shading (GPR's 2σ uncertainty).

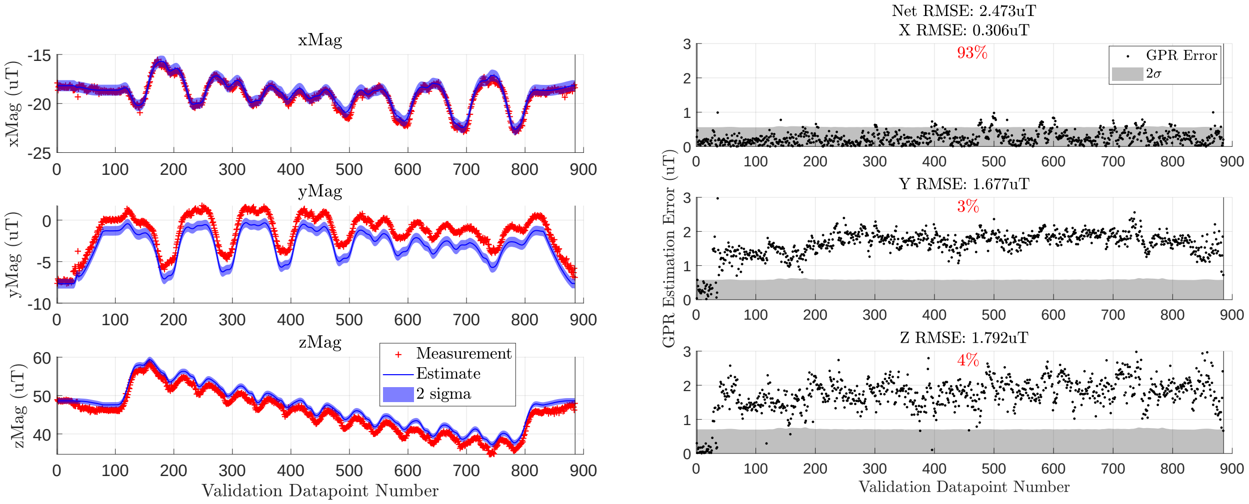
In Figure 2.8, the GPR map is trained on 2Hz-downsampled observations from a single flight (t1_09) and validated against 10Hz-downsampled observations from three flight tests from the same t1 flight series. These three validation examples are generally representative of the different flight-by-flight magnetic field anomalies we have seen. Figure 2.8b shows a relatively large steady-state prediction error throughout, Figure 2.8c has a steady-state error that changes suddenly partway through the flight, while Figure 2.8d is an ideal case where the mapping and validation flights have agreeable measurements.

This is why we refer to this anomaly as "flight-by-flight *variation*". The type of GPR error we see here across consecutive flights in the same test segment (e.g., t1_XX) does not seem to be cleanly distributed around some mean bias value. Instead, repetition t1_09 had similar magnetic observations as t1_15 (Figure 2.8d) yet starkly different observations than t1_05 (Figure 2.8b). Further complicating the matter are cases like t1_04 (Figure 2.8c) with a time-varying bias.

There is evidence in Figure 2.8 that suggests these magnetic anomalies are caused by the ESCs and motors. The initial black dots in Figures 2.8b and 2.8c have low errors before the anomalous bias takes effect. These beginning points are observations gathered when the drone is at rest on the ground (before it has taken off) and the bias sets in, if at all when the quadrotor is in the air. This suggests that before the motors and propellers are spinning, the observations from t1_05 and t1_04 are similar to those from t1_09.

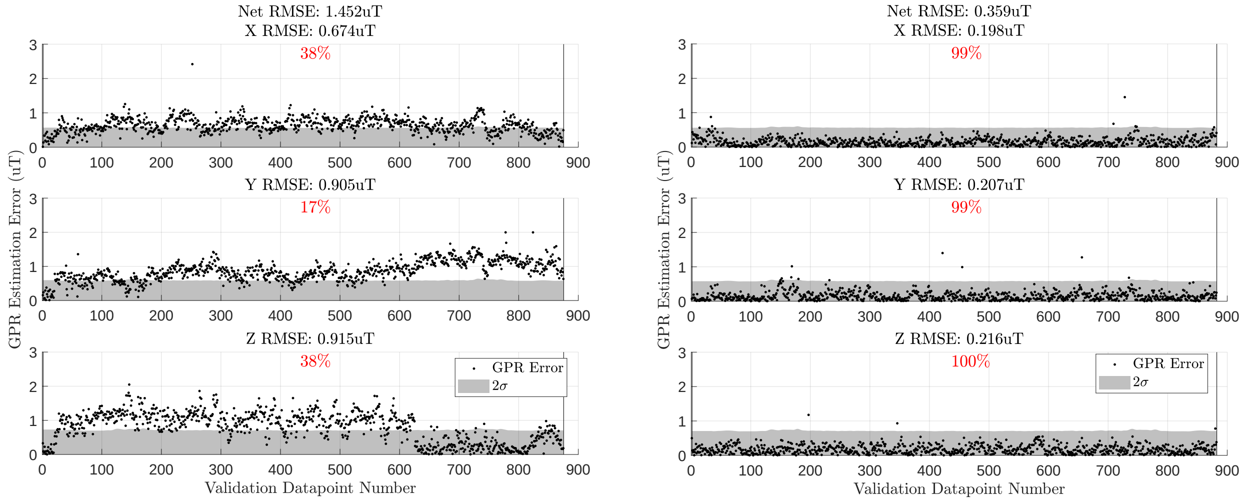
This insight begs the obvious experiment of constraining the motion of the quadrotor and sampling the magnetometer with and without the propellers spinning. Unfortunately, we could not safely conduct such a test with our current setup, but we plan to test this in the future.

Finally, it is possible that the ambient magnetic field in our flight arena is changing over time. To check this, placed a stationary RM3100 in our workspace for 3+ hours sampling data once every 10 seconds (0.1Hz). The raw data from this experiment is shown in Figure 2.9. Here, we still have regular, spurious measurements (like in Figure 2.5) but the main signal does not change by more than $0.15\mu T$ within each respective component of the measured magnetic field. This is much smaller than the $1.8\mu T$ of variation we see in the Z component of Figure 2.8b. Thus, we believe the flight-by-flight variations are due primarily to UAV-induced noise and not by time-varying changes in the ambient magnetic field.



(**a**) S_2=2cm. Trained on t1_09 (#04). Validated on t1_05 (#14). Same data as Figure 2.8b

(**b**) S_2=2cm. Trained on t1_09(#04). Validated on t1_05(#14). Same data as Figure 2.8a



(**c**) S_2=2cm. Trained on t1_09(#04). Validated on t1_04(#14).

(**d**) S_2=2cm. Trained on t1_09(#04). Validated on t1_15(#02).

Figure 2.8: Flight-by-flight variations shown through three validation tests.

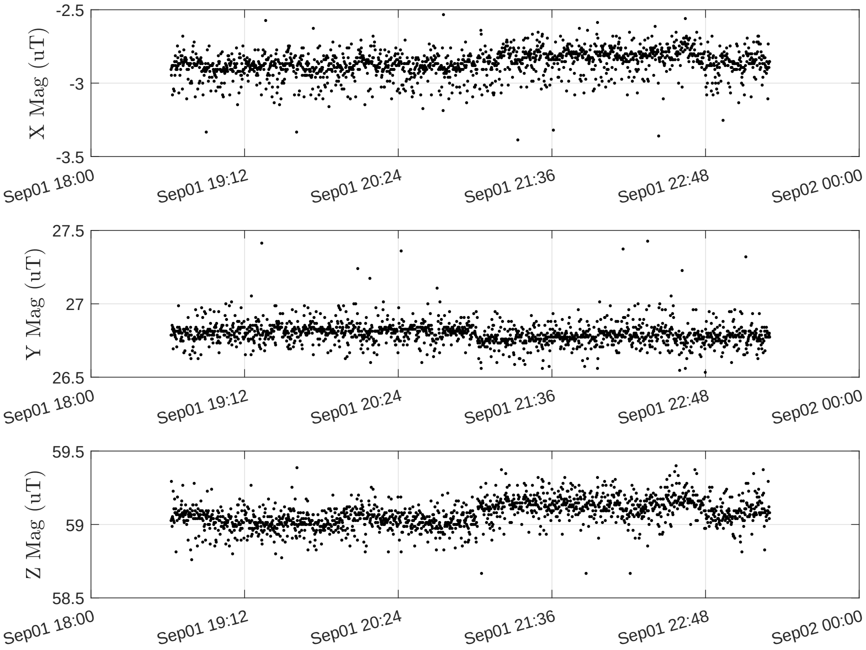


Figure 2.9: Stationary RM3100 data sampled at 0.1Hz over 3+ hours during t6_XX flight tests. x, y or z components of the magnetic field never change by more than $0.15\mu T$ respectively.

2.5.2.3 LiPo Batteries and Magnetic Variation

From the three cases in Figure 2.8, it is tempting to conclude that differences in batteries are the cause of these flight-by-flight anomalies. Here, we see that batteries #04 (used for t1_09) and #02 (used for t1_15) yield agreeable observations. However, from Figures 2.8b and 2.8c, it seems that battery #14 (used for t1_05 and t1_04) is disagreeable with the other two.

To investigate this idea, we flew 60 repetitions of the 1.5m-scanning trajectory from Figure 2.4 over four test sessions (t1_XX, t2_XX, t3_XX, and t4_XX) each corresponding to a different value of S_2 (2cm, 4cm, 6cm, and 8cm). For these tests, we used the same three batteries (#02, #04, and #14) and flew consecutive repetitions of the scanning trajectory for each battery. For each test session, we trained the GPR map on the *first* battery #04 flight and validated the map on all flight tests of the same test session. Note that the data in Figure 2.8 (tests t1_04, t1_05, t1_09, and t1_15) are four of the S_2=2cm repetitions used in this larger analysis.

The results of this 60-flight analysis are summarized as a scatter plot in Figure 2.10. The horizontal position of each point is meant to signify the value of S_2 as 2cm, 4cm, 6cm, or 8cm. Any horizontal deviations from these discrete values are only for visual clarity and do not reflect any deviation in the actual value of S_2 when the data was gathered. The vertical position is the vector RMSE of all three GPRs (Equation 2.9). Finally, the colors and symbols identify which battery was used for each repetition with a green $+$ for battery #04, an orange $\triangle$ for #02, and purple $\square$ for #14.

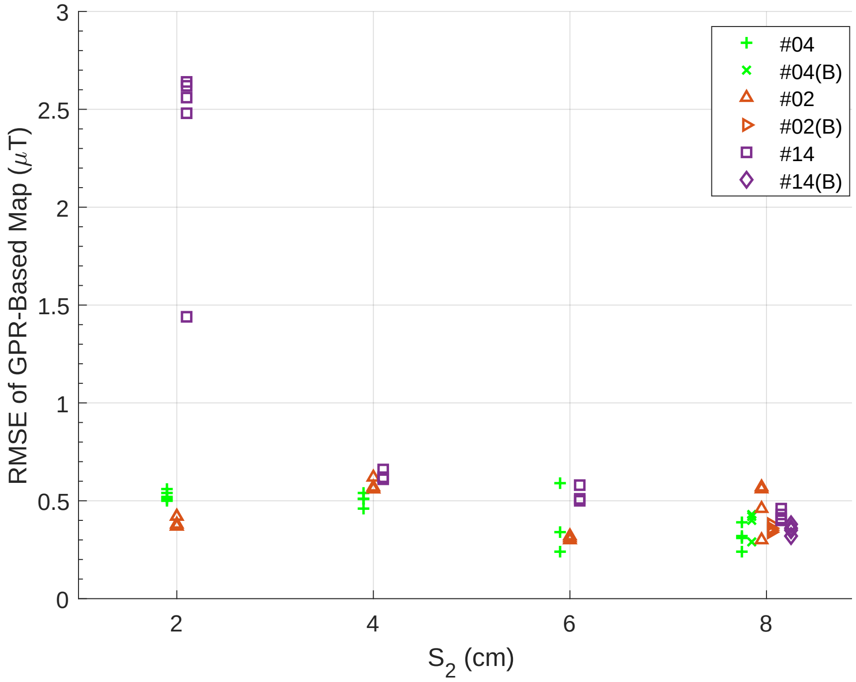


Figure 2.10: S_2 versus vector GPR RMSE. At each S_2, the map is trained on *one* flight from battery #04 (green +). For S_2={2,4,6,8}cm there are N={5,4,3,8} reps per battery.

At S_2 = {2, 4, 6, 8}cm, each battery flew N = {5, 4, 3, 8} repetitions of the 1.5m scanning trajectory. For S_2 = 8cm, each battery started at full charge, flew four consecutive reps, was re-charged, then flew four more repetitions. We present the second charge of each battery in the S_2=8cm dataset with a rotated $+$. $\triangle$, and $\square$ respectively.

Figure 2.10 shows two things. First, increasing S_2 makes the observations of each flight more consistent with one another. Since the GPR-based map (for each S_2) is trained on a battery #04 flight, the scatter plot really shows us the magnitude of the flight-by-flight variations relative to an arbitrary battery #04 flight. As we increase S_2, the variations tend towards a difference of about 0.3μT to 0.5μT.

Note that for each value S_2, there is a single green $+$ that is validated on the same observations used to train the GPR map. Even so, we do not see *any* of the validation tests going much below 0.3μT meaning this may be something of a lower limit on how accurate our GPR maps can be with our testing platform which would include inaccuracies due to sensor noise (a function of our 200Hz sampling rate) and any electromagnetic noise the quadrotor generates.

The second realization is that, unlike what the three examples in Figure 2.8 may suggest, there is not a clear relationship between battery choice and flight-by-flight variation. Although battery #14 (purple $\square$) tends to be most disparate from battery #04 at S_2 of 2cm, 4cm, and 6cm, we see its observations fall in line with the rest by S_2 = 8cm.

We take this moment to point out a flawed conclusion we stated in our previous work (Note: the data and analysis in Chapter 3 of this dissertation was done before any of the work discussed in this chapter). In Figure 3.10b (Chapter 3), observations from a rectangular trajectory (the first $\sim$500 points) are validated on a GPR-based map. In that plot, there is about 2μT of an error on each of the x and z GPR validation sets except before takeoff and after landing. When originally working on the work for Chapter 3, we hypothesized that this error in x and z was caused by a pitching misalignment in the magnetometer for that single flight (relative to the pitch angle for all the other flights). However, if this were the case, there likely would not be the brief moments of agreement before takeoff and after landing as shown in Figure 3.10b. With what we have learned since, we now believe this was actually an example of these flight-by-flight variations.

For the work in Chapter 3, these anomalies may not have happened as often due to us using a different instance of the M330 quadrotor. Alternatively, it could be that the nature of attitude estimation (the focus of Chapter 3) is fairly robust to 2-3μT variations between flights if such anomalies do not significantly change the *angle* of the ambient magnetic field (which has a magnitude of 40-70μT in our workspace). Thus, it is likely that these variations were simply less salient in Chapter 3 given our previous application of attitude estimation.

We believe that time-varying magnetometer calibration could address the flight-by-flight variation issues presented here. In ref. [68], they use measurements of the electric current near high-powered devices to estimate and offset vehicle-induced magnetic fields from their measurements. However, our vehicle does not have *any* electric current sensors, so such a time-varying calibration solution is out of reach for our platform. Instead, we address the problem with our "compromise" map.

2.5.3 Accuracy of Compromise Map

In this section, we propose a solution to the flight-by-flight variations that incorporate data from several training flights into a single magnetic field map. Since each flight has a chance of giving a different bias in the measured magnetic field, the GPR-based map will overfit if trained on only a single flight test. Thus, we instead, use n_2 observations from many flights that span the working volume to create an "intermediate map" that optimizes hyperparameters $\{\mathbf{\Theta}_x^*(D_x^{n_2}),\ \mathbf{\Theta}_y^*(D_y^{n_2}),\ \mathbf{\Theta}_z^*(D_z^{n_2})\}$ *and predicts with* the n_2 observations. The compromise map then uses the intermediate map's estimates at n_1 user-selected locations to predict the magnetic field using only n_1 points. The goal of this section is to see how the accuracy of the compromise map varies as a function of n_1; specifically, taking after ref. [19], to see how the spatial density of the n_1 training points effects map accuracy.

Table 2.1: Separate flights are used for *training* and *validation*. All magnetic field observations were downsampled to 2Hz for *training* and *validation*. Number of observations each training flight contributes to the map is listed.

Flight ID	**Flight Description**	**Training [2 Hz]** (# Observations)	**Validation [2 Hz]** (# Observations)
t6_00	Lower Four Alts.	571	–
t6_01	Upper Four Alts.	580	–
t6_03	Scan-γ	442	–
t6_21	Scan-$\epsilon^{\perp}$	408	–
t6_04	Lower Four Alts.	–	573
t6_05	Upper Four Alts.	–	573
t6_06	Scan-γ	–	441
t6_20	Scan-ϵ	–	384

Table 2.2: Multi-altitude map error. Trained on $n_2 = 408$ observations from flight test t6_21 (Scan-$\epsilon^{\perp}$).

Flight ID and Description	GPR RMSE (μT)				Error within 2σ (%)		
	Norm	*X*	*Y*	*Z*	*X*	*Y*	*Z*
t6_04 (Lower Alts.)	0.595	0.256	0.284	0.456	74.69	87.09	52.88
t6_05 (Upper Alts.)	0.998	0.633	0.541	0.551	6.81	19.72	32.11
t6_06 (Scan-γ)	0.642	0.362	0.318	0.424	56.01	86.62	68.25
t6_20 (Scan-ϵ)	0.556	0.229	0.293	0.413	56.25	47.40	39.06

2.5.3.1 Multi-Altitude Trajectories

This section uses data from eight different flight tests listed in Table 2.1 where the first four flights are typically used for *training* the intermediate/compromise map while the remaining four are used for *validation*. The trajectories listed in Table 2.1 are all multi-altitude tests where each single-altitude slice is the same trajectory from Figure 2.4. “Lower Four alts.” flies at altitudes Z = [-0.5, -0.75, -1.0, -1.25]m, “Upper Four alts.” at Z = [-1.5, -1.75, -2.0, -2.25]m, “Scan-γ” at Z = [-0.5, -1.375, -2.25]m, while “Scan-ϵ” and “Scan-$\epsilon^{\perp}$” fly at Z = [-0.75, -1.5, -2.0]m. The idea here is to gather redundant observations throughout the working volume to help our GPR map learn the flight-by-flight variations.

Additionally, Table 2.1 lists the number of 2Hz downsampled observations from each flight. The number of 4Hz and 10Hz observations are *approximately* $2\times$ and $5\times$ the values listed in Table 2.1.

Table 2.3: GPR error for multi-altitude *intermediate* map trained on $n_2 = 2001$ observations from tests t6_00, t6_01, t6_03, and t6_21.

Flight ID and Description	GPR RMSE (μT) *Norm—X—Y—Z*				Error within 2σ (%) *X—Y—Z*		
t6_04 (Lower Alts.)	0.466	0.144	0.145	0.419	98.43	98.08	85.17
t6_05 (Upper Alts.)	0.713	0.403	0.447	0.382	45.55	23.73	90.05
t6_06 (Scan-γ)	0.378	0.184	0.155	0.292	97.96	96.83	96.83
t6_20 (Scan-ϵ)	0.406	0.111	0.222	0.321	99.48	92.97	93.75

2.5.3.2 Intermediate Map Accuracy

To start, we demonstrate the value of adding observations from many flights to a GPR-based map. We call GPR maps that are trained on observations from multiple flights an *intermediate map*.

Table 2.2 gives the performance of a multi-altitude magnetic field map trained on $n_2 = 408$ (2Hz downsampling) observations from flight test t6_21 and validated on observations from four different flight tests. The table is split into three major columns that give the details of the validation flight and the vector RMSE (μT) of all three GPRs along with the RMSE of each x, y, and z GPR respectively. The third major column quantifies how often (as a percentage) an error data point lies within the 2σ uncertainty of the respective GPR. The performance metrics from the second and third major columns are the same as those listed in the grayscale plots of Figure 2.8 (with different training and validation flights in this section).

The underlying issue is the map will overfit to the observations gathered from t6_21. In Section 2.5.2, we showed that there are flight-by-flight variations in the measured magnetic field. By training a GPR-based magnetic field map on a single flight, we prevent the map from learning to account for these flight-by-flight variations. Note that training the map on $n_2 = 2038$ observations (10Hz downsampling) from t6_21 gives norm RMSE values of 0.629μT, 1.034μT, 0.751μT, and 0.552μT when validating on t6_04, t6_05, t6_06, and t6_20 respectively. Thus, simply increasing the number of training samples from a single flight test does not improve the map's performance.

By comparison, Table 2.3 trains the GPR hyperparameters (and performs inference) on $n_2 = 2001$ observations (2Hz downsampling) from four different flight tests: t6_00, t6_01, t6_03, and t6_21. Note, from Table 2.1, that we are now training and validating on one instance each of all four types of flights conducted for this analysis: lower four altitudes, upper four altitudes, scan-γ, and scan-ϵ. By comparing Table 2.2 to 2.3, we see that adding observations from a *variety* of flights uniformly reduces RMSE and increases the frequency that error falls within each GPR's 2σ error (usually caused by an increase in each GPR's uncertainty).

Additionally, training on $n_2 = 9999$ observations (10Hz downsampling) from the four training flights gives norm RMSE values of 0.497μT, 0.686μT, 0.406μT, and 0.408μT when validating on t6_04, t6_05, t6_06, and t6_20 respectively. Again, a bit to our surprise in this second case, adding more observations from the same ensemble of training flights does not necessarily improve the map's performance. However, it is clear that training on multiple flights (Table 2.3) is better than training on a single flight (Table 2.2).

Aside from the 2Hz vs. 10Hz comparisons done in this section, we do not aim to directly address ways to reduce the cost of training hyperparameters. Instead, we focus on the inference cost $\mathcal{O}\left({n_2}^2\right)$ with what we call a *compromise map*.

The idea here is to query the intermediate map (which uses n_2 observations for inference) at n_1 user-selected locations $\boldsymbol{X}^{n_1} \in \mathbb{R}^{3\times n_1}$ in the working volume. The intermediate map's estimates of the magnetic field at these n_1 locations give us a set of "measurements" $\hat{\boldsymbol{M}}^{n_1} \in \mathbb{R}^{3\times n_1}$ which (together with $\boldsymbol{X}^{n_1}$) form the "observations" used by the compromise map to perform inference. Note that the compromise map will use the same hyperparameters from the intermediate map for its inference. Section 2.4.4 gives a more detailed explanation of the process of creating a compromise map from an intermediate map.

2.5.3.3 Intermediate Map vs. Compromise Map

We now compare the accuracy of the intermediate map, which uses $n_2 = 2001$ observations (2Hz downsampling) to perform inference, to that of the "compromise" map which uses only $n_1 = 511$ inference points. Recall that both use the same sets of hyperparameters optimized over $n_2 = 2001$ observations.

The $n_1 = 511$ user-selected locations for this analysis are chosen as follows. Points are distributed evenly through the [-2, 2]m x axis span, [-1.5, 1.5]m y axis span, and [-2.25, -0.5]m z axis span of the working volume. A compromise map location is selected every 0.5m, 0.5m, and 0.25m for the x, y, and z axes respectively. In total, this gives 504 locations within the working volume. The remaining 7 points are evenly spaced from the ground to an altitude of 0.5m so the compromise map has some observations during the takeoff and landing sequence (which, for all our flights, are above the origin). The number $n_1 = 511$ is an important constraint for another toolbox we used for position localization. Thus, this spatial discretization (0.5m × 0.5m × 0.25m) became a common test state in our work.

We can see these trends better in Figure 2.11 which is similar in style to the grayscale plots from Figure 2.8, but with the subplots as three columns rather than as three rows. This format fits more figures on a single page to more easily compare the intermediate and compromise maps.

Figure 2.11 gives plots for the intermediate map on the left column and the compromise map on the right. Here, we see the compromise map has both *quantitatively* and *qualitatively* similar performance to the intermediate map despite using nearly a fourth of the points for inference. By comparing the norm RMSE values (in the titles) across the two columns of Figure 2.11, we see the two maps are never off by more than 0.013μT (13 nanoTesla) in norm RMSE. This difference is within the noise of the magnetic field measurements of our platform.

For reference, training a GPR map on 408 observations (2Hz downsample) from t6_21 and validating this map on the same 408 observations gives RMSE values of (0.196μT, 0.089μT, 0.108μT, 0.136μT) for norm, x, y, and z RMSE values respectively. This means that on our flight platform, even when the map has overfit for the exact observations it will be validated on, it still cannot discern differences of 0.013μT.

Thus, we take the differences between the intermediate map ($n_2 = 2001$) and the compromise map ($n_1 = 511$) in Figure 2.11 to be negligible and assert that their RMSE performance is effectively identical.

Of course, a lot of valuable information can be lost by simply comparing RMSE values. However, visual comparison across the two columns of Figure 2.11 further emphasizes that both the intermediate and compromise maps yield very similar prediction results. The key difference is in the uncertainty of the two maps.

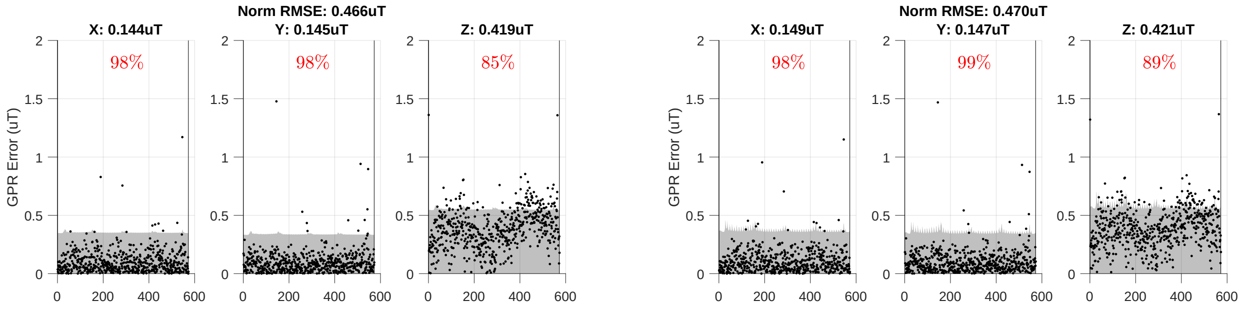
It's easiest to see this in row three (Figures 2.11d vs. 2.11e) where the compromise map (right) has spikier gray shading (2σ uncertainty) than the intermediate map (left). The compromise map's increased uncertainty is primarily due to our choice of prediction locations $\boldsymbol{X}^{n_1} \in \mathbb{R}^{3\times n_1}$ which have 504 points selected from a (0.5m × 0.5m × 0.25m) spatial discretization of our working volume. Recall that each altitude of our validation trajectories traverses lanes separated by 0.25m (Figure 2.4). Thus, anytime the compromise map is queried at a location between the (0.5m × 0.5m) x-y points, it will report a higher uncertainty in its predicted magnetic field estimate.

The overall increased uncertainty of the compromise map also explains why the respective GPR errors more frequently fall within two standard deviations of their respective uncertainties. Said another way: the plots on the right column of Figure 2.11 will almost always have higher red percentages than the comparative GPR on the left column *given our selected spatial discretization of* (0.5m × 0.5m × 0.25m).

In general, we expect the overall uncertainty of the compromise map to decrease with increased spatial density in our selection of locations for $\boldsymbol{X}^{n_1}$. The next section will further investigate how the spatial density of the points $\boldsymbol{X}^{n_1}$ affect the performance of the compromise map.

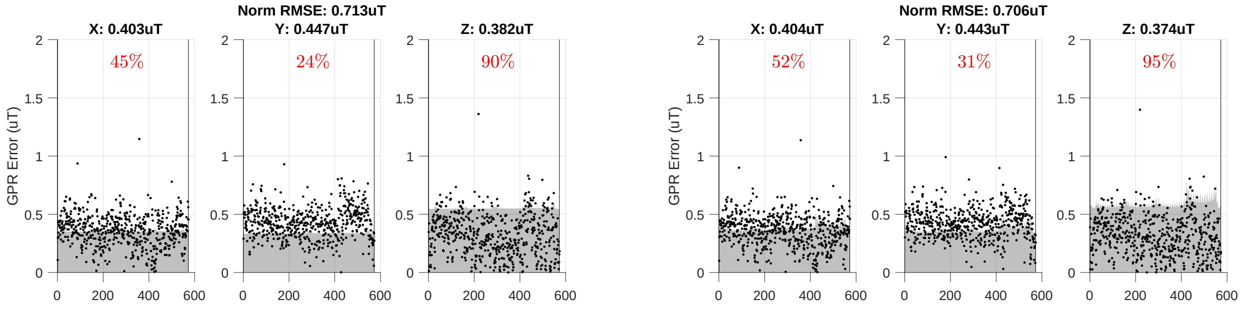
2.5.4 Compromise Map - Spatial Density Analysis

This section seeks to understand how the accuracy of the compromise map changes as a function of the spatial density of the locations used in $\boldsymbol{X}_1$. Generally, we expect GPR error to increase, the number of error points captured by 2σ uncertainty to increase, and the computation time to *decrease* as the training set points become more sparse. In this section, we will only analyze the first of these (GPR accuracy).



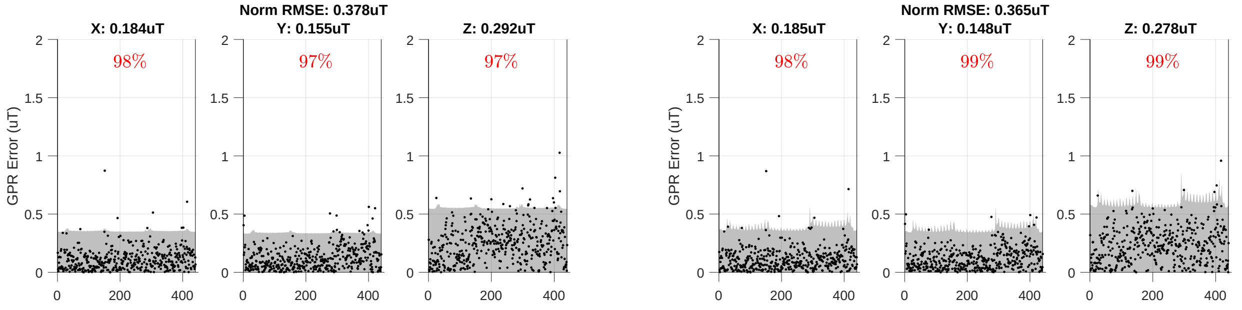
(**a**) n_2 for t6_04. No outliers.

(**b**) n_1 for t6_04. No outliers.



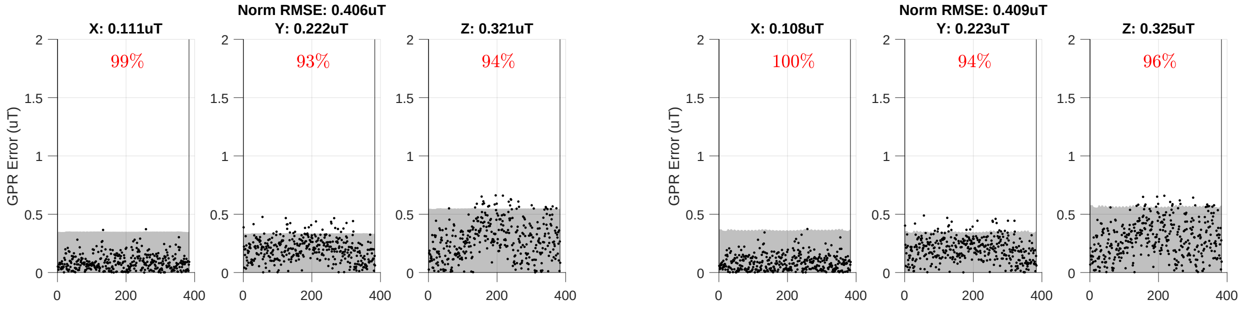
(**c**) n_2 for t6_05. Outliers $\{[2.89], [], [2.42]\}\mu$T.

(**d**) n_1 for t6_05. Outliers $\{[2.49],[],[2.39]\}\mu$T.



(**e**) n_2 for t6_06. Outliers $\{[2.41],[],[]\}\mu$T.

(**f**) n_1 for t6_06. Outliers $\{[2.36], [], []\}\mu$T.



(**g**) n_2 for t6_20. No outliers.

(**h**) n_1 for t6_20. No outliers.

Figure 2.11: Performance of the intermediate map (left column) and compromise map (right column). Outliers larger than 2μT are listed for $\{ [\mathcal{GPR}_x], [\mathcal{GPR}_y], [\mathcal{GPR}_z] \}$ in each subplot.

In the last section, our $n_1 = 511$ compromise training points came from a custom spatial density of (0.5m $\times$ 0.5m $\times$ 0.25m) which gave 504 locations, plus an additional 7 locations to have observations along the takeoff and landing segment for each flight. In this section, we will work with a uniform spatial density in all directions ($S \times S \times S$). Additionally, to simplify matters a bit, we will use a naive linear spacing of $[min : S : max]$ for the locations along each respective spatial axis of the lab. This means that different spatial densities S can yield the same *number* of points n_1 but at different locations. Recall that our working volume has spatial limits of [-2, 2]m in x, [-1.5, 1.5]m in y, and [-2.25, -0.5]m in z.

In this study, we vary S from 0.2m to 1m to emulate the study done in ref. [19]. However, our study will quantify the RMSE of the magnetic field map rather than relying exclusively on visual comparisons of the map. For the 17 values of S tested (in increasing order), n_1 = [3031, 1775, 931, 655, 447, 259, 259, 199, 133, 112, 97, 97, 79, 67, 47, 47, 47]. Recall, seven of the n_1 values listed here are for the takeoff and landing sequence.

Figure 2.12 shows the norm RMSE of the compromise map (validated on t6_04, t6_05, t6_06, t6_20) as a function of the spatial density term S. A zoomed version of the initial segment of data is embedded in the same figure. Here we see that the norm RMSE is fairly constant for values $S \leq 0.5$m with a small spike in norm RMSE at $S = 0.45$m. This spike is caused by our naive linear spacing which occasionally causes poor sampling along altitudes. At $S = 0.45$m, we get altitudes of z = [-2.25, -1.8, -1.35, -0.9]m. Recall from Section 2.5.3.1 that some of our validation flights have measurements as low as z = -0.5m.

Combining these sources of information, we can see that the black-$\times$ (t6_06; Scan-γ) and red-$\square$ (t6_04; Lower four alts) suffer the worst increases at $S = 0.45$m. Since the compromise map's observations only go as low as $z = -0.9$m, both these trajectories have higher GPR RMSE when validated at their $z = -0.5$m observations. Meanwhile, blue-$\circ$ (t6_20; Scan-ϵ) shows little change at $S = 0.45$m since all its observations are away from the extremes of the working volume. The other transient spikes in Figure 2.12 (e.g., $S = 0.6$m) are caused by similar z-axis sampling from our naive linear spacing. t6_04 and t6_06 continue to be the most sensitive to certain values S when the map has no observations near $z = -0.5$m.

Despite these transient spikes, there are two clear trends we can see from Figure 2.12. First, for values of $S \leq 0.55$m (with $S = 0.45$m as an exception), the norm RMSE is relatively insensitive to changes in S. This agrees quite well with the ref. [19] where they show (visually) that their magnetic field map is qualitatively similar for values of S = [0.2, 0.4, 0.6]m. In [19], their analysis led them to use 0.6m as a standard distance between observations in another experiment they conduct.

For comparison, we briefly summarize the spatial density of observations used by some other indoor magnetic field mapping works. Some works map along a single axis ($p = 1$) such as [22] that separates observations by 0.04m and 0.1m on a ground robot and chest-mounted pedestrian

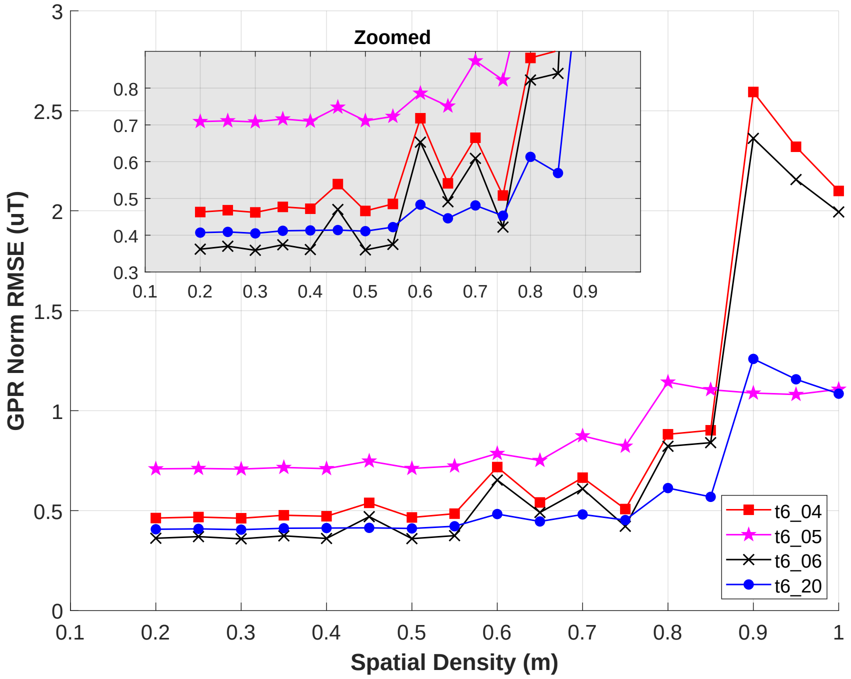


Figure 2.12: GPR norm RMSE fairly constant for $S \leq 0.55$m. Flights t6_04 (red $\square$) and t6_06 (black $\times$) have observations at $z = -0.5$m causing higher errors for certain values S.

setup respectively through hallway networks of up to 350m in length. Additionally, [21] makes a $p = 1$ map with 0.2m between each observation and [29] make a $p = 1$ map with 0.1m between observations.

For $p = 2$, planar maps [12] create a constant-altitude planar map using $S = 0.305$m separation in x and y for their coarse grid and $S = 0.005$m for fine grid spacing. This mapping is done in a 2.1m×2.1m room. Ref. [13] uses a type of occupancy grid for their $p = 2$ magnetic field SLAM solution with grid size 0.05m×0.05m. The solution ignores all previous magnetic field measurements further than 0.5m from the vehicle's current position. Ref. [15] makes $p = 2$ maps with lanes separated by 0.38m. Ref. [20] creates $p = 2$ maps at different heights along hallways in three separate buildings. The authors use a longitudinal separation distance of about 0.4m (varies per building) and a vertical separation of 0.08m between their observations.

Finally, Ref. [16] makes a $p = 3$ "map" of a hallway (no explicit interpolation of observations) with (0.33m $\times$ 0.45m $\times$ $S^?$) where the vertical spacing $S^?$ is not explicitly listed in their work, and [39] makes $p = 3$ maps with $S = 0.5$m uniform spacing in all axes. Chapter 3 of this dissertation uses planar lawnmower patterns like in Figure 2.4, but with 0.5m spacing in the y axis (instead of the 0.25m spacing used in this chapter). Furthermore, we used a vertical spacing of 0.75m in our training sets to investigate the accuracy of $p = 3$ magnetic field maps when interpolating between and extrapolating outside of pre-mapped altitudes.

In short, for 1D, 2D, or 3D *indoor* magnetic field maps, we have not seen any previous works using separation distances S larger than 0.6m. The exceptions are [19] where the authors explicitly study a reasonable upper limit on S and our work in Chapter 3 which studies the interpolative and extrapolative limits of $p = 3$ GPR maps.

We believe the empirical upper bound of $S = 0.6$m described above might be driven by proximity to walls and the floor. If this is the case, $S \leq 0.6$m might only be the upper bound for indoor regions that are within a couple meters from some wall or the floor. We acknowledge that this includes nearly every room or hallway in most buildings, but a large, multi-story atrium might be an exception. Portions of such a large, open, yet indoor volume may have lower spatial variation and can be accurately mapped with values $S > 0.6$m. The size of the two works that explicitly study spatial density indoor magnetic field mapping (3m $\times$3m$\times$2.2m for [19] and 4m$\times$3m$\times$2.25m in this chapter) might be too small (i.e., too close to the walls and floor) to test this idea.

It is also possible that the floor, and not walls, are the dominating factor for spatial variation in indoor magnetic fields. In [20], Hanley et al. show that that the magnetic field near the floor (within 0.5m from the ground) of three university buildings is most distinct from the field at other altitudes up to 2m. This might explain why the other works in our spatial density summary, which mostly use ground robots and pedestrians that remain "near" the floor, have $S \leq 0.6$m.

The second conclusion, the complement of the first, is that values of $S \geq 0.6$m start to show steady increases in norm RMSE. It is not clear what trends hold for $S > 1$m, but such separation distances are somewhat degenerate given the size of our working volume. Further, it is unlikely for the norm RMSE values to return to their $S = 0.2$m for values of $S > 1$m (which is what we are most interested in).

This analysis is what led us to use $S = 0.5$m for the x-y separation distance for the $n_1 = 511$ compromise map used elsewhere in this chapter. We chose $S = 0.25$m for z because we found (empirically) that the GPR's performance is most sensitive to omissions in observed altitudes. Recent works on indoor magnetic field mapping have shown similar trends: that indoor magnetic fields change noticeably as a function of altitude [23, 20]. For our analysis, similarly for [23, 20], it is not clear how much this altitude-based sensitivity is driven by the fact that all our flights (training and validation) are comprised of many single-altitude slices.

This section has shown that our compromise map can perform accurate magnetic field inference with fewer prediction points $n_1 < n_2$. Remember the goal here was to allow the GPR-based magnetic field map to learn the flight-by-flight variations across several training flights without incurring high computation costs when performing inference. However, some validation flights like t6_05 (Figures 2.11c and 2.11d) still have steady-state errors that even the compromise map (or the intermediate map) is unable to resolve given a set of training flights.

In the following section, we compare two methods of mapping the *norm* of the magnetic field in our workspace to investigate if composing the three outputs from a $3 \to 3$ map gives a less-accurate norm estimate than training a specialized $3 \to 1$ map on the magnetic field norm.

2.5.5 Creating a $3 \to 1$ map from estimates of a $3 \to 3$ map

This section compares two ways of estimating the norm of the magnetic field in a workspace. The first takes the three estimates from our $3 \to 3$ map and computes the norm of these estimates $\sqrt{\hat{m}_x^2 + \hat{m}_y^2 + \hat{m}_z^2}$. The alternative is to create a new $3 \to 1$ map, called GPR_{nrm}, trained on the norm of our measurements

$$\tilde{y}_{nrm} = \sqrt{\tilde{y}_x^2 + \tilde{y}_y^2 + \tilde{y}_z^2} \in \mathbb{R} \tag{2.12}$$

to create hyperparameter ($D_{nrm}^{n_2}$) and inference ($D_{nrm}^{n_1}$) sets. The scalar predictions from this new GPR are defined as $\hat{m}_{nrm} \in \mathbb{R}$.

The goal here is to see if a specialized $3 \to 1$ map can more accurately estimate the magnetic field norm than a composition of the estimates from a $3 \to 3$ map. This is an important distinction since some works create maps on special components of the ambient magnetic field like a $2 \to 2$ map on the horizontal and vertical components of the field [72, 73] or a $1 \to 1$ map on the heading (or declination) of the field [21]. Here, we aim to check if composing the three estimates from a $3 \to 3$ map accumulates error in estimating the norm of the field in ways that a specialized $3 \to 1$ map, which estimates the norm directly, would not.

The procedure for training hyperparameters and creating the $n_1 = 511$ compromise map for our new GPR_{nrm} is the same as described in Section 2.4.4, but with measurements as defined in Equation 2.12.

The RMSE accuracy of our maps requires a prediction from a GPR (or a set of GPRs) at location $\boldsymbol{r}^*$ and a measurement taken at the same location. We refer to the prediction given by our typical $3 \to 3$ map as our *vector* prediction

$$\hat{\boldsymbol{m}}_{vec}(\boldsymbol{r}^*) = \begin{bmatrix} \hat{m}_x(\boldsymbol{r}^*) \\ \hat{m}_y(\boldsymbol{r}^*) \\ \hat{m}_z(\boldsymbol{r}^*) \end{bmatrix} = \begin{bmatrix} \mathcal{GPR}_x(\boldsymbol{r}^*) \\ \mathcal{GPR}_y(\boldsymbol{r}^*) \\ \mathcal{GPR}_z(\boldsymbol{r}^*) \end{bmatrix}. \tag{2.13}$$

Now we define two error terms meant to compare the accuracy of estimating the magnetic field norm in our workspace. One is for the accuracy of our $3 \to 3$ vector map (vec) while the other is for the $3 \to 1$ norm map (nrm).

Table 2.4: Norm (nrm) GPR versus vector (vec) GPR on estimating norm of magnetic field.

Flight ID and Description	*RMSE* (μT)	
	e_{vec_nrm}	e_{hrz_nrm}
t6_04 (Lower Four Alts.)	0.391	0.390
t6_05 (Upper Four Alts.)	0.256	0.257
t6_06 (Scan-γ)	0.245	0.246
t6_20 (Scan-ϵ)	0.301	0.301

$$\begin{aligned} e_{vec\_nrm} &= \sqrt{\hat{m}_x^2 + \hat{m}_y^2 + \hat{m}_z^2} - \tilde{y}_{nrm} \\ e_{nrm\_nrm} &= \hat{m}_{nrm} - \tilde{y}_{nrm}. \end{aligned} \tag{2.14}$$

The error terms in Equation 2.14 are used to create RMSE metrics (Equation 2.8) for the accuracy of each respective GPR against some validation set.

Table 2.4 computes RMSE on the error metrics of Equation 2.14 for the four validation flights we used in the previous section. From this, we see that both methods of estimating the norm of the magnetic field are equivalent. There is never more than 1nT of difference between the two RMSE values which is well within the noise floor of the PNI RM3100 (even without accounting for noise generated by the drone).

Our result is reassuring in that there is no need to train a fourth GPR GPR_{nrm} if a user wants the three vector components *and* the norm of the magnetic field in their workspace. It is not immediately clear if this result extends to angle-based compositions of the magnetic field map like declination [21] or inclination of the ambient field.

Our last tool in this chapter, which we call the "consistency metric", uses the frequency of error falling with a GPR's 2σ interval to indicate if a validation flight is consistent with a given GPR map.

2.5.6 Consistency Check

Despite our efforts to reduce magnetic noise from motors and ESCs (Sections 2.5.1 and 2.5.2) and incorporate flight-by-flight variations into a compromise map (Section 2.5.3), there may still be flight tests with significantly different magnetic field observations than what a given GPR map can predict. Such cases can arise from the flight-by-flight variations of the same vehicle (e.g., Figures 2.11c and 2.11d), but may also come into play if different platforms are used for mapping and utilization (e.g., GPR map trained on drone data but utilized for indoor pedestrian localization). As such, it is important for users to know when they can rely on the predictions from their GPR-based magnetic field map, and when it may be better to leverage a different sensing modality for their state estimates.

For this, we introduce the concept of "consistency" between a GPR-based map and a validation flight. Given a $p \to m$ magnetic field map, we say that a validation flight is *consistent* with the map if all m GPRs respectively capture 96% of their error within two standard deviations (2σ) of their uncertainty.

Of course, the 96% threshold is inspired by the expected number of points that fall within two standard deviations of a univariate normal (Gaussian) distribution. Additionally, we note that a 96% threshold works well for our test platform, our choice of GPR kernel (squared exponential), the way we optimize hyperparameters (minimization of log marginal likelihood; Section 2.3.1), and our target application of position localization. Although these kernel and optimization methods are common, we have not investigated how GPR-based maps generated with different kernels or different hyperparameters behave in regard to our notion of consistency.

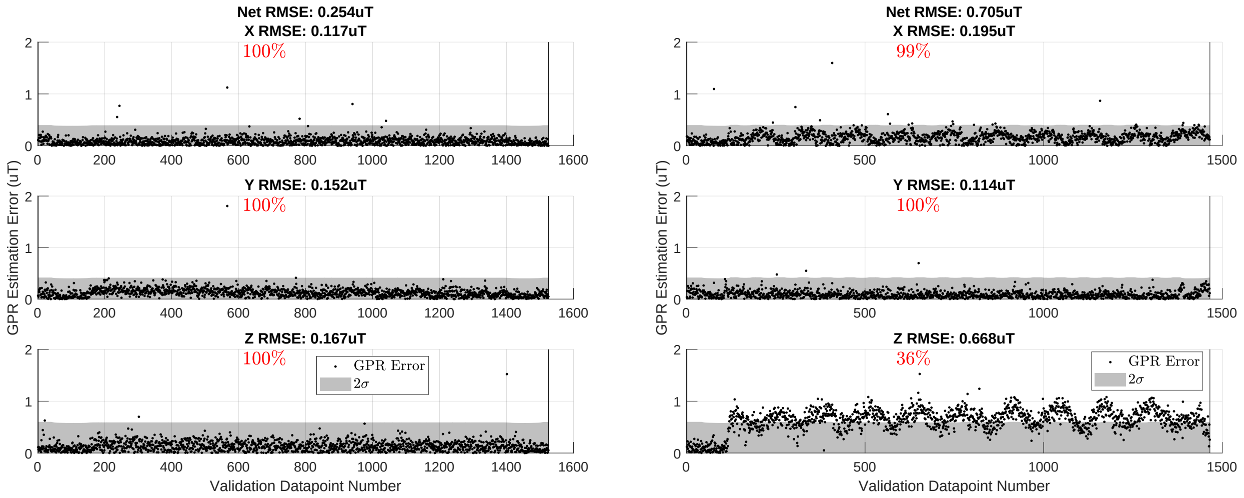
To demonstrate, we create a $n_1 = 511$ compromise map of 2Hz-downsampled observations from the following seven flight tests: t6_0, t6_1, t6_3, t6_4, t6_5, t6_6, t6_16. We validate this compromise map on 10Hz samples from four flight tests (t6_09, t6_11, t6_15, t6_18) to illustrate how we use the consistency metric. Note that we are training and validating on different t6_XX tests compared to the previous sections. The training flights used here are from an analysis of magnetic field position localization which we will explore more in Chapter 4.

Figure 2.13 shows the GPR's performance on the four selected validation flights. We start with a case that easily passes our consistency test (Figure 2.13a). Test t6_11 (Figure 2.13b) fails the consistency test since the $\mathcal{GP}_z$ does not capture enough of its error within its 2σ uncertainty. Note that even though x and y GPRs have a great understanding of their error, the poor performance of the $\mathcal{GP}_z$ alone means that test t6_11 is inconsistent with this $3 \to 3$ map. However, if the user needed only the outputs of the $\mathcal{GP}_x$ and $\mathcal{GP}_y$ (say for localization using magnetic field heading [21]), then the poor performance of $\mathcal{GP}_z$ may be irrelevant and the map could still be useful for such an application.

Figure 2.13c shows an example of the flight-by-flight variation changing partway through a validation flight. In this case, the first 1220 validation points easily pass our consistency metric for all three GPRs. From 1220 to 1335, things get worse and $\mathcal{GP}_x$ barely fails our consistency test. After data point 1335, all three GPRs fail the consistency check.

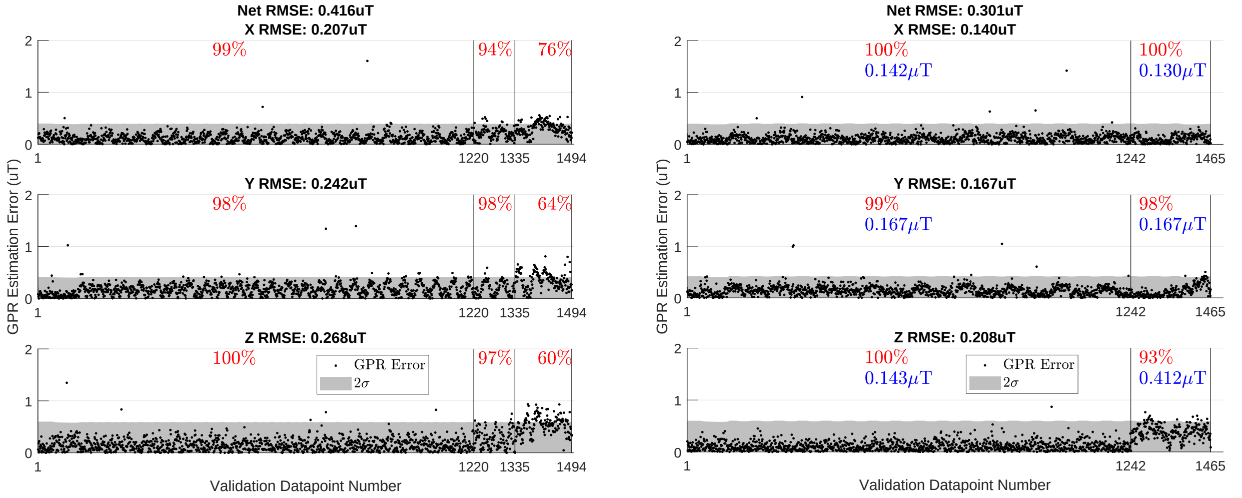
Note that the 3+ hours of 0.1Hz data from our stationary "corner" magnetometer in Figure 2.9 was gathered throughout the t6_XX test flights. As such, we believe the change at index 1220 of Figure 2.13c is from the UAV-induced noise and not a change in the ambient magnetic field.

This example allows us to emphasize an important point about how we use the consistency metric. Since most of test t6_15 (up to point 1220) is consistent with the compromise map, computing our 2σ percentage across the entire validation set gives (97%, 95%, 95%) for each GPR respectively.



(**a**) t6_09. No hidden outliers.

(**b**) t6_11. No hidden outliers.



(**c**) t6_15. Outliers $\{[2.16, 2.56], [2.10], []\}\mu$T.

(**d**) t6_18. No hidden outliers.

Figure 2.13: Consistency metric (red percentage) across four validation flights. Some error points are larger than 2μT and are cut off in the graphs. Such outliers are listed for $\{$ $[\mathcal{GPR}_x]$, $[\mathcal{GPR}_y]$, $[\mathcal{GPR}_z]$ $\}$ in the respective subplot caption. Empty brackets indicate no hidden outliers for the respective $\mathcal{GPR}$.

Working with these numbers alone (without having Figure 2.13c for context), it's tempting to say all of t6_15 is consistent with the map (although barely). The reality is that since the flight-by-flight variations can change bias values partway through a single flight test, it is sometimes necessary to see if *a portion* of a flight test fails the consistency test rather than checking the 2σ metrics across the whole flight.

Finally, validating on t6_18 gives Figure 2.13d which is another gray area in our use of the consistency check. Here, the first 1242 validation points undoubtedly pass the consistency test. However, the remaining points fail the consistency check on the z-component of the magnetic field. Note that the $\mathcal{GP}_y$ plot in Figure 2.13d also changes behavior near index 1380, but our analysis focuses on $\mathcal{GP}_z$ since it fails the consistency check before $\mathcal{GP}_y$ does.

96% should be taken as a "rule of thumb" rather than a magic threshold where something fundamentally different occurs with our construction of magnetic field maps. With this in mind, it is important to consider how the map will be used when deciding whether a validation test should be rejected on the basis of the consistency check alone. Here, for t6_18 in Figure 2.13d, we also include the RMSE values (in blue) for all points before and after 1242 to illustrate the accuracy of our compromise map before and after the mid-flight bias shift.

For our analysis in Chapter3, though we did not realize it at the time, these flight-by-flight variations were still present. Yet, we were still able to use our magnetic field maps (at the time, trained on a single flight test rather than our new compromise map method) to perform good attitude estimates. We believe this is due to the nature of attitude estimation which compares the angle of a measured vector against its corresponding reference vector. In this case, the flight-by-flight variations did not significantly change the *angle* of the magnetic field reference vectors given by our GPR maps.

However, our preliminary results on magnetic field position localization show that inconsistencies like those at the end of t6_15 (Figure 2.13c) and the end of t6_18 (Figure 2.13d) is enough to confuse a particle filter and cause noticeably larger estimation errors. The use of GPR-based magnetic field maps for position localization (via a particle filter) has been the motivation for much of this chapter. However, if the intended application is more robust to "small" inaccuracies of the GPR map, then a more lenient consistency check threshold may work.

2.6 Conclusions and Future Work

This chapter presents practical methods to address vehicle-induced noise when creating and validating GPR-based magnetic field maps with a UAV. First, we showed that certain portions of a vehicle's PID controller can increase the measurable magnetic noise when commanding the motors and ESCs on a flight vehicle.

Next, we introduced the flight-by-flight magnetic field variations present on our quadrotor and presented two ways to address them. The first involves moving the magnetometer further from the electronics on the vehicle while the second aims to teach the GPR-based map about the flight-by-flight variations through a "compromise" map. We show that our compromise map has similar accuracy to the "intermediate" map which uses four times as many reference points to estimate the magnetic field through in our working volume. We found that a compromise map trained on observations spaced no more than $S = 0.55$m give comparable accuracy.

Finally, we note that our proposed methods of reduction and incorporation do not encompass all possible mapping errors. As such, we introduce a "consistency check" as a rule of thumb to help experimentalists quickly assess if magnetic field observations from a particular test dataset are consistent with their magnetic field map. Later in Chapter 4, we use these new mapping techniques to demonstrate 3D position localization of a UAV using indoor magnetic field maps. Per Section 2.5.5, we aim to conduct studies similar to [72, 22, 73] that compare the position localization accuracy of magnetic maps with different output dimensions m.

Additionally, investigating time-varying magnetic fields (due to the position of large, metallic objects like elevators and doors) is important and can be complicated by vehicle-induced magnetic variations that may be confused for a change in the ambient magnetic field. Throughout this dissertation, we assume the indoor magnetic field is constant. This tends to be a good assumption unless large metal objects (like doors and elevators) are moved in the mapped area [22, 10, 48]. As such, we aim to extend our methods into time-varying magnetic field mapping that uses our consistency check (Section 2.5.6) to decide when a portion of a map may be out of date and needs to be updated.

Finally, a component-wise causal analysis of observed flight-by-flight variations may better help us understand measurable bias and noise sources to modify or eliminate entirely. For example, it is possible that simply using different brands/models of ESCs and motors could make the UAV more magnetically quiet.

CHAPTER 3

Using GPR-based Maps to Improve Attitude Estimation in Non-Constant Magnetic Fields

3.1 Introduction

The data and analysis used in this chapter were done before any work for Chapter 2 began. As such, this chapter (Chapter 3) does not use a compromise map (Section 2.5.3) to avoid overfitting. nor does it account for the UAV-induced flight-by-flight variations (Section 2.5.2.2) in any other way. The focus on this chapter is on using local magnetic field maps for IMU-based attitude estimation.

This chapter presents a pipeline to generate and validate 3D magnetic field maps, interpret the spatial variation of such maps via a new visual representation, and extract the value of magnetic field maps by leveraging them for three-degree-of-freedom (3DOF) attitude estimation. The maps utilize Gaussian process regression (GPR) to represent the magnetic field based on observations gathered during training flights. The GPR-based maps can then interpolate to provide estimates of the magnetic field at locations that were not directly sampled in the training set. These interpolated estimates of the field serve as a source of magnetic field reference vectors to estimate the attitude of a flight vehicle.

There are two opposing methods of using indoor magnetic fields in the literature. When the goal is to estimate attitude or heading using an indoor magnetic field, some works aim to remove or mitigate anomalies in the magnetic field [41, 42, 43] or reduce the impact of the magnetometer's effect on attitude estimates if disturbances are detected [44, 45, 46, 47, 48]. In this paradigm, Earth's magnetic field is seen as the signal, while any disturbances or anomalies are treated as noisy perturbations to be rejected. This methodology assumes that large perturbations are temporary and can have increased estimation errors if disturbances exist for long durations [41]. However, when performing position localization or full-pose (position and attitude) estimation [18, 13, 14, 29, 19, 21, 12, 22, 15], these "disturbances" become a signal whose spatial variation is needed to distinguish one location from another.

This chapter performs attitude estimation by mapping and leveraging magnetic field disturbances as a complement to works that use magnetic fields for position localization. Other works have shown that using magnetometers for attitude estimation indoors requires more care than outdoor spaces. Here, we investigated if creating magnetic field maps can solve this problem. To test this, we conducted a set of flight experiments both indoors and outdoors to compare the spatial variation of their respective magnetic fields and show how this spatial variation drives the need for mapping. In this chapter, the utility of a magnetic field map was measured by its accuracy in representing the local magnetic field and its ability to enable accurate 3DOF attitude estimates. As such, five distinct flight trajectories were flown to analyze the value of both indoor and outdoor magnetic field maps under different flight conditions that excite the roll, pitch, and yaw of a flight vehicle. The outline of our chapter is summarized in Figure 3.1.

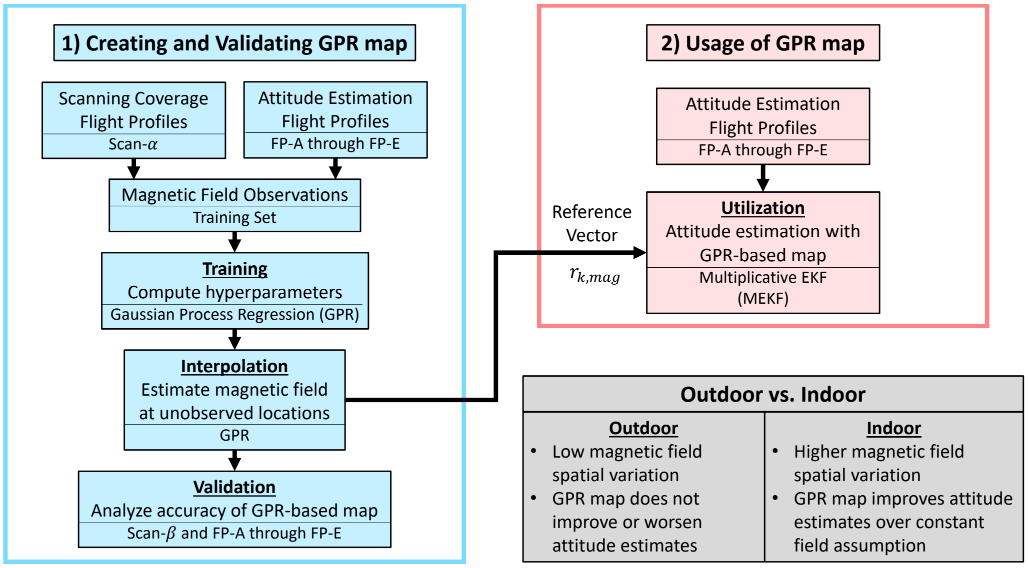


Figure 3.1: Outline of the main methods of this chapter. Several flight profile (FP) trajectories are used to train, validate, and utilize a Gaussian process regression (GPR)-based map.

The contributions of this chapter are as follows. First, we propose a new way of visualizing 3D magnetic fields that utilizes the interpolative power of GPR-based maps to show the planar direction and downward magnitude of the magnetic field at various altitudes. Next, we present different flight patterns and demonstrate how they can be used to train, validate, and utilize 3D magnetic field maps. Although this dissertation uses GPRs as the interpolative tool of the magnetic field maps, other interpolation methods might still find value in these flight patterns and our training, validation, and utilization techniques. Finally, to our knowledge, no other work has analyzed the accuracy of 3DOF attitude estimation using a single gyroscope, accelerometer, and magnetometer in an environment with the type of spatial variation in the magnetic field in which we tested. With an accurate magnetic field map, one can attain accurate 3DOF attitude estimates in locations with the spatial variation needed to perform magnetic field position localization. We understand that this last contribution comes with many qualifiers, and we will further discuss similar works in the literature and how our work is distinct from theirs in the Related Works Section (Section 3.2).

The remainder of this chapter is structured as follows. Section 3.2 of this chapter outlines related work before mathematical preliminaries are introduced in Section 3.3. The experimental setup and testing procedure are explained in Section 3.4. Finally, Section 3.5 presents results and discussion followed by a conclusion in Section 3.6.

3.2 Related Work

Previous publications have shown that modalities such as vision [74, 75] and wireless technologies [9] can be used for indoor navigation with demonstrations of these methods on UAVs [76, 77, 37]. However, these modalities add equipment to the test vehicle and sometimes require changes to the workspace (e.g., fiducial markers, ultra-wideband transmitters, etc.) for them to work effectively. Meanwhile, most modern IMUs have magnetometers that can leverage the magnetic fields already present in every building on Earth. Learning to extract value from indoor magnetic fields can improve the indoor navigation capabilities without additional equipment on or off the test vehicle.

Using magnetic fields as a modality is not without its own problems. Previous works have shown that the movement of doors, elevators [22], tool chests [10], and even whiteboards [48] can distort the magnetic field. However, these distortions have been shown to be very localized, while the remainder of the indoor magnetic field has been shown to vary slowly in time [22, 12]. Additionally, creating magnetic field maps of large volumes can be computationally expensive [19] if the user is not careful in modeling the magnetic field [10]. Despite this, several previous publications have leveraged the magnetic field of indoor spaces to improve position and attitude estimates for mobile robots such as [22, 21] with one dimension of translation, [12, 13, 14, 15, 29, 20] in planar motion and [18, 19] for three dimensions of translation. These works showed that with a magnetic field map of a workspace, it is possible to localize the position of a robot within the given magnetic field map. For localization to work, however, there must be sufficient spatial variation in the magnetic field for any algorithm to distinguish any one location from another.

To our knowledge, there is not yet a consistent method of defining spatial variation in the magnetic field. Reference [41] quantified the change in the magnetic field norm; Reference [12] discussed changes in each component of the magnetic field; Reference [20] used the mean and standard deviation of the change in the magnetic field norm as position changes. By contrast, Vallivaara et al. discussed spatial variation in a binary sense, by assessing if there was sufficient variation to localize or perform SLAM, which can depend on the parameters in the estimation algorithm [14, 13]. References [2, 3] were works on outdoor magnetic field navigation for fishing vessels and red-spotted newts, respectively. These works described the spatial variation as a gradient of change in the magnetic field strength or direction over some distance. This chapter presents a method to visualize the spatial variance of the magnetic field in a workspace and qualitatively compared the spatial variance in different areas.

Unlike position localization, attempts to estimate attitude indoors tend towards detecting and rejecting perturbations in Earth's magnetic field [41, 42, 43]. However, the methods from these works do not easily extend to full 3DOF roll, pitch, and yaw estimation and require multiple magnetometers. References [44, 45, 46, 47, 48] and others mitigated the effect of disturbances by adjusting the credibility of the magnetometer's or accelerometer's measurements when distortions in either were detected. Such works proposed different methods of detecting disturbances, but generally assumed that the disturbance was brief and the magnetic field would return back to some expected value. Reference [49] relaxed this assumption by requiring only that the norm of the magnetic field remain constant over small time periods. However, Reference [49] did not provide 3DOF attitude estimation metrics of their indoor tests to understand how well this method works for indoor navigation. Reference [48] compared their disturbance rejection method to that of [49] and other works with brief, yet large, magnetic field disturbances.

In this chapter, we performed 3DOF attitude estimation using a gyroscope, an accelerometer, and a single magnetometer despite magnetic field disturbances from buildings and human-made structures. Using magnetic field maps, we leveraged the spatial variation of the field to assist in 3DOF attitude estimation rather than rejecting these features as disturbances. Such capabilities are important for full pose estimation using magnetometers indoors since spatial variation in the magnetic field is desirable for position localization. Although we drew a distinction in our method from the disturbance rejection paradigm, our mapping-based methods can still be paired with the disturbance rejection methods if a magnetic field map is locally inconsistent in a region near doors, elevators, or other mobile ferromagnetic objects. Time-varying magnetic field mapping was addressed in [10], but not in this dissertation.

Previous works that generated magnetic field maps can be split into two groups. Some performed simultaneous localization and mapping (SLAM) and were able to construct a magnetic field map as they estimated the pose of their vehicle [18, 13, 14]. Others first had their robot traverse the workspace to learn the magnetic field and construct a map, then utilized this map in a separate trial to perform state estimation or localization [29, 19, 21, 12, 22, 15, 20]. Our work adopted the philosophy of the latter group by performing a set of *training* test flights to construct a magnetic field map followed by a separate set of *utilization* flights meant to assess the accuracy of our attitude estimates by leveraging an existing map. We analyzed the accuracy of our maps by using *validation* test flights and compared two methods of training a GPR-based magnetic field map.

There are different methods of representing magnetic field maps with varying complexities depending on the degrees of freedom of the state to estimate. A magnetic field map with an input of dimension p and output of dimension m is denoted as a $p \to m$ map. Generally, p depends on how the vehicle can traverse in space (e.g., $p = 2$ for ground-based, wheeled robots and $p = 3$ for flying vehicles or pedestrian localization), while m depends on how many components of the magnetic field are to be tracked. Typically, $m = 1$ when the map gives the norm of the magnetic field for each position, and $m = 3$ when all three components of the magnetic field are tracked. The methods of representing the magnetic field tend to differ based on the size of p. Note that the works that tried to detect and mitigate perturbations in Earth's magnetic field did not create local magnetic field maps [43, 42, 41, 44, 45, 46, 47].

Reference [21] constructed a $1 \to 1$ map of the magnetic field's declination using least squares to interpolate and create unique signatures at select locations of interest. Reference [22] constructed a $1 \to 1$ map of the magnetic field norm via linear interpolation. Reference [22] preceded [14, 13] in using Gaussian processes to create $2 \to 3$ maps. References [15, 78] constructed $2 \to 3$ and $2 \to 2$ maps respectively using bilinear interpolation of the four nearest observations gathered at evenly spaced locations. Reference [12] compared the position localization accuracy of a $2 \to 2$ map against a $2 \to 3$ magnetic field map by sampling the magnetic field at discrete locations. However, Reference [12] did not interpolate the magnetic field between observed locations, which made their "map" differ from other maps cited in this discussion. In [29, 19], Akai and Ozaki constructed $2 \to 3$ and $3 \to 1$ maps, respectively, inside a building to demonstrate the use of Gaussian processes maps in large volumes. Finally, Reference [18] created a $3 \to 3$ map also using Gaussian processes.

Linear interpolation is intuitive when $p = 1$ [22], but the bilinear interpolation for $p = 2$ used in [15, 78] requires magnetic field observations to be gathered at points on an evenly spaced grid throughout the workspace. Such a constraint is restrictive in how magnetic field map observations are gathered (Section 3.5.4), is time-consuming for $p = 3$ [19], and does not allow for extrapolation of points outside of those found in the training set (Section 3.5.1). It is not clear how the $p = 1$ least-squares-based feature selection in [21] could extend to $p = 2$ or $p = 3$. By contrast, Gaussian processes are a prominent way of representing local magnetic fields for $p = 2$ or $p = 3$ as per [14, 13, 29, 19, 18]. Many GPR-based applications appear to build on the methods of [67]. Gaussian process regression has no hard constraints on the spatial layout of observations, is able to extrapolate and estimate at locations outside of those found in the training set, as well as provides uncertainty metrics on the confidence of a provided magnetic field estimate. However, GPR comes with a high computational cost as a function of the training set size [10, 19]. Due to these listed advantages and their use in other magnetic field mapping works, we used three independent Gaussian processes to represent a $3 \to 3$ magnetic field map.

Kalman filters are a common way to conduct state estimation of dynamic systems: particularly when the rate of change of certain states can be measured or modeled well. Since Kalman filters assume the state and sensor measurements are all Gaussian, particle filters can outperform Kalman filters if parts of the system are non-Gaussian. However, particle filters come at an increased computational cost to track multiple hypotheses of the state, while Kalman filters only work with a single hypothesis. Methods such as TRIAD and derivatives of Wahba's problem are for attitude *determination* and do not incorporate system dynamics to integrate current estimates forward in time. In this chapter, we used a multiplicative extended Kalman filter (MEKF) which uses quaternions to represent attitude, is more careful to respect the constraints of rotations than a typical extended Kalman filter, and has been used on many platforms with an IMU [79, 80, 47, 81, 82, 83]. Here, we used the MEKF derived from [84, 85] to perform attitude estimation with a gyroscope, accelerometer, and magnetometer.

3.3 Mathematical Formulation

This section introduces the mathematical preliminaries used throughout the chapter. First, we summarize the multiplicative extended Kalman filter (MEKF) from [84, 85] that we used to perform attitude estimation from the vehicle's sensors. Next, we show how the Gaussian processes regression (GPR) method from [67] can be used to generate maps of the magnetic field in an area of interest. Finally, the magnetometer calibration method from [68] is introduced with a nine-parameter sensor model.

We point out a change in notation from Chapter 2. The variables $\tilde{\boldsymbol{z}}_{k,accl}$ and $\tilde{\boldsymbol{z}}_{k,mag}$ will now refer to accelerometer and magnetometer measurements. Additionally, $\boldsymbol{r}_{k,accl}$ and $\boldsymbol{r}_{k,mag}$ are reference vectors for the accelerometer and magnetometer respectively. In Chapter 2, $\boldsymbol{r}$ was a location in 3D space.

3.3.1 Multiplicative Extended Kalman Filter for 3DOF Attitude Estimation

The MEKF is a variation on the extended Kalman filter (EKF) that computes attitude error as the quaternion product of the true attitude and the estimated attitude [84, 85]. As with the EKF, the MEKF propagates the current attitude forward in time with a prediction step and uses measurements to correct this prediction in the update step. Estimates of the gyroscope (gyro) bias are used to correct for gyro drift. The formulation of the MEKF presented here was from Chapter 7.1 of [84]. A quaternion was used to represent vehicle global attitude. Instead of updating the four-element quaternion at each step, the MEKF tracks a three-element parameterization of the attitude error used to adjust the global attitude quaternion. As recommended by Section 6.2 of [85], gyro measurements were used for the prediction step instead of a dynamics model of the vehicle. An accelerometer and a magnetometer were used for the correction step.

The MEKF estimates vehicle attitude relative to a world-fixed frame with a quaternion $\hat{\boldsymbol{q}}_k$ and a gyroscope bias parameter $\hat{\boldsymbol{\beta}}_k$ at discrete time step k. During the prediction step, the gyro measurement was used to integrate the quaternion estimate to the next discrete time step $\hat{\boldsymbol{q}}_{k+1}^{-} = \boldsymbol{\Omega}(\hat{\boldsymbol{\omega}}_k, \Delta t)\hat{\boldsymbol{q}}_k^{+}$ where superscript $^{-}$ indicates a value immediately after the prediction step and $^{+}$ is the value immediately after the update step. During the update step, the MEKF uses a six-component state vector $\Delta\hat{\boldsymbol{x}}_k = \begin{bmatrix}\delta\hat{\boldsymbol{\alpha}}_k^\top & \Delta\hat{\boldsymbol{\beta}}_k^\top\end{bmatrix}^\top$, where $\delta\hat{\boldsymbol{\alpha}}_k$ is the three-component attitude estimate error vector and $\Delta\hat{\boldsymbol{\beta}}_k = \boldsymbol{\beta}_k - \hat{\boldsymbol{\beta}}_{\boldsymbol{k}}$ is a three-component gyro bias estimate error vector, to correct the prediction. In effect, the update step computes $\hat{\boldsymbol{q}}_{k+1}^{+} = \hat{\boldsymbol{q}}_{k+1}^{-} + \frac{1}{2}\boldsymbol{\Xi}(\hat{\boldsymbol{q}}_{k+1}^{-})\delta\hat{\boldsymbol{\alpha}}_k$ for the attitude and $\hat{\boldsymbol{\beta}}_{k+1}^{+} = \hat{\boldsymbol{\beta}}_k^{-} + \Delta\hat{\boldsymbol{\beta}}_k^{-}$ for the gyro bias.

As with all Kalman filters, the MEKF also keeps track of its estimate uncertainty with a 6×6 covariance matrix $\boldsymbol{P}_k$ where the six-dimensional estimate is assumed to be drawn from a Gaussian distribution $\Delta\hat{\boldsymbol{x}}_k \sim \mathcal{N}(\Delta\hat{\hat{\boldsymbol{x}}}_k, \boldsymbol{P}_k)$.

3.3.1.1 MEKF Prediction Step

The MEKF prediction step uses gyro data to march the current estimate forward. Gyroscope values are modeled as:

$$\tilde{\boldsymbol{\omega}}_k = \boldsymbol{\omega}_k + \nu \qquad \nu \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_k) \tag{3.1}$$

where $\boldsymbol{\omega}_{\boldsymbol{k}}$ is the true angular velocity and $\boldsymbol{Q}_k$ is the process noise matrix. The MEKF first offsets the measured gyro value by current gyro bias estimate $\hat{\boldsymbol{\omega}}_{\boldsymbol{k}} = \tilde{\boldsymbol{\omega}}_k - \hat{\boldsymbol{\beta}}_k$ where $\tilde{\boldsymbol{\omega}}_k$ is the raw gyro measurement and $\hat{\boldsymbol{\omega}}_{\boldsymbol{k}}$ was offset by our bias estimate $\hat{\boldsymbol{\beta}}_k$. We then updated the quaternion and gyro bias with:

$$\begin{aligned}\hat{\boldsymbol{q}}_{k+1}^{-} &= \boldsymbol{\Omega}(\hat{\boldsymbol{\omega}}_k, \Delta t)\hat{\boldsymbol{q}}_k^{+}\\ \hat{\boldsymbol{\beta}}_{k+1}^{-} &= \hat{\boldsymbol{\beta}}_k^{+}\end{aligned} \tag{3.2}$$

where $\boldsymbol{\Omega}(\hat{\boldsymbol{\omega}}_k, \Delta t)$ was defined in Chapter 7.1.2 of [84] and depends only on the bias-adjusted gyro measurement $\hat{\boldsymbol{\omega}}_{\boldsymbol{k}}$ and the amount of time given by discrete time step Δt. Note the gyro bias estimate is unchanged during the prediction step.

The covariance matrix becomes:

$$\boldsymbol{P}_{k+1}^{-} = \boldsymbol{\Phi}_k \boldsymbol{P}_k^{+} \boldsymbol{\Phi}_k^\top + \boldsymbol{\Upsilon} \boldsymbol{Q}_k \boldsymbol{\Upsilon}^\top \tag{3.3}$$

where $\boldsymbol{\Phi}_k = \boldsymbol{\Phi}_k(\hat{\boldsymbol{\omega}}_k, \Delta t)$ is a state transition matrix that depends on $\hat{\boldsymbol{\omega}}_k$ and Δt, while $\boldsymbol{\Upsilon}$ is a constant matrix. $\boldsymbol{Q}_k = \boldsymbol{Q}_k(\sigma_u, \sigma_v, \Delta t)$ is a process noise matrix that depends on gyroscope noise parameters σ_u and σ_v and the change in time Δt. All three of these matrices were defined in Chapter 7.1.2 of [84].

3.3.1.2 MEKF Update Step

The update step receives accelerometer $\tilde{\boldsymbol{z}}_{k,accl}$ and magnetometer $\tilde{\boldsymbol{z}}_{k,mag}$ measurements, compares these to their respective “reference vectors” in world frame $\boldsymbol{r}_{k,accl}$ and $\boldsymbol{r}_{k,mag}$, and uses the difference between the measurement and reference vector to update the estimate. Reference vectors are effectively what we expected each sensor to measure given the current vehicle state estimate. It was assumed that the gyro data used in the prediction step be measured at the same time as the accelerometer and magnetometer data. On our test UAV, the magnetometer was sampled less frequently than the accelerometer and gyroscope. As such, we used a sequential update method explained in Chapter 7.1.3 of [84] so that the matrices used in the update step need not change size depending on the availability of magnetometer data.

First, we define a rotation matrix $\boldsymbol{A}(\boldsymbol{q}_k)$ that transforms a vector represented in the world frame to the same vector represented in the vehicle body frame. We then define the Kalman gain:

$$\boldsymbol{K}_k = \boldsymbol{P}_k^- \boldsymbol{H}_k^\top \left[\boldsymbol{H}_k \boldsymbol{P}_k^- \boldsymbol{H}_k^\top + \sigma_{\_} \boldsymbol{I}_{3\times3}\right]^{-1} \tag{3.4}$$

where $\boldsymbol{I}_{3\times3}$ is the 3×3 identity matrix, and the 3×6 sensitivity matrix $\boldsymbol{H}_k$ is:

$$\boldsymbol{H}_k = \left[\left[\boldsymbol{A}(\hat{\boldsymbol{q}}_k^-)\boldsymbol{r}_{k,\_}\right]^\times \quad \boldsymbol{0}_{3\times3}\right] = \frac{\partial h_{\_}}{\partial(\Delta \boldsymbol{x})} = [H_{\boldsymbol{\alpha}} \quad H_{\boldsymbol{\beta}}]. \tag{3.5}$$

$\sigma_{\_}$ is a measurement noise parameter that depends on which sensor is being used (i.e., σ_{accl} or σ_{mag}) and $\boldsymbol{r}_{k,\_}$ is the corresponding accelerometer or magnetometer reference vector. $\boldsymbol{H}_k$ is a Jacobian matrix of the measurement model:

$$\begin{aligned} \tilde{\boldsymbol{z}}_{k,mag} &= h_{mag}(\boldsymbol{q}_k) + \eta_{mag} \\ \tilde{\boldsymbol{z}}_{k,accl} &= h_{accl}(\boldsymbol{q}_k) + \eta_{accl} \end{aligned} \tag{3.6}$$

where

$$\begin{aligned} \eta_{mag} &\sim \mathcal{N}(0,\, \sigma_{mag}\boldsymbol{I}_{3\times3}) \\ \eta_{accl} &\sim \mathcal{N}(0,\, \sigma_{accl}\boldsymbol{I}_{3\times3}). \end{aligned} \tag{3.7}$$

In general, h_{mag} varies as a function of position if the magnetic field is not constant. However, this has no effect on the Jacobian $\boldsymbol{H}_k = \frac{\partial h_{\_}}{\partial(\Delta \boldsymbol{x})}$ since position is not part of the state $\boldsymbol{\Delta x}$. The sensor-based subscript (i.e., $_{mag}$ or $_{accl}$) was omitted from $\boldsymbol{H}_k$ for brevity. Finally, the operator $[\boldsymbol{a}]^\times$ takes a vector $\boldsymbol{a} = [a_x \; a_y \; a_z]^\top$ as the input and returns a skew-symmetric cross-product matrix where:

$$[\boldsymbol{a}]^{\times} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \quad S.T. \quad \boldsymbol{a} \times \boldsymbol{b} = [\boldsymbol{a}]^{\times} \boldsymbol{b}. \tag{3.8}$$

The estimate error state is then updated with:

$$\begin{aligned} \boldsymbol{\epsilon}_{k,\_} &= \tilde{\boldsymbol{z}}_{k,\_} - \boldsymbol{A}(\hat{\boldsymbol{q}}_k^-)\boldsymbol{r}_{k,\_} \\ \Delta\hat{\boldsymbol{x}}_k^+ &= \Delta\hat{\boldsymbol{x}}_k^- + \boldsymbol{K}_k \left[\boldsymbol{\epsilon}_{k,\_} - \boldsymbol{H}_k \Delta\hat{\boldsymbol{x}}_k^-\right] \end{aligned} \tag{3.9}$$

where $\boldsymbol{\epsilon}_{k,\_}$ is often denoted the "residual" or "innovation" and $\tilde{\boldsymbol{z}}_{k,\_}$ is either the accelerometer or magnetometer measurement. $\Delta\hat{\boldsymbol{x}}_k^+$ is used to update the estimate quaternion $\hat{\boldsymbol{q}}_k^+$ and gyro bias estimate $\hat{\boldsymbol{\beta}}_k^+$ as specified in Chapter 7.1.1 of [84]. Finally, the covariance matrix is updated with:

$$\boldsymbol{P}_k^+ = \left[\boldsymbol{I}_{6\times 6} - \boldsymbol{K}_k \boldsymbol{H}_k\right] \boldsymbol{P}_k^- . \tag{3.10}$$

3.3.2 Gaussian Process Regression

The purpose of the GPR-based map is to take observations around the flight space and interpolate these measurements to predict the magnetic field at any point in and around the workspace. When estimating attitude, the GPR-based map is queried at the vehicle's location to provide a magnetic field reference vector $\boldsymbol{r}_{k,mag}$ to the MEKF.

Gaussian process regression uses a "training set" of observations to make predictions of the magnetic field. The predictions tend to only be accurate near positions used in the training set. Each observation has a 3D position and the three components of the magnetic field observed at position ($\tilde{B}_x$, $\tilde{B}_y$, $\tilde{B}_z$). Throughout this dissertation, we used three independent Gaussian processes, each with a 3D input (the position) and a scalar prediction as the output (the x, y, or z component of the magnetic field). As such, for a training set with n observations, we defined the design matrix $\boldsymbol{X} \in \mathbb{R}^{3\times n}$ as a collection of the n 3D positions and defined vectors $\tilde{\boldsymbol{y}}_x$, $\tilde{\boldsymbol{y}}_y$, $\tilde{\boldsymbol{y}}_z \in \mathbb{R}^n$ as the corresponding magnetic field observations for $\tilde{B}_x$, $\tilde{B}_y$, and $\tilde{B}_z$, respectively.

In [67], Rasmussen and Williams defined a Gaussian process as a distribution over functions written as:

$$f(\boldsymbol{x}) \sim \mathbb{GP}\left(\boldsymbol{0},\, k(\boldsymbol{x}, \boldsymbol{x}')\right) \tag{3.11}$$

where $\boldsymbol{0}$ is the zero-mean function, k is a covariance function (or kernel), and $\boldsymbol{x}$, $\boldsymbol{x'} \in \mathbb{R}^3$ are the 3D positions. The covariance function quantifies correlation, or similarity, in observations based on their spatial proximity. Throughout this dissertation, we used three separate Gaussian processes $\mathbb{GP}_x$, $\mathbb{GP}_y$, and $\mathbb{GP}_z$ that have a zero mean function and a squared exponential covariance function. Given a training dataset and a set of hyperparameters, each Gaussian process can be queried at a 3D location to provide a prediction of a single component of the magnetic field along with the uncertainty of each prediction. The "map" of the magnetic field is actually these three Gaussian processes with their corresponding training sets $D_x = (\boldsymbol{X}, \tilde{\boldsymbol{y}}_x)$, $D_y = (\boldsymbol{X}, \tilde{\boldsymbol{y}}_y)$, and $D_z = (\boldsymbol{X}, \tilde{\boldsymbol{y}}_z)$ and their respective optimized hyperparameters.

The hyperparameters are adjustable terms from the covariance function. With a squared exponential covariance function, the hyperparameters are a length scale and a signal standard deviation. Each of the three Gaussian processes had distinct hyperparameters computed by minimizing the negative log marginal likelihood over their corresponding training datasets. Hyperparameter optimization was performed using the `gpml-matlab` toolbox ^{1}.

In this dissertation, we refer to the three Gaussian processes as a single object (e.g., "map", "magnetic field map", "GPR-based map", "GPR") since they were used together to predict the full magnetic field vector at a desired location. These predictions are used by the MEKF as magnetic field reference vectors $\boldsymbol{r}_{k,mag}$ and to assess the accuracy of the GPR-based maps via a validation set per Sections 3.5.1 and 3.5.4. Finally, the predictions are used to visualize the magnetic field of our workspaces in Sections 3.5.2.

3.3.3 Magnetometer Measurement Model

To calibrate the magnetometer on our flight vehicle, we used the model and procedure from [68]. Measurements from a three-axis magnetometer can differ from the actual magnetic field because of noise, sensor bias, non-unit scaling factors, and nonorthogonalities in the alignment of the magnetometer's measurement axes. These imperfections are captured by the following model:

$$\tilde{B}_x = aB_x + x_0 + \eta_x \tag{3.12}$$

$$\tilde{B}_y = b(\tilde{B}_y \cos(\rho) + B_x \sin(\rho)) + y_0 + \eta_y \tag{3.13}$$

$$\begin{aligned}\tilde{B}_z = c(&B_x \sin(\lambda) + B_y \sin(\phi)\cos(\lambda) \\ &+ B_z \cos(\phi)\cos(\lambda)) + z_0 + \eta_z\end{aligned} \tag{3.14}$$

where ($\tilde{B}_x$, $\tilde{B}_y$, $\tilde{B}_z$) are measured values, (B_x, B_y, B_z) are the true magnetic field, (a, b, c) are scaling factors, (x_0, y_0, z_0) are constant bias, (η_x, η_y, η_z) are zero-mean noise, and (ρ, λ, ϕ) are angles quantifying the nonorthogonality of the axes (Figure 1 in [68]). The goal is to estimate

$\alpha = [a, b, c, x_0, y_0, z_0, \rho, \lambda, \phi]^\top$ by minimizing:

$$\begin{aligned} \Delta B &= B_R^2 - B^2 \\ &= B_R^2 - (B_x^2 + B_y^2 + B_z^2) \\ &= B_R^2 - g(\tilde{B}_x, \tilde{B}_y, \tilde{B}_z, \alpha) \end{aligned} \tag{3.15}$$

where B_R is the magnitude of the magnetic field reference vector obtained from the GPR-based map and $g()$ is a nonlinear function obtained by re-arranging Equations (3.12) through (3.14) for (B_x, B_y, B_z). The estimated α is obtained by iterative nonlinear least squares with $\alpha_0 = [1, 1, 1, 0, 0, 0, 0, 0, 0]^\top$ as an initial condition.

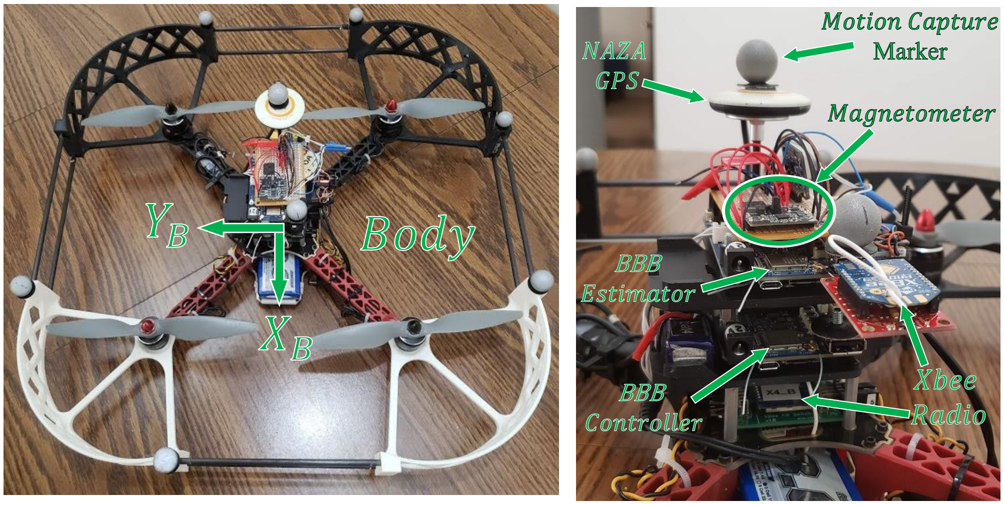
3.4 Experimental Setup and Procedure

The purpose of our experiments was to gather sensor data to create magnetic field maps, assess their accuracy in representing the local magnetic field, and analyze their ability to enable the 3DOF attitude estimation of a UAV. To this end, we created coverage flight patterns (Scan-α and Scan-β) to study different ways of gathering observations to train magnetic field maps and investigate the interpolation and extrapolation capabilities of GPR-based maps. Then, we flew five different flight profiles (FP-A through FP-E) that distinctly excited the roll, pitch, and yaw of the flight vehicle to provide different test cases for attitude estimation. These profiles were flown both indoors and outdoors to investigate how spatial variation in the magnetic field affects the utility of our magnetic field maps.

3.4.1 Equipment, Facilities, and Setup

Our experimental flight vehicle, shown in Figure 3.2, is a multirotor equipped with a gyroscope, accelerometer, and magnetometer. The multirotor used here were introduced in Section IV of [86] and flies an MPU9250 IMU (gyro, accelerometer, magnetometer) and BMP280 barometer. Since the work in [86], a second BeagleBone Blue microprocessor, a PNI RM3100 magnetometer, and a DJI Naza-M V2 GPS unit were added to the vehicle. In addition, the flight controller was changed to `rc_pilot_a2sys` modified from the open-source `rc_pilot` [3].

Figure 3.3 shows the major components of the experimental setup. The flight vehicle was equipped with two BeagleBone Blue (BB-Blue) microprocessors both running RC pilot. The controller BB-Blue uses pregenerated trajectories as guidance and uses a linear quadratic regulator (LQR) controller to maneuver the vehicle along the desired trajectory. The controller used in this chapter was a modified version of the LQR controller from [87]. For navigation, the MPU9250 IMU has a built-in processor that provides roll and pitch estimates, while motion capture position and yaw estimates are streamed from the ground station to the vehicle via wireless XBee radios during each flight.



(**a**) Vehicle body frame (**b**) Flight vehicle components

Figure 3.2: Flight vehicle.

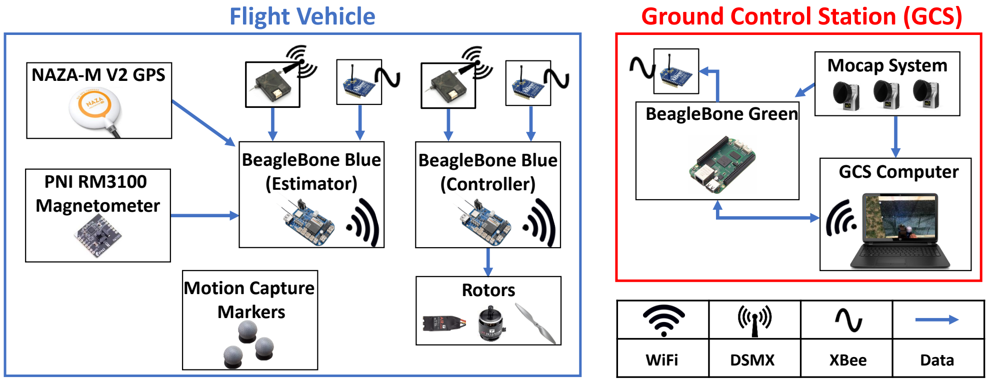


Figure 3.3: Diagram of the devices used to the gather experimental data.

The estimator BB-Blue logs data from the NAZA-M V2 GPS, PNI RM3100 magnetometer, MPU9250 IMU (gyro and accelerometer), BMP280 barometer, and Qualisys motion capture with a timestamp from when each respective value was logged. The MPU9250 IMU and BMP280 barometer are included with each BeagleBone Blue and are thus not included individually in Figure 3.3. The IMU was sampled at 200Hz, the magnetometer at approximately 144Hz, the barometer at 20Hz, and the GPS at 4Hz. The Qualisys (outdoor) motion capture pose estimates were computed at 100Hz, while the OptiTrack (indoor) system provides ground truth at 120Hz. The GPS and barometer data were not used in this chapter.

The LQR controller from [87] did not use integrator terms on position, altitude, or yaw angle. The controller was not tuned rigorously for this analysis, so there tended to be a slight, but slow decrease in altitude during each flight. In addition, the guidance method assumed the processor updates the outputs at 200Hz consistently, which is not true in reality. The result is that the desired position, velocity, and attitude fed to the controller is sometimes behind a location the drone has already passed. This causes some undesirable motion from a control perspective, but actually presents interesting oscillatory features for our MEKF to track. For each flight, the vehicle was placed near the motion capture origin with the $+x$ axis of the vehicle's body frame roughly aligned with the $+x$ axis of the motion capture origin. To prevent large, unpredictable motion during takeoff, the controller's initial reference xy position and heading were the initial position and heading of the vehicle before takeoff

The Network Time Protocol (NTP) was used to synchronize all data-gathering processors to ensure proper temporal correspondence between sensor data gathered onboard the vehicle and the ground truth from the motion capture system. There was a $\sim$40ms delay when wirelessly transmitting the motion capture pose from the ground station to the flight vehicle via Xbee radio. In order to correlate onboard sensor data to motion capture ground truth values, all BeagleBone processors (the two BB-Blues on the vehicle and the BB-Green on the ground station per Figure 3.3) used NTP to synchronize their local clock with the ground station's clock during the initial setup. Each BeagleBone then recorded their NTP clock offset from the ground station clock at the start and end of each flight trial. These offsets were linearly interpolated for all other points throughout the flight to serve as a method of synchronization between sensor data captured onboard the vehicle and motion capture data captured on the ground station *before* it was sent over the wireless XBee radio channel and subject to the 40ms delay. Note, after the work in this chapter, we came to use `chrony`[2] to solve this synchronization issue with less post-processing.

For the outdoor and indoor flight test sessions, a single gyro and accelerometer calibration was performed during the initial setup to remove any large offsets in either sensor. Calibration bias and scaling factors were automatically applied to all raw values for both sensors as a feature of the MPU9250 IMU. The outdoor dataset used in this chapter was not originally meant for an analysis of magnetic fields. As such, no magnetometer calibration procedure was performed before the outdoor flight data were gathered. Since the indoor data were gathered specifically for this chapter, we used the calibration method from [68]. Table 3.1 shows the scaling factor, bias, and nonorthogonality metrics for the PNI RM3100 applied to the magnetometer measurements for the indoor flight tests. A perfect sensor would have scale factors of 1 and bias and nonorthogonalities of 0. Even though we did not have a calibrated magnetometer for the outdoor readings, Table 3.1 gives us confidence that the RM3100 provides good measurements with uncalibrated errors that are small relative to the 50 μT field at the outdoor flight space.

Table 3.1: Calibration parameters of RM3100 magnetometer from indoor flight tests.

Scale Factor (-)			**Bias (μT)**			**Non-Orthogonalities (°)**		
a	b	c	x_0	y_0	z_0	ρ	λ	ϕ
1.01	0.955	0.942	−1.26	−2.46	3.24	0.182	2.28	−0.118

All outdoor tests were conducted at the University of Michigan's outdoor netted drone facility, M-Air, equipped with 30 Qualisys motion capture cameras for ground truth position and attitude. The indoor flight tests were conducted in the Robot Fly Lab in the Ford Motor Company Robotics Building equipped with 8 OptiTrack motion capture cameras for the ground truth (Figure 2.2). The outdoor data were collected in a single August 2020 evening in M-Air. The indoor data were gathered in March 2021 inside the Robot Fly Lab. The usable volume for the indoor flight lab space is smaller than that of M-Air. As such, the flight profiles used outdoors (Section 3.4.2) covered larger areas than their indoor equivalents (Section 3.4.3)

For the outdoor flights, the motion capture origin was set at the center of M-Air, which was surveyed to be 42°17'39.95257" N, 83°42'37.59818" W. According to Google Earth, the center of M-Air has an altitude of 270m above sea level. M-Air has pillars that align with geographic north–south and east–west. Using these pillars, the $+x$ axis of the motion capture system was aligned with geographic north (by visual inspection) and the $+z$ motion capture axis aligned with the gravity vector using a leveling tool.

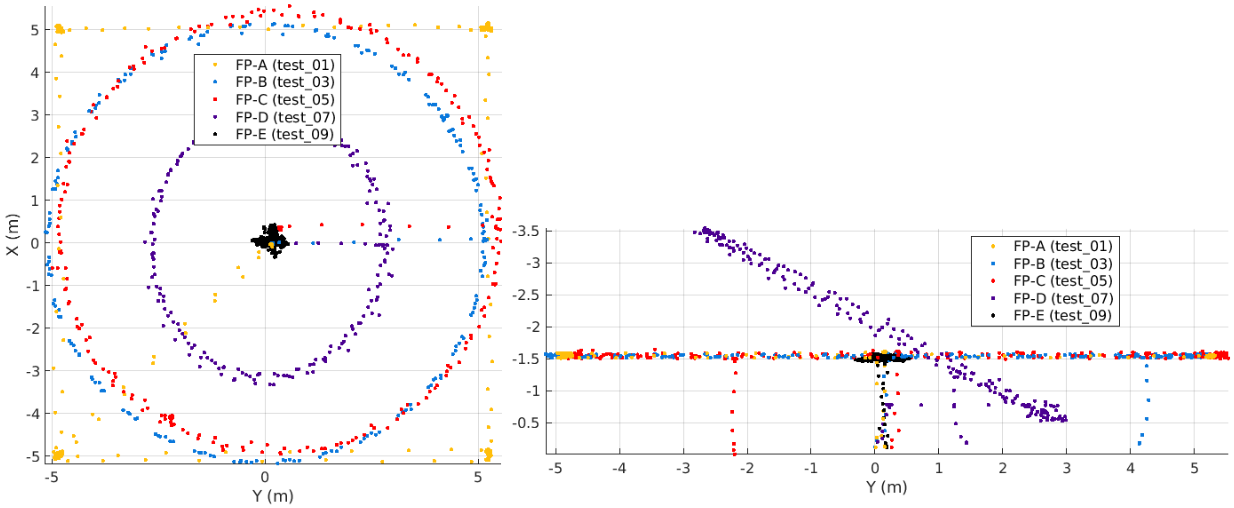
For indoor flights, the $+x$ axis of the motion capture origin was aligned with the long axis of the indoor flight space. There is no special relationship between the indoor fly lab's selected origin and Geographic North. To achieve consistency between flights, we marked the position and orientation of the desired motion capture origin on the ground. The $+x$ axis of the *indoor* motion capture origin was aligned with this tape marker via visual inspection while the $+z$ axis of the motion capture system was aligned with the gravity vector using a leveling tool.

Motion capture position and attitude estimates are given with respect to a virtual rigid body defined from a set of motion capture markers on the vehicle. The body frame for the virtual vehicle is the centroid of all motion capture markers used to define the rigid body and does not necessarily coincide with the center of mass of the actual vehicle. For outdoor data, the orientation of the virtual rigid body was aligned with the physical vehicle through select placement of motion capture markers, a feature of the Qualisys motion capture system. The indoor OptiTrack motion capture system has no feature to align the virtual rigid body with the physical vehicle, but our motion capture targets were only placed once for all indoor flights, so any systematic bias was consistent in all data gathered. To characterize the performance of our attitude estimator under different flight maneuvers, each flight test followed a trajectory from one of the flight profiles detailed below.

3.4.2 Outdoor Flight Profiles

To investigate the value of magnetic field maps for attitude estimation, the multirotor flew five different flight profiles, each meant to excite roll, pitch, and yaw in different ways. Flight Profile A (FP-A) is a square trajectory where the vehicle hovers at each corner of the square. This flight profile offers large translations with constant altitude and heading. Flight Profiles B and C are level circle trajectories with large, continuous translations, constant altitude, and a desired heading that is always tangent to the circle. FP-B has a slower tangential speed than FP-C. Flight Profile D is an inclined circle similar to B and C, but altitude is no longer constant. Finally, profile E is an angular oscillation test featuring small translation, but large oscillations in roll and pitch. All trajectories were flown at an altitude of 1.5m except FP-D (the inclined circle) with the desired altitude varied from 0.7m to 3.3m. Figure 3.4 shows the five outdoor flight profiles. All outdoor flight tests have two-digit identification numbers per Table 3.2.

FP-A flown outdoors has side length 10m. The vehicle is commanded to hover at each corner of the square for 10s followed by a 10s cubic spline translation to the next corner. From the takeoff point, the vehicle moves to the $-x, -y$ (southwest) corner, then traces a square by moving $+x$ (north), $+y$ (east), $-x$ (south), then $-y$ (west) two times before returning to the origin to land. Cubic spline velocities start and end at rest while position constraints change as appropriate to traverse the square resulting in a maximum commanded speed of 1.5m/s.



(**a**) xy View

(**b**) yz View

Figure 3.4: Five outdoor flight profiles.

Circular profiles FP-B and FP-C have turning radius 5m, while FP-D is inclined with respect to the horizontal plane by 30° and has turning radius 3m. FP-B and FP-C maintain an altitude of 1.5m, while the inclined circle's altitude varies from 0.7m to 3.3m. The tangential velocity of Profiles B and D are both “slow” at 1m/s, while Profile C has a commanded tangential velocity of 3m/s. For all three circular trajectories, the UAV heading is directed to follow a tangent to the circular path.

Flight Profile E follows a sinusoidal desired position in the x and y axes. The net result is small linear motion and fast changes in vehicle roll and pitch. The sinusoidal reference trajectory has a 0.3m amplitude with a frequency of 3Hz. In each trial of FP-E, the vehicle oscillates along the pitch axis ($\pm x$ direction) first for 90s followed by oscillations in the roll axis ($\pm y$ direction) for the same duration. The UAV is commanded to maintain a constant heading throughout the flight and remain stationary for 5s between roll and pitch oscillations to decouple motion in each axis.

Two three-minute trials of each flight profile were conducted with all sensors sampled per Section 3.4. One trial from each profile was used to create the GPR-based map (training data), while the other was used to perform attitude estimation with an MEKF (utilization data). The specific trials used for training and utilization are shown in Table 3.2. For the training data, all observations were downsampled to 1Hz before being fed into the GPR. In the case of the utilization data, the MEKF was given all sensor data at the respective sampling rates.

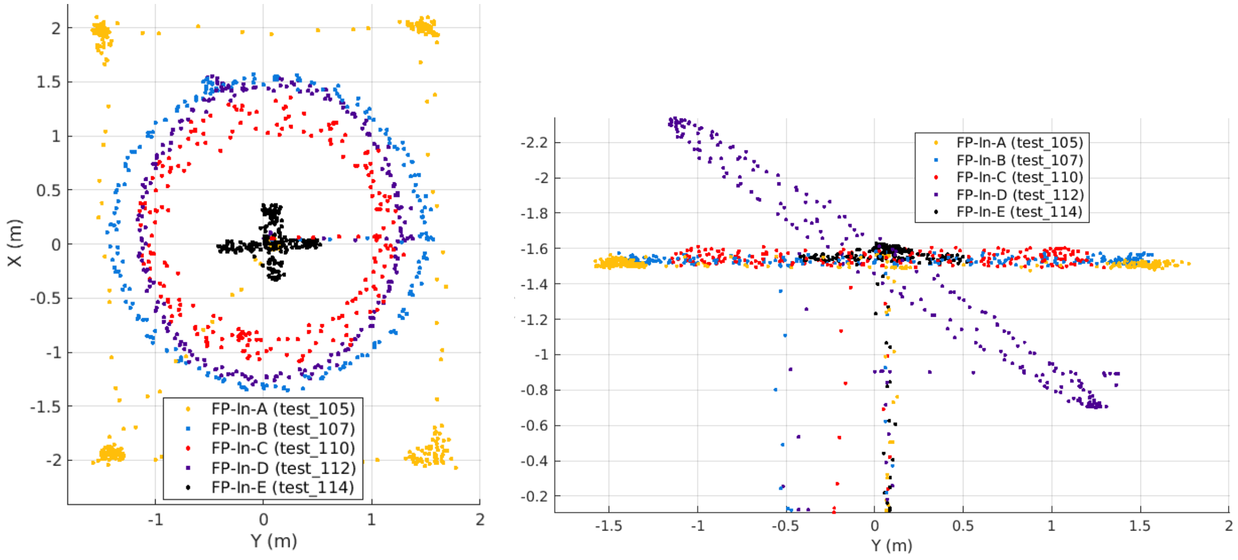
The flight profiles in this chapter were meant to capture different types of flight methods. FP-A has short periods of motion with long segments of stationary hover that emulate surveillance-based applications of multirotors. This trajectory enables the drone to experience the planar spatial variation of the workspace by covering the span of the workable area. FP-B was meant to emulate the continuous motion of fixed-winged vehicles, but we settled on a circular trajectory due to the space constraints of our workspace. In retrospect, an elongated and beveled rectangular circuit would allow for periods of straight flight with some banked turns to better represent fixed-winged flight. FP-C is a more aggressive version of FP-B, while FP-D was designed to capture the vertical spatial variation of the workspace. Finally, FP-E is a stress test of roll and pitch estimation, but not a typical maneuver one would expect a flight vehicle to perform.

Our vehicle and its controller are not able to safely execute aggressive maneuvers as seen in drone racing or in acrobatic drones. Though such cases would be good stress tests for the attitude estimation portion, they were not achievable with our current setup. As such, we used FP-C and FP-E as stand-ins for “aggressive” flight maneuvers within the capabilities of our system.

3.4.3 Indoor Flight Profiles

The five flight profiles from the outdoor dataset were repeated inside the Robot Fly Lab with the following differences. To distinguish between outdoor and indoor trajectories, let FP-A denote the outdoor level square, while FP-In-A the indoor variant; analogous notation was adopted for the other four flight profiles. The indoor profiles have smaller $x - y$ footprints to fit the indoor flight space, so more laps were traversed to achieve the desired three-minute flight duration.

FP-In-A is a rectangle with a 4m side length in the x direction and a 3m side length in y; it is otherwise the same as FP-A. FP-In-B (slow level circle) has a 1.5m radius, but is otherwise identical to FP-B. FP-In-C (fast level circle) has a 1m radius and a commanded tangential velocity of 1.5m/s, but is otherwise identical to FP-C. FP-In-D (slow inclined circle) has a 1.5m radius and 30° inclination, which bounds its commanded altitude between 0.85m and 2.15m. FP-In-E (stationary oscillations) is identical to the outdoor FP-E. Figure 3.5 shows the five indoor flight profiles. All flight tests conducted indoors have a three-digit number starting with a "1" per Table 3.3.

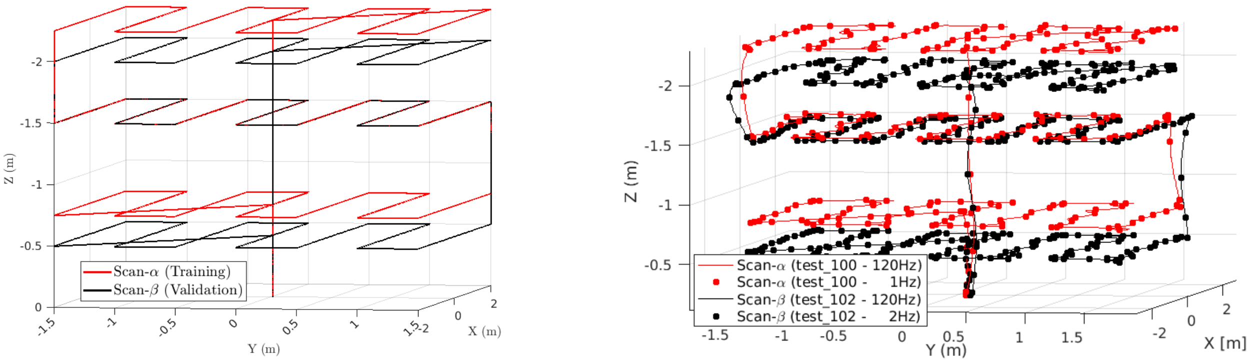


(**a**) xy View (**b**) yz View

Figure 3.5: Five indoor flight profiles.

Two additional trajectories (Scan-α and Scan-β) were flown indoors to train and validate the GPR-based magnetic field map. Both are a lawnmower coverage pattern in the $x - y$ plane at three different altitudes. Scan-α (training set) was flown at 0.75m, 1.5m, and 2.25m altitudes, while Scan-β (validation set) was flown at 0.5m, 1.5m, and 2m altitudes. The goal was for the validation set to have an altitude that forced the GPR to extrapolate (0.5m), interpolate (2m), and match (1.5m) the training set altitudes. At each altitude, the drone flew a lawnmower pattern in the $x - y$ plane by traversing from -2m to +2m in the x direction, then changing the y position from -1.5m to +1.5m in 0.5m increments. Similar coverage patterns have been used in other magnetic field mapping works [88, 89, 90]. Figure 3.6 shows the desired positions and actual flight position Scan-α (test_100) and Scan-β (test_102) trajectories. In Figure 3.6b, the lines show the ground truth position of the vehicle at 120Hz, while the dots are downsamples of the position used to train (downsampled to 1Hz) and validate (downsampled to 2Hz) the GPR-based map.

Two trials of each flight profile were conducted with all sensors being sampled at the aforementioned frequencies (Section 3.4). For the indoor data, we trained the GPR in two different ways. One methodology is to train the GPR using the Scan-α lawnmower pattern, and the alternative is to train the GPR using one trial each of the five indoor flight profiles (FP-In-A through FP-In-E), as summarized in Table 3.3. In both cases, we used both the Scan-β trajectory and separate tests of the FP-In flights to validate the map. Below, we use these two methods of training the GPR to analyze which yields more accurate maps (Sections 3.5.1 and 3.5.4). In general, training tests were downsampled to 1Hz, while validation sets were downsampled to 2Hz. The one exception is when training with FP-In-A, we downsampled to 2Hz to provide better coverage along the edges of the planar rectangle. The FP-In-A locations shown in Figure 3.5 were downsampled to 2Hz, while all others were at 1Hz.



(**a**) Desired positions of Scan-α (training) and Scan-β (validation) trajectories.

(**b**) Actual positions where the magnetic field was sampled during test_100 and test_102. The dots show the 1Hz and 2Hz downsampled locations, respectively.

Figure 3.6: Scan-α and Scan-β: coverage patterns to train and validate the magnetic field map for the indoor flights.

3.4.4 Creating the GPR-Based Magnetic Field Reference Map

To create the GPR-based magnetic field map, we first sampled the magnetic field at the desired locations to form a training dataset. For this, we split the flight tests into three sets: GPR *training*, GPR *validation*, and MEKF *utilization*. Table 3.2 shows which tests were used in each respective set for the outdoor dataset, while Table 3.3 shows analogous data for the indoor dataset. At the IMU 200Hz sampling rate, each three-minute flight generated over 40,000 datasets. However, many of these points are spatially redundant since the vehicle does not move very far in 1/200 of a second. In order to maintain efficient GPR execution, all training datasets were downsampled to 1Hz. The exception is FP-In-A, which was downsampled to 2Hz to provide better coverage on the edges of the planar rectangle. The outdoor GPR map uses 1157 observations from the five outdoor flight profiles. Each point in Figure 3.4 is an observation used to create the outdoor map. The indoor map was trained in two different ways, either using the $\sim$250 observations from one of the two Scan-α tests (Figure 3.6b) or using 1298 observations from the indoor flight profiles (Figure 3.5).

For each map, the observations that made the training set were used to train three separate Gaussian processes. Each observation has a 3D input (x, y, and z position of the observation) and a 1D output (a single component of the 3D magnetometer measurement). Once trained, each Gaussian process can be queried to give the x, y, or z component of the magnetic field at any desired position, including locations that were not directly observed in the training set. In effect, we used the GPR to interpolate and predict the magnetic field at unobserved locations in each flight space. Although three independent Gaussian processes performed regression for each of the x, y, and z components of magnetic field, we refer to the composition of these as a single unit (e.g., "GPR", "map", "GPR-based map", etc.).

Table 3.2: Outdoor flight profiles from Section 3.4.2 were split into a "training" set (to create the GPR map) and a "utilization" set (to analyze attitude estimation). Each flight profile contributes a single training flight and a single utilization flight.

	Training (1.0 Hz)	**Validation**	**Utilization (Max Sensor Frequency)**
FP-A	test_01	–	test_10
FP-B	test_03	–	test_11
FP-C	test_05	–	test_12
FP-D	test_07	–	test_13
FP-E	test_09	–	test_14

Table 3.3: The indoor flight profiles from Section 3.4.3 have dedicated flight trajectories for GPR training and validation. The indoor experimental profiles match the outdoor versions, but with smaller footprints to fit in the spatially constrained indoor motion capture arena.

	Training (1.0 Hz)	**Validation (2.0 Hz)**	**Utilization (Max Sensor Frequency)**
Scan-α	test_100 & test_101	–	–
Scan-β	–	test_102 & test_103	–
FP-In-A	test_105 (2Hz)	test_115	test_115
FP-In-B	test_107	test_116	test_116
FP-In-C	test_110	test_117	test_117
FP-In-D	test_112	test_119	test_119
FP-In-E	test_114	test_120	test_120

The open-source toolbox `gpml-matlab`[1] was used for training and querying the GPR. A squared exponential covariance function was used with a Gaussian likelihood. The hyperparameters (i.e., characteristic length scale, signal and noise standard deviations) for all three Gaussian processes were computed separately by minimizing the log marginal likelihood. The `gpml-matlab` toolbox implements the methods from [67].

3.4.5 Querying the WMM Magnetic Reference Map

The World Magnetic Model (WMM), built by the British Geological Survey group, models Earth's magnetic field [11]. A similar model called the International Geomagnetic Reference Field (IGRF) gives the same reference field values at M-Air within 0.01μT.

WMM uses observations of Earth's magnetic field from Earth-orbiting spacecraft and ground geomagnetic observations to construct a model of the magnetic field and how it will change over the course of five years. The model is most accurate on the year the new version is released, and accuracy degrades over five years until the next release is available. WMM2020 was released in December of 2020. Although our *outdoor* tests were performed in August 2020, we used WMM2020 rather than its predecessor, WMM2015. For this chapter, we used WMM to provide a magnetic reference field value $\boldsymbol{r}_{k,mag}$ that was assumed to be constant at all locations throughout the workspace since WMM gives the same magnetic field values for positions within 10m of the outdoor workspace origin. This is why we refer to WMM as our low-resolution source, while the GPR-based map is our high-resolution source of magnetic field reference vectors $\boldsymbol{r}_{k,mag}$.

WMM was not used as a source of reference vectors for the indoor analysis since its predictions were not accurate for the values measured inside. Instead, to emulate the low resolution of WMM, we queried the GPR at a single location indoors and assumed the magnetic field as constant for all points in the workspace.

3.5 Results and Discussion

In this section, we first assess the accuracy of the GPR-based magnetic field map by analyzing its error against a validation set. Since the outdoor dataset did not have a calibrated magnetometer, this accuracy analysis was only performed on the indoor Robotic Fly Lab data. Next, we present a novel method of visualizing 3D magnetic fields in Section 3.5.2 to highlight the spatial variation differences between the indoor and outdoor magnetic fields. This analysis was used to explain the trends seen in Section 3.5.3, where we demonstrate the value of magnetic field maps on 3DOF attitude estimation in regions with spatially varying magnetic fields. In Section 3.5.4, we compare the accuracy of GPR maps trained on the observations from the Scan-α coverage pattern against the maps trained from the indoor flight profiles. Finally, Section 3.5.5 summarizes how magnetometers can assist with roll and pitch estimates typically reserved for the accelerometer unit to handle.

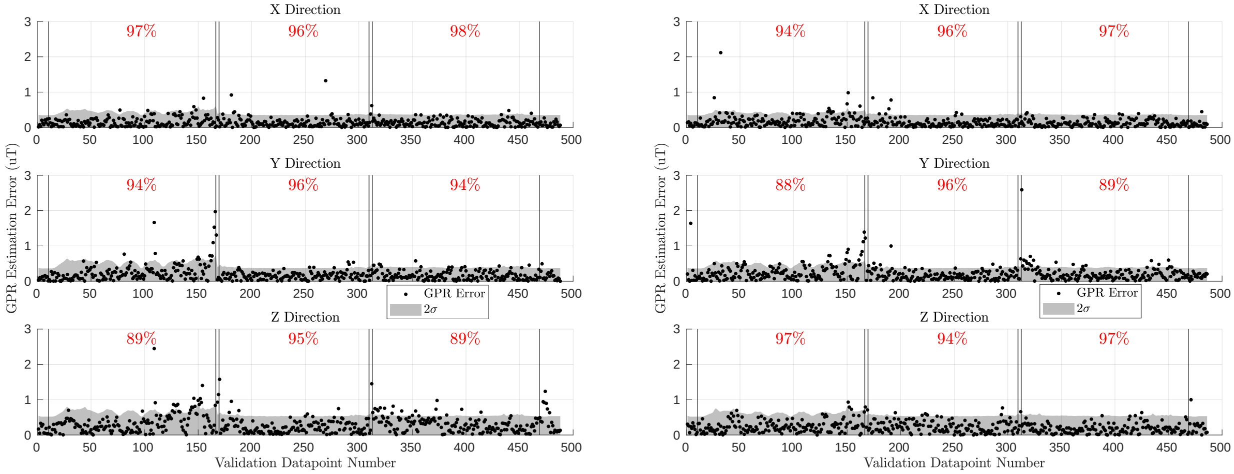
3.5.1 Accuracy of GPR-Based Magnetic Field Map

This section analyzes the accuracy of the GPR when used to map the magnetic field inside the Robotic Fly Lab. The GPR-based map is capable of accurately estimating the magnetic field even when extrapolating outside of and interpolating between observed regions.

To analyze the accuracy of the magnetic field map, we created a GPR-based map with observations from a Scan-α flight test and checked its accuracy against validation data from a Scan-β test (see Section 3.4.3). Per Section 3.4.4, the GPR for the indoor map was created by taking $\sim$250 observations from the Scan-α flight test_100 or test_101. Figure 3.7 shows the $\tilde{B}_x$, $\tilde{B}_y$, and $\tilde{B}_z$ estimation error at the $\sim$490 locations observed in the Scan-β validation flight tests (see Figure 3.6). The vertical lines in each subplot group the flight into three constant-altitude sections separated by a few observations during altitude changes. Observations at 0.5m are indexed from 10 to 166, 1.5m from Indices 169 to 309, and 2.0m from Indices 312 to 468. The segments from Indices 0 to 10 and 468 onward represent the takeoff and landing sequences, respectively. The black dots represent the GPR estimation error, while the gray fill shows two standard deviations (2σ) of the GPR's uncertainty at each location. The red percentage in each constant-altitude section quantifies how often the error is within the gray uncertainty region. Figure 3.7e,f show estimated and measured magnetic field values at each location, which were used to compute the errors shown in Figure 3.7a,d, respectively. Here, each red cross is a measurement from the validation set, while the blue line and shading are the GPR prediction and uncertainty, respectively. Figure 3.7e,f are qualitatively similar to the other two training/validation pairs.

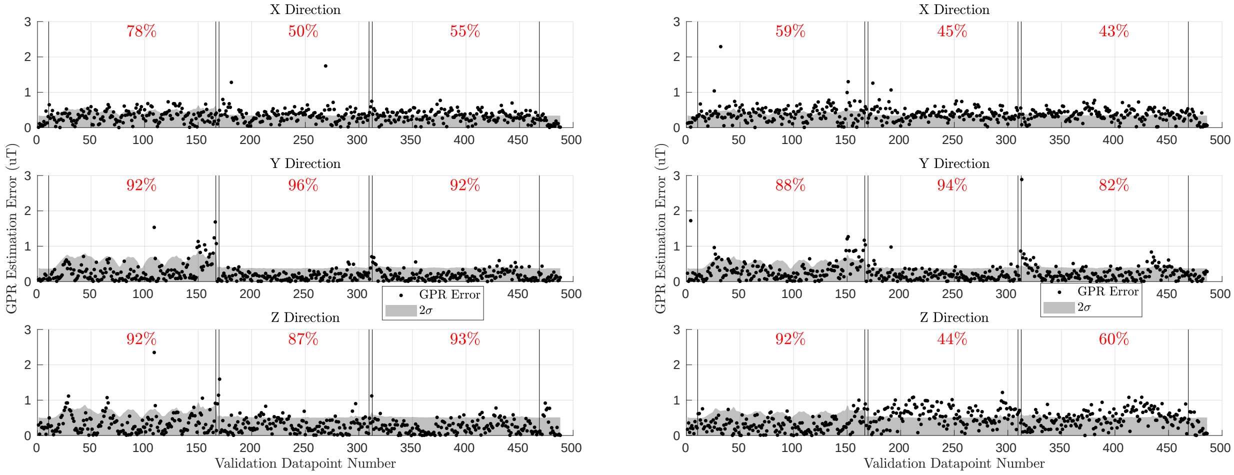
The GPR for each axis generally had less than 1μT of error over all validation measurements. The Scan-α training trajectory sampled points at altitudes 0.75m, 1.5m, and 2.25m, while the Scan-β validation set flew at 0.5m, 1.5m, and 2.0m. As such, the GPR must extrapolate at 0.5m (Indices 10 to 166) and interpolate at 2.0m (Indices 312 to 468) and directly predicts the results at an observed altitude of 1.5m (Indices 169 to 309).

Since each GPR's estimate is a Gaussian distribution, we expected their estimation error to be within two standard deviations 95% of the time. By analyzing the red percentages in Figure 3.7, training on test_100 (top row) yielded more reliable error and uncertainty metrics than test_101 (middle row). As such, test_100 was used to train the indoor GPR-based magnetic field map. Once the accuracy of using GPRs to serve as maps of magnetic field was characterized, the maps were used to analyze the spatial variation of the magnetic fields in our indoor and outdoor test arenas.



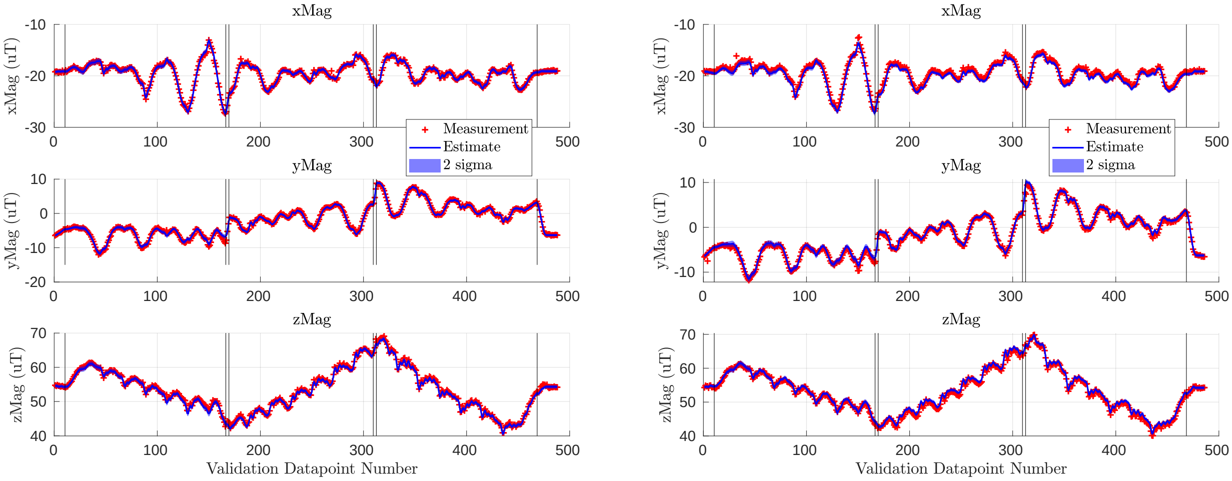
(**a**) Training with test_100. Validating with test_102.

(**b**) Training with test_100. Validating with test_103.



(**c**) Training with test_101. Validating with test_102.

(**d**) Training with test_101. Validating with test_103.



(**e**) Training with test_100. Validating with test_102.

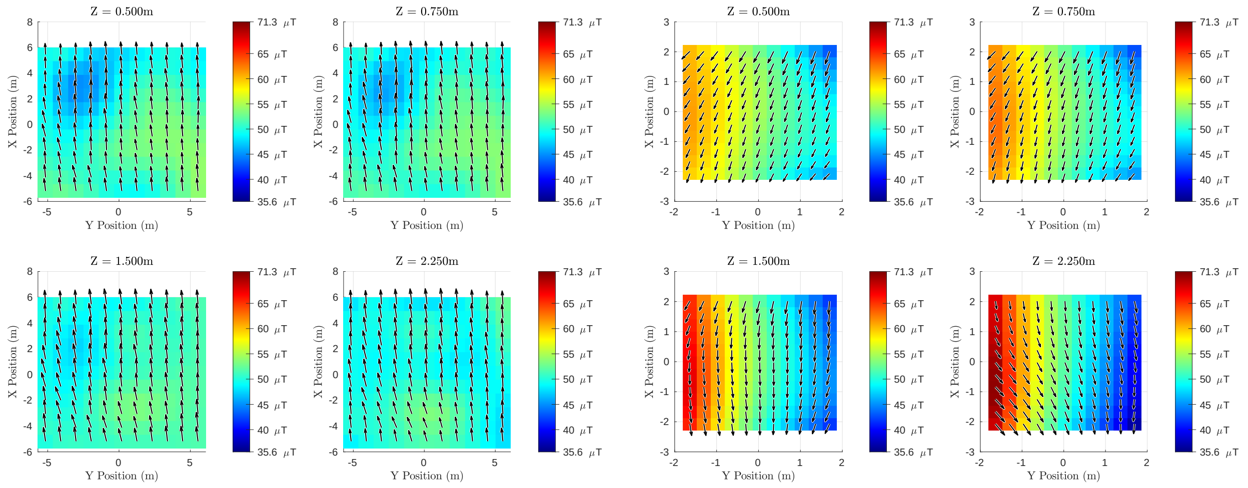
(**f**) Training with test_101. Validating with test_103.

Figure 3.7: GPR-based magnetic field map error on ∼490 validation points from Scan-β. In order, the large segments separated by vertical bars are at altitudes 0.5m, 1.5m, and 2m, respectively. For Figure 3.7a–d, the red percentage value shows how often the error is within the 2σ uncertainty for each altitude segment. Data is presented in the world frame.

3.5.2 Analyzing the Spatial Variation of Magnetic Fields

This section presents a new method to visualize 3D magnetic field maps and compares the magnetic field inside the Robotics Fly Lab to that of the outdoor M-Air facility. We show, qualitatively, that the magnetic field indoors has more spatial variation than outside.

Figure 3.8a shows the GPR prediction of the magnetic field at the outdoor facility for four different altitudes. Black arrows indicate the horizontal field component, while the color behind each arrow depicts $\tilde{B}_z$ at that location. In Figure 3.8a, the $+x$ direction is Geographic North. Figure 3.8b shows the GPR prediction for the indoor lab where Geographic North is approximately in the $-y$ direction. The outdoor magnetic field map (Figure 3.8a) was generated using 1157 observations of training data from the five outdoor flight profiles (Table 3.2), while the indoor map (Figure 3.8b) was generated using the 245 observations from test_100 (a Scan-α trajectory). The indoor map was qualitatively similar when trained on test_101.



(**a**) M-Air magnetic field. North is in the $+x$ direction.

(**b**) Fly Lab magnetic field when trained on test_100 (Scan-α). North is approximately in the -y direction.

Figure 3.8: Outdoor (**left**) and indoor (**right**) magnetic field at four different altitudes. Black arrows show the direction of the planar $(x - y)$ component of the magnetic field. The color behind each arrow shows the strength and direction of the z component of the magnetic field.

At each location, the black arrow gives the local planar heading (or declination) of the magnetic field, which is equivalent to the direction a compass would point to at that location. To most people, this is the most familiar part of the magnetic field, and it is used to estimate heading. Parsing just the black arrows gives users an idea of how the local Magnetic "North" changes throughout a workspace and how reliable heading estimates might be without any magnetic field mapping or disturbance rejection. The color behind each arrow gives the strength and direction of $\tilde{B}_z$. This component of the magnetic field is often ignored if the magnetometer is used only for estimating heading.

However, Figure 3.8b depicts features in $\tilde{B}_z$ that would be valuable for position localization, and in Section 3.5.5, we show how using $\tilde{B}_z$ can assist with roll and pitch estimates.

This method of visualization allowed us to analyze the relative strength of the $\tilde{B}_x$ and $\tilde{B}_y$ and the absolute strength of $\tilde{B}_z$. Unfortunately, these figures show no visual indication of how the horizontal ($\tilde{B}_x$–$\tilde{B}_y$) magnitude of the field compares to the vertical component ($\tilde{B}_z$). For the indoor flight space, such comparisons can be made using Figure 3.7e,f. For the 1157 observations in the outdoor training dataset, the range of measured values for each component of the magnetic field was $\tilde{B}_x \in [15,\ 24]\,\mu$T, $\tilde{B}_y \in [-7,\ 2]\,\mu$T, and $\tilde{B}_z \in [48,\ 54]\,\mu$T. Finally, there is no visual indication of the uncertainty of the GPR's estimate in our figures. However, as in [18, 10], one could use the transparency of the arrows and colors to indicate regions of lower certainty.

Returning to Figure 3.8a, note that the magnetic field at M-Air generally points towards Geographic North and into the ground. These characteristics remain fairly constant as a function of altitude. The total magnetic field strength varied from 51.6μT to 57.6μT for all points in the outdoor training dataset.

Figure 3.8b shows that the indoor magnetic field at 0.5m was generally westward but pointed more northwest as altitude increased. Recall that Geographic North is approximately in the $-y$ direction in this figure. Figure 3.7e,f show that $\tilde{B}_z$ is the largest component of the magnetic field indoors as well. The total magnetic field strength varied from 43.9μT to 73.3μT for all points in the indoor training dataset.

For the indoor magnetic field, $\tilde{B}_z$ had a larger range of values as altitude increased with both the smallest and largest values appearing at an altitude of 2.25m. At all altitudes, $\tilde{B}_z$ had a larger magnitude on the $-y$ side and a smaller magnitude on the $+y$ side. The $-y$ wall is a movable divider between the Robot Fly Lab and an adjacent lab space. This divider has reconfigurable panels that run along a metal track about 3m off the ground. We believe this metallic track and metallic portions of the panels were responsible for the strengthened z field on the $-y$ side of the indoor flight lab.

The magnetic field of the indoor workspace had much more spatial variation than the magnetic field outside. The indoor field also changed noticeably as a function of altitude, a trend that was recently analyzed in [20]. In the next section, we show how high-resolution GPR-based magnetic field maps can be used for attitude estimation in regions with spatially varying magnetic fields.

3.5.3 Comparison of Magnetic Field References

This section shows that the high-resolution GPR-based magnetic field map is a better source of reference magnetic field vectors $\boldsymbol{r}_{k,mag}$ when there is high spatial variation in the magnetic field. If the magnetic field has low spatial variation, the GPR's reference vectors yield similar attitude estimate errors assuming the field is constant everywhere.

Table 3.4 shows the root-mean-squared error (RMSE) in degrees of the MEKF's attitude estimator for all five flight profiles of the outdoor dataset. The "training" flight tests used to create the GPR-based magnetic field map were different from the "utilization" tests used here to evaluate attitude estimation error (Table 3.2).

For each flight test, the three table columns correspond to roll, pitch, and yaw RMSE in degrees. Estimated quaternions were converted to Euler angles using a 3-2-1 (yaw-pitch-roll) rotation sequence. The first column indicates the source of the magnetic field reference vectors. The first row labeled "Gyro/Accl Only" excludes magnetometer data; it uses only the gyro and accelerometer to perform attitude estimation. Yaw RMSE values were omitted here since heading is not observable with only gyro and accelerometer data. The second row is for WMM, which has such a low resolution that the reference vector is the same for all locations in the five flight profiles. The third row queries the GPR-based map at a single location and uses that single reference magnetic field vector regardless of the vehicle's actual location. The position $(0, 0, -1.5)$m was chosen since all flight profiles were centered around the origin and most had a fixed 1.5m desired altitude. The exceptions are FP-D and FP-In-D, the inclined circle trajectories. This row can be interpreted as using the GPR with the same low-resolution WMM offers. Finally, the last row, labeled "GPR", queries the GPR-based map at each time step to provide a specific magnetic field reference vector for each vehicle position to leverage the full interpolation capability of the GPR. We refer to this case as the "high-resolution" map and the other two, constant, reference vector methods as "low-resolution" alternatives. The bold value in each column corresponds to the reference vector source that yielded the lowest roll, pitch, or yaw RMSE.

Table 3.5 shows similar data, but now for the indoor dataset. Again, the "training" flight tests are different from the "utilization" tests used to evaluate the estimation error per Table 3.3. The format of Table 3.5 is the same as that described above. WMM's magnetic reference vector is not representative of the magnetic field inside the Robot Fly Lab, so WMM was not considered for indoor analyses.

Table 3.4: RMSE for each *outdoor* experimental test using different magnetic field reference sources. The first row excludes magnetometer data. For each test, the three columns show the RMSE for roll, pitch, and yaw, respectively, in degrees.

RMSE **(°)** ***Roll—Pitch—Yaw***	**FP-A** **test_10**			**FP-B** **test_11**			**FP-C** **test_12**			**FP-D** **test_13**			**FP-E** **test_14**		
Gyro/Accl Only	**1.6**	2.2	–	2.6	2.0	–	10	4.2	–	3.2	2.2	–	7.8	10	–
WMM	**1.6**	0.7	**2.3**	2.3	**1.5**	**2.2**	3.2	**1.4**	**4.8**	2.5	**1.4**	**2.2**	**1.8**	0.7	**1.9**
GPR (0,0,-1.5)m	1.7	**0.4**	2.6	2.5	1.7	2.4	3.4	1.6	5.1	2.7	1.6	2.5	1.9	**0.4**	2.2
GPR	1.7	0.7	2.7	**1.7**	1.9	4.5	**2.2**	1.8	5.5	**1.3**	2.5	3.7	**1.8**	0.5	4.7

Table 3.5: RMSE for each *indoor* experimental test using different magnetic field reference sources. The first row excludes magnetometer data. For each test, the three columns show the RMSE for roll, pitch, and yaw, respectively, in degrees.

RMSE **(°)** ***Roll—Pitch—Yaw***	**FP-In-A** **test_115**			**FP-In-B** **test_116**			**FP-In-C** **test_117**			**FP-In-D** **test_119**			**FP-In-E** **test_120**		
Gyro/Accl Only	1.1	0.8	–	2.5	1.5	–	10	6.7	–	2.2	1.6	–	1.5	1.3	–
GPR $(0, 0, -1.5)$m	2.6	4.7	9.0	2.5	3.2	11	2.9	3.6	13	4.9	3.8	11	1.1	0.6	5.0
GPR	**0.7**	**0.2**	**1.7**	**1.1**	**0.9**	**3.7**	**1.1**	**1.7**	**5.0**	**0.8**	**0.9**	**2.7**	**1.0**	**0.4**	**1.9**

For each flight test, the same sensor data were played back to the MEKF with the only difference being the sensors available or the source of magnetic field reference vectors. The MEKF parameters were tuned by hand, resulting in reasonable, but not "optimal", parameter values. The MEKF parameters were held constant across all indoor and outdoor flight datasets. Because of this, it is possible to have the overall lower RMSE values in Tables 3.4 and 3.5 if the MEKF is tuned more specifically for the indoor or outdoor environment or specific flight profiles.

For the outdoor dataset in Table 3.4, the high-resolution map ("GPR" row) yields comparable attitude estimates relative to the low-resolution magnetic field reference sources. The high-resolution map tended to give lower (but comparable) roll error than the alternatives. Aside from FP-A and FP-C, the outdoor flight tests had noticeably higher yaw error in the "GPR" row. We believe this increased yaw error was caused from not calibrating the RM3100 magnetometer before performing the outdoor flight tests. Since the horizontal component of the magnetic field is much smaller than $\tilde{B}_z$ (Section 3.5.2), small errors due to lack of calibration would more easily change the measured heading of the magnetic field. This poses more of a problem to our yaw estimates than roll and pitch [41]. We would expect such small calibration errors to have less impact on yaw estimates in a more equatorial part of the world where the horizontal component of the magnetic field is more dominant.

Table 3.5 shows that when testing indoors, with more spatial variation in the field, the high-resolution GPR-based map consistently allowed the MEKF to achieve lower errors in roll, pitch, and yaw estimates. The most significant outlier was observed from test_119 where using the high-resolution map reduced the error in all the Euler angles by at least one-third. We believe this was because test_119 was a trial of FP-In-D (inclined circle) and spent more time in regions of the magnetic field that differed significantly than what was observed at the reference $(0, 0, -1.5)$m state (see Figure 3.8b). In fact, similar trends occurred for test_115 and FP-In-A, the level rectangle that traced the perimeter of the flight space. In such cases, the constant ambient magnetic field assumption was inconsistent with the magnetometer's observations as the vehicle flew.

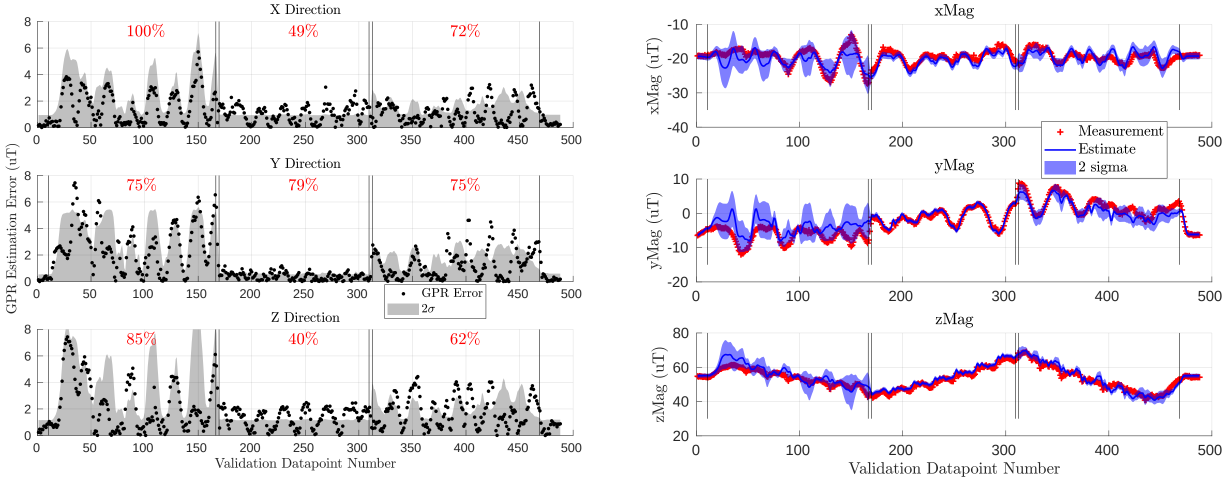
From Table 3.4, we see that when the magnetic field of a workspace had a low spatial variation, having a higher resolution map did not reduce the attitude estimation error in general, but also did not make attitude estimates generally worse. The exception was with yaw estimation: which was consistently worse when we constantly changed the magnetic reference vector as the vehicle moved. However, the relatively larger yaw error was likely due to a lack of magnetometer calibration for the outdoor flights in combination with a small horizontal magnetic field. In contrast, Table 3.5 shows a promising environment where the high resolution of the GPR-based maps can be utilized to improve vehicle attitude estimation. In short, high-resolution GPR-based magnetic field maps can significantly improve attitude estimation if there is high spatial variance, but do not perform noticeably better or worse than the low-resolution alternatives when there is low spatial field variation.

It is important to note that the outdoor magnetic field map was trained using the flight profiles, while the indoor map was trained using Scan-α. In the next section, we show that both methods of training yielded similar accuracy when validating on the flight profiles.

3.5.4 Comparison of GPR-Based Map Training Methods

This section analyzes how the locations of the observations used in the training set affected the accuracy of the GPR-based map by comparing the accuracy of a GPR trained on Scan-α trajectories against a GPR trained on indoor flight profiles. We show that training on Scan-α yielded mean absolute errors of 0.28μT, 0.27μT, 0.48μT, while training on the five indoor flight profiles gave mean absolute errors of 0.30μT, 0.28μT, and 0.48μT for B_x, B_y, and B_z, respectively, when validated on the indoor flight profiles. As such, the outdoor GPR-based map was sufficiently accurate for the analysis conducted in Section 3.5.3, and the conclusions drawn from Tables 3.4 and 3.5 are sound despite differences in the GPR training methodologies. Note that these metrics exclude validating on FP-In-A due to a sensor misalignment during test_105 (FP-In-A training flight).

Figure 3.9 shows the accuracy of the indoor magnetic field map when trained on five indoor flight profiles (FP-In) and validated on a Scan-β flight. The left plot is the GPR estimation error when validated on flight test_102, while the right plot shows the GPR estimates (blue line) against observations from the validation set (red cross). The results were qualitatively similar when validating on test_103. The blue and gray shaded regions depict two standard deviations of GPR uncertainty at each sampled location. The black vertical lines partition the figure into sections of constant altitude (0.5m, 1.5m, and 2m, respectively). We used samples from test_105, test_107, test_110, test_112, and test_114 as shown in Table 3.3 to train the GPR on the FP-In trajectories. Figure 3.7 shows similar metrics for the indoor GPR-based map when trained on the lawnmower coverage pattern Scan-α.



(**a**) Training with FP-In flights. Validating with test_102.

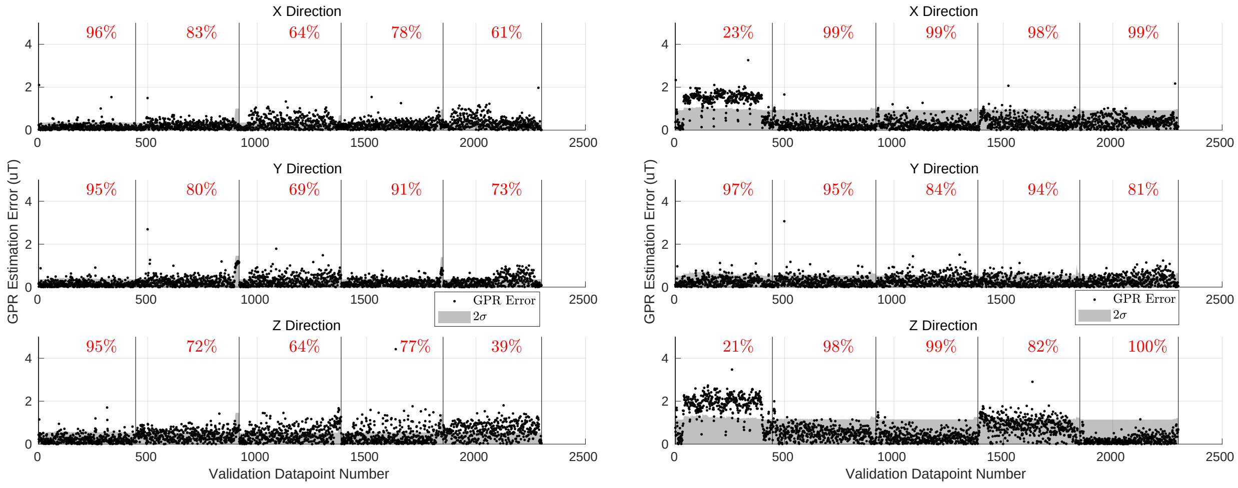
(**b**) Training with FP-In flights. Validating with test_102.

Figure 3.9: Accuracy of the indoor GPR-based magnetic field map when trained on the five indoor flight profiles (FP-In flights) and validated on test_102 (Scan-β). The large regions separated by the black vertical lines show areas of constant desired altitude for the Scan-β trajectory.

Figure 3.9 shows that the error on the 1.5m sweep was much lower than for the other two altitudes. This is because most of the flight profiles in the FP-In training set were trajectories at an altitude of 1.5m. As such, the GPR trained on FP-In had a good understanding of the magnetic field at this altitude, but not at 0.5m and 2.0m, the other Scan-β altitudes. This was expected since Gaussian processes tend to only provide accurate predictions near data found in their training set.

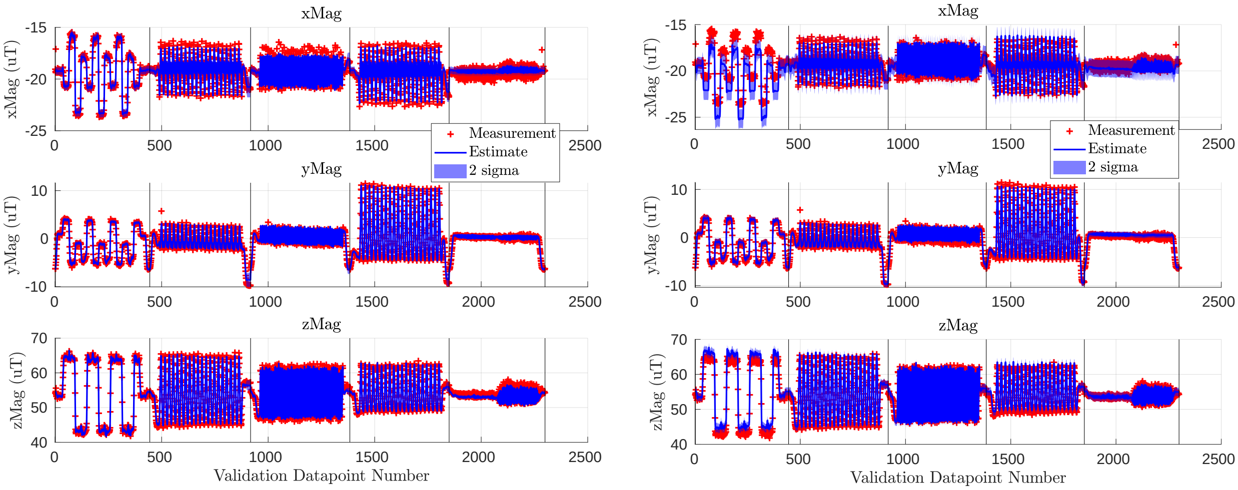
Above, we showed that when validating on Scan-β, training on Scan-α (Figure 3.7) yielded better results than training on FP-In trajectories (Figure 3.9). However, the MEKF did not use data from the Scan-β trajectories, so it was important to analyze the accuracy of the two mapping methodologies on FP-In flights as well.

Figure 3.10 shows the results of validation with the five FP-In validation flights (test_115, test_116, etc.; see Table 3.3). The left figure column shows the results when training on Scan-α, and the right column was trained on the five FP-In training flights. Here, in contrast to other similar plots, the vertical black lines separate each of the five FP-In validation flights (in order of FP-In-A to FP-In-E). We can see that when validating on the intended flight profiles, both training methodologies had a similar error. The exception was the steady-state error of B_x and B_z when validating on the level rectangle (test_115 of FP-In-A) in Figure 3.10b,d.



(**a**) Trained with test_100. Validated with FP-In flights.

(**b**) Trained and validated on FP-In flights.



(**c**) Trained with test_100. Validated with FP-In flights.

(**d**) Trained and validated on FP-In flights.

Figure 3.10: The accuracy of the indoor GPR-based magnetic field map when *validated* on the five indoor flight profiles (FP-In flights). The black vertical lines separate each of the five indoor flight profiles. The steady-state error in Figure 3.10b was likely caused by a misalignment with the magnetometer in test_105 (FP-In-A training flight).

Figure 3.10d shows that the steady-state shift for FP-In-A was consistent with the magnetometer being pitched slightly in the training flight (blue line) relative to the validation flight (red cross). Although there were no significant crashes during the indoor flight tests, there were some rough landings that occasionally dislodged some standoffs and screws. It is possible such a harsh landing misaligned the RM3100 magnetometer relative to the rest of the rigid body. Since we only saved the data from the successful flights, this misalignment must have happened on an unsuccessful flight just before test_105 but was fixed before test_107. Note: this hypothesis also explains the sinusoidal B_x and B_z error trends on the 1.5m altitude section of Figure 3.9a where the peaks in errors occurred when the Scan-β trajectory was on the outer rectangular perimeter of the flight space (i.e., overlapping most with test_105), and the dips occurred when Scan-β passed through the center region with the data from the flight tests that had a properly aligned magnetometer. Excluding FP-In-A, training on Scan-α yielded mean absolute errors of 0.28μT, 0.27μT, 0.48μT while training on the five indoor flight profiles gave mean absolute errors of 0.30μT, 0.28μT, and 0.48μT for B_x, B_y, and B_z, respectively.

Note, that the previous paragraph has an erroneous conclusion that was not discovered until we completed work on Chapter 2. With what we have learned since the work on Chapter 3 (this chapter), we now believe that the steady-state shift in FP-In-A is actually an example of flight-by-flight variations (Section 2.5.2.2.

Nevertheless, Figure 3.10 shows us that the two training methodologies yielded comparable results when validating on the FP-In flight profiles. Due to the inaccuracies of test_105 (FP-In-A) and that querying the GPR was faster when it was trained on fewer data, we used test_100 (Scan-α trajectory) to train the indoor GPR during the analysis of Section 3.5.3. However, we showed that training the indoor GPR on the five indoor flight profiles would likely yield similar results to Table 3.5.

This brings us to an important point. If the intended route through an indoor workspace is known a priori, then it is sufficient to construct a magnetic field map by gathering observations along the path. This kind of mapping methodology is common with self-driving cars where maps are only generated along roads the vehicle intends to traverse. However, training magnetic field maps from coverage patterns such as Scan-α will generally yield more accurate results throughout the workspace.

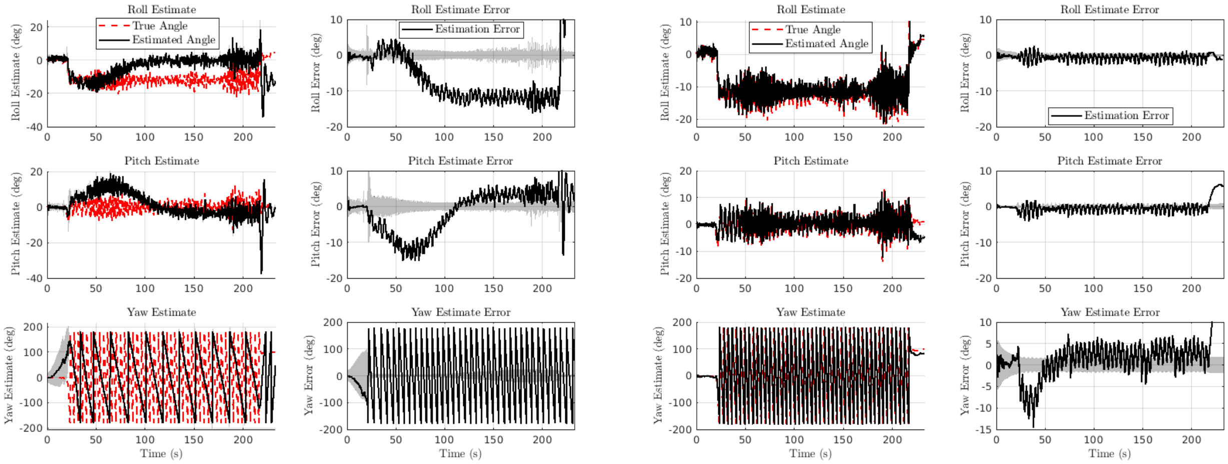
In summary, we used the indoor flight space to compare two mapping methodologies and showed that mapping with the Scan-α coverage pattern yielded a uniformly more accurate map when validating on Scan-β but had comparable accuracy when validating on the indoor flight trajectories. Since the outdoor GPR was trained using the five outdoor flight profiles, we can conclude that our outdoor map is sufficiently accurate near the locations of the five flight profiles.

3.5.5 Magnetometer Correction of Accelerometer Roll/Pitch Errors

This section shows that magnetometers can resolve accelerometer roll/pitch errors observed during maneuvers. For this analysis, we focused on test_117, a FP-In-C trajectory from the indoor dataset. Recall that all FP-In-C flights followed a level circle with a radius of 1m, a commanded tangential velocity of 1.5m/s, and a commanded heading that changed to keep the $+x$ body axis tangent with the circle. As such, in an ideal flight, the vehicle would experience 2.25m/s^2 of centripetal acceleration along its y axis.

Figure 3.11a shows the roll, pitch, and yaw estimates along with their respective errors when using only the gyroscope and accelerometer. A 20s into the flight, the vehicle accelerated from rest to achieve the 1.5m/s desired tangential velocity. Acceleration was measured primarily in the body x direction, which causes a transient MEKF pitch error around $t = 65$s. As the vehicle reached its maximum speed, the accelerometer measured centripetal acceleration in the body y-direction, causing a steady-state roll error over most of the flight.

This is a common problem for flight vehicles [91, 92, 93] and other platforms [44, 45, 46, 47] that use the accelerometers for attitude estimation. Typically, the accelerometer's measurement is discarded or de-emphasized when nongravitational accelerations or sensor anomalies are detected. Here, we circumvented the need to detect nongravitational accelerations by simply allowing the magnetometer to assist with roll and pitch estimation.



(**a**) Gyro and accelerometer only.

(**b**) Gyro, accelerometer, and magnetometer with full GPR.

Figure 3.11: Time series plots for roll, pitch, and yaw estimates of test_117 (fast level circle). Linear and centripetal accelerations caused pitch and roll errors (respectively) in Figure 3.11a that were corrected with the magnetometer in Figure 3.11b.

Figure 3.11b is the same flight, but now with the magnetometer. Here, we see that both the transient pitch error and the steady-state roll error were eliminated, demonstrating that the information from the magnetometer and the GPR-based map was able to improve roll and pitch estimates. This is important because, in many applications, the magnetic field is only used to provide yaw (heading) measurements, while roll and pitch are left to the accelerometer. However, with sufficient information of the magnetic field, magnetometers can also assist in correcting roll and pitch estimates.

It is important to qualify when magnetometer adjustment of roll and pitch estimates will be beneficial. Table 3.5 shows that for FP-In-C, roll and pitch estimates were improved even with a constant (low-res) magnetic field reference at all points. This improvement increased when using the high-resolution, full GPR. FP-In-E showed a similar trend since the magnetometer was able to offset errors in the attitude estimates caused by the acceleration of maneuvering. However, with FP-In-A, where the vehicle was hovering for much of the flight, the magnetometer with a low-res GPR map actually made roll and pitch estimates worse. This is because the indoor flight space had spatial variation in the magnetic field that was not captured by assuming a constant reference source. However, with the full GPR, we again see the magnetometer able to improve roll and pitch estimates past initial accelerometer estimates. Finally, such roll and pitch corrections may not be viable in equatorial parts of the planet where the horizontal component of the magnetic field is dominant.

3.6 Conclusions and Future Work

This chapter presented methodologies to train, validate, and visualize 3D GPR-based magnetic field maps and examined their value for indoor attitude estimation. We presented a new visualization technique to better understand the spatial variation of 3D magnetic fields and discussed its advantages and drawbacks. Our visualization technique relies on some interpolative tool, such as GPR, to estimate the magnetic field at locations not found in the training set. Next, we showed that in spaces with high spatial variation, high-resolution GPR maps can improve attitude estimates, but these maps did not yield significantly better or worse attitude estimates in spaces with low spatial variation. Finally, we demonstrated two ways of training GPRs and showed that training a GPR-based map using coverage flight trajectories such as Scan-α was generally better, though constructing a map using only observations along flight routes of interest was just as accurate in the vicinity of the flight route. We also provided an illustrative example of how magnetometers can assist with roll and pitch estimates in a workspace with high magnetic spatial variation.

In future work, it will be critical to better understand how much spatial variation an environment needs before it becomes advantageous or necessary to use magnetic field maps of the workspace. We plan to apply quantitative methods of describing the spatial variation of a workspace and investigate how workspaces with spatially varying magnetic fields can be used to improve position estimates. The goal is to perform full 6DOF position and attitude estimation using GPR-based magnetic field maps [18]. This would require an analysis of multiple indoor spaces as in [41, 20], which is currently difficult due to our dependence on motion capture cameras to construct our maps. Previous works have found success with using LiDAR or visual–inertial odometry to construct magnetic field maps without motion capture [29, 19, 15, 18, 20]. These methods are recommended to facilitate accurate magnetic field mapping in complex regions where nearby structures offer distinct mapping information for LiDAR and vision sensors. Furthermore, we aim to investigate how to better address the time-varying nature of magnetic fields [10] to more accurately represent the magnetic field near large moving ferromagnetic structures such as elevators and doors [22, 10, 48].

CHAPTER 4

Using Magnetic Field Gradients to Understand and Predict Position Localization Accuracy

4.1 Introduction

This chapter uses a GPR (Gaussian Process Regression)-based magnetic field map to estimate the location of a UAV (unmanned aerial vehicle) in an indoor workspace. Building upon Chapter 2, we use a *compromise* magnetic field map as our fingerprinting map. In general, one would expect a region with low spatial gradient of a signal to have lower position localization accuracy. Intuitively, if the signal does not change as the vehicle moves around, it becomes difficult to tell if the vehicle has moved or not. As such, this chapter analyzes the relationship between gradients in the ambient magnetic field of the workspace against the position localization accuracy of a particle filter.

To perform this analysis, we create a magnetic field map and find the locations within the map with the highest and lowest magnetic field gradient. These locations are denoted as $\boldsymbol{r}^*_{max}$ and $\boldsymbol{r}^*_{min}$ respectively. Given a center point $\boldsymbol{r}^*$, we define four trajectories (hover, X Stride, Y Stride, and Z Stride) that fly through the center point. These stationary and single-axis stride trajectories serve as our "utilization" flights and allow us to contrast the accuracy of our particle filter near a gradient-rich location $\boldsymbol{r}^*_{max}$ versus a gradient-poor one $\boldsymbol{r}^*_{min}$.

The contributions of this chapter are as follows. First, we analyze the relationship between magnetic field gradients and the ability for a particle filter to localize position. We find that the magnetic field gradient in the x direction of our workspace is much lower than y or z gradients which results in higher x position localization errors than the other axes. Overall, the particle filter can estimate position within 0.2m across six of our eight utilization flight profiles. Next, we show how information about the ambient magnetic field can help determine if a particle filter's estimation error is caused by low spatial gradient in the magnetic field, error in the GPR map, or magnetic ambiguities in nearby areas. We demonstrate this by analyzing the time-series error of our particle filter through some of our utilization flights. To our knowledge, this is the first time a work has clearly demonstrated how ambiguities in the magnetic field and errors in mapping the ambient field can directly affect position estimates. Finally, we propose a new metric based on the magnetic field gradient to predict the accuracy of position estimates within the workspace. Our proposed metric (which we call a gradient localization metric) accounts for both the gradient of the magnetic field and any errors introduced by the magnetic field map from lessons learned in Chapter 2.

The remainder of this chapter is structured as follows. A summary of related works is presented next in Section 4.2 followed by an introduction to the mathematical models for our GPR map and particle filter in Section 4.3. Our workspace, UAV, and experimental procedure is then explained in Section 4.4. Finally, Section 4.5 presents the results of our gradient-based analysis followed by a brief discussion (section 4.6) and conclusion with ideas for future work (Section 4.7).

4.2 Related Work

This chapter seeks to compare the magnetic field gradient of an indoor workspace against the accuracy of a position estimator for an unmanned aerial vehicle. We do this using a two-step process that starts with an offline magnetic field mapping phase using Gaussian process regression (GPR). This is followed by a position localization phase with a particle filter. Other similar works refer to the offline mapping phase as “fingerprinting” and investigate how their fingerprinting map assists with indoor navigation.

In this chapter, like with other works, we create maps of the ambient magnetic field. However, fingerprinting has been done with other signals like radio [94], Wi-Fi [95, 96], and visible light [38]. There are also works that perform simultaneous localization and mapping (SLAM) using similar reference signals [25, 13, 97, 18, 98]. A more detailed breakdown of the related works in this field is presented in survey papers that summarize types of mapping and localization techniques for indoor navigation [99, 35, 100]. Most similar to our technique are those that use the magnetic field as the reference signal for their localization [13, 15, 29, 20, 18, 19, 101].

The position estimation accuracy of works that localize with the magnetic field range by a couple orders of magnitude. A direct comparison across all related works is difficult due to differences in the size of their workspaces and the varying sensors and algorithms used to compute ground-truth pose estimates.

Recent results from Almeida et al. in [30] give an average of 0.07m of 2D position error on a ground robot across three different test cases in an 8m×10m room. Their work has a 2D LIDAR sensor for ground-truth, uses a particle filter to estimate position, and presents estimation accuracy after their PF has converged (i.e., after the PF first achieves error within 0.1m). Similarly in 2013 Frassl, Robertson, et al. obtain 2D position estimates of 0.064m for their wheeled robot and 0.08m for a pedestrian walking through a 6m×10m room in [24]. They use the same room in a SLAM-based approach and obtain 0.01m to 0.2m of average error for foot-mounted pedestrian localization on the last 80% of their four test trials [25]. For both works, the authors use Vicon motion capture for ground-truth, a particle filter for position estimates, and initialize their PF with ground-truth position and attitude like in our work.

We suggest Refs. [30, 24, 25] provide the most direct and fair comparisons to our work here given the size of the respective workspaces and their method of computing position estimation error throughout the trajectory. This is despite differences in the respective vehicles (i.e., wheeled robots and pedestrians vs UAV). Finally, we mention [50] by Lee et al. which achieves 0.1m of minimum mean squared error (MMSE) across four trials using a SLAM-based magnetic field mapping method on a wheeled robot. Though [50] also tests in rooms of similar size (e.g., 5m×4m), they compute their position estimation error at $\sim$10 known ground-truth locations instead of throughout their trajectories.

In this chapter, we perform magnetic field mapping and position localization on an unmanned aerial vehicle (UAV), use a particle filter to estimate 3D position, and test within a 4m×3m×2.25m volume equipped with motion capture cameras for ground truth. Our sensors include an accelerometer, altimeter, and magnetometer. Here, we use the altimeter to demonstrate how another sensing modality can improve position estimates when the given magnetic field map is erroneous. Our method is able to estimate the UAV's position within 0.1m of root mean squared error (RMSE) for three of our eight flight test cases and within 0.2m for six of the eight flights. In Chapter 4, to estimate the position of a UAV in our indoor workspace. These position estimates errors in our magnetic map to investigate the cause of our position estimation errors.

The core focus of this chapter is to relate the gradients of the magnetic field to localization accuracy. In a workspace with a constant magnetic field in all locations, it would be impossible to localize position using the magnetic field alone. As such, we see the gradient of the magnetic field as a source of information and aim to analyze the relationship between the signal's spatial gradient and our ability to localize. To our knowledge, there are no other works that quantitatively compare the spatial gradient of an indoor magnetic field to the accuracy of their position localization. A closely related set of works, however, utilize gradients of a signal as their fingerprints and obtain less spatial ambiguities in their maps [51], better motion tracking through a workspace [52], or better agreement in fingerprints across different devices [53].

Finally, this chapter provides a thorough investigation to the source of position estimation errors through our flight trajectories. Of particular interest, we discuss how we learned to distinguish between PF estimation errors due to low spatial gradients (i.e., insufficient magnetic field information to distinguish nearby locations), errors in the magnetic field map, or ambiguities in the ambient magnetic field that allow for the same signal to be measured at multiple locations.

In Figure 7 of [72], LeGrand and Thrun show an example of their PF using the magnetic field to estimate the position of pedestrian motion within a mapped indoor workspace. Their PF is consistently inaccurate at a particular portion of the circular trajectory, but their analysis does not investigate if this is caused by ambient magnetic field (insufficient information to properly localize), errors in their map, or spatial ambiguities in the field. Our focus on evaluating the accuracy of our maps (Chapter 2) and the study of the gradient of the field (in this chapter) allows us to better understand the underlying cause of our PF's localization error across our flight trajectories.

4.3 Mathematical Preliminaries

This section briefly introduces notation on querying Gaussian-processes-regression-based magnetic field maps, how we use the GPR-based map to compute the spatial gradient of the local magnetic field, and finally how we construct our particle filter to estimate the UAV's position.

4.3.1 Gaussian Process Regression

Gaussian Process Regression (GPR) is a machine learning tool that can be used to estimate a signal given a set of noisy measurements. The representation of GPRs used in this dissertation was learned from [67]. Our magnetic field maps use GPRs to create a $3 \to 3$ mapping from each 3D location in the workspace to a 3D magnetic field vector estimate.

A GPR is defined over some input dimension p and an output dimension m which we will denote as a $p \to m$ map. It is common for a GPR to output a single dimension $m = 1$. Thus, to construct our $3 \to 3$ map that estimates the full magnetic field vector at each location in our workspace, we use three separate GPRs ($\mathcal{GP}_x$, $\mathcal{GP}_y$, $\mathcal{GP}_z$).

In this dissertation, each GPR takes a 3D location $\boldsymbol{r} \in \mathbb{R}^3$ as input ($p = 3$) and returns an estimate of the x, y, or z component of the magnetic field at $\boldsymbol{r}$ [4]. Though each GPR outputs an expected value and variance for its estimate at each location, we most frequently need to refer to the expected value. Thus, the expected value of $\mathcal{GP}_x$ will be denoted as $\hat{m}_x(\boldsymbol{r}) = GPR_x(\boldsymbol{r})$ and similarly for the expected value of $\mathcal{GP}_y$ and $\mathcal{GP}_z$ respectively. With this, the estimate magnetic field vector at some location $\boldsymbol{r}$ is

$$\hat{\boldsymbol{m}}(\boldsymbol{r}) = GPR(\boldsymbol{r}) = \begin{bmatrix} \hat{m}_x(\boldsymbol{r}) \\ \hat{m}_y(\boldsymbol{r}) \\ \hat{m}_z(\boldsymbol{r}) \end{bmatrix} = \begin{bmatrix} GPR_x(\boldsymbol{r}) \\ GPR_y(\boldsymbol{r}) \\ GPR_z(\boldsymbol{r}) \end{bmatrix}. \tag{4.1}$$

This chapter uses the "compromise mapping" technique which computes kernel hyperparameters over observations from several "training" flight tests but ultimately performs inference by comparing the target location $\boldsymbol{r}$ to a set of n_1 user-selected locations called the "inference set". For us, the number of inference points n_1 is important due to a constraint in a Python toolbox (`GPyTorch`

---

[4] Note the variable $\hat{\boldsymbol{r}}$ in Chapter 3 was a reference vector. In this chapter, it represents a 3D position vector.

[102]) we used. A more formal and detailed introduction to our compromise mapping technique is given in Section 2.5.4.

4.3.2 Point-wise Gradient Metrics

At each location $\boldsymbol{r} \in \mathbb{R}^3$ in the workspace, we can compute a point-wise gradient metric along the three spatial directions of the lab's ground frame. Let the GPR-based map return a predicted magnetic field vector $\hat{\boldsymbol{m}} = [\hat{m}_x, \hat{m}_y, \hat{m}_z]^\top$ at each location $\boldsymbol{r}$ as defined in Equation 4.1.

Using a finite difference to approximate the spatial change of the magnetic field as

$$\frac{\partial \hat{m}_y(\boldsymbol{r})}{\partial x} \approx \frac{\hat{m}_y(\boldsymbol{r} + \boldsymbol{\delta x}) - \hat{m}_y(\boldsymbol{r})}{|\boldsymbol{\delta x}|} \tag{4.2}$$

with vectors

$$\boldsymbol{\delta x} = \begin{bmatrix} \delta s \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{\delta y} = \begin{bmatrix} 0 \\ \delta s \\ 0 \end{bmatrix}, \quad \boldsymbol{\delta z} = \begin{bmatrix} 0 \\ 0 \\ \delta s \end{bmatrix} \tag{4.3}$$

as a function of some distance parameter δs, we can compute the following spatial gradient matrix at each point

$$\begin{aligned} \frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial \boldsymbol{r}} &= \begin{bmatrix} \frac{\partial \hat{m}_x(\boldsymbol{r})}{\partial x} & \frac{\partial \hat{m}_x(\boldsymbol{r})}{\partial y} & \frac{\partial \hat{m}_x(\boldsymbol{r})}{\partial z} \\ \frac{\partial \hat{m}_y(\boldsymbol{r})}{\partial x} & \frac{\partial \hat{m}_y(\boldsymbol{r})}{\partial y} & \frac{\partial \hat{m}_y(\boldsymbol{r})}{\partial z} \\ \frac{\partial \hat{m}_z(\boldsymbol{r})}{\partial x} & \frac{\partial \hat{m}_z(\boldsymbol{r})}{\partial y} & \frac{\partial \hat{m}_z(\boldsymbol{r})}{\partial z} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x} & \frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial y} & \frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial z} \end{bmatrix}. \end{aligned} \tag{4.4}$$

This chapter is particularly interested in the *spatial* gradient of the magnetic field. For this, we isolate the columns of Equation 4.4 to quantify how the magnetic field vector changes with motion in the x $\left(\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right)$, y $\left(\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial y}\right)$, and z $\left(\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial z}\right)$ directions respectively. Later in this chapter, we show how the norm of these spatial gradient terms (e.g., $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right|$) compares to position estimation error.

In addition, it is convenient to define the norm of the magnetic field gradient

$$|\nabla \hat{\boldsymbol{m}}(\boldsymbol{r})| = \sqrt{\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right|^2 + \left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial y}\right|^2 + \left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial z}\right|^2} \in \mathbb{R} \tag{4.5}$$

where we slightly abuse the notation of ∇ since we do not define Equation 4.4 as a gradient *vector*, but instead as a nine-component matrix.

4.3.3 Particle Filter

This chapter uses a particle filter (PF) to estimate vehicle position $\hat{\boldsymbol{r}}$ and velocity $\hat{\boldsymbol{v}}$. To the author's knowledge, even with a given magnetic field map, there are no models that relate an error in magnetic field readings to an update in position estimate. Such a correction term would be necessary to implement any variation of the update step in the Kalman Filter. Instead, we rely on evaluating the credibility of many hypotheses generated from a particle filter that estimates the position and velocity of the vehicle.

In this chapter, we use ground truth attitude estimates from motion capture to isolate the relationship between magnetic field gradients and position localization. Indoor attitude estimation using IMUs in the face of spatially-varying magnetic fields has been demonstrated in other works [23, 48, 49].

Particle filters keep track of a set of N particles each with a state estimate $\hat{\boldsymbol{X}}^i$ and a weight w^i where $i \in [1, N]$. The weight of particle w^i is a measure of how closely said particle reflects the true state of the actual flight vehicle. Particle weights are all non-negative and sum to one $\sum_i w^i = 1$. Our application requires a six-dimensional state vector $\hat{\boldsymbol{X}}_k = \begin{bmatrix} \hat{\boldsymbol{r}}_k^\top & \hat{\boldsymbol{v}}_k^\top \end{bmatrix}^\top$ where $\boldsymbol{r}_k$ is 3D position of the vehicle, $\boldsymbol{v}_k$ is the 3D velocity, both at timestamp k, and the caret symbol $\hat{}$ denotes that this is an estimate of the true state $\boldsymbol{X}_k$.

There are three main steps to state estimation using a particle filter: the process model, measurement model, and resampling. These steps are summarized visually in Figure 4.1.

4.3.3.1 Process Model

For each particle i, the process model

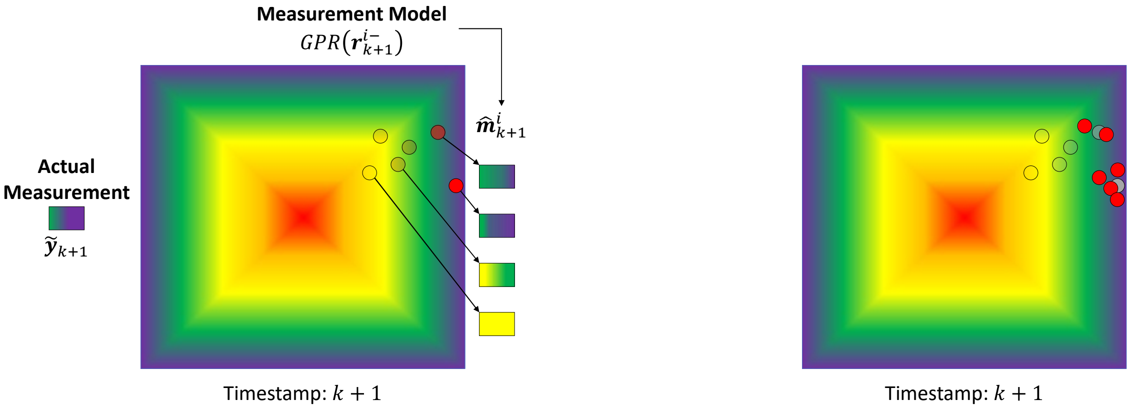
$$\hat{\boldsymbol{X}}_{k+1}^{i-} = f\left(\hat{\boldsymbol{X}}_k^i, \tilde{\boldsymbol{u}}_k\right) + \boldsymbol{\eta}^i \tag{4.6}$$

takes the current state estimate $\hat{\boldsymbol{X}}_k^i$, along with some sensor reading $\tilde{\boldsymbol{u}}_k$, to integrate each particle forward in time. During the process model step, each particle is perturbed by a sample from a zero-mean Gaussian distribution $\boldsymbol{\eta}^i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q})$. Each particle's weight remains constant through the process model $w_{k+1}^{i-} = w_k^i$ where the superscript $^-$ denotes a quantity just after the process model step. This step is performed each time a new sensor measurement $\tilde{\boldsymbol{u}}_k$ is available and $\tilde{\boldsymbol{u}}_k$ is assumed to be resolved in the world frame.

We use an *inertial* process model where the input is taken as the accelerometer measurement ($\tilde{\boldsymbol{u}}_k = \tilde{\boldsymbol{a}}_k$) resolved in the world frame (via motion capture attitude measurements). With this, the

(**a**) Initial state with $N = 6$ particles (red circles) in a 2D workspace (colorful background).

(**b**) Process model moves all particles. Each then gets its own unique noise $\boldsymbol{\eta}^i$.



(**c**) Measurement model is used to evaluate the credibility (compute the weight w^i) of each particle.

(**d**) Resampling creates N new particles near more credible hypotheses.

Figure 4.1: Visualization of particle filter with $N = 6$ particles (red circles) in a 2D workspace with a spatially varying magnetic fields (colorful background).

inertial process model is

$$\hat{\boldsymbol{X}}_{k+1}^i = \begin{bmatrix} \hat{\boldsymbol{r}}_{k+1}^i \\ \hat{\boldsymbol{v}}_{k+1}^i \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{r}}_k^i + \hat{\boldsymbol{v}}_{k+1}^i \Delta t \\ \hat{\boldsymbol{v}}_k^i + \tilde{\boldsymbol{a}}_k \Delta t \end{bmatrix} + \boldsymbol{\eta}_{inertial}^i \tag{4.7}$$

where we note that the propagation for $\hat{\boldsymbol{r}}_{k+1}$ depends on $\hat{\boldsymbol{v}}_{k+1}$ and not $\hat{\boldsymbol{v}}_k$. Additionally, $\Delta t = 0.005$s based on the update rate of our software, $\boldsymbol{\eta}^i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_{inertial})$. The process model is run each time a new accelerometer measurement becomes available.

We note that it is uncommon to only use a low-cost accelerometer for the process model. Additionally, our multirotor UAV is subject to high vibrations from the propellers which make it difficult to extract just the portion of the measured acceleration responsible for the change in linear velocity. Since this chapter focuses on the relationship between magnetic field gradients and position localization, we are alright with having poor dead reckoning as it forces the PF to estimate position using the magnetic field. We acknowledge our position estimates would likely be better with a higher-fidelity process model, some vibration damping on our vehicle, or even a better accelerometer, but we leave such analyses for future work.

4.3.3.2 Measurement Model

The measurement model is used to update the weight w^i of each particle to evaluate how credibly it reflects the true state $\boldsymbol{X}_k$. We use one measurement model for the magnetic field and another model for each particle's altitude.

The magnetometer measurement model $w_k^{i+} = h_{mag}\left(\hat{\boldsymbol{r}}_k^{i-}, \tilde{\boldsymbol{y}}_k, w_k^{i-}\right)$ updates the weight of each particle by comparing the actual magnetometer measurement $\tilde{\boldsymbol{y}}_k$ to that of the predicted magnetic field $\hat{\boldsymbol{m}}_k^i = GPR(\hat{\boldsymbol{r}}_k^{i-})$ for a particle i at position $\hat{\boldsymbol{r}}_k^{i-}$.

$$\begin{aligned} w_k^{i+} &= h_{mag}\left(\hat{\boldsymbol{r}}_k^{i}, \tilde{\boldsymbol{y}}_k, w_k^{i}\right) \\ &= w_k^{i-}\, p_{mag}(\tilde{\boldsymbol{y}}_k^i|\hat{\boldsymbol{r}}_k^i) \\ p_{mag}(\tilde{\boldsymbol{m}}_k^i|\hat{\boldsymbol{r}}_k^i) &\propto \exp\left\{-\frac{1}{2}\left(\Delta\boldsymbol{m}_k^i\right)^\top \boldsymbol{R}_{mag}^{-1}\left(\Delta\boldsymbol{m}_k^i\right)\right\} \end{aligned} \tag{4.8}$$

where $\Delta\boldsymbol{m}_k^i = (\hat{\boldsymbol{m}}_k^i - \tilde{\boldsymbol{y}}_k)$ is the primary indicator of each particle's credibility with this measurement model and $\boldsymbol{R}_{mag}$.

The altimeter measurement model

$$\begin{aligned} w_k^{i+} &= h_{alt}\left(\hat{\boldsymbol{r}}_k^{i}, \tilde{z}_k, w_k^{i}\right) \\ &= w_k^{i-}\, p_{alt}(\tilde{z}_k|\hat{\boldsymbol{r}}_k^i) \\ p_{alt}(\tilde{z}_k|\hat{\boldsymbol{r}}_k^i) &\propto \exp\left\{-\frac{1}{2}\frac{(-{}_z\hat{\boldsymbol{r}}_k^i - \tilde{z}_k)^2}{\sigma_{alt}^2}\right\} \\ {}_z\hat{\boldsymbol{r}}_k^i &= \begin{bmatrix}0 & 0 & 1\end{bmatrix}\hat{\boldsymbol{r}}_k^i \end{aligned} \tag{4.9}$$

uses the altimeter's measurement $\tilde{z}_k$ and the scalar, z-component of the particle's position estimate ${}_z\hat{\boldsymbol{r}}_k^i$ (negated in Equation 4.9 because of the North-East-Down frame). Finally, σ_{alt}^2.

Both h_{mag} and h_{alt} are used to update weights whenever a new measurement from the respective sensor is available. Depending on their sampling frequencies, there will be iterations where neither, one, or both measurement models are used at a particular timestamp k.

Lastly, since both measurement models depend only on position, the true velocity is not *directly* observable. However, since velocity influences the position of the vehicle through the process model $f()$, it is possible the particle filter can make sense of the vehicle's motion with the magnetic field and altimeter alone.

4.3.3.3 Resampling

Resampling creates N new particles based on the states and weights from the current set of particles. The main idea is stop propagating particles that are not accurately reflecting the true state of the flight vehicle (i.e., have low weight w^i) and instead consider new hypotheses that are close to the most credible particles. We use "low-variance resampling" [103], sometimes called "systematic resampling" [104], which selects N new particles by drawing a single random number. New samples are drawn the effective number of particles

$$N_{\text{eff}} = \frac{1}{\sum_{i=1}^{N}(w_k^i)^2} \tag{4.10}$$

drops below the threshold $N_{\text{eff}} < 0.75N$ at which point all weights are reset to $w_k^i = 1/N$. The newly resampled particles have their states perturbed by zero-mean Gaussian noise with covariance $\boldsymbol{P}_{resample}$.

Given the spatial boundaries of the workspace, we can enforce some constraints on the position of each particle. When creating our magnetic field maps, we design flight trajectories that use the "training" boundaries that span ±2m in x, ±1.5m in y, and [-2.25, -0.5]m in z. Similarly, the "utilization" flight trajectories use boundaries ±1.75m, ±1.25m, and [-2, -0.75]m which are set 0.25m *inwards* from the training boundaries. Finally, we define the "constraint" boundaries as ±2.25m, ±1.75m, and [-2.5, 0]m set 0.25m *outward* from the training boundaries; except the upper-bound z term which is the ground floor.

Sometimes, primarily due to the unobservability of velocity, the particle filter's velocity estimate will cause all the particles to drift off outside the physical bounds of the workspace. To remedy this, we enforce the constraint boundaries in two ways. If a particle violates any of the constraint boundaries, its weight is set to zero to prevent it from contributing to our state estimate or being a candidate for resampling. This, inevitably, reduces N_{eff} and can accelerate the need for a resampling step.

In the edge case where all particles violate some constraint boundary in the same timestamp, we can no longer maintain the invariant weight property $\sum_{i=1}^{N} w_i = 1$. Here, we must create new particles inside the boundaries without losing too much of the state history each particle encodes. To achieve this, we modify the position of all particles, but only along the axes in which they violate a constraint. In this case, the particle's respective position component is moved by $\delta c = 0.25$m inward from the constraint boundary. Additionally, since constraint violations are often due to unbounded velocity estimation errors, we set the particle's respective velocity term to zero. Finally, all weights are reset to $w^i = 1/N$ since we have effectively created a new set of particles. Tables 4.1 and 4.2 illustrate an example of this edge case for $N = 3$.

Table 4.1: Example of $N = 3$ particles all having violated a spatial constraint. Bolded values violate a constraint.

${}_x\boldsymbol{r}^i$	${}_y\boldsymbol{r}^i$	${}_z\boldsymbol{r}^i$	${}_x\boldsymbol{v}^i$	${}_y\boldsymbol{v}^i$	${}_z\boldsymbol{v}^i$	w^i
3.2	1.1	-2.1	0.75	0.12	0.13	0.15
1.7	**1.9**	-2.2	0.1	0.3	0.12	0.8
1.8	**2.1**	**-2.7**	0.2	0.5	-0.2	0.05

Table 4.2: $N = 3$ particles from Table 4.1 moved within constraint bounds. Bolded values are modified states.

${}_x\boldsymbol{r}^i$	${}_y\boldsymbol{r}^i$	${}_z\boldsymbol{r}^i$	${}_x\boldsymbol{v}^i$	${}_y\boldsymbol{v}^i$	${}_z\boldsymbol{v}^i$	w^i
2	1.1	-2.1	**0**	0.12	0.13	**0.33**
1.7	**1.5**	-2.2	0.1	**0**	0.12	**0.33**
1.8	**1.5**	**-2.25**	0.2	**0**	**0**	**0.33**

4.3.3.4 State Estimate

The estimated position $\hat{\boldsymbol{r}}_k$ given by the particle filter at timestep k is simply the weighted sum of all N particles

$$\hat{\boldsymbol{r}}_k = \mathbb{E}[\hat{\boldsymbol{r}}_k^i] = \sum_{i=1}^{N} \hat{\boldsymbol{r}}_k^i w_k^i \tag{4.11}$$

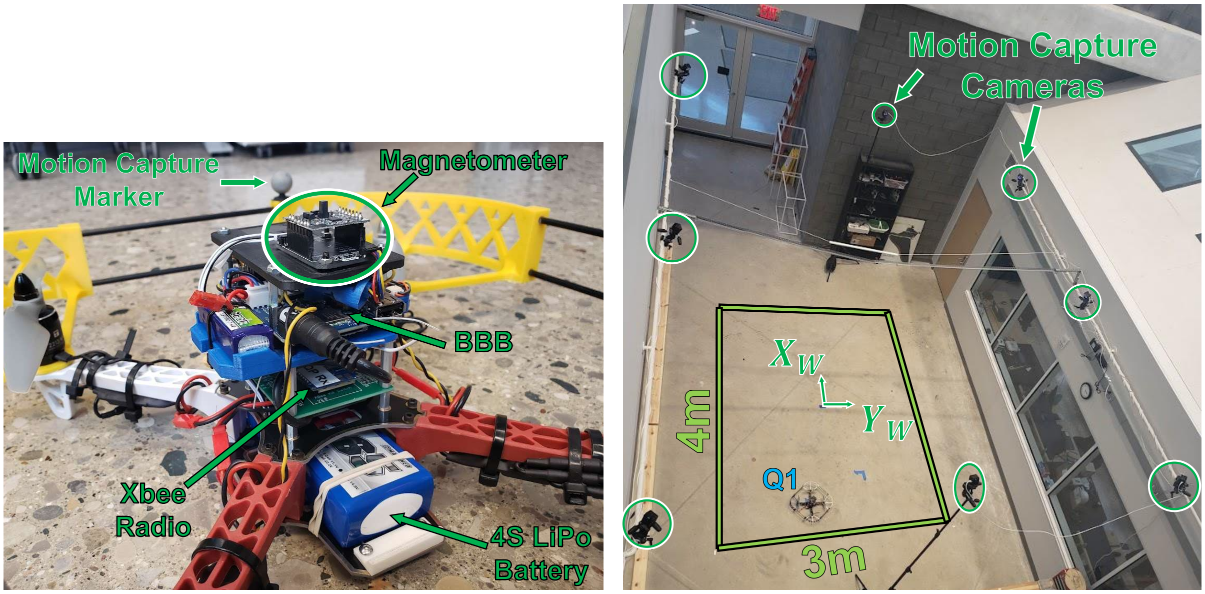
and similarly for the velocity estimate $\hat{\boldsymbol{v}}_k = \mathbb{E}[\hat{\boldsymbol{v}}_k^i]$.

4.3.4 Magnetometer Calibration

This chapter uses the same two-step calibration process described in Section 2.3.2 . It is based on the model and iterative least-squares solver from [68] with the two-step calibration algorithm from [69].

Ultimately, we seek to solve for the set of parameters $\theta = [\theta_a, \theta_b, \theta_c, \theta_{x_0}, \theta_{y_0}, \theta_{z_0}, \theta_\rho, \theta_\lambda, \theta_\phi]^\top$ with bias terms ($\theta_{x_0}, \theta_{y_0}, \theta_{z_0}$), scaling factors ($\theta_a, \theta_b, \theta_c$), and non-orthogonality terms ($\theta_\rho, \theta_\lambda, \theta_\phi$).

We gather magnetometer calibration data outdoors, just south of the University of Michigan's outdoor netted drone facility M-air. This location is far enough away from any buildings that the magnetic field strength is constant over a few meters and should be accurately reflected by the World Magnetic Model (WMM). The ambient magnetic field reference term used in our calibration is $B_R = 53.1351\mu T$ as taken from a WMM online calculator for M-Air's location (42.294431°N, 83.710442°W, and 270m above sea level) on the date of our flight experiments (September 1st, 2022).



(**a**) Main components of our flight vehicle Q1.

(**b**) The Robot Fly Lab in the Ford Motor Company Robotics Building at the University of Michigan.

Figure 4.2: Q1: The flight vehicle used for experiments in this chapter.

4.4 Experimental Procedure

All flight tests for this chapter were conducted in the Robot Fly Lab in the Ford Motor Company Robotics Building at the University of Michigan as shown in Figure 4.2. The workspace has eight Optitrack motion capture cameras that provide ground-truth position and attitude data at 120Hz. The workable volume of the workspace is 4m × 3m × 2.25m in x, y, and z respectively. We use the ground-truth attitude data from motion capture to rotate all magnetic field observations into the world frame. This allows us to better isolate the relationship between magnetic field gradients and position localization.

Finally, we use a specific naming convention for flight tests that allows the reader to associate our results with the relevant publicly-accessible data. Each flight test is denoted as "tY_XX" where *Y* refers to the flight test *series* and *XX* is the two-digit ID of the flight test in that series. As such, all the data used in this chapter is from the t6 series of data gathered on September 1st, 2022 meaning all flight tests will be identified as "t6_XX". More information is available in Appendix A.

4.4.1 Boundaries of Working Volume

Given the dimensions of our workable volume, we define a set of boundaries. Notice from Figure 4.2 that our world frame is a North-East-Down frame. The *training* boundaries range from ±2m, ±1.5m, and [-2.25, -0.5]m for x, y, and z respectively.

The *utilization* boundaries span ±1.75m, ±1.25m, and [-2.0, -0.75]m for x, y, and z respectively. The utilization boundaries are set 0.25m inward from each extremum of the training boundaries to ensure that our GPR never has to extrapolate when queried for the particle filter's measurement model (Equation 4.8). Early in our work, we incorrectly thought that respecting these utilization boundaries would prevent some of the noise we investigated in 2.5.2.2.

Finally, the *constraint* boundaries range from ±2.25m, ±1.75m, and [-2.5, 0]m for x, y, and z respectively. Since dead reckoning with our process model (Equation 4.7) is actually quite bad, these constraints help reject particles that are outside of the physical boundaries of our workspace.

4.4.2 Flight Profiles

The flight trajectories for this chapter are split into *training* and *utilization* flights. Unlike Chapter 3, we do not have a separate set of flight tests to validate the magnetic field maps. Instead, the utilization tests are each checked against the GPR map to ensure the training/utilization pair of flights are "consistent" in their magnetic field measurements.

Magnetic field observations from a flight are "consistent" with the GPR-based map if the flight's validation error is within two standard deviations (2σ) of the map's uncertainty bounds 96% of the time or more. This consistency check must pass for each of the three GPR_x, GPR_y, and GPR_z maps respectively. The concept of a consistency check was introduced in Section 2.5.6 to prevent the flight-by-flight variations from corrupting our analysis of magnetic field gradients and position localization. In Appendix B, we show the percentage of times each GPR captures its prediction error within 2σ for all eight utilization flights.

4.4.2.1 Training Trajectories

Training trajectories provide observations used to train the hyperparameters of our $3 \rightarrow 3$ magnetic field map and create the "compromise map" as explained in Section 2.5.3. As one might expect, the training flights span the training boundaries defined in the previous section (Section 4.4.1).

All training trajectories are based around single-altitude lawnmower trajectory shown in Figure 4.3. Here, the drone is commanded to stride 4m in the x direction, translate 0.25m in y, then cover the span in the x direction again. This continues until the 4m × 3m planar region is covered at which point the drone changes altitude and repeats the planar trajectory.

Each of the three training trajectories repeat this planar lawnmower pattern across different sets of altitudes. The first of these training flight trajectories gathers data at the lower altitudes {0.5m, 0.75m, 1.0m, 1.25m}, the second observes the higher altitudes {1.5m, 1.75m, 2.0m, 2.25m}, while the third trajectory (Scan-γ) gathers observations at altitudes {0.5m, 1.375m, and 2.25m}. Scan-γ

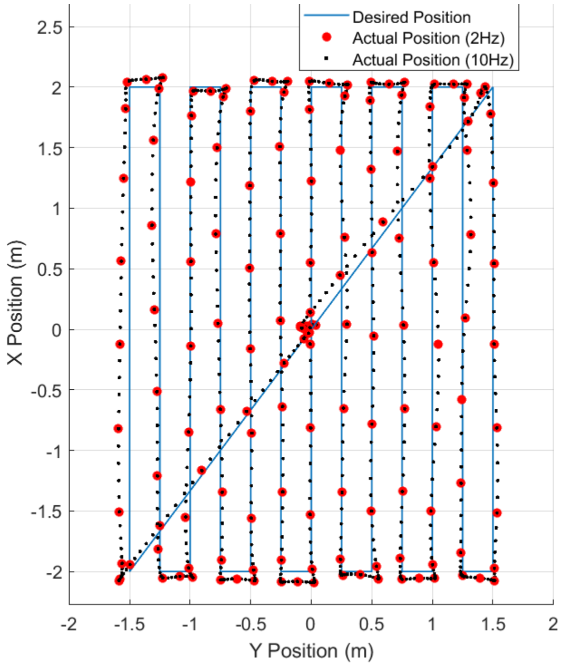


Figure 4.3: Single-altitude lawnmower pattern used in all *training* trajectories.

is similar to the trajectories we used in Chapter 3 while the other two are designed to give a uniform 3D grid of observations with 0.25m separation.

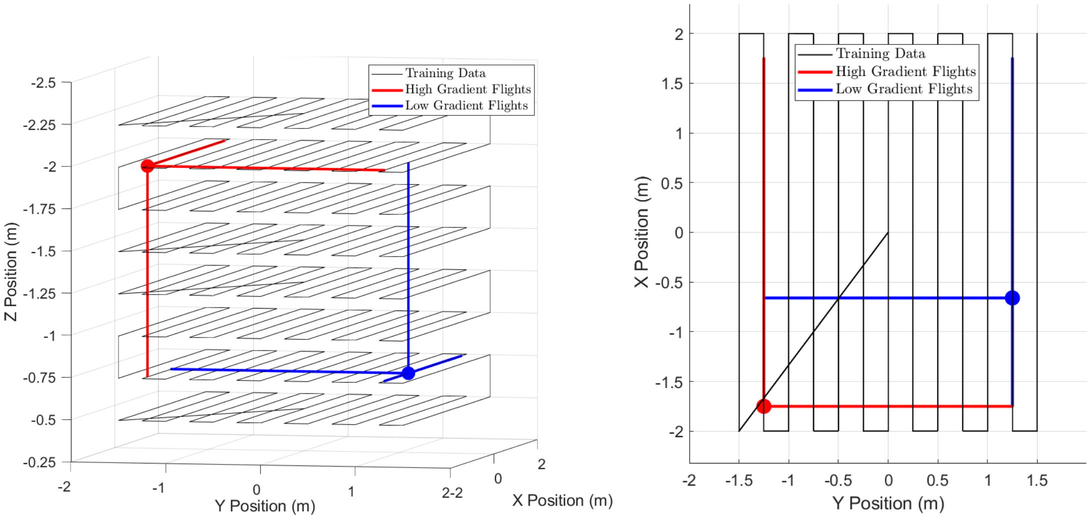
4.4.2.2 Utilization Trajectories

The utilization flights have commanded positions 0.25m inward from any extrema of the training boundaries that respect the utilization boundaries (Section 4.4.1).

All utilization trajectories are defined with respect to some center point $\boldsymbol{r}^*$. Given a center point, we either hover at the given point or define x, y, or z single-axis strides through the center point. This gives four trajectories (hover, X Stride, Y Stride, and Z Stride) per center point $\boldsymbol{r}^*$. Each utilization trajectory spends 120 seconds hovering or striding along its designated axis.

The goal of the utilization flights is to test our central hypothesis: how does the gradient of the magnetic field effect our ability to estimate the position of the UAV? As such, we define a center point at the location within our utilization boundaries with the highest magnetic field gradient norm ($\boldsymbol{r}^*_{max}$) and the location with the lowest gradient norm ($\boldsymbol{r}^*_{min}$). This gives us a total of eight utilization trajectories with two center points and four trajectories for each center point.

Given a GPR-based magnetic field map, we compute the spatial gradients (as in Equation 4.4) at 5400 locations ($30 \times 30 \times 6$ for x, y, and z respectively) within the *utilization boundaries*. With this, the point with the largest magnetic field gradient norm is $\boldsymbol{r}^*_{max}$ = (-1.75, -1.25, -2.0)m while the lowest gradient norm is at $\boldsymbol{r}^*_{min}$ = (-0.66, 1.25, -0.75)m.



(**a**) Isometric view

(**b**) $x - y$ (top-down) view

Figure 4.4: The eight utilization trajectories. $\boldsymbol{r}^*_{min}$ in blue and $\boldsymbol{r}^*_{max}$ in red. Utilization boundaries are 0.25m inward from training boundaries.

Figure 4.4 shows an isometric and top-down (X-Y) view of our eight utilization trajectories with respect to the lawnmower training trajectories. Here, $\boldsymbol{r}^*_{max}$ is in red with its X, Y, and Z strides passing through it. Similarly, $\boldsymbol{r}^*_{min}$ and its trajectories are shown in blue. The thin black lines are the training locations. Note how the eight utilization trajectories respect the utilization boundaries and are always set 0.25m inward from the edges defined by the lawnmower training locations.

The hover trajectories allow us to study how magnetic field gradients influence position localization when the vehicle is effectively stationary. This allows us to test the baseline hypothesis and see if flying in locations with higher magnetic field gradients leads to higher position localization accuracy.

4.4.3 Data Collection

Our flight vehicle "Q1" is equipped with a RM3100 magnetometer that is sampled at 200Hz, a MPU9250 IMU (gyroscope, accelerometer, and magnetometer) sampled at 200Hz, and a VL53L1X laser-based altimeter sampled at 4Hz. This dissertation made no use of the magnetometer on the MPU9250 relying solely on the RM3100 for all magnetic field measurements. For our use, the VL53L1X is useful given we fly indoors (little ambient infrared light from the sun) and at low altitudes.

Additionally, the Optitrack motion capture system provides ground truth position and attitude of our vehicle at 120Hz. Ground truth *attitude* data is used to rotate all sensor measurements into a common world frame while ground truth *position* is used for training our GPR maps and evaluating the accuracy of the particle filter's position estimates. The particle filter is only given the UAV's initial ground-truth position to sample a prior distribution of particles.

Table 4.3 lists all the t6-series flights used in this chapter. Observations from seven flights are used to train the magnetic field map (via our compromise map technique). Here, we fly two trials of each training trajectory (Section 4.4.2.1) to train the map. A single $\boldsymbol{r}^*_{max}$ flight is included in the training data to make the utilization flight t6_08 consistent with the resultant map (i.e., both t6_08 and t6_16 were inconsistent with the compromise map, so observations from t6_16 were added to the map training process to resolve this issue).

The remaining eight flight trials in Table 4.3 are the $\boldsymbol{r}^*_{min}$ and $\boldsymbol{r}^*_{max}$ trajectories used to test the accuracy of our particle filter. We also check that each utilization flight is consistent with the GPR map (Appendix B). Recall that a validation flight is "consistent" with a given GPR-based map if the flight's validation error is within two standard deviations (2σ) of the map's uncertainty bounds 96% of the time or more. This must hold true for each GPR (GPR_x, GPR_y, and GPR_z) respectively to pass the consistency check.

When gathering flight data, if a utilization flight failed the consistency check, we flew another repetition of the same trajectory to see if the flight-by-flight variation of the new trial would be more agreeable with the given map. Many repetitions of the $\boldsymbol{r}^*_{max}$ hover trajectory were inconsistent with the map, so we instead added observations from t6_16 to the map rather than flying additional repetitions.

4.4.4 Creating and Querying the GPR Magnetic Field Map

In this chapter, we create what we call a "compromise map" which was first introduced in Section 2.5.3. This method allows us to train the hyperparameters of our GPR-based map on n_2 observations from multiple flight tests. Then, to keep inference time low, we create an "intermediate map", query this intermediate map at n_1 user-selected locations which are then used as the inference set for the final, compromise map.

We use `gpml-matlab`[1] to optimize the hyperparameters of our squared exponential kernels, to create the visualizations of our magnetic field maps (Figures 4.5 and 4.6 introduced later), and to validate each utilization flight to enforce our consistency check (Figure B.1).

Our particle filter is built upon the `particles` [5] open-source Python library. The original version of this code was not setup for robotics-based navigation so we, locally, made many modifications to accommodate our application.

Running a Python-based particle filter with a Matlab-based GPR map proved to be cumbersome. As such, we also use `GPyTorch` [102] as the backbone of our GPR-based map used in the particle

---

[5] `particles` was created by Nicolas Chopin and Omiros Papaspiliopoulos `https://particles-sequential-monte-carlo-in-python.readthedocs.io/en/latest` accessed on January 2022

Table 4.3: List of flights used for *training* and *validation/utilization*. All magnetic field observations were downsampled to 2Hz for *training*. Number of observations each training flight contributes to the map is listed.

Flight Number	**Flight Description**	**Training [2 Hz]** (# Observations)	**Validation [200 Hz] & Utilization [200 Hz]**
t6_00	Lower Four Alts.	571	–
t6_04	Lower Four Alts.	573	–
t6_01	Upper Four Alts.	580	–
t6_05	Upper Four Alts.	573	–
t6_03	Scan-γ	442	–
t6_06	Scan-γ	441	–
t6_16	$\boldsymbol{r}^*_{\text{max}}$ Hover	305	–
t6_09	$\boldsymbol{r}^*_{\text{min}}$ Hover	–	✓
t6_08	$\boldsymbol{r}^*_{\text{max}}$ Hover	–	✓
t6_18	$\boldsymbol{r}^*_{\text{min}}$ X Stride	–	✓
t6_17	$\boldsymbol{r}^*_{\text{max}}$ X Stride	–	✓
t6_13	$\boldsymbol{r}^*_{\text{min}}$ Y Stride	–	✓
t6_19	$\boldsymbol{r}^*_{\text{max}}$ Y Stride	–	✓
t6_15	$\boldsymbol{r}^*_{\text{min}}$ Z Stride	–	✓
t6_14	$\boldsymbol{r}^*_{\text{max}}$ Z Stride	–	✓

filter's measurement model (Equation 4.8). The hyperparameters computed by `gpml-matlab` are ported to `GPyTorch` and both use the same n_1 inference points. The two toolboxes have comparable magnetic field estimation error when validating over our utilization flights if we use no more than $n_1 = 511$ inference points in `GPyTorch`.

For $n_1 \geq 512$, `GPyTorch` uses a different set of mathematical operations to interpolate using Gaussian process regression. For reasons we have not been able to identify or solve, the predictions given by `GPyTorch` are significantly worse with $n_1 \geq 512$. As such, we use at most $n_1 = 511$ inference points when constructing our compromise map.

The $n_1 = 511$ user-selected locations for this analysis are chosen as follows. Points are distributed evenly through the ±2m x axis span, ±1.5m y axis span, and [-2.25, -0.5]m z axis span of the working volume (i.e., the training boundaries). A target location for the compromise map is selected every 0.5m, 0.5m, and 0.25m for x, y, and z axes respectively. In total this gives 504 locations within the working volume. The remaining seven points are evenly spaced from the ground to an altitude of 0.5m so the compromise map has some observations during the takeoff and landing sequence (which, for all our flights, are above the origin).

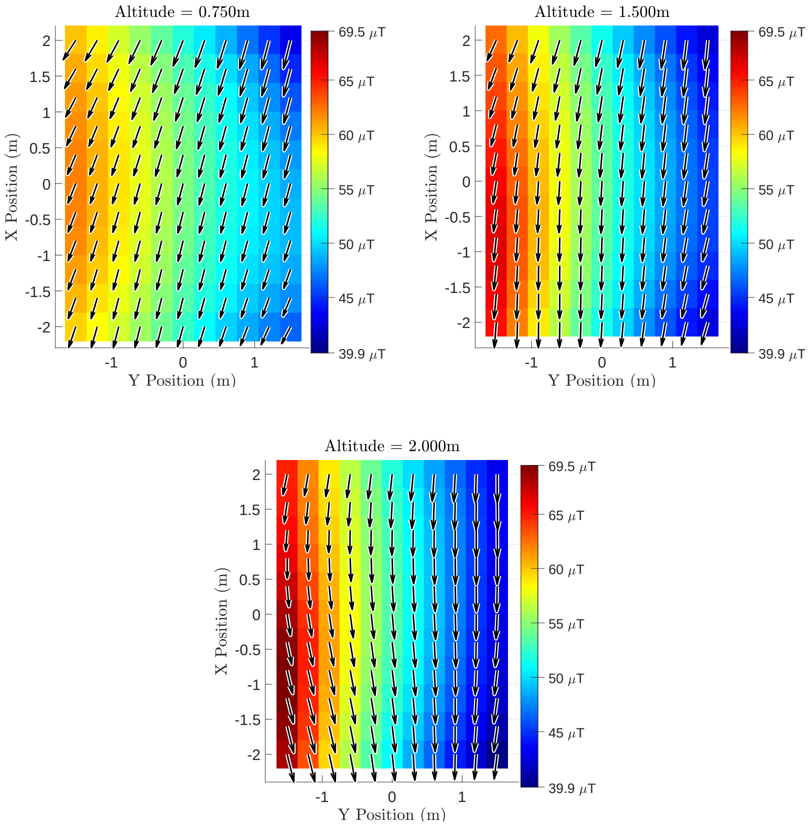


Figure 4.5: Magnetic field map of indoor flight space. Black arrows show the direction of the planar $(X - Y)$ component of the magnetic field. The color behind each arrow shows the strength and direction of the Z component of the magnetic field.

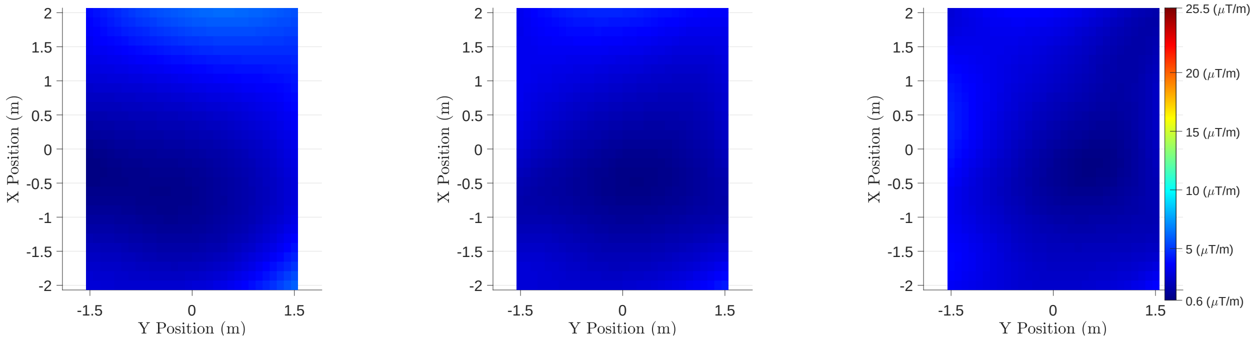
4.5 Results

This section starts with visualizations of the magnetic field map in the flight lab and visualizations of the spatial gradient norms $\left|\frac{\partial \tilde{\boldsymbol{B}}(\boldsymbol{r})}{\partial x}\right|$, $\left|\frac{\partial \tilde{\boldsymbol{B}}(\boldsymbol{r})}{\partial y}\right|$, $\left|\frac{\partial \tilde{\boldsymbol{B}}(\boldsymbol{r})}{\partial z}\right|$ and $|\nabla \hat{\boldsymbol{m}}(\boldsymbol{r})|$. This is done to contextualize the $\boldsymbol{r}^*_{min}$, and $\boldsymbol{r}^*_{max}$ locations chosen for the utilization flight trajectories introduced in Section 4.4.2. Next, Section 4.5.2 compares the accuracy of PF position localization against the gradient of the magnetic field. Finally, in Section 4.5.3, we present a gradient localization metric meant to predict the accuracy of position localization using the gradient of the magnetic field.

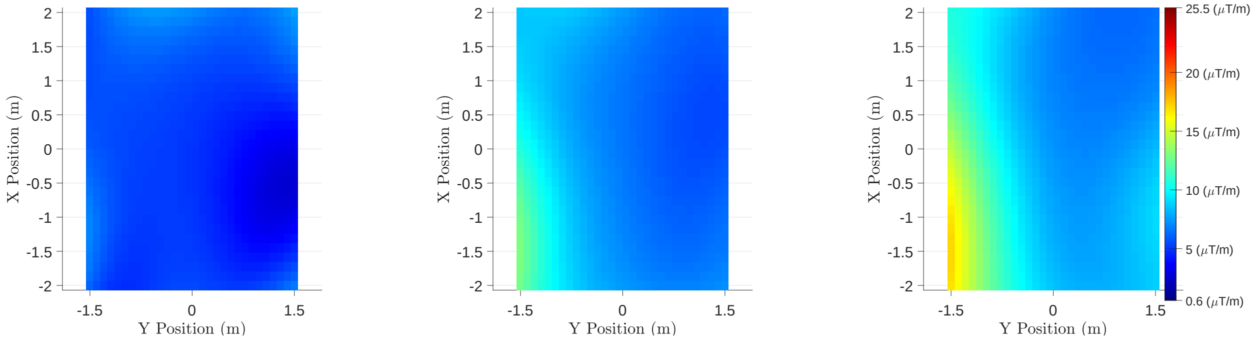
4.5.1 Magnetic Field Map and Gradients

This section presents a visualization of the *compromise* magnetic field map and its gradients to demonstrate how the locations $\boldsymbol{r}^*_{\text{min}}$ and $\boldsymbol{r}^*_{\text{max}}$ were selected for the utilization trajectories.

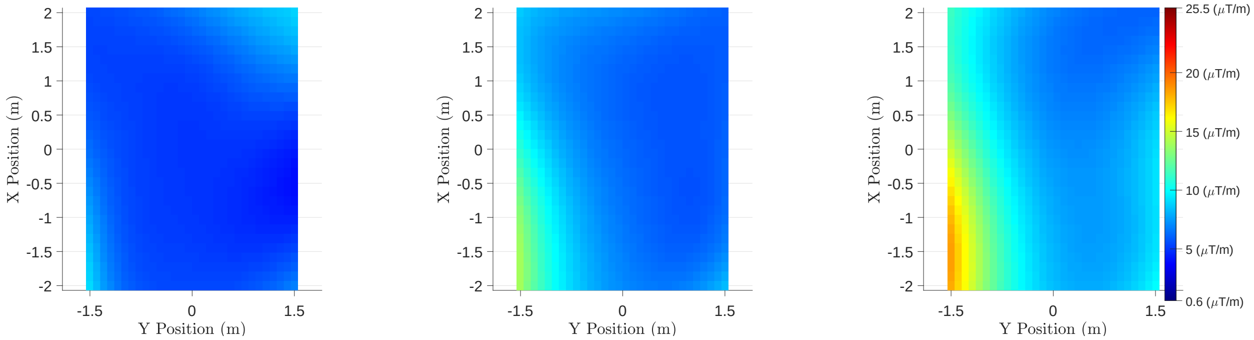
The indoor magnetic field map is shown in Figure 4.5 at altitudes 0.75m, 1.5m, and 2.0m. Note that $\boldsymbol{r}^*_{min}$ is at $z =$ -0.75m and $\boldsymbol{r}^*_{max}$ is at $z =$-2.0m (Figure 4.4). This visualization tool is the same



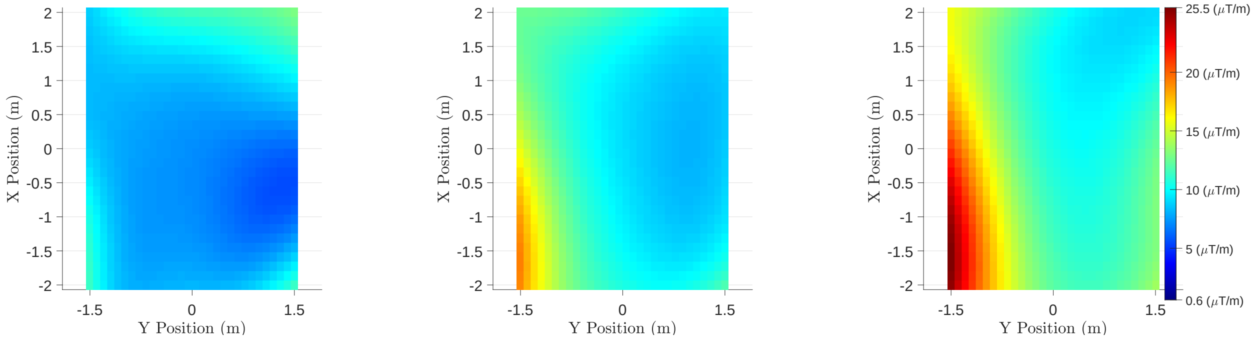
(a) $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right|$ at 0.75m altitude. **(b)** $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right|$ at 1.5m altitude. **(c)** $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right|$ at 2.0m altitude.



(d) $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial y}\right|$ at 0.75m altitude. **(e)** $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial y}\right|$ at 1.5m altitude. **(f)** $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial y}\right|$ at 2.0m altitude.



(g) $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial z}\right|$ at 0.75m altitude. **(h)** $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial z}\right|$ at 1.5m altitude. **(i)** $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial z}\right|$ at 2.0m altitude.



(j) $|\nabla \hat{\boldsymbol{m}}(\boldsymbol{r})|$ at 0.75m altitude. **(k)** $|\nabla \hat{\boldsymbol{m}}(\boldsymbol{r})|$ at 1.5m altitude. **(l)** $|\nabla \hat{\boldsymbol{m}}(\boldsymbol{r})|$ at 2.0m altitude.

Figure 4.6: Magnetic field gradient in the flight workspace.

as that introduced in Chapter 3 where the black arrows show the direction of the $x - y$ component of the magnetic field while the color behind each arrows depicts the magnitude and direction of the z component of the field. Even though the data for this chapter was collected a year after the analysis of Chapter 3, the map from Chapter 3 is qualitatively similar to that of this chapter for the common altitudes depicted.

Figure 4.6 shows the norm of the x, y, z, and magnitude of the magnetic field *gradient* at the same altitudes from Figure 4.5. We compute the magnetic field gradient at 900 locations per altitude using the finite difference method from Equations 4.2 and 4.3 with $\delta s = 0.05$m. The four rows in Figure 4.6 depict $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right|$, $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial y}\right|$, $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial z}\right|$, and $|\nabla \hat{\boldsymbol{m}}(\boldsymbol{r})|$ respectively (from Equations 4.4 and 4.5).

By comparing the first three rows of Figure 4.6, we see that the magnetic field changes the least when moving in the x direction (Figures 4.6a - 4.6c) and has much higher spatial gradients when moving in the y and z directions (Figures 4.6d - 4.6i). These trends can be seen qualitatively by the arrow direction and z component colors in Figure 4.5.

The locations $\boldsymbol{r}^*_{\text{min}}$ and $\boldsymbol{r}^*_{\text{max}}$ are the locations of the minimum and maximum norm of the spatial gradient (Figures 4.6j-4.6l) within the limits of the utilization bounds of $\pm$1.75m for x, $\pm$1.25m for y, and [-2.0, -0.75]m for z. Together, these rules give $\boldsymbol{r}^*_{min}$ = (-0.66, 1.25, -0.75)m and $\boldsymbol{r}^*_{max}$ = (-1.75, -1.25, -2.00)m with a $|\nabla \hat{\boldsymbol{m}}(\boldsymbol{r})|$ value of 5.61μT/m and 21.88μT/m respectively.

4.5.2 Magnetic Field Gradient Impact on Position Localization

This section analyses how the accuracy of our particle filter (Section 4.3.3) relates to the gradient of the magnetic field in our workspace. We begin by presenting the RMSE of the particle filter (PF) against the eight utilization trajectories, continue by analyzing why it is difficult to draw conclusions by looking only at the RMSE across the whole flight, then present the particle filter's accuracy at select timestamps of each utilization flight that isolate each flight trajectory's intended behavior.

Table 4.4 shows the RMSE of the particle filter (in meters) for all eight utilization flights and across the two measurement models presented in Section 4.3.3.2. The table is a bit dense with information, so we take a moment to describe its organization.

The three *major columns* (separated by double vertical bars) give the ID of each utilization flight, the particle filter RMSE using just the magnetometer measurement model, and finally the PF RMSE with both magnetometer and altimeter measurement models. Within the two numerical major columns, each *minor column* gives the norm, x, y, and z RMSE values of the particle filter respectively. Here, we isolate the norm RMSE with a single vertical bar since it is used most often in the rest of this chapter.

The table is split into four *major rows* that present RMSE for Hover, X Stride, Y Stride, and Z Stride trajectories respectively. The minor rows within a major row distinguish trajectories at $\boldsymbol{r}^*_{min}$

Table 4.4: *Full-flight* RMSE (takeoff and landing included). Each RMSE value is the mean of 20 repetitions of the particle filter per flight test. Bolded values indicate lower norm RMSE between $\boldsymbol{r}^*_{min}$ and $\boldsymbol{r}^*_{max}$ while underlines indicate lower norm RMSE across the two measurement models.

RMSE* (m)** ***Norm—X—Y—Z	**Inertial: Mag**				**Inertial: Mag + Alt**			
(t6_09) $\boldsymbol{r}^*_{min}$: Hover	**0.07**	0.03	0.06	0.01	**0.07**	0.03	0.06	0.01
(t6_08) $\boldsymbol{r}^*_{max}$: Hover	0.15	0.15	0.03	0.03	$\underline{0.09}$	0.09	0.02	0.02
(t6_18) $\boldsymbol{r}^*_{min}$: X Stride	**0.11**	0.07	0.08	0.03	$\underline{\mathbf{0.10}}$	0.06	0.07	0.03
(t6_17) $\boldsymbol{r}^*_{max}$: X Stride	0.40	0.39	0.02	0.11	$\underline{0.22}$	0.22	0.02	0.05
(t6_13) $\boldsymbol{r}^*_{min}$: Y Stride	0.15	0.14	0.05	0.02	$\underline{0.14}$	0.13	0.05	0.02
(t6_19) $\boldsymbol{r}^*_{max}$: Y Stride	**0.09**	0.09	0.02	0.03	**0.09**	0.08	0.02	0.02
(t6_15) $\boldsymbol{r}^*_{min}$: Z Stride	0.53	0.48	0.10	0.19	$\underline{0.52}$	0.50	0.13	0.05
(t6_14) $\boldsymbol{r}^*_{max}$: Z Stride	**0.08**	0.07	0.04	0.02	**0.08**	0.07	0.04	0.02

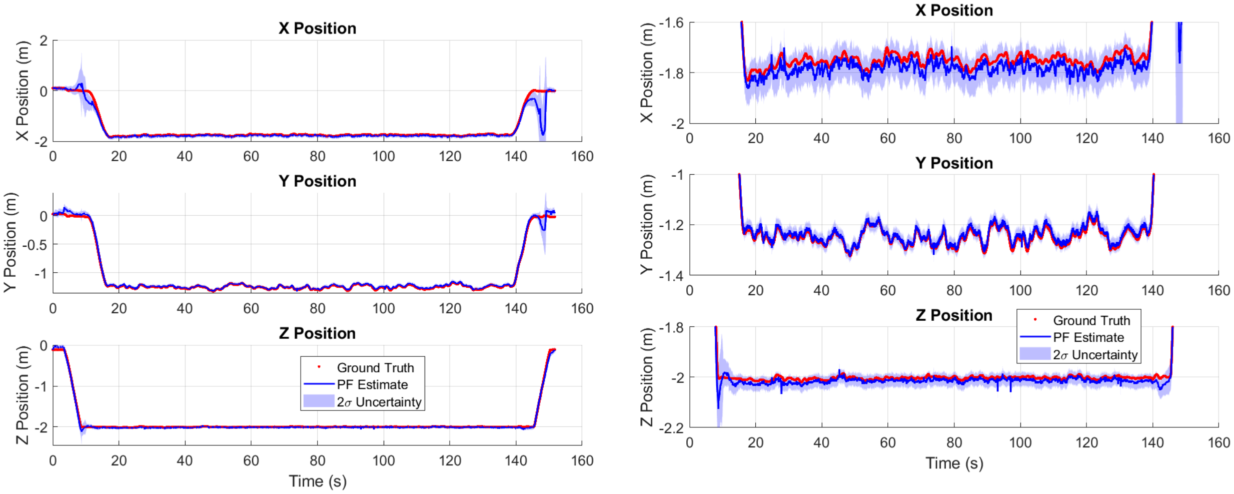
(location with the lowest magnetic field gradient in the workspace) versus $\boldsymbol{r}^*_{max}$ (highest magnetic field gradient).

Additionally, **bold** values indicate lower norm RMSE for either $\boldsymbol{r}^*_{min}$ or $\boldsymbol{r}^*_{max}$ within a major row. Since a higher magnetic field gradient should make position localization easier, we expect trajectories around $\boldsymbol{r}^*_{max}$ will have lower RMSE than those around $\boldsymbol{r}^*_{min}$. Similarly, underlined values indicate which measurement model yields a lower norm RMSE for each of the eight utilization flights. Note that if both measurement models have the same RMSE (to the hundredths place), *neither* will be underlined. Here, we expect the additional information from the altimeter to improve the particle filter's performance and reduce RMSE. Finally, (though not indicated with any typographical emphasis), we expect X position error to be higher than Y or Z since gradients in the X direction are much smaller (Figure 4.6).

In short, **bold** compares norm RMSE between $\boldsymbol{r}^*_{min}$ and $\boldsymbol{r}^*_{max}$ within a major row, and underlined compares norm RMSE across the two measurement models.

Note that for each of the eight utilization flights, we ran 20 repetitions of the particle filter since we draw samples from random distributions in our process model (Equation 4.7) and resampling steps.

From Table 4.4, the only expectation that is met is that x RMSE values tend to be higher than y or z RMSE. When comparing $\boldsymbol{r}^*_{max}$ against $\boldsymbol{r}^*_{min}$, both Hover and X Stride trajectories have noticeably lower RMSE at $\boldsymbol{r}^*_{min}$ than at their higher gradient counterparts. Further, there is a nearly-even split in cases where the altimeter measurement model outperforms the magnetometer-only model while it seems to have no improvement in the other cases. Finally, we note that t6_15 on both measurement models and t6_17 with just the magnetometer have significantly higher norm RMSE values than any



(**a**) t6_08 ($\boldsymbol{r}^*_{max}$ Hover) full flight.

(**b**) t6_08 ($\boldsymbol{r}^*_{max}$ Hover) zoom on stationary hover section.

Figure 4.7: t6_08 particle filter results. Takeoff and landing sequence have large, and inconsistent, error profiles.

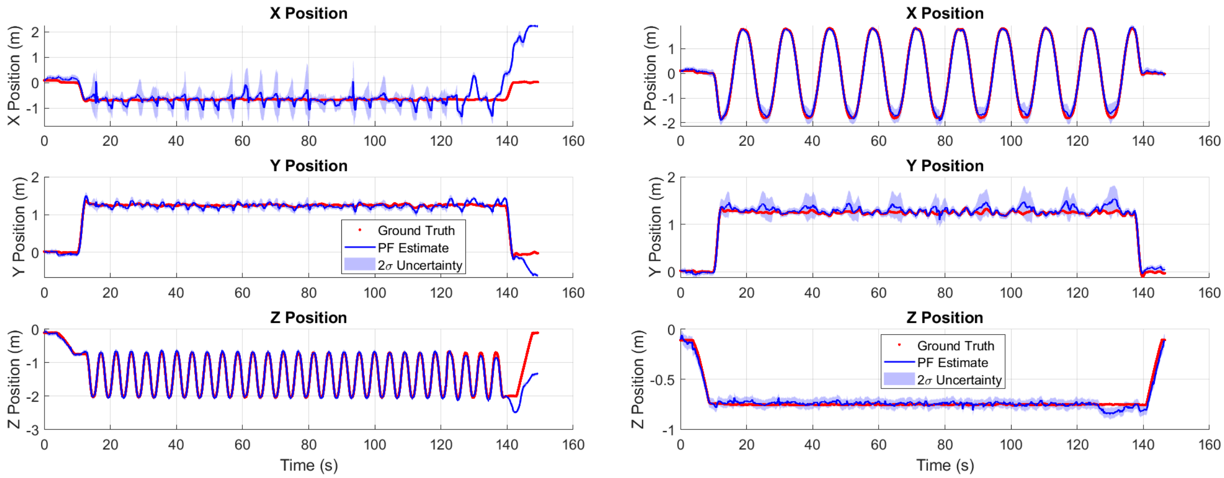
other cases on the table. We turn our investigation to these outliers to better understand the lack of trends in Table 4.4.

4.5.2.1 Accounting for Takeoff and Landing

Figure 4.7 shows the time-series results of one repetition of the particle filter on t6_08 ($\boldsymbol{r}^*_{max}$ Hover). Even though t6_08 is not an outlier in Table 4.4, the time-series plot shows clear issues during the takeoff and landing sequences of this flight trajectory. In fact, the particle filter's norm RMSE during just the hover portion of t6_08 is 0.04m on average which is much lower than the 0.15m norm RMSE from Table 4.4.

Each of the particle filter's 20 repetitions had slightly different error profiles during the takeoff and landing portions of t6_08, but all consistently gave 0.04m norm RMSE during the 120s of hovering. This issue of varying error profiles during takeoff and landing holds true for many of the other utilization trajectories. For this, we propose computing particle filter accuracy only during the main 120s portion of each utilization flight. The particle filter's performance on this *region of interest* will be presented in a table after we discuss one additional adjustment.

In Appendix B, we show that t6_18 and t6_15 became inconsistent with the compromise GPR map during the last $\sim$15% of their respective flights. t6_15 has more severe GPR inconsistencies and the consequence of that on the particle filter's ability to localize is shown in Figure 4.8a. Here, we see a regular pattern of error from t =15s to t =130s (after takeoff and before GPR inconsistency) and a clear degradation of localization accuracy after 130s. Note that the particle filter's accuracy starts to degrade before the landing sequence begins at 140s and gets significantly worse during



(**a**) t6_15. Accuracy diverges (130s) before landing sequence (139s).

(**b**) t6_18. Accuracy diverges (126s) before landing sequence (137s).

Figure 4.8: Tests t6_15 and t6_18 have slightly shorter regions of interest due to sudden mapping inconsistencies (Figure B.1).

landing. The accuracy degradation at 130s (out of a 149.59s total flight time) lines up with the GPR inconsistency in Figure B.1g in Appendix B. Figure 4.8a shows similar behavior just before the landing sequence in test18, but with less severe degradation in position localization accuracy.

For this, we propose a slightly different truncation scheme for t6_18 and t6_15 where the takeoff sequence is removed from the start, and the last 15% of these flights are removed to avoid the GPR inconsistency segments. The remaining six utilization flights will, instead, truncate the takeoff and landing sequences only.

Table 4.5 lists the indices (and equivalent timestamps) for each flight's region of interest. Note that the main flight time, indicated by the difference between upper t_{ub} and lower-bound t_{lb} timestamps, is $(t_{ub} - t_{lb})$ $\geq$120s for all flights except t6_18 and t6_15. Finally, note that all 20 repetitions for each of the eight utilization flights still simulate the entire flight profile (from t =0s to the last timestep) meaning any errors in position estimation during takeoff are still propagated throughout the flight. The proposed truncations are only for values we report in the following tables and plots.

4.5.2.2 Particle Filter Accuracy

Table 4.6 shows the particle filter's performance for each flight computed only within the respective region of interest. Similar to Table 4.4, bolded values indicate lower norm RMSE between $\boldsymbol{r}^*_{min}$ and $\boldsymbol{r}^*_{max}$ while underlined entries have lower norm RMSE between the two measurement models.

Recall our three expectations with the data in this table. First, trajectories at $\boldsymbol{r}^*_{max}$ should have

Table 4.5: Region of interest for each utilization flight indicated by discrete timestamp k and continuous timestamp t (in seconds). Lower bound is defined as end of takeoff sequence while upper bound is either start of landing sequence, or start of GPR inconsistency (t6_18 and t6_15).

Test #	k_{lb}	k_{ub}	t_{lb} (s)	t_{ub} (s)
(t6_09) $\boldsymbol{r}^*_{min}$: Hover	3556	27600	17.78	138
(t6_08) $\boldsymbol{r}^*_{max}$: Hover	3525	27530	17.625	137.65
(t6_18) $\boldsymbol{r}^*_{min}$: X Stride	2480	25127	12.4	125.635
(t6_17) $\boldsymbol{r}^*_{max}$: X Stride	2555	27437	12.775	137.185
(t6_13) $\boldsymbol{r}^*_{min}$: Y Stride	2475	27588	12.375	137.94
(t6_19) $\boldsymbol{r}^*_{max}$: Y Stride	2443	27545	12.215	137.725
(t6_15) $\boldsymbol{r}^*_{min}$: Z Stride	2505	25972	12.525	129.86
(t6_14) $\boldsymbol{r}^*_{max}$: Z Stride	3763	29054	18.815	145.27

Table 4.6: *Truncated* Position RMSE. Each value here is the mean of 20 repetitions of the particle filter per flight test. Bolded values indicate lower norm RMSE between $\boldsymbol{r}^*_{min}$ and $\boldsymbol{r}^*_{max}$ while underlines indicate lower norm RMSE across the two measurement models. Compare to Table 4.4.

RMSE **(m)** ***Norm—X—Y—Z***	**Inertial: Mag**				**Inertial: Mag + Alt**			
(t6_9) $\boldsymbol{r}^*_{min}$: Hover	0.07	0.03	0.07	0.01	0.07	0.03	0.07	0.01
(t6_8) $\boldsymbol{r}^*_{max}$: Hover	**0.04**	0.03	0.01	0.01	**0.04**	0.03	0.01	0.01
(t6_18) $\boldsymbol{r}^*_{min}$: X Stride	**0.10**	0.07	0.07	0.02	**0.10**	0.07	0.07	0.02
(t6_17) $\boldsymbol{r}^*_{max}$: X Stride	0.42	0.40	0.02	0.12	$\underline{0.22}$	0.22	0.01	0.05
(t6_13) $\boldsymbol{r}^*_{min}$: Y Stride	0.16	0.15	0.06	0.02	$\underline{0.15}$	0.14	0.05	0.02
(t6_19) $\boldsymbol{r}^*_{max}$: Y Stride	**0.09**	0.08	0.02	0.03	$\underline{\mathbf{0.08}}$	0.07	0.02	0.02
(t6_15) $\boldsymbol{r}^*_{min}$: Z Stride	$\underline{0.17}$	0.16	0.06	0.03	0.26	0.25	0.05	0.03
(t6_14) $\boldsymbol{r}^*_{max}$: Z Stride	**0.08**	0.07	0.04	0.01	**0.08**	0.07	0.04	0.01

lower PF error (i.e., should be bolded) since it should be easier to detect a change in position if there is a higher magnetic field gradient. Second, we expect entries on the right-most column to have lower RMSE (i.e., should be underlined) since adding altimeter information should improve the state estimate. Note that if both measurement models have the same RMSE (to the hundredths place), *neither* will be underlined. Finally, x RMSE values should be larger than y or z since the spatial gradient in the x direction is much smaller (Figure 4.6).

Table 4.4 shows that our first hypothesis holds true for three of the four trajectory types. However, for X Stride, we see lower error at the location with lower gradient. The reason for this is depicted in Figure 4.9 which shows the time-series position estimate of a single t6_17 ($\boldsymbol{r}^*_{max}$ X Stride) repetition. Figure 4.9a uses just the magnetometer while Figure 4.9b adds in the altimeter.

In Figure 4.9a, we see cyclical errors in both x and z position estimates which occurred for all 20 repetitions of t6_17 when using only the magnetometer. In essence, when using only the magnetic field to evaluate the credibility of each particle (Equation 4.8), the PF consistently, and erroneously, believes the drone is changing altitude at the +x side of the trajectory. However, when we add in the altimeter (Figure 4.9b), the PF's estimate aligns more closely with the ground-truth data. Although not depicted here, as the gain σ^2_{alt} decreases, the x and z PF estimates get closer to their respective ground-truth values in Figure 4.9b. Essentially, as the altimeter is given higher credibility than the magnetometer, the z estimate improves which, in turn, prevents the x position estimate from diverging. Though it is clear that adding the altimeter helps in this case, we will continue to explain why the magnetometer-only estimates (Figure 4.9a) have these errors.

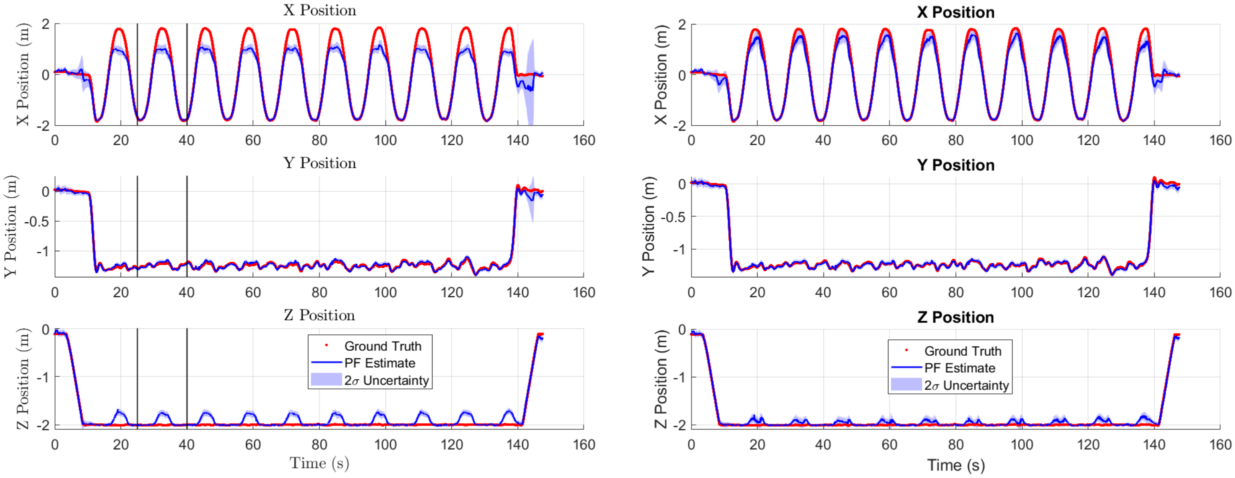
The cyclical position estimation errors in Figure 4.9a occur because of inaccuracies in the GPR map. Figure 4.9c plots the GPR map's error $GPR(\boldsymbol{r}_k) - \tilde{\boldsymbol{y}}_k$ across all 200Hz measurements of flight t6_17 ($\boldsymbol{r}^*_{max}$ X Stride). Here, we see cyclical mapping errors with GPR_x and GPR_y that temporally align with the x and z position errors in Figure 4.9a.

It is important to note two things about the magnetic field measurement model from Equation 4.8. First, it gives higher weight (more credibility) to particles that *minimize* $|\Delta \boldsymbol{m}^i_k| = |\hat{\boldsymbol{m}}^i_k - \tilde{\boldsymbol{y}}_k|$. Second, $\Delta \boldsymbol{m}^i_k$ is dependent on the GPR's understanding of the ambient magnetic field since $\hat{\boldsymbol{m}}^i_k = GPR(\hat{\boldsymbol{r}}^i_k)$. In essence, when using only the magnetometer measurement model, our particle filter aims to find particles whose positions $\hat{\boldsymbol{r}}^i_k$ give a GPR estimate $\hat{\boldsymbol{m}}^i_k(\hat{\boldsymbol{r}}^i_k)$ that most closely agrees with the actual measurement $\tilde{\boldsymbol{y}}_k$.

For this, it is helpful to define two metrics

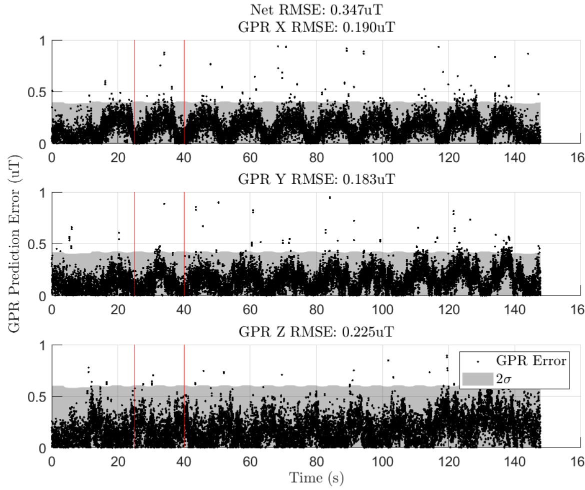
$$\begin{aligned} \Delta \boldsymbol{m}^{PF}_k &= GPR(\hat{\boldsymbol{r}}_k) - \tilde{\boldsymbol{y}}_k = \Delta \hat{\boldsymbol{m}}_k \\ \Delta \boldsymbol{m}^{GT}_k &= GPR(\boldsymbol{r}_k) - \tilde{\boldsymbol{y}}_k \end{aligned} \tag{4.12}$$

where $\Delta \boldsymbol{m}^{PF}_k$ is associated with the GPR's prediction at the particle filter's position estimate ($\hat{\boldsymbol{r}}_k$)



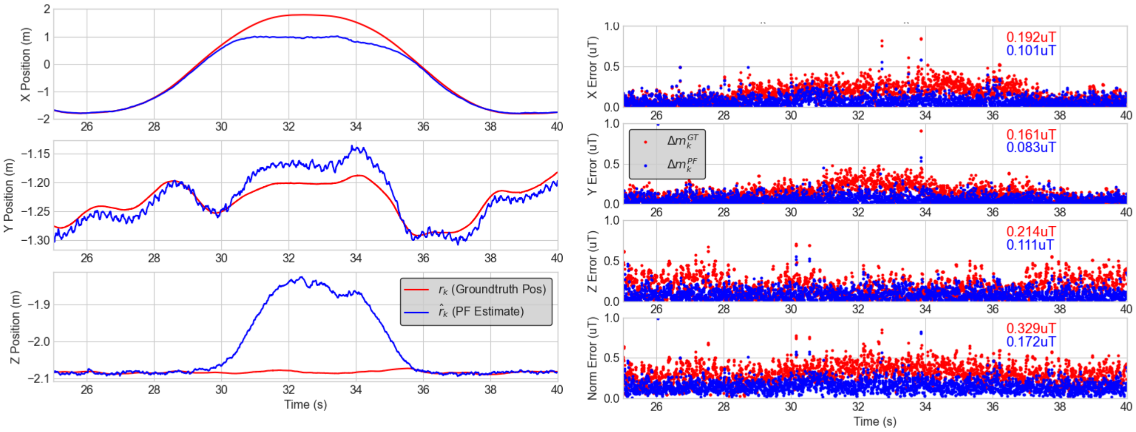
(**a**) t6_17: Only magnetometer. t =25s - 40s highlighted.

(**b**) t6_17: Magnetometer and altimeter. $\sigma_{alt}^2 = d = 0.005m^2$.



(**c**) t6_17: GPR error. Poor PF estimates occur during higher x and y RMSE. t =25s - 40s highlighted.

Figure 4.9: Erroneous x and z position estimates (4.9a) are casued by GPR error (4.9c). Adding an altimeter (4.9b) reduces the likelihood of our PF choosing the erroneous, alternate path.



(**a**) Target region of t6_17 with erroneous PF estimate.

(**b**) $\Delta \boldsymbol{m}_k^{GT}$ and $\Delta \boldsymbol{m}_k^{PF}$ at target region.

Figure 4.10: t6_17 from timestamps t =25s - 40s. Poor PF estimates occur when x and y GPRs have higher RMSE.

while $\Delta \boldsymbol{m}_k^{GT}$ uses the GPR's prediction at the ground-truth location ($\boldsymbol{r}_k$). Note that this "new" $\Delta \boldsymbol{m}_k^{PF}$ is identical to the term $\Delta \hat{\boldsymbol{m}}_k$ defined earlier with Equation 4.8. The re-definition helps to disambiguate the following analysis.

Figure 4.10 shows a zoomed-in segment $t = 25$s - 40s of t6_17 (the same timestamps highlighted by vertical bars in Figures 4.9a and 4.9c). On the left, Figure 4.10a shows the ground truth position ($\boldsymbol{r}_k$) in red and particle-filter-estimated position ($\hat{\boldsymbol{r}}_k$) in blue for t6_17 ($\boldsymbol{r}^*_{max}$ X Stride). Similarly, Figure 4.10b shows $|\Delta \boldsymbol{m}_k^{GT}|$ in red and $|\Delta \boldsymbol{m}_k^{PF}|$ in blue for each respective location given in Figure 4.10a. The numeric values within each subplot gives the x, y, z and norm RMSE across the plotted datapoints.

Figure 4.10b shows us that $|\Delta \boldsymbol{m}_k^{PF}| < |\Delta \boldsymbol{m}_k^{GT}|$ for the segments where the PF's estimate diverges from the ground truth position. Said another way, the particle filter has foundd a path ($\hat{\boldsymbol{r}}_k$) within the workspace that minimizes $|\Delta \boldsymbol{m}_k^i|$ better than the true trajectory ($\boldsymbol{r}_k$). This explains why all 20 repetitions of the particle filter show this same cyclical error on the +x side of t6_17.

Note, however, that this alternate path requires a change in altitude. This is why adding in the altimeter measurement model (Figure 4.9b) improved performance for both the x and z components of the particle filter's estimate. Evidently, there are no other *nearby* alternative paths where $|\Delta \boldsymbol{m}_k^{PF}| < |\Delta \boldsymbol{m}_k^{GT}|$ at a height of $z = -2$m. Decreasing $|\Delta \boldsymbol{m}_k^{GT}|$ by improving map accuracy (which in our case is caused by noise from the UAV) might improve PF estimates without the altimeter.

This analysis is only possible because we make, validate, and leverage our magnetic field maps. To our knowledge, this is the first time a work has clearly demonstrated how ambiguities in the magnetic field (and errors in mapping the magnetic field) can directly affect position estimates and distinguish such errors from low gradients in the ambient field. The closest we have seen is in Figure 7 of Ref. [72] where they show similar divergences with a circular trajectory and attribute the periodic errors to a "bimodality" in the PF estimate. Their PF temporarily gives similar credibility to two different trajectories simultaneously, but the net estimate ($\hat{\boldsymbol{r}}_k = \mathbb{E}[\hat{\boldsymbol{r}}_k^i]$) is far from the ground truth. However, in our case, the PF simply *prefers* an alternate path over the truth since $|\Delta \boldsymbol{m}_k^{PF}| < |\Delta \boldsymbol{m}_k^{GT}|$ along this erroneous trajectory.

In summary, Table 4.6 allows us to analyze our three hypotheses. First, $\boldsymbol{r}^*_{max}$ trajectories typically have lower error than their $\boldsymbol{r}^*_{min}$ counterparts. The one exception to this, t6_17, is because an alternate viable path near the $\boldsymbol{r}^*_{max}$ X Stride trajectory and GPR errors that make this alternate path more credible via Equation 4.8.

Second, adding the altimeter measurement model either has no effect on accuracy (in four cases), improves accuracy (in three cases) or makes position estimation accuracy worse (t6_15). For test t6_15, adding the altimeter makes x RMSE significantly worse than using just the magnetic field measurement model. We believe that the VL53L1X low-cost altimeter has some timing issues that produces an irregular time delay from the ground-truth altitude as shown in Figure 4.11. Here, the zero-order hold of 4Hz the altimeter measurements (black dots) are shown along with the ground truth altitude (red line). Comparing the delay between a red, ground-truth datapoint and the nearest temporal point of the altimeter leads to inconsistent delays. Near $t = 32$s the delay is about 200ms and decreases to 10ms near $t = 39$s. The irregularity in the time delay made it difficult to account for this anomaly in pre-processing and ultimately effects our position estimates for the t6_15 ($\boldsymbol{r}^*_{min}$ Z Stride) trajectory more than the constant-altitude flights. We believe increasing the sensor sampling rate from 4Hz (which we use in this chapter) to something higher might improve the situation. However, since the focus of this chapter is on the relationship between magnetic field gradients and position localization, we leave such altimeter-based investigations for future work.

Finally, Table 4.6 appears to confirm our hypothesis that the smaller magnetic field gradient in the x direction makes our x RMSE metrics larger than y or z RMSE. Note that t6_09 ($\boldsymbol{r}^*_{min}$ Hover) is the one exception to this trend. We see from Figures 4.6a and 4.6d that $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right|$ and $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial y}\right|$ values at $\boldsymbol{r}^*_{min}$ = [-0.66, 1.25, -0.75]m are comparable at 2.30μT/m and 2.67μT/m respectively. We expect this is the reason x position error is not necessarily higher than y error since t6_09 hovers at this location for its entire flight. All other trajectories that move through $\boldsymbol{r}^*_{min}$ will regularly visit locations where $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial y}\right| > \left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right|$.

In the following section, we further analyze our first hypothesis (relationship between magnetic field gradient and position localization accuracy) by introducing a new metric.

4.5.3 Gradient Localization Metric as a Localizability Metric

In this section, we introduce a metric that might indicate the accuracy of position localization given only a magnetic field map of the area. The idea is to enable users to first survey their

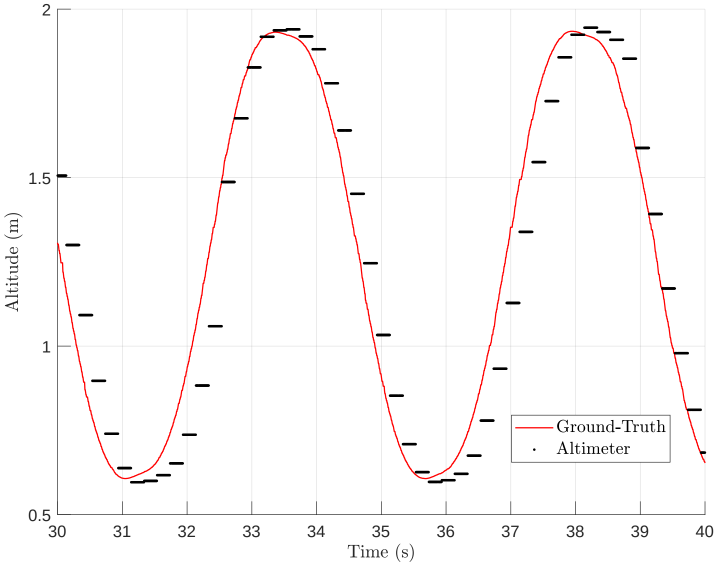


Figure 4.11: The low-cost VL53L1X altimeter has inconsistent timing delays when sampled at 4Hz. The delay near $t = 32$s is 200ms but decreases to 10ms near $t = 39$s.

workspace and conclude if it is worthwhile to add magnetometers to their sensor fusion method based on how much spatial variation is present in their workspace.

Our newly proposed quantity is called the gradient localization metric (GLM) defined as a ratio of two terms. The numerator or "signal"

$$signal_k = |\nabla \hat{\boldsymbol{m}}(\boldsymbol{r}_k)|\, \delta s \tag{4.13}$$

is computed at timestamp k, $|\nabla \hat{\boldsymbol{m}}(\boldsymbol{r})|$ is defined in Equation 4.5, and δs is the distance parameter from Equation 4.3. The goal is to have our numerator to be related to the magnetic field gradient, but for it to have units of μT. Here, we use $\delta s = 0.05$m.

The gradient term is normalized by a "noise" term taken as the magnitude of the GPR error at the location $\boldsymbol{r}_k$

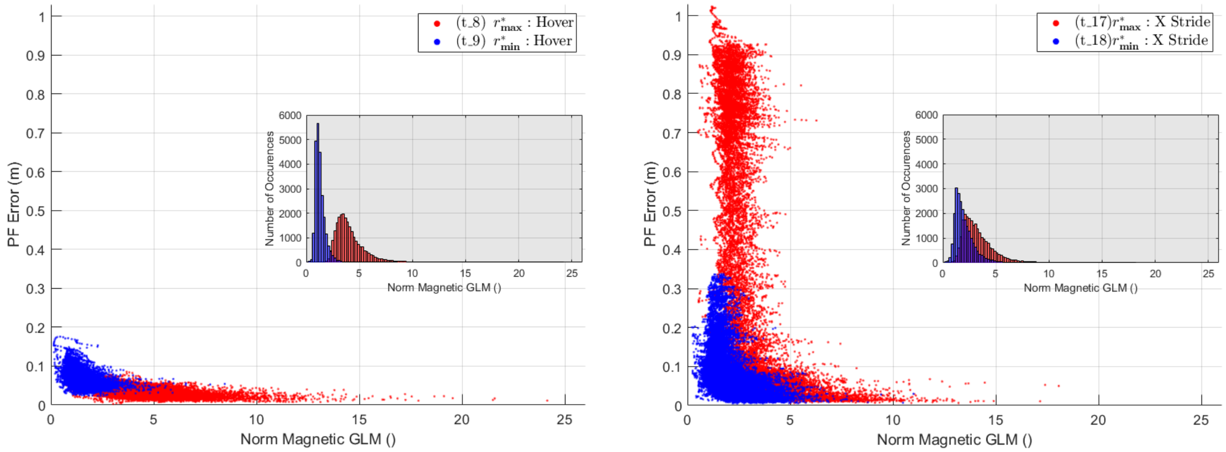
$$noise_k = |\Delta \boldsymbol{m}_k^{GT}| \tag{4.14}$$

where $\Delta \boldsymbol{m}_k^{GT}$ is defined in Equation 4.12. In essence, since our gradient terms are based on our GPR, we felt it important to account for any inaccuracies the GPR map introduces. Note that $\Delta \boldsymbol{m}_k^{GT}$ will include magnetometer sensor noise and any UAV-induced magnetic noise (Section 2.5.2.2).

Combining these two gives us a unitless, scalar quantity

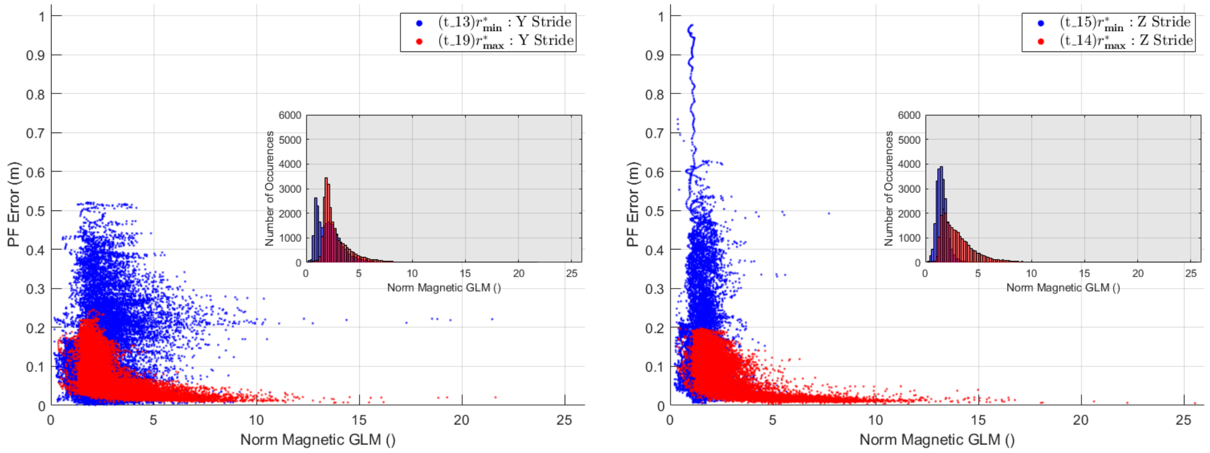
$$GLM_k = \frac{\mathbb{E}[signal_{k-4:k}]}{\mathbb{E}[noise_{k-4:k}]}. \tag{4.15}$$

Our noise term (i.e., GPR error) is inconsistently close to zero at some timestamps (see Figure B.1



(**a**) Hover trajectories.

(**b**) X Stride trajectories.



(**c**) Y Stride trajectories.

(**d**) Z Stride trajectories.

Figure 4.12: Particle filter error against our GLM. One point is plotted for each discrete timestamp of one repetition of the particle filter's estimate for a utilization flight. Generally, particle filter error decreases as GLM increases. The histograms help depict overlap between blue and red datapoints.

in Appendix B) which causes some GLM_k values to be arbitrarily large in a way that is not useful for our use here. As such, we use a moving average filter with window size 5 for both signal and noise terms to smooth out these spurious occurrences. Note, this window size is the same used in our pre-processing step of our map-generation procedure (Section 2.4.3).

Now, to understand the relationship between GLM and our ability to localize, we compare the PF's position estimation error to the GLM at each timestamp of a flight. As before, we only study the behavior of our PF in the region of interest (Table 4.5). In addition, it is difficult to conduct this analysis with the altimeter since it will, to some degree, make our PF estimate independent of $\left|\frac{\partial \hat{m}(r)}{\partial z}\right|$ in a way that is difficult to quantify. As such, this section uses only the magnetometer and entirely excludes the altimeter. Finally, to do this concisely, we isolate our study to one of the 20 repetitions of our PF per flight, but the results are similar across the 20 reps (for the target region of interest).

The result is shown in Figure 4.12 where the eight utilization flights are split into pairs for Hover, X Stride, Y Stride, and Z Stride trajectories respectively. In each subplot, the PF error and GLM_k for the corresponding $\boldsymbol{r}^*_{min}$ trajectory is shown in blue, while the $\boldsymbol{r}^*_{max}$ points are in red. Many of the points overlap and we struggled to find an appropriate transparency value to adequately represent this. As such, we include a histogram of the GLM for each pair to help the reader visualize how much overlap exists between the datasets. We hand-selected which dataset (red or blue) appeared in front to keep the visualization as "intuitive" as possible (i.e., the reader will likely be able to infer the trends of the occluded points with the given ordering and the histograms).

For the most part, the results of Figure 4.12 show a promising trend that higher GLM values correspond to lower PF error. Most of the $\boldsymbol{r}^*_{max}$ trajectories (red) have a max error near 0.2m. Additionally, their max errors are typically lower than their $\boldsymbol{r}^*_{min}$ counterparts. The exception here is t6_17 ($\boldsymbol{r}^*_{max}$ X Stride) with red dots in Figure 4.12b which is the flight discussed earlier that had an "alternate viable path" that confuses the particle filter when only the magnetometer is used (Figure 4.10).

It is difficult to visually compare the density of red or blue points in any particular region of these plots. The embedded histograms help to some extent, but visually combining information from both types of plots is non-trivial. Additionally, the histograms can only give insight on the horizontal spread of the points and tell nothing about the distribution along the vertical (PF error) axis. The RMSE values in Table 4.6 give a sense of the average PF error for each of the eight utilization flights. Note, however, that the data in Table 4.6 is aggregated over 20 repetitions for each utilization flight, while Figure 4.12 depicts only one of the 20 repetitions.

There are two problematic aspects of our GLM in regard to predicting the PF localization accuracy. First, and probably most importantly, our PF error is a function of our low-cost sensor set, selected process and measurement models, and our particle filter's gains. Our analysis does not enable us to disentangle how much of our results are a function of our implementation (which may have, inadvertently, been overfit to this particular analysis). Similarly, we do not know if there are more clear trends between GLM and PF error that appear with better dead-reckoning (using just a low-cost accelerometer for a process model is uncommon and typically not expected to give good results) or higher-quality sensors. Finally, the gains in our particle filter were hand-tuned and held constant across all eight utilization flights. Since we cannot claim our gains are optimal in any regard, it is possible there are some gains that will yield a clearer trend. It's also important to note that our results inherit any problems from our inability to completely remove magnetic-field noise from our UAV (Figure B.1 in Appendix B).

Second, Figure 4.12 shows that there are no clear minima or maxima at any given GLM value across our utilization flights. The large error values in the red dots of Figure 4.12b are due to the alternate viable path on t6_17 (Figure 4.10). Similarly, the thread of high-error points in the blue dots of Figure 4.12d are due to the GPR map's inconsistencies towards the end of t6_15 (Figure 4.8a). A good localization metric should give some bound on error from above or below, but it seems ours cannot do that with the data presented (keeping in mind the caveats listed in the previous paragraph). In short, these low-GLM-high-error cases make it difficult to draw a conclusion when juxtaposed to t6_09 ($\boldsymbol{r}^*_{min}$ Hover) and t6_18 ($\boldsymbol{r}^*_{min}$ X Stride) that also have low GLM but have much lower max PF error (low-GLM-low-error cases).

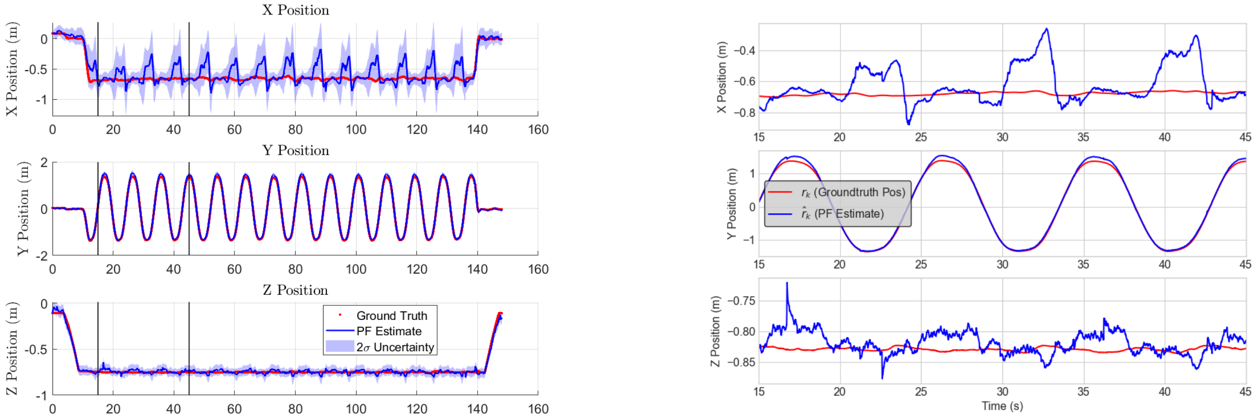
We take a moment to discuss t6_13 ($\boldsymbol{r}^*_{min}$ Y Stride) with blue dots in Figure 4.12c since its error profile seems insensitive to GLM. The main anomaly here is the presence of points at 0.21m PF error across a wide range of GLM values. To analyze further, it is best to discuss with the plots in Figure 4.13. Figure 4.13a shows the full t6_13 flight data (in red) along with one of the 20 repetitions of the PF (in blue). Figure 4.13b highlights three Y Stride periods to show that the x estimation error correlates well with particular segments of the flight trajectory. Here, we also see that there is some y estimation error at the +y side of the trajectory.

The reason we see a PF error of 0.21m for many GLM values in Figure 4.12c is because the high x Position error in 4.13b (e.g., t=20s to t=25s) is maintained across many y positions. This means the PF error remains around 0.2m while the magnetic field gradient changes as a function of y position (Figure 4.6j).

Unlike t6_17 (Figure 4.10), the reason for the t6_13's x position error is not due to the same *type* of alternate viable path. Figure 4.13c shows that $|\Delta \boldsymbol{m}_k^{GT}|$ and $|\Delta \boldsymbol{m}_k^{PF}|$ have comparable values when the PF's x position error is high. We think the higher x position estimation accuracy here is because the x spatial gradient ($\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right|$) is very low along portions of the $\boldsymbol{r}^*_{min}$ Y Stride trajectory; effectively the horizontal line at x = -0.66 in Figure 4.6a. On this horizontal line, $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right|$ ranges from $0.85 \mu T/m$ on the -y (left) extreme to $2.31 \mu T/m$ on the +y (right) side. This correlates well with Figure 4.13b where the highest x position error is on the -y portion of the trajectory (with low $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right|$) and the lowest x error is on the +y side.

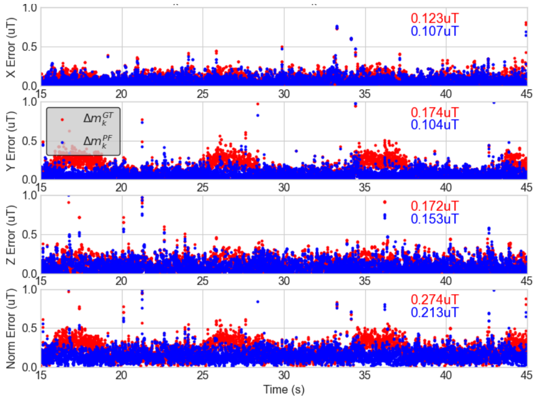
In short, the PF is better able to resolve the x position of the drone when in a region with higher $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right|$. However, the y position error on the +y portion of the trajectory is best explained as an alternate viable path (Figure 4.13c at t=26s) since $|\boldsymbol{m}_k^{PF}|$ is clearly lower than $|\Delta \boldsymbol{m}_k^{GT}|$.

Ultimately, we see that that increasing GLM *generally* decreases the mean and variance of the PF's error. Though this is well-supported by our data, it is not that surprising since it matches our intuition on the relationship between spatial gradient and position localization accuracy. And since our metric fails to quantitatively bound (from above or below) the amount of error we should expect at various GLM values, our results mostly serve to confirm our intuition. Figure 4.12 does show that higher GLM correlates well with low PF error, but the histograms show that GLM values >5 are uncommon for our workspace. Thus, we do not feel there are enough datapoints with GLM



(**a**) t6_13: Full time series (only magnetometer).

(**b**) t6_13: Zoomed to t=[15, 45]s.



(**c**) t6_13: Comparable $\Delta \boldsymbol{m}_k^{GT}$ and $\Delta \boldsymbol{m}_k^{PF}$ during high X position error.

Figure 4.13: t6_13 ($\boldsymbol{r}^*_{min}$ Y Stride) position estimation errors caused both by low $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right|$ and by errors in GPR map. In Figure 4.13c, the max of all outliers larger than $1\mu T$ are $1.90\mu T$, $1.76\mu T$, $2.61\mu T$, and $2.69\mu T$ for X, Y, Z, and norm errors respectively (across either GT or PF datasets).

>5 to make any solid conclusions. In future work, we may artificially increase the magnetic field gradient in our workspace by adding permanent magnets near desired trajectories to increase the number of high-GLM points.

One notable thing is that our lower-gradient-higher-error hypothesis seems well supported on both a macro level (x RMSE values are generally higher than y or z RMSE in Table 4.6) and a micro level (higher x position error during lower $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right|$ in Figure 4.13b). There is certainly promise in what we have uncovered, but improvements are needed if we want a useful and predictive localizability metric.

4.5.3.1 How Averaging can Remove GLM trends

Our initial attempt at constructing a GLM involved averaged over the entire flight by way of the following equation

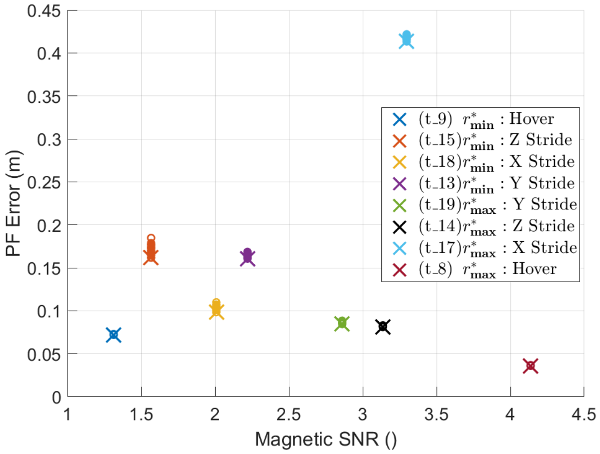


Figure 4.14: PF accuracy against GLM$_{avg}$ metric. PF accuracy computed over region of interest (Table 4.5).

$$GLM_{avg} = \frac{1}{n_v}\sum_{i=1}^{n_v}\left(\frac{signal_i}{noise_i}\right) \tag{4.16}$$

where n_v is the number of individual datapoints in the utilization flight. We thought this averaging was a good way to capture the various gradient values seen by the non-stationary trajectories.

Figure 4.14 shows a scatter plot of all 20 repetitions of the PF for each of the eight utilization flights. The horizontal axis is our GLM$_{avg}$ metric averaged over 2Hz downsampled observations from each respective utilization flight and vertical axis is the norm RMSE of the PF (which itself is an averaged metric of the PF's performance). Each of the 20 repetitions are plotted as a circle for the respective utilization flight and an × is redundantly plotted over the repetition with the *lowest* norm RMSE. Figure 4.14 uses only the magnetometer for position estimation. The legend lists each utilization flight in their order of appearance in the plot (i.e., in increasing order of GLM$_{avg}$).

The PF norm error values depicted here are computed over the "region of interest" as defined in Table 4.5 (i.e., they exclude takeoff and landing sequences as explained in Section 4.5.2.2). Thus, for each utilization flight, the average of the 20 circles in Figure 4.14 is the same as the norm RMSE quantities listed in Table 4.6.

From Figure 4.14, we see that the four $\boldsymbol{r}^*_{min}$ trajectories all have lower GLM$_{avg}$ values than the four $\boldsymbol{r}^*_{max}$ flights. As we expect, $\boldsymbol{r}^*_{min}$ Hover and $\boldsymbol{r}^*_{max}$ Hover have the lowest and highest GLM$_{avg}$ of the eight utilization flights since they both loiter at the locations with lowest and highest gradients.

Figure 4.14 shows that our particle filter can estimate position within 0.2m for seven of the eight utilization flights (within each respective region of interest from Table 4.5). The primary issue with this visual representation is that it seems to imply that PF localization accuracy is fairly insensitive to our GLM$_{avg}$ metric. From Figure 4.14 we see very low PF RMSE at both ends of our GLM$_{avg}$ span (for the two hover trajectories) while the values in the center are not necessarily decreasing as

GLM$_{avg}$ increases. For this, GLM$_{avg}$ does not achieve its intended goal of serving as a localizability metric to predict how accurately one could estimate position given just a magnetic field map of a workspace. We think there are two reasons worth considering here.

First, which is the focus of this section, is to emphasize issues caused by averaging GLM and PF RMSE over an entire flight. Though this may be appropriate for our stationary (hover) trajectories, it can blur together small-timescale trends like in Figure 4.13b where we showed that x position error is higher when $\left|\frac{\partial \hat{m}(\boldsymbol{r})}{\partial x}\right|$ is lower. Also note that both our GLM (Equation 4.15) and GLM$_{avg}$ (Equation 4.16) metrics aggregate all three magnetic gradients (for $\delta\boldsymbol{x}$, $\delta\boldsymbol{y}$, and $\delta\boldsymbol{z}$) into a single metric. Similarly, the norm PF error we report combines the three PF error metrics (x, y, and z). The net effect could simply be that we have removed too much spatial (x, y, z PF error), directional (x, y, z components of mag gradient), and temporal (averaging over the flight) information in the GLM$_{avg}$ plot for there to be a discernible trend.

Additionally, averaging over an entire flight effectively gives us fewer points to compare against. In Figure 4.12, a single PF repetition gave 20,000+ points for each utilization flight. With the averaging approach in Figure 4.14, each utilization gives at most 20 points (since we report 20 reps of the PF's performance for each utilization flight), but most repetitions are redundant. This leads to our next consideration.

It is possible that there *is* a trend in the GLM$_{avg}$ plot of Figure 4.14, but it could be quite noisy (like those in Figure 4.12). If this were the case, eight averaged utilization flights might not provide enough information to recover the overall trend. Imagine picking eight red points from Figure 4.12b that can represent the trend we see there. This idea could be tested by performing more flights at various GLM$_{avg}$ values throughout the workspace. If there is a noisy trend in the GLM$_{avg}$ metric, one obscuring factor could be our flight-by-flight variations (Section 2.5.2.2) which we could investigate with more iterations of the same eight utilization flights. In addition, some outdoor flights (which should have lower GLM$_{avg}$) would be informative to investigate how the trend extends to values of GLM$_{avg}$ < 1. Alternatively, simulating a UAV moving through an artificial magnetic field with particular GLM$_{avg}$ values could improve our understanding and allow for specific GLM$_{avg}$ values.

4.6 Discussion

This chapter investigated methods to use magnetic field maps for indoor state estimation. The primary analysis was done with post-processing to support benchmarking and statistical characterization of results.

Ideally, we want design tools that allow us to bound the accuracy of our position estimates prior to conducting flight tests. The proposed gradient localization metric (GLM) can characterize position estimation accuracy given only the ambient magnetic field of a workspace. Additionally, our spatial gradient analysis can be used in conjunction with additional sensing modalities that might be supplemented with magnetic field data. For example, given knowledge of how small $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial x}\right|$ is relative to $\left|\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{r})}{\partial z}\right|$, in future work one might mount the VL53L1X time of flight sensor in the x direction rather than in z.

In follow-on work, it would be useful to employ real-time analysis on whether the PF should trust or discredit magnetometer measurements. From Table 4.6, we see that adding the altimeter improves PF performance on t6_17 ($\boldsymbol{r}^*_{max}$ X Stride) but degrades performance on t6_15 ($\boldsymbol{r}^*_{min}$ Z Stride). Having a real-time indicative metric that flagged the altimeter as the more credible source on t6_17 and the magnetometer as the more reliable measurement for t6_15 would help the PF get the best out of both sensing modalities. Related works in this topic [44, 45, 46, 47, 48] propose adjusting the covariance terms of each sensing modality (e.g., $\boldsymbol{R}_{mag}$ and σ^2_{alt}) in real time, showing such adjustments can yield better attitude estimates than using fixed covariance parameters.

Testing such a method for our PF would require alternative sensing modalities for x and y position estimates. This could be achieved by adding a camera or LIDAR sensor to give more spatial information in x, y, and z. More simply, one could add VL53L1X laser range finders to measure the distance to nearby walls in the x and y directions. In essence, developing a real-time indicative metric that trades between different sensing modalities would require UAV upgrades beyond the scope of this work.

In this chapter, we rely on ground-truth attitude estimates to rotate sensor measurements into a common world frame. This allows us to better investigate the relationship between magnetic field gradients and position estimation without needing to account for errors in an attitude estimator. In future work, we would like to either construct a full pose state estimator (building on our attitude estimation work from Chapter 3) or study how using fewer output dimensions on our GPR map (i.e., $3 \to 2$ and $3 \to 1$ maps) effect our overall position estimates. Such maps would reduce our PF's reliance on ground-truth or alternative attitude estimates but are likely to create more spatial ambiguities in the magnetic field and produce more viable path options.

In closing, we present preliminary results on using a $3 \to 1$ map for 3D position estimates in Table 4.7. Such a map would need *no* knowledge of the UAV's attitude to be used but clearly has higher errors than when using a $3 \to 3$ map (Table 4.6). Further work is required to evaluate whether higher RMSE terms are due exclusively to the presence of alternative viable paths or if other factors are involved.

Table 4.7: Truncated position RMSE for 3 → 1 map (rotationally invariant). Preliminary results as motivation for future work.

RMSE **(m)** ***Norm—X—Y—Z***	**Inertial: Mag**				**Inertial: Mag + Alt**			
(t6_09) $\boldsymbol{r}^*_{min}$: Hover	**0.78**	0.33	0.41	0.57	**0.28**	0.27	0.06	0.01
(t6_08) $\boldsymbol{r}^*_{max}$: Hover	2.56	2.54	0.28	0.15	3.19	3.17	0.34	0.06
(t6_18) $\boldsymbol{r}^*_{min}$: X Stride	1.32	0.68	0.52	1.00	**0.56**	0.55	0.11	0.02
(t6_17) $\boldsymbol{r}^*_{max}$: X Stride	**0.96**	0.85	0.19	0.41	1.34	1.32	0.17	0.05
(t6_13) $\boldsymbol{r}^*_{min}$: Y Stride	1.51	1.38	0.22	0.59	**0.96**	0.96	0.09	0.03
(t6_19) $\boldsymbol{r}^*_{max}$: Y Stride	**1.06**	1.04	0.14	0.13	1.27	1.26	0.14	0.05
(t6_15) $\boldsymbol{r}^*_{min}$: Z Stride	**0.55**	0.42	0.25	0.26	**0.84**	0.80	0.25	0.12
(t6_14) $\boldsymbol{r}^*_{max}$: Z Stride	1.30	1.25	0.16	0.32	0.97	0.95	0.13	0.12

4.7 Conclusions and Future Work

This chapter analyzed the relationship between the gradient of the ambient magnetic field and the accuracy of position localization. We showed that the magnetic field gradient in the x direction is much lower than y or z directions in our workspace which corresponds to lower x position estimation accuracy. This relationship (lower gradient, higher position estimation error) is demonstrated in complete flight tests (Table 4.6) and for particular time intervals within a single flight (Figure 4.13). Overall, we find our particle filter (PF) can estimate the position of our UAV within 0.2m norm RMSE for seven of our eight utilization trajectories (Table 4.6). Note that we achieve accuracy within 0.2m for six of the eight trajectories when including takeoff and landing sequences (Table 4.4).

Next, we showed how spatial ambiguities in the ambient magnetic field allow for multiple plausible hypotheses of the UAV's position when using only magnetic field for navigation. These ambiguities, combined with small errors in the magnetic field map, create errors in the particle filter's position estimates. In addition, we are able to use our map to determine if the PF's errors are due to low magnetic field gradient or errors in the map (Figure 4.13).

Finally, we propose two "localizability" metrics meant to predict how accurately one can estimate position given just information about the ambient magnetic field. We show that our metrics have promise but require more development and analysis to be useful in predicting position estimation accuracy.

In the future, we would like to test our methods in different locations. A goal would be to find indoor spaces with $GLM > 5$ (Figure 4.12) and others with $GLM_{avg} < 1$ (Figure 4.14) that allow us to further investigate our localizability metrics. In addition, we want to reduce the amount of high-frequency noise and the flight-by-flight variations (Figure B.1 in Appendix B) that may be obscuring trends in our localizability metrics.

This analysis would lend itself well to simulation. Though the opening of this dissertation asserted there are not best practices on how to simulate indoor magnetic fields, we at least have upper and lower bounds on how much spatial variation to simulate. Simulation can be used to further study our localizability metrics at arbitrary GLM values and also disentangle our analysis from the flight-by-flight variations that create errors in our GPR map. In simulation, we can begin with a perfect map and a magnetically quiet drone, then isolate how gradients in the ambient magnetic field relate to position localization accuracy.

CHAPTER 5

Conclusion and Future Work

This dissertation has developed and analyzed methods to extract and utilize information about the magnetic fields measured by low-cost magnetometers. A primary goal was to show how magnetometer information can be leveraged to improve indoor navigation.

To accomplish this, we developed methods of creating accurate indoor magnetic field maps and demonstrated their value by estimating the position and attitude of a UAV. Our techniques were designed with practical implementation in mind and were experimentally validated on a UAV with real sensor data. We hope our approach, methods, and insights are valuable to those looking to implement our techniques.

Before our work, there was not a clear connection between two paradigms of utilizing indoor magnetic field measurements. Those wishing to estimate attitude indoors saw Earth's magnetic field as the true signal and distortions caused by materials in buildings as a nuisance. Meanwhile, those aiming to estimate position relied on the spatial variation in the magnetic field to track motion and distinguish one location from another. In this dissertation, we showed how local magnetic field maps are the bridge between these two paradigms that allow position estimation techniques to leverage the spatial variation in the field and simultaneously provide reference vectors for an attitude estimator.

In our work, magnetic field measurements serve as an information supplement not a full replacement for other indoor navigation sensing systems. We have shown how magnetometer data can be processed to map an indoor environment and in turn contribute navigation information. The sections below summarize our primary conclusions and present ideas for future work.

5.1 Dissertation Conclusions

Chapter 2 showed how to use GPR (Gaussian process regression) to map the magnetic distortions of indoor environments. Using a UAV made gathering observations faster and more consistent but resulted in undesired magnetic disturbances. We showed how to reduce UAV disturbances and how to teach the GPR-based map about these disturbances to improve indoor pose estimation accuracy.

Chapter 3 showed how to estimate 3DOF attitude of a UAV using GPR-based magnetic field maps and a multiplicative extended Kalman filter (MEKF). Here, we found that magnetic field maps can improve attitude estimation by a factor of two in regions of high magnetic field spatial variation. However, in environments with lower spatial variation in the magnetic field, our GPR-based maps were not required since a constant-field assumption is more appropriate.

Finally, in Chapter 4, we used a particle filter to estimate the 3D position of a UAV using an accelerometer, altimeter, and magnetometer. Using ground-truth attitude estimates to isolate the accuracy of position estimation, we were able to localize the UAV's position within 0.2m for six of our eight test cases. For reference, these tests were done within a 4m×3m×2.25m test arena. Additionally, we showed there is a correlation between the magnetic field gradient and the accuracy of our position estimates. Essentially, the gradient of the magnetic field can serve as an indicator of how accurately position estimates *could* be in a mapped region.

5.2 Future Work

Below are recommendations for future work motivated by this dissertation.

5.2.1 Investigation of UAV-Induced Magnetic Biases

As presented Section 2.5.2.2, our UAV injects time-varying magnetic biases into the measured magnetic field. Recall that we reduced the impact of these biases by distancing our magnetometer from the electronics and ultimately presented our compromise map as a way to incorporate the flight-by-flight variations into the map directly.

However, it would be useful to understand which components are the dominant actors in these time-varying biases. One way to investigate this is to constrain the UAV in a way that allows the motor and propellers to spin without the UAV moving. Next, one could measure the magnetic field near particular electronics (motors, electronic speed controllers, microcontroller, etc...) when the UAV is powered off and when it is operating at different throttles.

One complicating factor is that the time-varying biases we show in Section 2.5.2.2 did not appear to correlate with certain throttle settings or particular portions of the UAV's trajectory. For this, it might be worth investigating the voltage regulation on the BeagleBone Blue: specifically on its output 3.3V line the RM3100 uses as a reference voltage for its output measurements.

5.2.2 Simulated Magnetic Fields

This dissertation relied entirely on actual magnetometer measurements of two different environments. Though this offered insight on what we can expect from indoor and outdoor magnetic fields, it brought along a set of problems primarily caused by magnetic noise from our UAV.

Our goal with Chapters 3 and 4 was not to claim we have the best attitude or position estimators, but instead to analyze trends on how spatial variation of the magnetic field changes our resultant pose estimates. Throughout these two chapters, our investigations were impacted by the UAV-induced noise or lack of data in regions with specific amounts of spatial variation.

Simulations offer promise to investigate both. Here, we could study how discrepancies such as sensor noise, UAV-induced noise, or unmapped, mobile ferromagnetic objects effect our pose estimates. Additionally, we can simulate ambient magnetic fields with specific magnetic field gradients that allow us to understand how spatial variation relates to the value of maps for attitude estimation (Chapter 3) and our ability to estimate position (Chapter 4).

5.2.3 Full 6DOF pose estimation using indoor magnetic field maps.

Our work always focused on estimating three degrees of freedom at a time and used motion capture for the other degrees of freedom. In the future, we would like to implement full 6DOF pose estimation and take steps towards having it run in real-time on our UAV. Testing pose estimation accuracy with different output dimensions m (for a $p \to m$ map) is a good way to incrementally reduce the impact of ground-truth attitude on position estimates. Additionally, we would like to revisit 3DOF attitude estimation (Chapter 3) with our new mapping and calibration tools to see if our estimates can be made better by simply improving the magnetic field map.

5.2.4 Sensor fusion with other sensing modalities

The central goal of this dissertation is to demonstrate how to extract value out of the ambient magnetic field to improve position and attitude estimates. However, we do not intend to express that IMUs and magnetic field maps *alone* should be used for navigation. In the future, we would like to fuse IMU data with magnetic field maps and other sensing modality information (e.g., vision, LIDAR, Wi-Fi fingerprinting).

For this, it would be valuable to have real-time heuristics/metrics that indicate which sensing modality is most reliable at any given location and time. This way, the state estimator can ignore certain sensing modalities if their measurements become less useful (i.e., camera if lighting changes significantly or magnetometer if near an elevator). For the magnetic field, some of this burden can be shared by the magnetic map itself (return high uncertainty of the predicted magnetic field in areas where the magnetic field tends to change regularly).

5.2.5 Alternating Current (AC) Magnetic Field Mapping and Localization.

Our work has focused entirely on the direct current (DC) magnetic field measured by a RM3100 magnetometer. However, by performing a frequency analysis on our measurements, it is possible to create spatial maps of the magnetic fields at higher frequencies. Some works [39] do this using a UAV and show that the spatial variation in the DC magnetic field is not the same as higher-frequency modes of the magnetic field. As such, it might be possible to leverage both the DC and AC fields in conjunction for position localization. Effectively, we could get additional information from the same sensor if it is sampled sufficiently fast. For example, power grids typically run at 50 or 60Hz. If the sampling rate of the onboard magnetometer was sufficiently high to measure these AC magnetic noise sources, the location and orientation of power grid signals could be used for pose estimation.

5.2.6 Robust utilization of magnetic field maps

This dissertation focused entirely on time-invariant magnetic field mapping, however there are many mobile objects that can distort the measured field. There are three main sources that create the DC magnetic field inside buildings. Earth and the non-moving structural components in the building will change slowly over time or not at all. The third component deals with the motion of doors, elevators, and some furniture. It is important to consider the impact of such mobile objects in a workspace and how the map can either update its estimates based on the state of some objects (e.g., elevator gives its real-time position to the magnetic field map allowing the map to update its estimates near the elevator) or simply return a higher uncertainty near regularly moving objects.

Additionally, it is important to consider that most users will not want to (or be able to) calibrate their magnetometers in the same way as those who created a building's magnetic field map. As such, it is important to consider how indoor navigation with magnetic fields can be done when magnetometers may output different measurements given the same signal.

In addition, some phone cases and accessories may further distort the measured magnetic field. For example, some cell phone cases contain a permanent magnet to attach the phone to metallic objects and take group pictures with ease. Such phone cases would clearly present a challenge to localizing within a magnetic field map, especially if the attached permanent magnet is saturating the onboard magnetometer. There are many works like [28] on crowdsourcing indoor magnetic field maps that are thinking about this sort of problem and developing methods to deal with the discrepancies across several magnetic sensors.

Finally, it is important to consider how to store and share magnetic field maps of rooms and buildings. The compromise map method we present here could be promising since the hyperparameters can be shared along with the fixed set of n_1 user-selected locations in the room/building of interest. Here, it might be prudent to learn lessons on how to store and share maps from spatial mapping works that use cameras and LiDAR.

APPENDIX A

Flight Test Name Reference

The, rather obtuse, naming convention for flight tests in this dissertation is meant to serve the reader in associating our results with the relevant data in our dataset. Recall that each flight test is denoted as "tY_XX" where *Y* refers to the flight test *series* and *XX* is the two-digit ID of the flight test in that series. The Table A.1 associates each series numbers *Y* used in this dissertation with their respective directories in our dataset.

Note, Chapter 3 was completed and published in Ref. [23] before we created this global naming convention for our flight tests. The version of Chapter 3 presented in this dissertation is largely identical to Ref. [23] and thus does not have the same numbering system. Nonetheless, we include their flight data as series t7, t8, and t9 in this table for reference.

Table A.1: Data used in this dissertation.

Test Series (tY)	**Dataset Directory**	**Use In This Dissertation**
t1	20220617_batteryAndMagTesting	$S_2 = 2\text{cm}$
t2	20220622_batteryAndMagPartTwo	$S_2 = 4\text{cm}$
t3	20220627_verifyingSufficientMagDist	$S_2 = 6\text{cm}$
t4	20220707_8cm_magAndBatt	$S_2 = 8\text{cm}$
t5	20220503_maggieAndEstibonFlights	'noisy' versus 'quiet' PID controller gains
t6	20220901_researchData	Intermediate and compromise maps. Consistency metric.
t7	20200817_night_flight	Outdoor attitude estimation flights for Chapter 3.
t8	20210318_flight_test	Outdoor attitude estimation flights for Chapter 3.
t9	20210423_mag_cali	Magnetometer calibration data used for t8 flight series in Chapter 3.

APPENDIX B

GPR Map Consistency with Utilization Flights

This section shows the GPR map's error for all eight utilization flights. The plots here depict the difference between the actual measured magnetic field $\tilde{\boldsymbol{y}}_k$ and the GPR's predicted magnetic field $\mathcal{GPR}(\boldsymbol{r}_k)$ at the ground-truth UAV position $\boldsymbol{r}_k$ for some timestamp k.

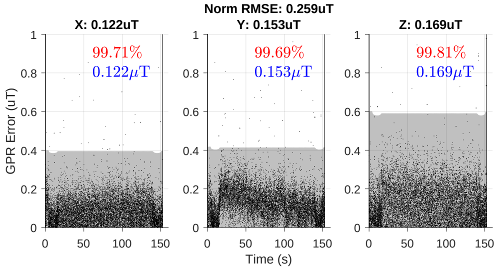
Figure B.1 depicts the maps' accuracy for all 200Hz of magnetometer measurements from each utilization flight. In each plot, the black dots are $|\tilde{\boldsymbol{y}}_k - GPR(\boldsymbol{r}_k)|$ for all timestamps k, the gray shading is two standard deviations (2σ) of each respective GPR's uncertainty, the red percentages are the fraction of times error falls within the 2σ uncertainty, and the blue text is the RMSE for each GPR.

Here, we see that all utilization flights are *consistent* with the GPR map because each X, Y, and Z GPR captures its error within two standard deviations of its uncertainty at least 96% of the time respectively. t6_15 ($\boldsymbol{r}^*_{min}$ Z Stride) is an exception and fails the consistency check (see red percentages).

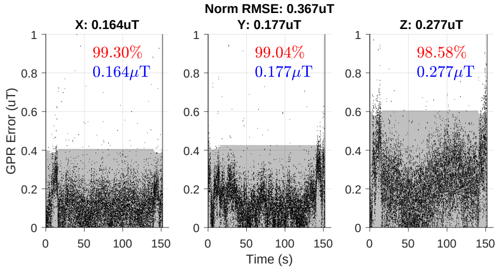
As shown in Figure B.1g, t6_15 ($\boldsymbol{r}^*_{min}$ Z Stride) has a sudden increase in X, Y, and Z GPR error starting around $t =$130s. Nothing of note occurred at this moment in t6_15 so we believe this is simply another case of a change in the drone-induced magnetic noise discussed in Section 2.5.2.2. t6_18 (Figure B.1c) has a similar shift in the Z-component of its magnetic field around $t =$125s.

In Section 4.5.2.1, we truncate t6_15 and t6_18 before their landing sequences because of a sudden change in the bias of the measured or predicted magnetic field. The justification for this can be seen in Figures B.1g and B.1c. As mentioned in Section 2.5.6, applying the consistency check across an entire flight can be misleading and it can be useful to isolate the analysis to particular portions. The truncation for t6_15 and t6_18 were chosen based on the approximate timestamps where each respective GPR_z failed to capture 96% of its error within two standard deviations of its uncertainty.

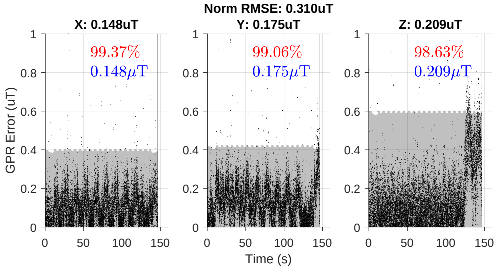
When gathering the flight data for Chapter 4, we checked to ensure each utilization flight was consistent with the compromise map. In fact, t6_16 ($\boldsymbol{r}^*_{max}$ Hover) was added to the training set specifically to make t6_08 (another $\boldsymbol{r}^*_{max}$ Hover) consistent with our map. For t6_15, we noticed the change in the magnetic field measurement at the end of the flight, but since GPR_z captured the error 95% of the time, we chose not to fly the trajectory again thinking it was close enough to our 96% threshold. We did not notice the change in measured magnetic field on t6_18 during flight time since it passed the consistency criterion for all three GPRs.



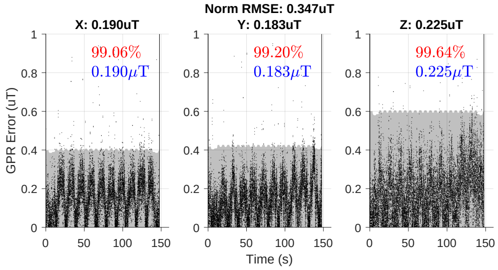
(**a**) t6_09 ($\boldsymbol{r}^*_{min}$ Hover). Max of outliers: $\{2.79, 2.25, 2.61\}\mu$T.



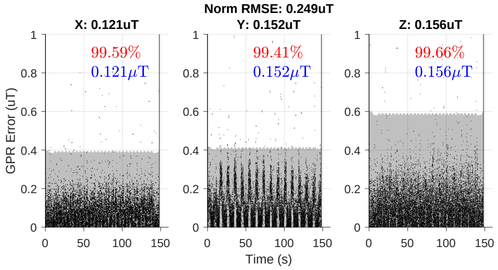
(**b**) t6_08 ($\boldsymbol{r}^*_{max}$ Hover). Max of outliers: $\{2.36, 2.37, 1.64\}\mu$T.



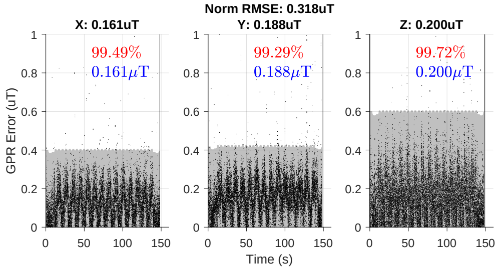
(**c**) t6_18 ($\boldsymbol{r}^*_{min}$ X Stride). Max of outliers: $\{2.54, 2.06, 1.77\}\mu$T.



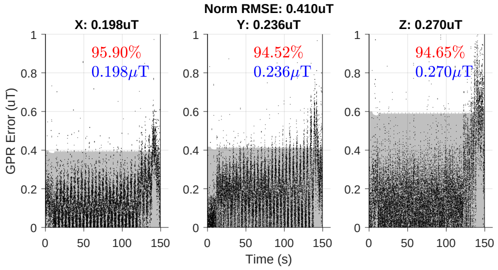
(**d**) t6_17 ($\boldsymbol{r}^*_{max}$ X Stride). Max of outliers: $\{2.69, 2.10, 3.15\}\mu$T.



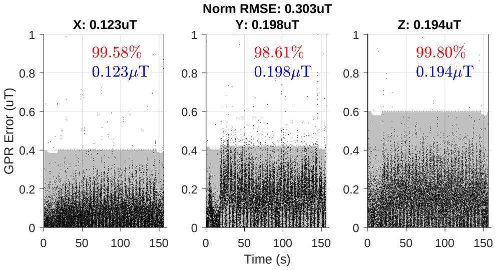
(**e**) t6_13 ($\boldsymbol{r}^*_{min}$ Y Stride). Max of outliers: $\{2.48, 2.41, 2.42\}\mu$T.



(**f**) t6_19 ($\boldsymbol{r}^*_{max}$ Y Stride). Max of outliers: $\{2.61, 2.75, 2.71\}\mu$T.



(**g**) t6_15 ($\boldsymbol{r}^*_{min}$ Z Stride). Max of outliers: $\{2.56, 2.35, 2.30\}\mu$T.



(**h**) t6_14 ($\boldsymbol{r}^*_{max}$ Z Stride). Max of outliers: $\{2.45, 2.52, 2.58\}\mu$T.

Figure B.1: Accuracy of the GPR map against the eight utilization flights. Limiting the vertical axis to 1μT maximum hides many outliers. The *max* error of these outliers is listed as $\{\ \mathcal{GPR}_x,\ \mathcal{GPR}_y,\ \mathcal{GPR}_z\ \}$ in the caption of each subplot.

BIBLIOGRAPHY

[1] D. J. Yeong, G. Velasco-Hernandez, J. Barry, and J. Walsh, "Sensor and Sensor Fusion Technology in Autonomous Vehicles: A Review," *Sensors*, vol. 21, p. 2140, mar 2021.

[2] C. Tyrén, "Magnetic Anomalies as a Reference for Ground-speed and Map-matching Navigation," *Journal of Navigation*, vol. 35, pp. 242–254, may 1982.

[3] J. B. Phillips, M. J. Freake, J. H. Fischer, and C. S. Borland, "Behavioral titration of a magnetic map coordinate," *Journal of Comparative Physiology A*, vol. 188, pp. 157–160, mar 2002.

[4] H. E. Soken and S. ichiro Sakai, "Attitude estimation and magnetometer calibration using reconfigurable TRIAD+filtering approach," *Aerospace Science and Technology*, 2020.

[5] M. Maaref, J. Khalife, and Z. M. Kassas, "Aerial Vehicle Protection Level Reduction by Fusing GNSS and Terrestrial Signals of Opportunity," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, pp. 5976–5993, sep 2021.

[6] Z. Xu, M. Haroutunian, A. J. Murphy, J. Neasham, and R. Norman, "An Underwater Visual Navigation Method Based on Multiple ArUco Markers," *Journal of Marine Science and Engineering*, vol. 9, p. 1432, dec 2021.

[7] V. Usenko, J. Engel, J. Stuckler, and D. Cremers, "Direct visual-inertial odometry with stereo cameras," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1885–1892, IEEE, may 2016.

[8] R. Hartley, M. Ghaffari, R. M. Eustice, and J. W. Grizzle, "Contact-aided invariant extended Kalman filtering for robot state estimation," *The International Journal of Robotics Research*, vol. 39, pp. 402–430, mar 2020.

[9] F. Zafari, A. Gkelias, and K. K. Leung, "A Survey of Indoor Localization Systems and Technologies," *IEEE Communications Surveys 'I&' Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.

[10] A. Solin, M. Kok, N. Wahlström, T. B. Schön, and S. Särkkä, "Modeling and interpolation of the ambient magnetic field by Gaussian processes," *IEEE Transactions on Robotics*, vol. 34, pp. 1112–1127, sep 2015.

[11] N. G. M. Team and B. G. Survey, "World magnetic model 2020," 2020.

[12] B. Li, T. Gallagher, A. G. Dempster, and C. Rizos, “How feasible is the use of magnetic field alone for indoor positioning?,” in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 1–9, IEEE, nov 2012.

[13] I. Vallivaara, J. Haverinen, A. Kemppainen, and J. Roning, “Magnetic field-based SLAM method for solving the localization problem in mobile robot floor-cleaning task,” in *2011 15th International Conference on Advanced Robotics (ICAR)*, pp. 198–203, IEEE, jun 2011.

[14] I. Vallivaara, J. Haverinen, A. Kemppainen, and J. Roning, “Simultaneous localization and mapping using ambient magnetic field,” in *2010 IEEE Conference on Multisensor Fusion and Integration*, pp. 14–19, IEEE, sep 2010.

[15] Z. Wu, M. Wen, G. Peng, X. Tang, and D. Wang, “Magnetic-Assisted Initialization for Infrastructure-free Mobile Robot Localization,” in *2019 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM)*, pp. 518–523, IEEE, nov 2019.

[16] B. Brzozowski, K. Kazmierczak, Z. Rochala, M. Wojda, and K. Wojtowicz, “A concept of UAV indoor navigation system based on magnetic field measurements,” *3rd IEEE International Workshop on Metrology for Aerospace, MetroAeroSpace 2016 - Proceedings*, pp. 636–640, 2016.

[17] B. Brzozowski and K. Kazmierczak, “Magnetic field mapping as a support for UAV indoor navigation system,” in *2017 IEEE International Workshop on Metrology for AeroSpace (MetroAeroSpace)*, pp. 583–588, IEEE, jun 2017.

[18] M. Kok and A. Solin, “Scalable Magnetic Field SLAM in 3D Using Gaussian Process Maps,” *2018 21st International Conference on Information Fusion (FUSION)*, pp. 1353–1360, apr 2018.

[19] N. Akai and K. Ozaki, “3D magnetic field mapping in large-scale indoor environment using measurement robot and Gaussian processes,” in *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, vol. 2017-Janua, pp. 1–7, IEEE, sep 2017.

[20] D. Hanley, A. S. De Oliveira, X. Zhang, D. H. Kim, Y. Wei, and T. Bretl, “The Impact of Height on Indoor Positioning with Magnetic Fields,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, 2021.

[21] S. Suksakulchai, S. Thongchai, D. Wilkes, and K. Kawamura, “Mobile robot localization using an electronic compass for corridor environment,” in *SMC 2000 Conference Proceedings. 2000 IEEE International Conference on Systems, Man and Cybernetics. 'Cybernetics Evolving to Systems, Humans, Organizations, and their Complex Interactions' (Cat. No.00CH37166)*, vol. 5, pp. 3354–3359, IEEE, 2000.

[22] J. Haverinen and A. Kemppainen, “Global indoor self-localization based on the ambient magnetic field,” *Robotics and Autonomous Systems*, vol. 57, no. 10, pp. 1028–1035, 2009.

[23] P. E. Kuevor, J. W. Cutler, and E. M. Atkins, “Improving Attitude Estimation Using Gaussian-Process-Regression-Based Magnetic Field Maps,” *Sensors*, vol. 21, p. 6351, sep 2021.

[24] M. Frassl, M. Angermann, M. Lichtenstern, P. Robertson, B. J. Julian, and M. Doniec, "Magnetic maps of indoor environments for precise localization of legged and non-legged locomotion," *IEEE International Conference on Intelligent Robots and Systems*, pp. 913–920, 2013.

[25] P. Robertson, M. Frassl, M. Angermann, M. Doniec, B. J. Julian, M. G. Puyol, M. Khider, M. Lichtenstern, and L. Bruno, "Simultaneous localization and mapping for pedestrians using distortions of the local magnetic field intensity in large indoor environments," *2013 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2013*, 2013.

[26] G. Liu, B. Yu, L. Huang, L. Shi, X. Gao, and L. He, "Human-Interactive Mapping Method for Indoor Magnetic Based on Low-Cost MARG Sensors," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, 2021.

[27] B. Gozick, K. P. Subbu, R. Dantu, and T. Maeshiro, "Magnetic maps for indoor navigation," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 12, pp. 3883–3891, 2011.

[28] A. Ayanoglu, D. M. Schneider, and B. Eitel, "Crowdsourcing-Based Magnetic Map Generation for Indoor Localization," in *2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 1–8, IEEE, sep 2018.

[29] N. Akai and K. Ozaki, "Gaussian processes for magnetic map-based localization in large-scale indoor environments," *IEEE International Conference on Intelligent Robots and Systems*, vol. 2015-December, pp. 4459–4464, 2015.

[30] D. Almeida, E. Pedrosa, and F. Curado, "Magnetic Mapping for Robot Navigation in Indoor Environments," *2021 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2021*, 2021.

[31] Y. Zheng, S. Li, K. Xing, and X. Zhang, "Unmanned aerial vehicles for magnetic surveys: A review on platform selection and interference suppression," *Drones*, vol. 5, no. 3, 2021.

[32] R. Versteeg and M. McKay, "Feasibility Study for an Autonomous UAV -Magnetometer System – Final Report on SERDP SEED 1509:2206," Tech. Rep. September, Idaho National Laboratory, sep 2007.

[33] T. Koyama, T. Kaneko, T. Ohminato, T. Yanagisawa, A. Watanabe, and M. Takeo, "An aeromagnetic survey of Shinmoe-dake volcano, Kirishima, Japan, after the 2011 eruption using an unmanned autonomous helicopter," *Earth, Planets and Space*, vol. 65, pp. 657–666, jun 2013.

[34] G. Vasiljevic, D. Martinovic, M. Batos, and S. Bogdan, "Validation of two-wire power line UAV localization based on the magnetic field strength," *2022 International Conference on Unmanned Aircraft Systems, ICUAS 2022*, pp. 434–441, 2022.

[35] Y. Li, S. Zahran, Y. Zhuang, Z. Gao, Y. Luo, Z. He, L. Pei, R. Chen, and N. El-Sheimy, "IMU/magnetometer/barometer/mass-flow sensor integrated indoor quadrotor UAV localization with robust velocity updates," *Remote Sensing*, vol. 11, no. 7, pp. 1–22, 2019.

[36] S. Zahran, A. M. Moussa, A. B. Sesay, and N. El-Sheimy, “A New Velocity Meter Based on Hall Effect Sensors for UAV Indoor Navigation,” *IEEE Sensors Journal*, vol. 19, no. 8, pp. 3067–3076, 2019.

[37] J. Tiemann, A. Ramsey, and C. Wietfeld, “Enhanced UAV indoor navigation through SLAM-Augmented UWB Localization,” *2018 IEEE International Conference on Communications Workshops, ICC Workshops 2018 - Proceedings*, pp. 1–6, 2018.

[38] G. Niu, J. Zhang, S. Guo, M.-O. Pun, and C. S. Chen, “UAV-Enabled 3D Indoor Positioning and Navigation Based on VLC,” in *ICC 2021 - IEEE International Conference on Communications*, pp. 1–6, IEEE, jun 2021.

[39] P. Lipovský, K. Draganová, J. Novotňák, Z. Szőke, and M. Fil̆ko, “Indoor mapping of magnetic fields using uav equipped with fluxgate magnetometer,” *Sensors*, vol. 21, no. 12, 2021.

[40] J. Blažek, P. Lipovský, F. Heško, and D. Repčík, “Electromagnetic image of small UAV in very low frequency range,” *Journal of Electrical Engineering*, vol. 69, no. 6, pp. 438–441, 2018.

[41] M. H. Afzal, V. Renaudin, and G. Lachapelle, “Multi-magnetometer based perturbation mitigation for indoor orientation estimation,” *Navigation, Journal of the Institute of Navigation*, vol. 58, no. 4, pp. 279–292, 2011.

[42] G. Yin and L. Zhang, “Magnetic heading compensation method based on magnetic interferential signal inversion,” *Sensors and Actuators A: Physical*, vol. 275, pp. 1–10, jun 2018.

[43] M. Karimi, E. Babaians, M. Oelsch, T. Aykut, and E. Steinbach, “Skewed-redundant Hall-effect Magnetic Sensor Fusion for Perturbation-free Indoor Heading Estimation,” *Proceedings - 4th IEEE International Conference on Robotic Computing, IRC 2020*, pp. 367–374, 2020.

[44] N. Gao, M. Y. Wang, and L. Zhao, “A novel robust Kalman filter on AHRS in the magnetic distortion environment,” *Advances in Space Research*, vol. 60, no. 12, pp. 2630–2636, 2017.

[45] A. M. Sabatini, “Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing,” *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 7, pp. 1346–1356, 2006.

[46] Y. S. Suh, “Orientation estimation using a quaternion-based indirect Kalman filter with adaptive estimation of external acceleration,” *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 12, pp. 3296–3305, 2010.

[47] M. Ghobadi, P. Singla, and E. T. Esfahani, “Robust attitude estimation from uncertain observations of inertial sensors using covariance inflated multiplicative extended kalman filter,” *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 1, pp. 209–217, 2018.

[48] T. Michel, P. Genevès, H. Fourati, and N. Layaïda, "Attitude estimation for indoor navigation and augmented reality with smartphones," *Pervasive and Mobile Computing*, vol. 46, pp. 96–121, jun 2018.

[49] V. Renaudin and C. Combettes, "Magnetic, acceleration fields and gyroscope quaternion (MAGYQ)-based attitude estimation with smartphone sensors for indoor pedestrian navigation," *Sensors (Switzerland)*, vol. 14, no. 12, pp. 22864–22890, 2014.

[50] S.-M. Lee, J. Jung, S. Kim, I.-J. Kim, and H. Myung, "DV-SLAM (Dual-Sensor-Based Vector-Field SLAM) and Observability Analysis," *IEEE Transactions on Industrial Electronics*, vol. 62, pp. 1101–1112, feb 2015.

[51] J. Xu, Z. Yang, H. Chen, Y. Liu, X. Zhou, J. Li, and N. Lane, "Embracing Spatial Awareness for Reliable WiFi-Based Indoor Location Systems," in *2018 IEEE 15th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, pp. 281–289, IEEE, oct 2018.

[52] J. Kuang, X. Niu, P. Zhang, and X. Chen, "Indoor Positioning Based on Pedestrian Dead Reckoning and Magnetic Field Matching for Smartphones," *Sensors*, vol. 18, p. 4142, nov 2018.

[53] H. Xie, T. Gu, X. Tao, H. Ye, and J. Lv, "MaLoc," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, (New York, NY, USA), pp. 243–253, ACM, sep 2014.

[54] D. Hanley, A. B. Faustino, S. D. Zelman, D. A. Degenhardt, and T. Bretl, "MagPIE: A dataset for indoor positioning with magnetic anomalies," *2017 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2017*, vol. 2017-Janua, pp. 1–8, 2017.

[55] I. Ashraf, S. Din, S. Hur, G. Kim, and Y. Park, "Empirical Overview of Benchmark Datasets for Geomagnetic Field-Based Indoor Positioning," *Sensors*, vol. 21, p. 3533, may 2021.

[56] M. Laska, T. Schulz, J. Grottke, C. Blut, and J. Blankenbach, "Vi-slam2tag: Low-effort labeled dataset collection for fingerprinting-based indoor localization," in *2022 IEEE 12th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 1–8, 2022.

[57] M. Angermann, A. Friese, M. Khider, B. Krach, K. Krack, and P. Robertson, "A Reference Measurement Data Set for Multisensor Pedestrian Navigation with Accurate Ground Truth," *European Navigation Conference-GNSS*, no. May 2014, 2009.

[58] S. Bozkurt, A. Yazici, S. Gunal, U. Yayan, and F. Inan, "A novel multi-sensor and multi-topological database for indoor positioning on fingerprint techniques," in *2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA)*, pp. 1–7, IEEE, sep 2015.

[59] P. Kasebzadeh, G. Hendeby, C. Fritsche, F. Gunnarsson, and F. Gustafsson, "IMU dataset for motion and device mode classification," *2017 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2017*, pp. 1–8, 2017.

[60] S. C. Simon Brodeur, “Create: Multimodal dataset for unsupervised learning and generative modeling of sensory data from a mobile robot,” 2018.

[61] G. Mendoza-Silva, P. Richter, J. Torres-Sospedra, E. Lohan, and J. Huerta, “Long-Term WiFi Fingerprinting Dataset for Research on Robust Indoor Positioning,” *Data*, vol. 3, p. 3, Jan 2018.

[62] N. Zhu, M. Ortiz, V. Renaudin, C. Ichard, and S. Ricou, “Dataset of the intermediate competition in challenge MALIN: Indoor–outdoor inertial navigation system data for pedestrian and vehicle with high accuracy references in a context of firefighter scenario,” *Data in Brief*, vol. 34, p. 106626, feb 2021.

[63] C. Sanchez-Belenguer, E. Wolfart, A. Casado-Coscolla, and V. Sequeira, “RISEdb: a Novel Indoor Localization Dataset,” in *2020 25th International Conference on Pattern Recognition (ICPR)*, pp. 9514–9521, IEEE, jan 2021.

[64] P. Barsocchi, A. Crivello, D. La Rosa, and F. Palumbo, “A multisource and multivariate dataset for indoor localization methods based on WLAN and geo-magnetic field fingerprinting,” in *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 1–8, IEEE, oct 2016.

[65] J. Torres-Sospedra, D. Rambla, R. Montoliu, O. Belmonte, and J. Huerta, “UJIIndoorLoc-Mag: A new database for magnetic field-based localization problems,” in *2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 1–10, IEEE, oct 2015.

[66] D. Hanley, A. S. Dantas de Oliveira, X. Zhang, D. H. Kim, Y. Wei, and T. Bretl, “The impact of height on indoor positioning with magnetic fields,” 2020.

[67] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[68] J. C. Springmann and J. W. Cutler, “Attitude-Independent Magnetometer Calibration with Time-Varying Bias,” *Journal of Guidance, Control, and Dynamics*, vol. 35, pp. 1080–1088, jul 2012.

[69] H. Wu, X. Pei, J. Li, H. Gao, and Y. Bai, “An improved magnetometer calibration and compensation method based on levenberg–marquardt algorithm for multi-rotor unmanned aerial vehicle,” *Measurement and Control*, vol. 53, pp. 276–286, 3 2020.

[70] L. H. Regoli, M. B. Moldwin, M. Pellioni, B. Bronner, K. Hite, A. Sheinker, and B. M. Ponder, “Investigation of a low-cost magneto-inductive magnetometer for space science applications,” *Geoscientific Instrumentation, Methods and Data Systems*, vol. 7, no. 1, pp. 129–142, 2018.

[71] M. Romano, A. Ye, J. Pye, and E. Atkins, “Cooperative multilift slung load transportation using haptic admittance control guidance,” *Journal of Guidance, Control, and Dynamics*, pp. 1–14, 7 2022.

[72] E. Le Grand and S. Thrun, “3-Axis magnetic field mapping and fusion for indoor localization,” in *2012 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pp. 358–364, IEEE, sep 2012.

[73] D. Vandermeulen, C. Vercauteren, and M. Weyn, “Indoor localization Using a Magnetic Flux Density Map of a Building Feasibility study of geomagnetic indoor localization,” *AMBIENT 2013 : The Third International Conference on Ambient Computing, Applications, Services and Technologies Magnetic*, pp. 42–49, 2013.

[74] G. Desouza and A. Kak, “Vision for mobile robot navigation: a survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 2, pp. 237–267, 2002.

[75] F. Bonin-Font, A. Ortiz, and G. Oliver, “Visual Navigation for Mobile Robots: A Survey,” *Journal of Intelligent and Robotic Systems*, vol. 53, pp. 263–296, nov 2008.

[76] Y. Lu, Z. Xue, G.-S. Xia, and L. Zhang, “A survey on vision-based UAV navigation,” *Geo-spatial Information Science*, vol. 21, pp. 21–32, jan 2018.

[77] J. Delmerico, T. Cieslewski, H. Rebecq, M. Faessler, and D. Scaramuzza, “Are we ready for autonomous drone racing? the UZH-FPV drone racing dataset,” *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2019-May, pp. 6713–6719, 2019.

[78] D. Navarro and G. Benet, “Magnetic map building for mobile robot localization purpose,” *ETFA 2009 - 2009 IEEE Conference on Emerging Technologies and Factory Automation*, pp. 30–33, 2009.

[79] R. C. Leishman and T. W. McLain, “Multiplicative extended Kalman filter for relative rotorcraft navigation,” *Journal of Aerospace Information Systems*, 2015.

[80] D. P. Koch, D. O. Wheeler, R. W. Beard, T. W. McLain, and K. M. Brink, “Relative multiplicative extended Kalman filter for observable GPS-denied navigation,” *International Journal of Robotics Research*, 2020.

[81] J. Hall, N. Knoebel, and T. McLain, “Quaternion attitude estimation for miniature air vehicles using a multiplicative extended kalman filter,” in *Record - IEEE PLANS, Position Location and Navigation Symposium*, 2008.

[82] N. Filipe, M. Kontitsis, and P. Tsiotras, “Extended Kalman Filter for spacecraft pose estimation using dual quaternions,” *Journal of Guidance, Control, and Dynamics*, 2015.

[83] P. Martin and E. Salaün, “Generalized Multiplicative Extended Kalman Filter for Aided Attitude and Heading Reference System,” in *AIAA Guidance, Navigation, and Control Conference*, (Reston, Virigina), American Institute of Aeronautics and Astronautics, aug 2010.

[84] J. L. Crassidis and J. L. Junkins, *Optimal Estimation of Dynamic Systems*. Chapman and Hall/CRC, oct 2011.

[85] F. L. Markley and J. L. Crassidis, *Fundamentals of Spacecraft Attitude Determination and Control*. New York, NY: Springer New York, 2014.

[86] M. Romano, P. Kuevor, D. Lukacs, O. Marshall, M. Stevens, H. Rastgoftar, J. Cutler, and E. Atkins, "Experimental Evaluation of Continuum Deformation with a Five Quadrotor Team," in *2019 American Control Conference (ACC)*, pp. 2023–2029, IEEE, jul 2019.

[87] J. Paredes, P. Sharma, B. Ha, M. Lanchares, E. Atkins, P. Gaskell, and I. Kolmanovsky, "Development, implementation, and experimental outdoor evaluation of quadcopter controllers for computationally limited embedded systems," *Annual Reviews in Control*, 2021.

[88] T. Lee, M. Majji, and P. Singla, "Data-driven modeling for magnetic field variations using the GLO-MAP algorithm," *Computers and Geosciences*, vol. 144, no. July 2018, p. 104549, 2020.

[89] G.-X. Liu, L.-F. Shi, S. Chen, and Z.-G. Wu, "Focusing Matching Localization Method Based on Indoor Magnetic Map," *IEEE Sensors Journal*, vol. 20, pp. 10012–10020, sep 2020.

[90] S. A. Rahok and Ozaki Koichi, "Odometry correction with localization based on landmarkless magnetic map for navigation system of indoor mobile robot," in *2009 4th International Conference on Autonomous Robots and Agents*, pp. 572–577, IEEE, feb 2009.

[91] W. Youn, H. Choi, A. Cho, S. Kim, and M. B. Rhudy, "Accelerometer Fault-Tolerant Model-Aided State Estimation for High-Altitude Long Endurance UAV," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 1–1, 2020.

[92] M. Euston, P. Coote, R. Mahony, Jonghyuk Kim, and T. Hamel, "A complementary filter for attitude estimation of a fixed-wing UAV," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 340–345, IEEE, sep 2008.

[93] R. G. Valenti, I. Dryanovski, and J. Xiao, "Keeping a good attitude: A quaternion-based orientation filter for IMUs and MARGs," *Sensors (Switzerland)*, vol. 15, no. 8, pp. 19302–19330, 2015.

[94] J. O'keefe and J. Etele, "Ultrasonic localization of a quadrotor using a portable beacon," *AIAA Scitech 2019 Forum*, no. January, pp. 1–16, 2019.

[95] S. Yiu, M. Dashti, H. Claussen, and F. Perez-Cruz, "Wireless RSSI fingerprinting localization," *Signal Processing*, vol. 131, pp. 235–244, 2017.

[96] W. Sun, M. Xue, H. Yu, H. Tang, and A. Lin, "Augmentation of Fingerprints for Indoor WiFi Localization Based on Gaussian Process Regression," *IEEE Transactions on Vehicular Technology*, vol. 67, pp. 10896–10905, nov 2018.

[97] X. Wang, C. Zhang, F. Liu, Y. Dong, and X. Xu, "Exponentially weighted particle filter for simultaneous localization and mapping based on magnetic field measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 7, pp. 1658–1667, 2017.

[98] P. Mirowski, T. K. Ho, Saehoon Yi, and M. MacDonald, “SignalSLAM: Simultaneous localization and mapping with mixed WiFi, Bluetooth, LTE and magnetic signals,” in *International Conference on Indoor Positioning and Indoor Navigation*, pp. 1–10, IEEE, oct 2013.

[99] C. Laoudias, A. Moreira, S. Kim, S. Lee, L. Wirola, and C. Fischione, “A Survey of Enabling Technologies for Network Localization, Tracking, and Navigation,” *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 3607–3644, 2018.

[100] T. Kim Geok, K. Zar Aung, M. Sandar Aung, M. Thu Soe, A. Abdaziz, C. Pao Liew, F. Hossain, C. P. Tso, and W. H. Yong, “Review of Indoor Positioning: Radio Wave Technology,” *Applied Sciences*, vol. 11, p. 279, dec 2020.

[101] B. Kim, M. Kwak, J. Lee, and T. T. Kwon, “A multi-pronged approach for indoor positioning with WiFi, magnetic and cellular signals,” in *2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 723–726, IEEE, oct 2014.

[102] J. R. Gardner, G. Pleiss, D. Bindel, K. Q. Weinberger, and A. G. Wilson, “Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration,” in *Advances in Neural Information Processing Systems*, 2018.

[103] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. Cambridge, Mass.: MIT Press, 2005.

[104] F. Gustafsson, “Particle filter theory and practice with positioning applications,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 25, pp. 53–82, jul 2010.

[105] D. Kishkinev, F. Packmor, T. Zechmeister, H. C. Winkler, N. Chernetsov, H. Mouritsen, and R. A. Holland, “Navigation by extrapolation of geomagnetic cues in a migratory songbird,” *Current Biology*, 2021.

[106] S. Trimpe and R. D’Andrea, “Event-based state estimation with variance-based triggering,” in *Proceedings of the IEEE Conference on Decision and Control*, pp. 6583–6590, IEEE, dec 2012.

[107] G. Alag, “Evaluating the performance of intel realsense t265 xsens technologies b.v.,” March 2020.

[108] C. Li, L. Mo, and D. Zhang, “Review on UHF RFID Localization Methods,” *IEEE Journal of Radio Frequency Identification*, vol. 3, no. 4, pp. 205–215, 2019.