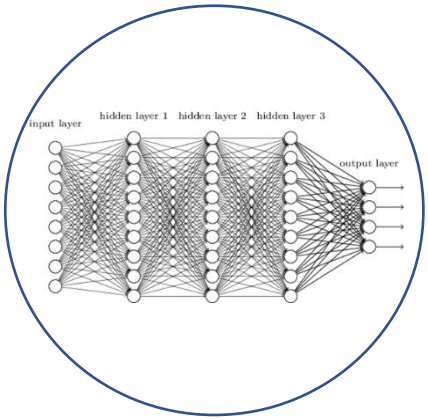
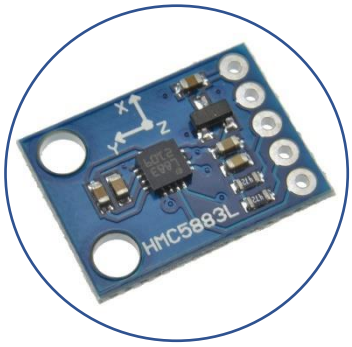
Engineer internship thesis March-August 2022

Magnetic navigation - On-line calibration of airborne magnetometer

Nathan Laoué





DGA Maîtrise de
l'information - BP 7 -
35998 RENNES Cedex 9

ELISA Aerospace Hauts de France ELISA Aerospace Bordeaux
48 rue Raspail 176 impasse du Forestier
02100 SAINT QUENTIN - France 33127 SAINT JEAN D'ILLAC - France
Tel. 03 23 68 06 11 - contact@elisa-aerospace.fr Tel. 06 48 95 76 72 - contactbdx@elisa-aerospace.fr

Acknowledgements

I would like to thank Arnaud Lepers, my tutor at SYSNAV, for having trusted me on this project but also for having guided me throughout this internship. His advice was fundamental to the good progress of the internship as well as his benevolence.

I would also like to thank Laure Deletraz and Charly Faure, my tutors in the IA2P department, who advised me throughout this internship. Their experience in the field of AI allowed me to discover new methods. I thank them for taking the time to answer my many questions.

I would also like to thank David Traore, head of the SYSNAV department, for welcoming me to the SYSNAV department and supporting me throughout my internship.

I would also like to thank Alain Droniou, head of the IA2P department, for trusting me as a trainee in his department.

I would also like to thank the many people in the SYSNAV department and in the DIA trade for the rich exchanges and the many tips they gave me.

Summary

DISSEMINATION CLAUSES : 3

Acknowledgements 4

Contents 5

Table of illustrations 8

Table of Tables 8

Glossary 10

Rating 11

Context 12

Presentation of the company 13

1- DGA-MI 13

2- SYSNAV Department 13

3- IA2P Department 14

Magnetic Navigation: Compensation by Artificial Intelligence 15

1- Introduction 15

2- State of the art 16

2-1. Understanding the Earth's magnetic field 16

2-1.1. Earth's core field 16

2-1.2. Lithospheric field 17

2-1.3. External effects 17

2-1.4. Summary of components 17

2-2. Measuring devices 18

2-2.1. Vector magnetometer 18

2-2.2. Scalar magnetometer 19

2-3. Magnetic Anomaly Cards 19

2-4. Components of the magnetic measurement 20

2-5. Compensation for measures 21

2-5.1. Tolles-Lawson model 21
2-5.2. Offsetting manoeuvres 22
2-5.3. Limitations 22
3- Data set 23
3-1. MIT dataset 23
3-2. Flight path visualization 25
3-3. Pre-selection of data 28
3-3.1. Correlation methods 28
3-3.2. Mutual information 29
3-3.3. XGBoost 31
3-3.4. New data representation 32
3-3.5. Selected data for the models 34
4- Compensation by neural networks 35
4-1. Automatic learning 35
4-2. Models used 36
4-3. Data pre-processing 37
4-4. Training and model parameters 38
5- Results 40
5-1. Application of Tolles-Lawson on our data 41
5-2. Comparison of models 45
6- Prospects for improvement 49
6-1. Data 49
6-2. Models 50
7- Conclusion 50
Bibliography 53
Netography 55
Appendix A - Tolles-Lawson Equations and Resolution 56

Appendix B - List of data in the dataset 63

Appendix C - Neural Networks 66

C.1 - Multilayer Perceptron 66

C.2 - A neuron 66

C.2.1 - Activation function 67

C.2.2 - Learning 68

C.3 - Convolutional neural networks 69

C.3.1 - Convolutional layer 69

C.3.2 - Pooling layer 70

C.4 - Recurrent neural networks 71

Summary 74

Abstract 75

Table of illustrations

Figure 1 - Trajectory 1002 (left) and 1003 (right). 25
Figure 2 - Trajectory 1004 (left) and 1005 (right). 26
Figure 3 - Trajectory 1006 (left) and 1007 (right). 26
Figure 4 - Magnetic maps of paths 1002, 1003, 1004, 1005, 1006, 1007. 27
Figure 5 - Correlation between the uncompensated magnetometer 1 and the data set. 28
Figure 6 - Correlation between magnetometer 5 without compensation and non-magnetic data from the dataset 29
Figure 7 - Mutual information between the data set and the uncompensated magnetometer 1. 30
Figure 8 - Mutual information between the non-magnetic data and the uncompensated magnetometer 5. 31
Figure 9 - Importance of variables according to XGBoost. 32
Figure 10 - Correlation between non-magnetic data and magnetometer readings 34
Figure 11 - Evolution of the learning rate. 40
Figure 12 - Offset maneuvers 1002.02 (purple) and 1002.20 (green). 41
Figure 13 - Comparison of magnetometers for Tolles-Lawson compensation with square maneuver and regularization parameter=0 43
Figure 14 - Prediction and error on MLP Flight 1007. 46
Figure 15 - Prediction and error on CNN Flight 1007. 46
Figure 16 - Prediction and error on LSTM Flight 1007. 46
Figure 17 - Sectional performance of MLP (left) and CNN (right) for Flight 1007. 47
Figure 18 - 2D view of position-dependent compensation errors for the CNN. 48
Figure 19 - 3D view of position-dependent compensation errors for the CNN. 48
Figure 20 - Importance of training data for model prediction. 49
Figure 21 - Measured magnetic field vectors of the Earth and the aircraft. 58
Figure 22 - Architecture of an RNN. Source: stanford.edu 72

Table of tables

Table 2.1: Components of the total magnetic field 18

Table 3.1: Position of magnetometers in the aircraft..... 23

Table 3.2: Summary of dataset data 24

Table 4.1: Major differences between the different models. 36

Table 4.2: Training parameters for each model. 39

Table 5.1: Comparison of compensation maneuvers, regulation parameter=0 42

Table 5.2: Comparison of trimming operations, regulation parameter=0.00025. 43

Table 5.3: Comparison of the compensation maneuvers, regulation parameter=0.0025 44

Table 5.4: Comparison of trimming operations, regulation parameter=0.025. 44

Table 5.5: Comparison of trimming operations, regulation parameter=0.25. 44

Table 5.6: Comparison of the compensation error of the models used. 45

Glossary

USAF	United States Air Force
GPS	Global Positioning System
GAO	Government Accountability Office
GIFI	International Geomagnetic Reference Field
MAD	Magnetic Anomaly Detection
MIT	Massachusetts Institute of Technology
DAF-MIT AI Accelerator	Air force Massachusetts Institute of Technology Artificial Intelligence Accelerator
SGL	Sanders Geophysics Ltd
INS	Inertial Navigation System
CNN	Convolution Neural Network
RNN	Recurrent Neural Network
MLP	Multilayer Perceptron
BP	Backpropagation
BPTT	Backpropagation Through Time
LSTM	Long Short-Term Memory
GRU	Gated recurrent Unit
TL	Tolles-Lawson

Rating

a	A scalar
$\boldsymbol{a}$	A vector
$\boldsymbol{A}$	A matrix
$\boldsymbol{I}$	Identity matrix
$\dot{y}$	Newton differentiation

Context

Magnetic navigation is one of the oldest forms of navigation, thanks to the compass, which has been used for almost 1000 years to find the North. Nowadays, with the arrival of satellites and GPS, it is much less common to use a compass to find your way. The GPS system is reliable and allows to have a very good precision on the position. But GPS can be jammed. So this is a major problem that means you can't be sure that the GPS system is available. Many military missions are carried out in environments where the GPS system can be jammed, potentially putting human lives at risk.

It is in this context that navigation using the Earth's magnetism has regained strong interest in recent years. The magnetic field is available 24 hours a day, 7 days a week, it depends little on the weather and it is impossible to jam without being very close to the target. It is therefore a very interesting means of navigation for military applications.

Magnetic navigation had been set aside until now mainly because it was very difficult to compensate for the magnetic effects of the carrier. A linear compensation method called Tolles-Lawson has been mainly used since the 1950s but it is not very effective in the case of magnetometers that are strongly noisy by the carrier, i.e. carried in the cabin. Solutions such as a boom at the rear of the carrier are used today but this is not viable for example in the case of a fighter aircraft.

In 2016, Major A. Canciani of the USAF showed that it was possible to obtain good precision performance with magnetic navigation. Today, many avenues are being explored to overcome this problem of compensating for the effects of the carrier, in particular by methods based on artificial intelligence, which are known to be able to model non-linear effects.

The aim of this internship is to study different architectures of artificial intelligence models to see if it is possible to surpass the Tolles-Lawson model. For that, a first part will be dedicated to the state of the art in order to understand the different sources of the magnetic field, how it is measured and how to compensate the measurements made. The next part will be dedicated to a presentation and exploration of the data used. This will be followed by a presentation of the models used and the training method for these models. Finally, these models will be compared to each other and to the Tolles-Lawson method. The last part will be reserved for the possible perspectives of improvements on the subject.

Presentation of the company

1- DGA-MI

DGA-MI or Direction Générale de l'Armement - Maîtrise de l'information is a DGA center located in Bruz, south of Rennes. It is a center of expertise for information and communication systems, cybersecurity, electronic warfare, guidance-navigation and missile systems. DGA-MI has over 60 years of expertise in a wide range of fields, from components to systems of systems. DGA-MI is involved in the entire cycle of armament operations, whether upstream with the characterization of the electromagnetic or cyber threat or downstream with support for the configuration of operational systems. The main services of the center are :

- Advice to the project management of armament operations
  - Knowledge of the threat and the state of the art
  - Technical specification of requirements
  - Risk management in development
  - Qualification and performance evaluation
- Support to the Forces
  - Support for the configuration and programming of systems for training and military operations
- Sovereignty developments
- Export control and support

DGA-MI also fulfils the fundamental missions of DGA, i.e., to equip armies in a sovereign manner, to prepare the future of defense systems, to promote European cooperation, to support exports and to contribute to the transformation of the Ministry of Defense.

2- SYSNAV Department

SYSNAV is one of the oldest departments of the site. The main role of the department is the expertise of navigation systems with several activities which are:

- Contribution to the design and optimization of navigation systems

- Use of evaluations
- Feedback
- Simulation & hybrid simulations

This represents a variety of skills, from sensor expertise to simulation, trajectography and signal processing.

3- IA2P Department

IA2P is one of the two departments in artificial intelligence at DGA-MI that deals mainly with machine learning methods on unstructured data. The department is structured according to the following 3 themes:

- AI for multimedia data processing
- AI for temporal signal processing
- AI for image processing from defense sensors

The IA2P department works mainly to help the other DGA professions to integrate AI into the armament programs carried out for the benefit of the different armies (land, sea, air) and in different fields (command, cyber security, intelligence, etc.).

Magnetic navigation: Compensation by Artificial Intelligence

1- Introduction

Several studies, mainly by the USAF, have shown that magnetic navigation could become a viable alternative to GPS. Magnetic navigation consists of using maps of magnetic anomalies in the earth's magnetic field that are correlated with magnetic measurements to obtain a position. The United States has also shown interest in this new means of navigation in a GAO report [1]. It thus seems to be a new emerging means of navigation and especially with many assets for the field of defense.

Magnetic navigation is passive, only the earth's magnetic field is received. It is also independent of position, the magnetic field measurement is accessible everywhere. Moreover, it is not dependent on the weather, and weather conditions have no impact on the magnetic measurements. Some special cases like thunderstorms could have an impact, but these are rarer cases. Finally, one of the great advantages of this means of navigation is its resistance to interference. The magnetic field decreases in $\frac{1}{r^3}$This means that to send jamming signals to measuring instruments would require very powerful equipment. In the case of aircraft, this is almost impossible since they fly far from the ground. The jamming equipment would have to be close to the aircraft in flight via, for example, another aircraft for the jamming to be effective. This is a very important advantage for the defense community as it makes it a reliable means of navigation in mission areas where GPS would not be available.

However, magnetic navigation also has its disadvantages. The measured magnetic field is a linear superposition of several magnetic fields. However, some of the measured fields are disturbances due to the measuring carrier. One of the challenges for magnetic navigation is to compensate for the noisy magnetic measurements by the carrier in order to have the closest possible measurement to the one available on the magnetic anomaly maps. This is what has been studied during this internship, several compensation techniques will be explored including neural network methods.

First, there will be a presentation of the state of the art on the magnetic field, its measurement and the compensation of the measurements. Then a presentation and exploration of the data set used. Then finally the models used in this work and their performances. Some perspectives on the future work to be done will close this work.

All the code used in this work is available online and accessible to all via the following link: https://github.com/Naatyu/MagNav

2- State of the art

The purpose of this chapter is to introduce the state-of-the-art method for magnetic navigation. For this purpose, it is important to understand what the Earth's magnetic field is composed of as well as its different sources. A comparison of the sources will help to determine which one seems to be the most suitable for magnetic navigation. A presentation of the measuring devices is also made in order to better understand the distinction between the different sensors. Magnetic anomaly maps will also be presented in order to understand their importance for magnetic navigation.

The Tolles-Lawson compensation method will be introduced as well as its current limitations.

2-1. Understanding the Earth's magnetic field

The total magnetic field of the earth is mainly composed of a linear superposition of 3 main sources coming from the earth's core, the earth's crust and external effects mainly due to space effects. The field used for magnetic navigation comes from the earth's crust.

2-1.1. Earth's core field

One of the main sources of the magnetic field is the Earth's core. This is due to the Earth's convection movements, which set in motion the elements of the Earth's core, 90% of which is liquid iron. This creates a dynamo effect which generates an electric field which in turn generates a magnetic field. The magnetic field is measured in Tesla (T) and the magnetic field generated by this dynamo effect is between 20000 and 60000 nT. It represents about 95% of the total magnetic field emitted by the Earth.

The IGRF (International Geomagnetic Reference Field) model is used to model the magnetic field of the Earth's core. It is able to model magnetic field wavelengths of the order of 3000 km and above. These wavelengths are too long to use the core magnetic field as a reference for

magnetic navigation. The magnetic field of the core varies from a month to several million years. The IGRF model is therefore updated every 5 years to compensate for these variations.

2-1.2. Lithosphere field

A small part of the magnetic field comes from the Earth's crust. This field is produced by the rocks below the Earth's surface and above the isothermal surface. They are below the Curie temperature at which rocks lose their magnetic properties. This permanent magnetization comes from minerals that are magnetized by induction and then cooled below the Curie temperature. It has an intensity in the order of a hundred nano Tesla and varies very little from year to year ($\approx 1\ nT$).

This makes it a very good candidate as a reference for magnetic navigation. This field is also called magnetic anomaly or crustal field and it strongly decreases with altitude however it remains usable as a reference even at about 10 km. The altitude acts as a low pass effect on the measurement. One of the disadvantages of the magnetic anomaly is its weak intensity compared to the field generated by the core. It is therefore more difficult to isolate.

2-1.3. External effects

The majority of the Earth's magnetic field comes from effects within the Earth. But there are other external effects to consider. These are effects due to the solar wind on the ionosphere. For example, the day/night cycle is one of the most common effects. It is also possible to have effects like the aurora borealis or even magnetic storms coming from the sun. Globally these effects are in the order of ten nanotesla. These effects are also called *diurnal effects*, *temporal variations* or *space weather effects*.

The measurement of these effects can be obtained from ground stations for a given area. In the case of an aircraft, this becomes more complicated as it would be necessary to be able to retrieve data from stations along the flight path. It is possible to make a rough model of these effects, mainly of the day/night cycle, but this remains imprecise. These effects must be compensated for in the measurement of the magnetic field to obtain the most accurate position possible.

2-1.4. Summary of components

Here is a table inspired by [2] which summarizes the different components of the field with their variations:

Table 2.1: Components of the total magnetic field. The measurements are based on the area of Ottawa, Ontario, Canada. The amplitudes vary depending on the position of the measurement but the order of magnitude is similar. The data in the table are similar to what would be obtained in France, the amplitude of the core field would be slightly lower ($\approx 45000\ nT$).

Component	Amplitude	Spatial Variation	Time Variation
Core field	50,000 nT	100 nT	1 nT/week
Earth's crust field	100 nT	200 nT	1 nT/year
External effects	10 nT	5 nT	10 nT/day

This table shows the difference between the different components of the total earth magnetic field. The field coming from the earth's crust is the best candidate as a reference for magnetic navigation because of its small variation in time. On the other hand, its intensity is quite low and therefore it is harder to extract from the total field. However, its spatial variations are strong unlike the other sources, which makes it easier to measure the magnetic anomaly.

2-2. Measuring devices

The devices used to measure the magnetic field are called magnetometers and there are two types, scalar or vector. Since the 1800s, magnetic measuring devices have been used for the discovery of mineral or oil deposits. It is with the second world war that their use evolved and that they were embarked in planes for the detection of submarines. Nowadays, these devices are mainly used by scientists to guide mining explorations or to create magnetic anomaly maps. They are also still used for magnetic anomaly detection (MAD).

2-2.1. Vector magnetometer

Vector magnetometers allow to measure the magnetic field in one direction. The most classical is to measure the field in three different directions:

$$\overrightarrow{B_t} = B_x\hat{\imath} + B_y\hat{\jmath} + B_z\hat{k}$$

(2.1)

With B_x, B_y, B_z the field magnitudes in the orthogonal directions $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$ of the reference plane. *Fluxgate* compasses are frequently used to perform the measurement. Three *Fluxgates* are mounted perpendicular to each other, which allows the direction of the magnetic field in space to be obtained. It is also possible to use a rotating magnetic induction magnetometer. It rotates at a known speed around an axis and allows us to obtain the measurement of the magnetic field in the plane of rotation.

One of the disadvantages of these magnetometers is that they are less accurate than vector magnetometers. They are generally not used as the sole means of measurement for magnetic navigation.

2-2.2. Scalar magnetometer

A scalar magnetometer measures the total magnetic field strength. The total magnetic field strength is written as follows:

$$\left|\overrightarrow{B_t}\right| = \sqrt{B_x^2 + B_y^2 + B_z^2}$$

(2.2)

The scalar magnetometer is only able to calculate $\left|\overrightarrow{B_t}\right|$. Nowadays, most measurements are made with nuclear resonance magnetometers. These are magnetometers that use the atomic properties of magnetic field sensitive gases to make their measurements [3]. A particular type of magnetic resonance magnetometer are the *alkali-vapor* magnetometers. They are also called optically pumped magnetometers and use different types of alkali vapor such as cesium or potassium. They are small and light instruments capable of operating alone. They can detect changes in the picotesla range and have an absolute measurement accuracy in the nanotesla range.

Scalar magnetometers are generally more accurate than vector magnetometers and are mainly used for most magnetic measurements today.

2-3. Magnetic anomaly cards

In order to be able to navigate using the Earth's magnetic field, one needs a good quality measurement but also reference maps to find one's way around. These maps reference the magnetic field of the Earth's crust at a certain position, which is perfectly suited for magnetic navigation. For this purpose, campaigns to measure the total magnetic field of the Earth are carried out, mainly from aircraft modified to reduce magnetic disturbance. Once the measurements of the magnetic field of the area are obtained, the field of the Earth's core is subtracted using the IGRF model, as well as the variations due to the ionosphere using ground stations. This allows us to obtain a magnetic anomaly map of the area. Other correction steps are also performed, but they correspond to map interpolation methods, which is of less interest to us in this case.

Some maps are freely available on the internet and have different resolutions. The resolution corresponds to the fineness of the mesh made during the measurements. The finer the mesh, the more accurate the map. The resolution of the map is very important in the case of magnetic navigation, some maps are global but their resolution is low and cannot be used as a reference. For example, the WDMAM (*World Digital Magnetic Anomaly Map*) is a global map of magnetic anomaly measurements but its resolution is too low to be used for navigation. Generally speaking, the closer to the ground a map is made with a tight mesh, the better the quality. By being close to the ground, it guarantees to correctly measure the differences of granularity in the ground, since the altitude acts as a low pass filter on our measurements.

Thanks to the magnetic anomaly maps, it is possible to know the anomaly measurements at a higher altitude, it is an *upward continuation*. This means that for navigation in an aircraft, flying above the altitude of the reference map is not a problem. The *chessboard* method of Cordell-Hildenbrand allows for example to know the measurement of anomalies at a higher altitude than the chart. These methods will not be detailed here because it is not the purpose of this work. It is also possible to know the measurements below the altitude of the measurement map. However, they are much more limited than the *upward continuation* methods and they introduce errors in the map. It is generally not recommended to use these methods for altitudes lower than about ten meters from the map altitude.

2-4. Components of magnetic measurement

To carry out magnetic navigation, the measurement of the magnetic anomaly is essential. However, as we have seen previously, the measuring devices are only capable of measuring the

total intensity of the magnetic field, and the magnetic field measured is composed of at least 4 sources:

$$\vec{B}_{mesure} = \vec{B}_{noyau} + \vec{B}_{anomalie} + \vec{B}_{diurne} + \vec{B}_{avion} \tag{2.3}$$

With $\vec{B}_{mesure}$ the total magnetic field measured by the magnetometer, $\vec{B}_{noyau}$ the measurement of the magnetic field from the Earth's core, $\vec{B}_{anomalie}$ the crustal magnetic field, $\vec{B}_{diurne}$ the magnetic field from external effects and $\vec{B}_{avion}$ the magnetic field emitted by the aircraft. Since in our dataset our measurements were made from an aircraft, the disturbing magnetic field was named $\vec{B}_{avion}$In addition, the altitude of the plane is sufficiently high, which implies that the magnetic field emitted by buildings or human sources on the ground is very weak, so that the only disturbance in the measurement comes from the plane.

So, we're trying to get $\vec{B}_{anomalie}$ but there is no way to measure it directly with a magnetometer. Since the total magnetic field is a linear superposition of other fields, it is possible to simply subtract $\vec{B}_{noyau}$ using the IGRF model. This is not a problem in real time because the magnetic navigation is used to re-calibrate an INS which gives us an approximate position, but of sufficient accuracy to use the IGRF model and get good performance. The component $\vec{B}_{diurne}$ component can be subtracted with ground stations or even with estimates of diurnal effects. This leaves us with $\vec{B}_{anomalie}$ and $\vec{B}_{avion}$. Only the measurement of the magnetic anomaly is needed, but it is noisy due to the aircraft disturbance, and there is no way to measure only the aircraft disturbance. It is therefore necessary to implement a method capable of modelling the disturbance field in order to obtain the magnetic anomaly measurement.

2-5. Compensation for measures

Once the magnetic field has been measured, it is often necessary to compensate for the effects of the carrier who took the measurement. This is where the Tolles-Lawson model comes in.

2-5.1. Tolles-Lawson model

In order to compensate for the magnetic disturbances of the aircraft on the measurement, a model capable of modeling these different effects is needed. The current state-of-the-art compensation is the Tolles-Lawson model [4, 5, 6]. This model uses measurements of the magnetic field made by a vector magnetometer to calculate coefficients to remove the effects of the aircraft from the measurements of a scalar magnetometer. This model was developed during World War II mainly for magnetic anomaly detection (MAD). It was then improved by Leliak who proposed to perform sinusoidal maneuvers over magnetic plains (i.e. little variation of the magnetic field) in order to highlight the disturbances emitted by the aircraft [7]. This method has been used since the 1960s and is still state of the art today. This study is mainly based on the model of [8] which is a modified version of Tolles-Lawson but the process remains essentially the same. A detailed development of the Tolles-Lawson model is available in Appendix A.

2-5.2. Compensating maneuvers

The Tolles-Lawson model needs magnetometer measurements to be able to calculate the coefficients of its model and as seen previously Leliak proposed to perform compensation maneuvers. One of the classical calibration patterns consists in performing a square in the air with pitch maneuvers of $\pm$ 5°roll of $\pm$ 10° and yaw of $\pm$ 5°. There is no order to respect and the angle values are not absolute, it is quite possible to make larger angles. However, it is not recommended to make smaller angles as this would decrease the amplitude of the aircraft disturbances.

Maneuvers are often square but other shapes are possible. For example, it is possible to perform cloverleaf or figure-of-eight maneuvers, which have the effect of covering more movements by the aircraft, but the impact on performance is not always positive.

2-5.3. Limitations

The Tolles-Lawson model is sufficient to remove disturbances in the case where the aircraft intended to take magnetic readings (limited aircraft movements, reduction of magnetic sources from the aircraft, etc.) and a boom located at the rear of the aircraft. The purpose of the pole is to move the magnetometer away from the aircraft and thus from disturbances. With this system it is possible to achieve an accuracy of about ten meters with magnetic navigation.

Even if this system works, it is not at all practical to use, especially for combat aircraft, which would be the first to benefit from this type of technology. So, it's not feasible to place the magnetometers on a boom. Only when the magnetometers are placed in the aircraft, they are strongly disturbed by it and Tolles-Lawson is not sufficient to reduce the effects which are strongly non-linear.

It is therefore necessary to implement a more advanced compensation method in order to be able to use onboard magnetometers for magnetic navigation.

3- Data set

In order to navigate using the Earth's magnetism, it is necessary to have magnetic field measurements from a vehicle and also to have the magnetic anomaly map of the measurement area. This chapter presents the magnetic anomaly maps and their creation in addition to the magnetic measurements used.

3-1. MIT dataset

In order to study the problem of compensating for the magnetic effects of the carrier, we will use a challenge from MIT. This challenge [9] aims at obtaining the best possible compensation performance on flight trajectories.

To create this dataset, the *Department of the Air Force Massachusetts Institute of Technology Artificial Intelligence Accelerator* (DAF-MIT AI Accelerator) partnered with Sanders Geophysics Ltd. (SGL) to collect data over the Ottawa area in Ontario, Canada. They collected data from five scalar magnetometers placed at different locations on the aircraft, a Cessna 208B Grand Caravan, as well as from three vector magnetometers also placed at different locations on the aircraft. The scalar magnetic field data in the dataset used were measured using an optically pumped magnetometer. The vector magnetic field measurements were made using a *Fluxgate* magnetometer. The following is a summary table of the magnetometers along with their positions from [9]:

Table 3.1: Position of magnetometers in the aircraft . The reference point is the front seat rail. X is positive in the direction of travel of the aircraft, Y is in the direction of the aircraft's left wing, perpendicular to X (the entrance door to the aircraft is on the left, on the same side as the wing) and Z is positive towards the top of the aircraft.

Sensor Name	Position	*X* (m)	*Y* (m)	*Z* (m)
Scalar Magnetometers				
Mag 1	Tail boom	-12.01	0	1.37
Mag 2	Forward cabin, behind the cockpit	-0.60	-0.36	0
Mag 3	Mid-cabin, close to INS	-1.28	-0.36	0
Mag 4	Rear of the cab, on the ground	-3.53	0	0
Mag 5	Rear of the cabin, on the ceiling	-3.79	0	1.20
Vector Magnetometers				
Stream B	Tail of the aircraft, at the base of the boom	-8.92	0	0.96
Flux C	Rear of the cabin, near the door	-4.06	0.42	0
Flow D	Rear of cab, opposite side of door	-4.06	-0.42	0

In order to study the effects of aircraft components, electrical data, aircraft orientation, INS data, and radar data are also available. A list of all the data is available in Appendix B. Thus 92 different variables are available to compensate for the noisy magnetic measurements over a total duration of 20 hours and 55 minutes. The measurements are made at a rate of 10 Hz.

However, the magnetometers located inside the aircraft are noisy due to the aircraft components. For this reason, a two-to-three-meter pole is placed at the back of the aircraft with a scalar magnetometer. As the magnetic field decreases very quickly, the simple fact of having a magnetometer a few meters away is enough to attenuate the magnetic effects emitted by the plane and thus obtain only the total magnetic field of the Earth. This measurement will be used as a reference to measure the compensation performance.

There are a total of 6 flights in the dataset, summarized in the table below:

Table 3.2: Summary of the dataset. Some sections are not available in the dataset made available by MIT because they are used as evaluation for the challenge. The number of sections presented here corresponds to the number of sections available.

Flight number	Flight sections	Flight time	Number of data	Frequency
1002	28	5h 45m 55s	207578	10 Hz
1003	10	4h 26m 42s	160030	10Hz

1004	21	2h 15m 38s	81408	10 Hz
1005	10	2h 16m 12s	81731	10 Hz
1006	8	3h 00m 31s	108318	10 Hz
1007	6	3h 10m 50s	114506	10 Hz

Flight 1002 has two trim maneuvers and different flight paths to highlight the effects of the aircraft. Flight 1003 consists mainly of figure-of-eight trajectories. Flights 1004 and 1005 are magnetic anomaly measurement trajectories for the purpose of creating anomaly maps. Flight 1006 has different compensation maneuvers and flight 1007 has different trajectories to highlight the different effects of the aircraft. These different flights will serve as a basis for this work and are sufficiently representative of real trajectories to estimate the performance of the compensation methods used on real cases.

It is also possible to add in-flight events such as radar activation, radio communication or flap lowering available in the dataset. However, they have not been taken into account in this work due to lack of time and also because some of these effects are reflected in the electrical measurements.

3-2. Flight path visualization

A first visualization of the trajectories allows to gain understanding on the data used. Figures 1 to 3 represent the trajectories used in the data with a different color for each flight section. These trajectories are close to Ottawa, Canada:

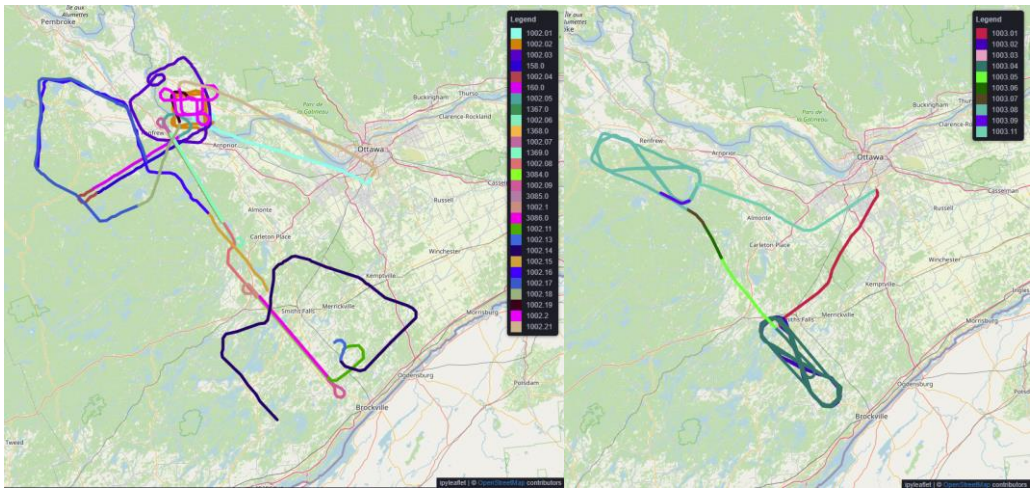


Figure 1 - Trajectory 1002 (left) and 1003 (right).

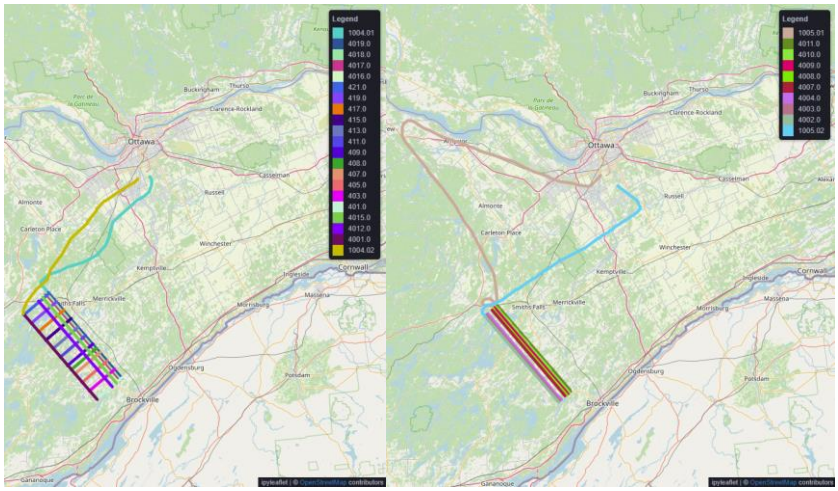


Figure 2 - Trajectory 1004 (left) and 1005 (right).

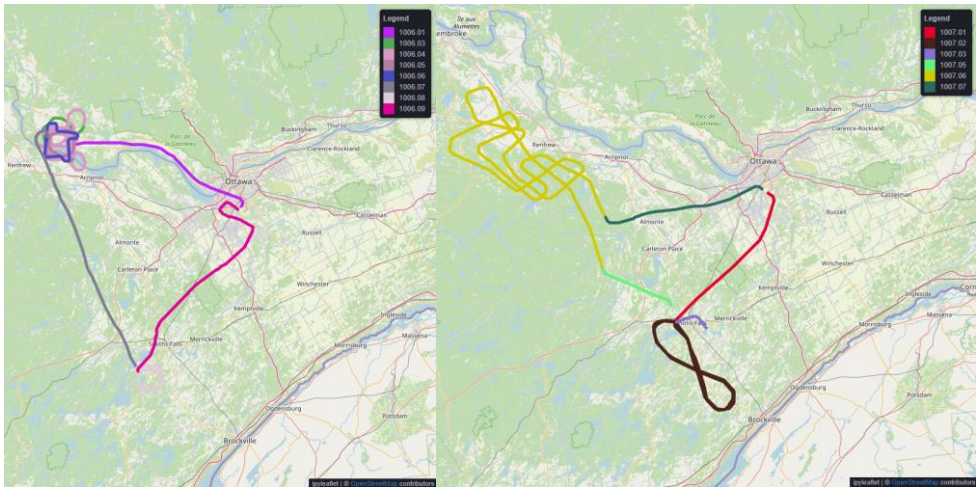


Figure 3 - Trajectory 1006 (left) and 1007 (right).

Almost all the flights were made at an altitude of between 800 and 400 meters. Three flights appear to be out of the ordinary. Flights 1004, 1005 and 1006 are not usual trajectories for an aircraft. Flights 1004 and 1005 are in fact magnetic map measurement flights, which explains the long lines close together. Flight 1006 represents several trim maneuvers made at different altitudes and with different shapes. It is also interesting to note that some sections of the trajectories are cut off. These sections have been removed by the creators of the dataset either due to erroneous measurements or in order to keep these sections as test data for the challenge. The most interesting flight seems to be flight 1007. It has different and slightly more complex trajectories than a simple straight line. Section 1007.04 serves as a test section for the challenge, which is why this section is not available.

It may also be interesting to visualize a magnetic anomaly map of the area. This will show if the area is subject to strong crustal field variations. The magnetic anomaly map of Canada, available online, provides the following graphs:

In general, the various flights appear to be in areas with relatively little crustal field variation.

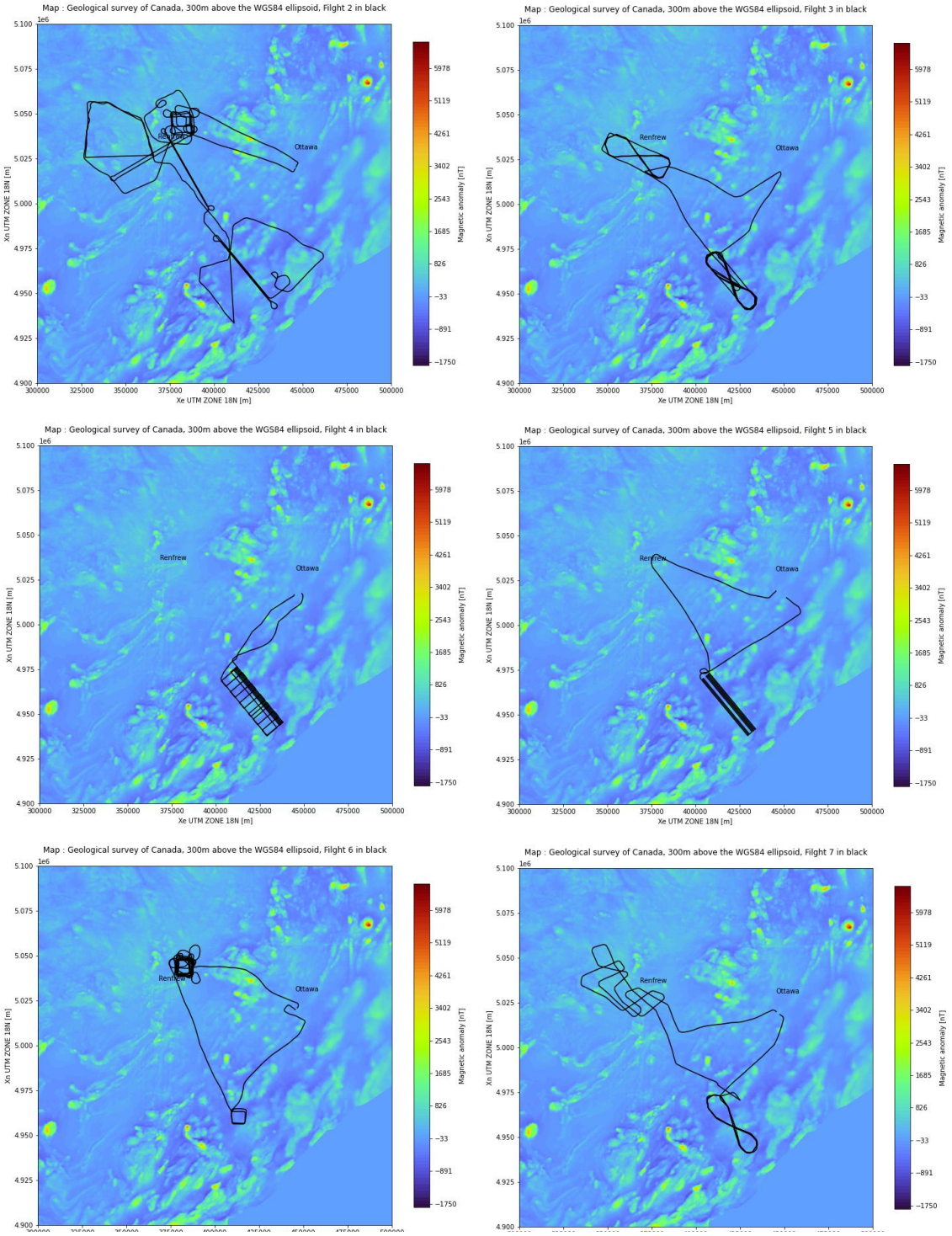


Figure 4 - Magnetic maps of paths 1002, 1003, 1004, 1005, 1006, 1007.

It is also possible to verify that the compensation maneuvers in flight 1002 are indeed over an area with very little magnetic variation.

3-3. Pre-selection of data

The dataset has 92 different variables. It seems interesting to make a first selection of variables before training our models. This will make the training faster and will limit the noise that some variables can bring.

3-3.1. Correlation methods

It is possible to find relationships between data using the Spearman correlation and the Pearson correlation. Pearson's correlation aims to find linear relationships between data and Spearman's correlation aims to find monotonic relationships. Even though these are fairly simple relationships, it allows us to make an initial selection of variables with strong relationships between them. By applying these two correlations between the complete data set and magnetometer 1, which represents the ground truth, independently for each time sample:

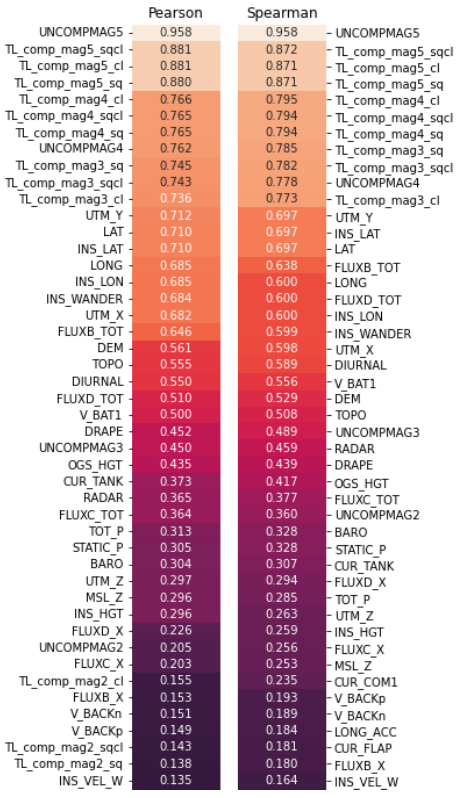


Figure 5 - Correlation between the uncompensated magnetometer 1 and the data set.

Only a portion of the data is displayed on this graph for readability. Magnetometer 5 and 4 are the most correlated magnetometers with magnetometer 1 for both correlation methods. By correlating magnetometer 5 with all the variables that do not represent magnetic measurements, it will be possible to see which variables have an impact on the magnetic field measurement inside the aircraft:

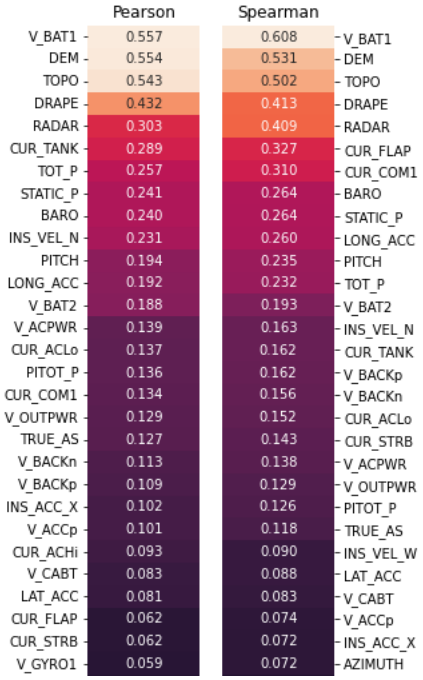


Figure 6 - Correlation between magnetometer 5 without compensation and non-magnetic data from the dataset

Only part of the data is displayed on this graph for readability reasons. For both correlation methods, the variable V_BAT1 corresponding to the number one battery of the aircraft near the cockpit seems to have a strong correlation with the magnetometer 5. It is interesting to note that the variables related to altitude are correlated with the magnetometer (TOPO, DRAPE, RADAR, TOT_P), due to the fact that the increase in altitude acts as a low pass filter on our measurements

With only these two correlation methods, it already seems interesting to use magnetometers 4 and 5 as well as the variable V_BAT1 as training data for the model.

3-3.2. Mutual information

Let's move on to another method capable of detecting more complex relationships, mutual information. Mutual information is a method capable of measuring the dependence between 2 variables. It is based on information theory and quantifies the amount of information brought

about on a variable by observing another variable [w1]. Applying this method between magnetometer 1 and the complete data set for each sample independently, it gives the following graph:

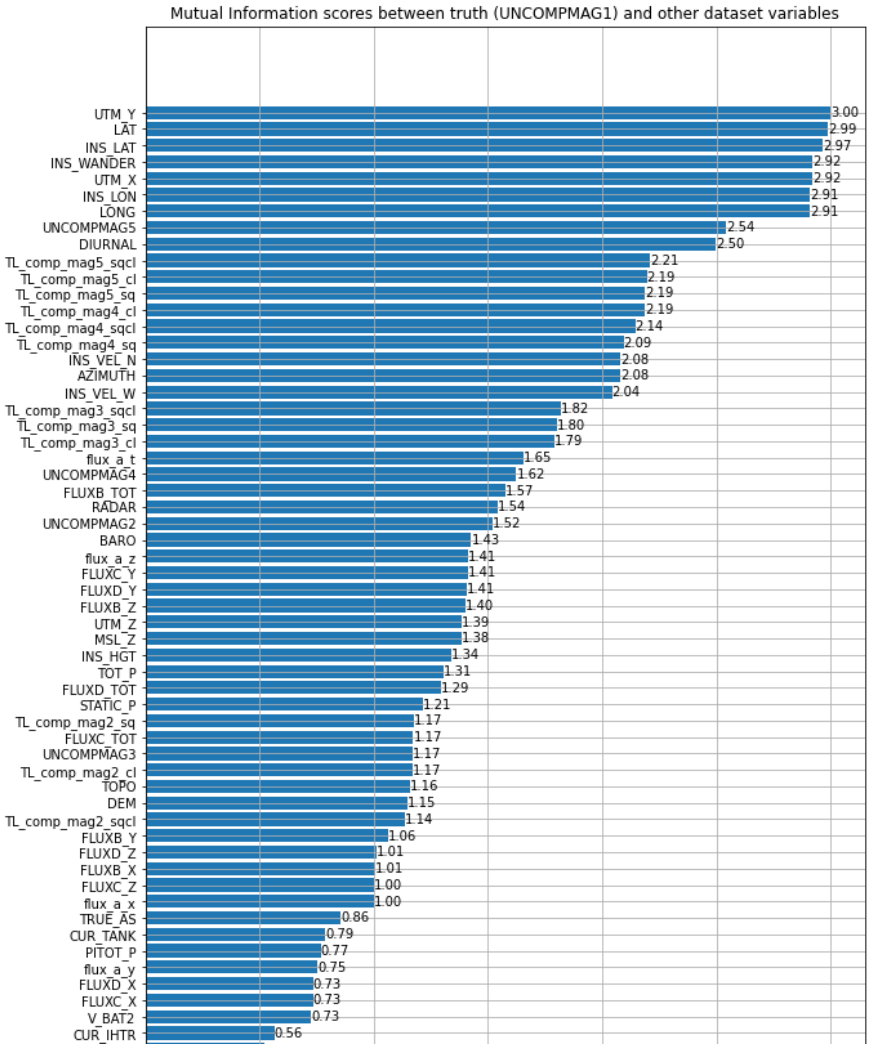


Figure 7 - Mutual information between the data set and the uncompensated magnetometer 1.

Only a part of the data is displayed on this graph for readability reasons. The most interesting variables have a mutual information greater than 1. As before, magnetometer 5 is the most correlated magnetometer with the ground truth which is magnetometer 1. So, let's reuse magnetometer 5 to compute the mutual information with the non-magnetic data of the dataset:

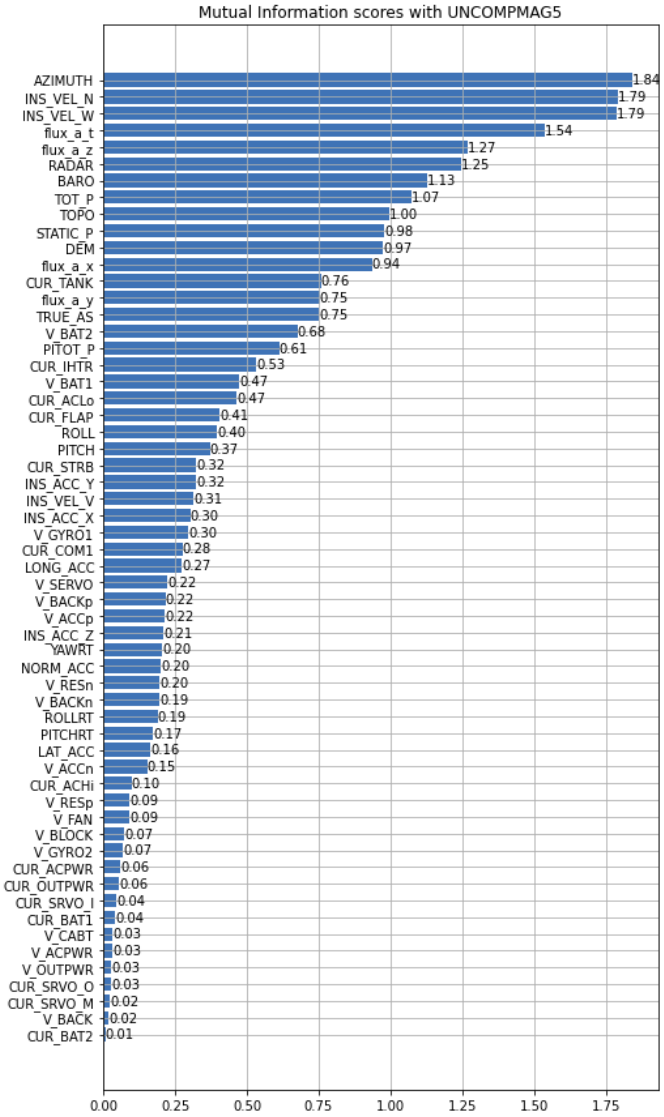


Figure 8 - Mutual information between the non-magnetic data and the uncompensated magnetometer 5.

It can be seen from the graph that the aircraft orientation and speed variables appear to share information with the magnetic measurement from magnetometer 5. This is also the case for the altitude variables. The "flux_a" variables are not considered because they correspond to the A vector magnetometer which is only available for flights 1006 and 1007. This magnetometer cannot be considered as a training variable for our model. Some electrical variables also seem to have an impact, such as "CUR_TANK". As a reminder, a description of all the variables in the dataset is available in Appendix B.

3-3.3. XGBoost

XGBoost is a machine learning algorithm based on decision trees. This model will be trained on our dataset and this allows us to see which variables have had the most impact on the

prediction of the model. Even if the prediction of the model is not very good, it allows us to have feedback on the interesting variables for the prediction and to compare the results with the previous methods and see if they seem to show the same trend. Here is the importance of the variables for the model:

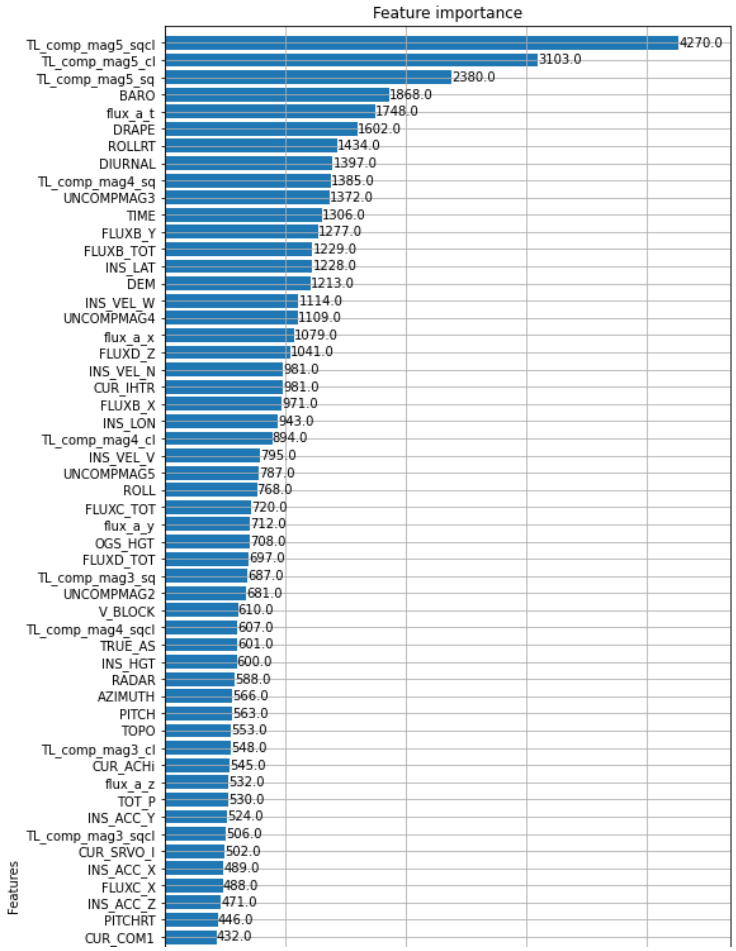


Figure 9 - Importance of variables according to XGBoost.

Only part of the data is displayed on this graph for readability reasons. The model was trained to predict the magnetic anomaly (the variable IGRFMAG1) at time t from the measurements in the dataset at time t. These measurements also include TL compensated magnetometers corrected by the IGRF model and diurnal effects. As with the previous methods, magnetometer 5 is the most useful variable for model prediction. Variables related to aircraft speed also stand out. An electrical variable also seems to have an impact, CUR_IHTR. The interesting variables are very close to those found with the previous methods.

3-3.4. New data representation

It is also possible to modify the dataset to try to get more information. To do this, the magnetometer measurements will be modified to obtain the residual between the magnetometers inside the aircraft and the magnetometer outside. Here is the transformation used:

$$residu = Mag\ N - Mag\ 1 \quad (3.1)$$

Where $Mag\ N$ is one of the four uncompensated magnetometers inside the aircraft and $Mag\ 1$ is the uncompensated magnetometer 1 on the boom at the rear of the aircraft. The same procedure will be applied for the magnetometers inside the compensated aircraft and the magnetometer at the rear of the compensated aircraft. Pearson and Spearman correlation methods will then be applied, at each time sample, between this residual and the non-magnetic variables in the dataset. The position-related variables are also excluded. These modifications result in the following graph:

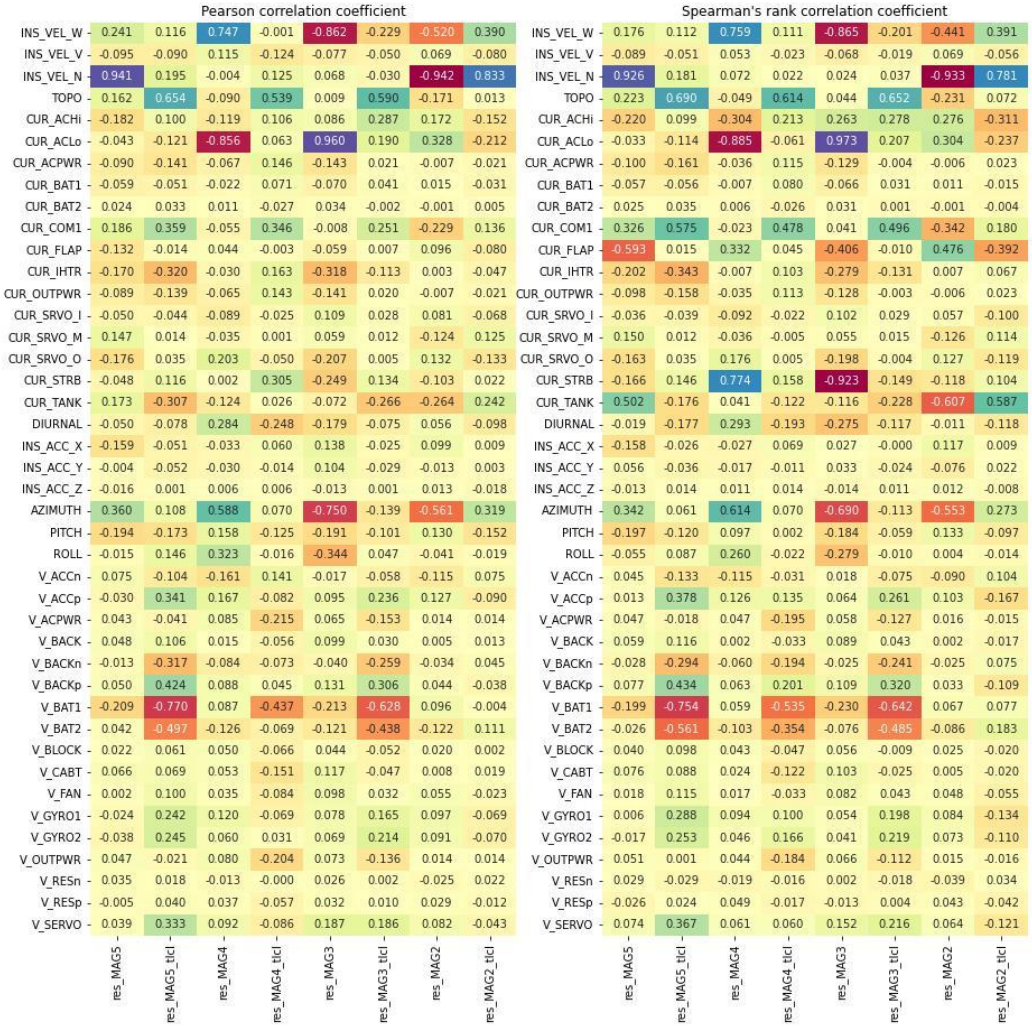


Figure 10 - Correlation between non-magnetic data and magnetometer readings

One of the first interesting things we notice is that the residuals of the compensated and uncompensated magnetometers do not show the same variables. The electrical variables V_BAT1, V_BAT2, CUR_ACLo and CUR_FLAP seem to be strongly correlated with this residual. The velocities calculated by the INS also seem to have a strong impact.

3-3.5. Selected data for the models

After having used different methods to highlight relationships between data, here is the list of data retained for training models:

- Magnetometers 4 and 5 - these are the two magnetometers that share the most information with magnetometer 1.
- INS calculated velocities - in many methods, INS calculated velocities share information with the magnetic field measurement.
- Altitude - altitude has a strong effect on the magnetic measurement and therefore it seems logical to use it as a training variable. Barometer measurements will be used
- Electrical elements - several electrical elements seem to have an impact on the magnetic measurement. The elements selected are V_BAT1 and V_BAT2, CUR_ACLo and CUR_FLAP as well as CUR_TANK and CUR_IHTR.
- Aircraft orientation - variables related to the orientation of the aircraft also appear to be of interest as they provide information on the movement of the aircraft in the Earth's field and it is known that the disturbance produced by certain components of the aircraft depends on the orientation of the aircraft in the Earth's magnetic field.

This variable selection allows to reduce the noise injected in the models but also to reduce the number of data that the model uses and thus to reduce the training time in order to avoid the plague of the dimension [w2].

4- Compensation by neural networks

In view of recent works on the subject [2, 10, 11, 12], neural networks seem to be an interesting alternative to the Tolles-Lawson model. This section aims at introducing the different networks used as well as the parameters and the learning method used. As a reminder, the aim of this work is not to find the most optimized architecture but to study if the architectures currently used for temporal data are relevant for this problem.

4-1. Automatic learning

Machine *learning is a* method based on mathematics and statistics that allows computers to learn from data without being explicitly programmed. *Deep* learning will only be used for this work. Deep learning is a sub-part of machine learning using neural network models.

Their strength comes mainly from the fact that they are able to model non-linear effects. A reminder on the different neural networks is available in appendix C. This is what allows us to

differentiate between neural networks and the Tolles-Lawson model on this problem. Tolles-Lawson is an insufficient method to model the non-linear effects that occur in an aircraft. Neural networks are known to be able to model these non-linear effects. It is notably thanks to non-linear activation functions that these networks are able to model non-linear effects.

The models used will therefore aim to predict the value of magnetometer 1, located at the end of the pole, using the data selected earlier in section 3.3. As here the value of magnetometer 1 is known and represents our ground truth, the problem is supervised.

4-2. Models used

In total 5 different models were used. First, MLPs were tested, then in a second step, networks more adapted to time series such as a CNN using 1D convolutional layers. Finally, RNNs were also tested, more particularly LSTMs. Most of the chosen architectures come from iterative testing. There were also some hyperparameter searches done by Optuna (python parameter optimization library), however there are still many ways to improve them. Generally speaking, models with few parameters seem to be the most efficient, they generalize better. Big models with many parameters get better performances on training data but they overlearn on these data and thus are much less efficient on test data. Here is the architecture chosen for each model:

- MLP - a network with 2 layers of 16 and 4 neurons.
- CNN - a network containing 2 convolutional layers with filters of size 16 and 32
- RNN - an LSTM type network with 2 LSTM cells

To have more details on the parameters of the models, please refer to the "src/models" folder on GitHub. These models will allow to test different classical architectures for time series processing. Each model aims to search for information in the data at different places. This is called inductive bias [13]. Here is a summary table to show the main differences between the models:

Table 4.1: Major differences between the different models.

Layer type	Data structure	Type of Relationship between the data	Network invariance
Fully connected neurons	Unitary	All	-
Convolutional	Grid	Local	Spatial invariance
Recurring	No time	Sequential	Time invariance

CNNs therefore tend to look for local relationships between data, whereas RNNs tend to look for sequential relationships. MLPs assume that all data are related, but this leads to a faster combinatorial explosion.

4-3. Data pre-processing

Before using the data to train the models, there are some pre-processing steps that can be done. The raw data can be used to train the models, but it is also possible to try to improve the performance with preprocessing of our data. The following preprocessing were used:

- TL compensation - application of TL compensation on our magnetometers to remove known aircraft magnetic effects.
- IGRF correction - when applying the IGRF model, this removes the variations due to the Earth's core field leaving only the crustal field and the aircraft disturbance field. In a real-time case, since the Earth's core field varies little with respect to position, the latitude and longitude calculated by the INS are sufficient to apply the IGRF model.
- Diurnal effects correction - these effects add noise to our measurements which adds error to the actual crustal field measurement.

These corrections come from physical knowledge of the subject. They allow us to remove what we already know how to model. Another type of pre-processing can be applied to accelerate the descent of the gradient but also to reduce the bias during training:

- Data scaling - scaling the data (via standardization or normalization) ensures that all data are of equal importance to the model. In the case where the orders of magnitude between different data are too large, the model tends to neglect data with small orders of magnitude.

Through the different parameter tests, it appeared that the models learn better with all the possible corrections (TL compensation, IGRF and diurnal effects correction) as well as a standardization of the data. These parameters were therefore kept.

4-4. *Training and model parameters*

All models were trained on a Lenovo ThinkStation P340 with an Intel CPU I5-10400@2.90GHz, an Nvidia GeForce GTX 1660 Ti GPU and 32 GB of RAM.

To train and validate the models under test, we need to split our data into a training set and a test set. However, care must be taken on how the data is split. In a single flight the aircraft may fly over the same position several times. This could lead to a potential leakage of the training data into the test data which is not desirable. Therefore, training data passing the same position should not be included in the test data. This is mainly the case for flight 1003 where two figure-of-eight sections overlap. The model may therefore rely on irrelevant features in the data for this flight which will result in good performance on this test set, but there is no guarantee that this will generalize to other flights. The best way to avoid this is to use an entire flight as test data for the models. This avoids putting data already seen in training into the test data but also avoids the fact that the trajectories are different.

To avoid leakage of training data into the test set, all models were tested on two different flights as follows:

- Training 1 - flights 1002, 1003, 1004 and 1006 for training and flight 1007 for testing.
- Training 2 - flights 1002, 1004, 1006 and 1007 for training and flight 1003 for testing.

This breakdown allows us to verify that the model performs well on data that it has never seen in training, but also that it is capable of obtaining good performance on different flights. Flights 1003 and 1007 are the flights with the most representative trajectories of real aircraft

flights. Flight 1005 has not been kept in order not to introduce too much data corresponding to measurement campaigns, which does not resemble at all the data of a classical flight.

As for the training parameters, here is a summary table for the different models. These parameters come from iterative explorations, they are the hyperparameters retained for the training of the models:

Table 4.2: Training parameters for each model.

Model	**Number of epochs**	**Weight regulation**	**Batch size**	**Initial learning rate**	**Size of a sequence**
MLP	25	$1e-2$	256	$1e-3$	5 time steps
CNN	25	$0.5e-2$	256	$1e-3$	80 time steps
LSTM	35	$0.5e-3$	256	$1e-3$	80 time steps

Regularization allows to restrict the weights in order to avoid the model learning to reconstruct only the training data. This allows the model to learn more generalized functions and thus a better performance on the test data. However, a too strong regularization leads the model to learn nothing. The numbers of epochs chosen correspond to the moment when the models have stopped learning and before the model overlearns. This method is called *early stopping* and avoids overlearning. The impact of the batch size is quite minimal here, values between 32 and 256 obtain similar results.

As our data correspond to time series, they will be injected into the model in the form of sequences:

$$séquence = \begin{bmatrix} V_{1_0} & \dots & V_{1_k} \\ \vdots & \vdots & \vdots \\ V_{n_0} & \dots & V_{n_k} \end{bmatrix} \tag{4.1}$$

With V_{n_k} the variable n at time step k. The number of variables corresponds to the data selected previously in section 3.3. As for the length of the sequence, it corresponds to the number of time steps injected in the model. For example, if our sequence has a length of 80, it means that we inject 80 steps of measurements for each variable in the model. As for the measurement of the reference magnetic field (the truth), it will correspond to the last time step of the sequence:

$$vérité = Mag1_k$$

(4.2)

This means that a sequence ending at time step k predicts the value of magnetometer 1 at time step k. As for the size of a sequence, a size between 80 and 160 time steps seems to give good results.

For all models, the cost function used for training is the mean error:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2$$

(4.3)

Thus, the predictions of the model $\hat{Y}$ are compared with Y the true values of the magnetic field from magnetometer 1.

The learning rate used, Lrfollows the following pattern:

$$Lr = Lr_{Epochs-1} \times 0.90^{Epochs}$$

(4.4)

This means that the learning rate decreases with the number of epochs:

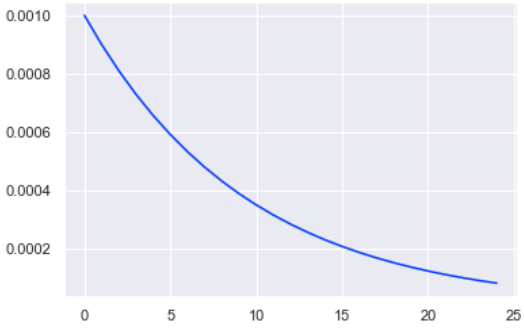


Figure 11 - Evolution of the learning rate.

This avoids making too large update steps towards the end of the training and therefore better converges to optimal parameters for the model.

5- Results

First, the Tolles-Lawson method will be applied to the data to serve as a basis for comparison. The different models will then be compared to each other and to the Tolles-Lawson compensation.

5-1. Application of Tolles-Lawson on our data

We will now apply this compensation method on our dataset. The entire code used is available online and accessible to all via the following link: https://github.com/Naatyu/MagNav. Only the results of the compensation for flight 1007, which is supposed to represent a slightly complex trajectory, are used in order to choose the compensation maneuver as well as the regularization parameter.

In order to perform a Tolles-Lawson trim, a flight performing the trim maneuvers is also required. Therefore, we will use sections 1002.02 and 1002.20:

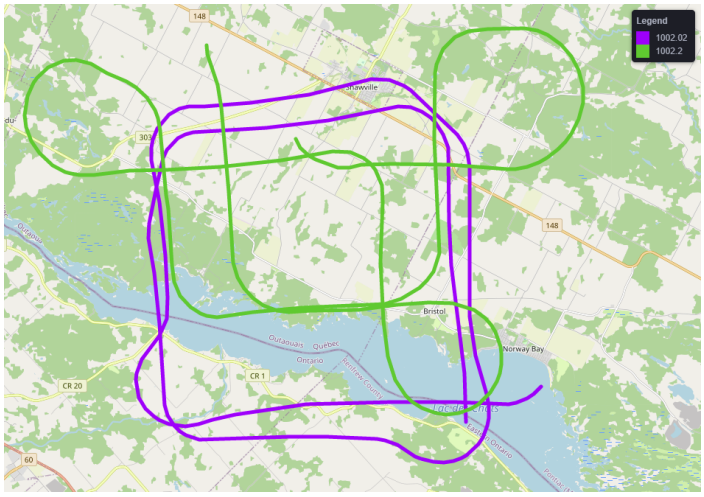


Figure 12 - Offset maneuvers 1002.02 (purple) and 1002.20 (green).

It can be seen that section 1002.02 is a square-shaped maneuver and section 1002.20 is a cloverleaf-shaped maneuver. Flight 1006 also has trim maneuvers but they were not included because flights 1006 and 1007 were added to the dataset later. These two maneuvers will be used as the basis for performing Tolles-Lawson compensation on all flights. An experimental maneuver was also studied, which simply consists of using a merge of the two sections as the Tolles-Lawson compensation maneuver. A 4th order Butterworth bandpass filter, with a first cut-off frequency of 0.1 Hz and a second cut-off frequency of 0.9 Hz, will be used for all compensations. To compare the different magnetometers, the root mean square error, i.e. the *RMSE,* is used over a complete flight. Here are the results of the compensation on flight 1007.

The COMPMAG1 data (magnetometer 1 compensation performed by SGL) from the data set is used as the reference compensation:

Table 5.1: Comparison of compensation maneuvers, regulation parameter=0.

Magnetometer	**RMSE [nT]**		
	Cloverleaf Maneuver (1002.20)	**Square Maneuver (1002.02)**	**Merged Maneuver (1002.02+1002.20)**
Mag 1	0.65	0.46	156.29
Mag 2	23540.45	17778.74	17318.59
Mag 3	133.40	163.19	172.95
Mag 4	146.01	178.35	141.58
Mag 5	43.29	46.15	133.28

Table 5.1 clearly shows the difference between the different magnetometers. Magnetometer 5 (located furthest back on the aircraft) clearly performs better than the others. It comes very close to the truth, i.e., magnetometer 1. Table 5.1 shows that for magnetometer 1, the errors are less than nT, which confirms that the compensation performed is very close to that of SGL. On the other hand, for the merged maneuver the performance is very poor. The error should be similar to the other maneuvers for magnetometer 1. The magnetometer 2 seems to be totally unusable. Figure 13 also shows a graph of all the magnetometers for a square maneuver:

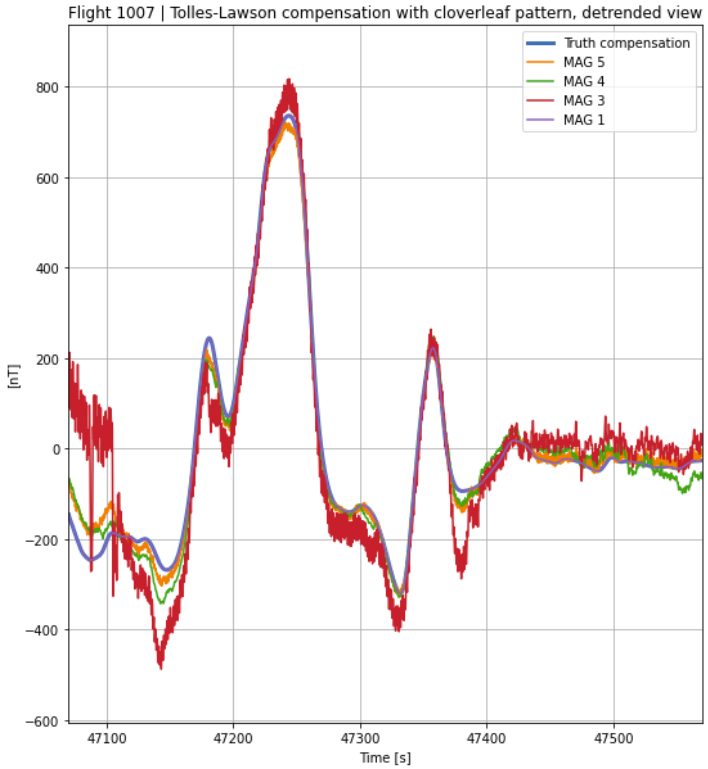


Figure 13 - Comparison of magnetometers for Tolles-Lawson compensation with a square maneuver and a regularization parameter=0

Figure 13 shows that magnetometer 5 (orange) is closest to the truth on this segment. Magnetometer 4 (green) is also fairly close to the truth on this segment. Magnetometer 3 (red) is much noisier here. Magnetometer 2 is not displayed due to its poor performance.

The best magnetometer obtains a performance around 40-45 nT. This is a good performance for a magnetometer placed inside the aircraft but not enough for magnetic navigation approaching 10 meters. Let's see if the performance changes for regularization parameters λ different from 0 in the Tolles-Lawson method:

Table 5.2: Comparison of the compensation maneuvers, regulation parameter=0.00025.

Magnetometer	**RMSE [nT]**		
	Cloverleaf Maneuver (1002.20)	**Square Maneuver (1002.02)**	**Merged Maneuver (1002.02+1002.20)**
Mag 1	0.19	0.22	32.39
Mag 2	17060.09	10611.39	14086.37
Mag 3	100.61	98.98	105.87
Mag 4	134.59	137.47	130.10
Mag 5	51.46	51.45	34.93

Table 5.3: Comparison of compensation maneuvers, regulation parameter=0.0025

Magnetometer	RMSE [nT]		
	Cloverleaf Maneuver (1002.20)	**Square Maneuver (1002.02)**	**Merged Maneuver (1002.02+1002.20)**
Mag 1	0.18	0.22	4.72
Mag 2	17170.91	10685.24	13869.15
Mag 3	100.00	98.53	99.03
Mag 4	134.69	137.62	134.76
Mag 5	51.19	51.37	47.63

Table 5.4: Comparison of compensation maneuvers, regulation parameter=0.025

Magnetometer	RMSE [nT]		
	Cloverleaf Maneuver (1002.20)	**Square Maneuver (1002.02)**	**Merged Maneuver (1002.02+1002.20)**
Mag 1	0.17	0.21	1.87
Mag 2	18171.89	11486.59	14532.58
Mag 3	100.21	99.08	98.53
Mag 4	134.86	138.21	135.96
Mag 5	51.08	51.75	50.17

Table 5.5: Comparison of compensation maneuvers, regulation parameter=0.25

Magnetometer	RMSE [nT]		
	Cloverleaf Maneuver (1002.20)	**Square Maneuver (1002.02)**	**Merged Maneuver (1002.02+1002.20)**
Mag 1	0.17	0.19	1.50
Mag 2	17476.41	11183	14882.53
Mag 3	114.50	110.80	107.79
Mag 4	134.55	138.38	136.36
Mag 5	50.37	51.24	50.50

It can be seen from these different tables that the regularization parameter improves the performance especially for the merged maneuver where magnetometer 1 is very close to the truth. Adding this parameter improves the error for all magnetometers except 5. In fact, 5-6 nT of error is added for the square and cloverleaf maneuvers. A value of 0.025 for the regularization parameter seems interesting because it gives the best performance for all magnetometers.

Despite the different parameters, the mean square error between the compensated magnetometer 1 and the best magnetometer, number 5, remains high for magnetic navigation. A compensation with an error close to 10 nanotesla would be ideal to have a position error of the order of 10 meters. To compare the TL method and a *deep learning* method, the compensation of flights 1007 and 1003 with a regularization parameter of 0.025 and a cloverleaf type compensation maneuver will serve as a basis for comparison.

5-2. Comparison of models

Once the models are trained with the selected parameters, the performance of each of these models will be compared to see if they are more efficient than the current TL method. Here is a summary table of the performance of the models:

Table 5.6: Comparison of the compensation error of the models used.

Model	RMSE [nT] vol 1003	RMSE [nT] vol 1007
Tolles-Lawson	64.44	51.08
MLP	**32.19**	28.20
CNN	36.84	**27.37**
LSTM	42.14	44.29

Overall, all the models provide improved performance over Tolles-Lawson. The most interesting models are the MLP and the CNN. Table 5.6 shows a 43% performance improvement over Tolles-Lawson for the CNN on flight 1007. The LSTM seems to perform less well than the other models but still provides an improvement in performance over the Tolles-Lawson model. It is interesting to plot the predictions of the models for flight 1007 to see their differences:

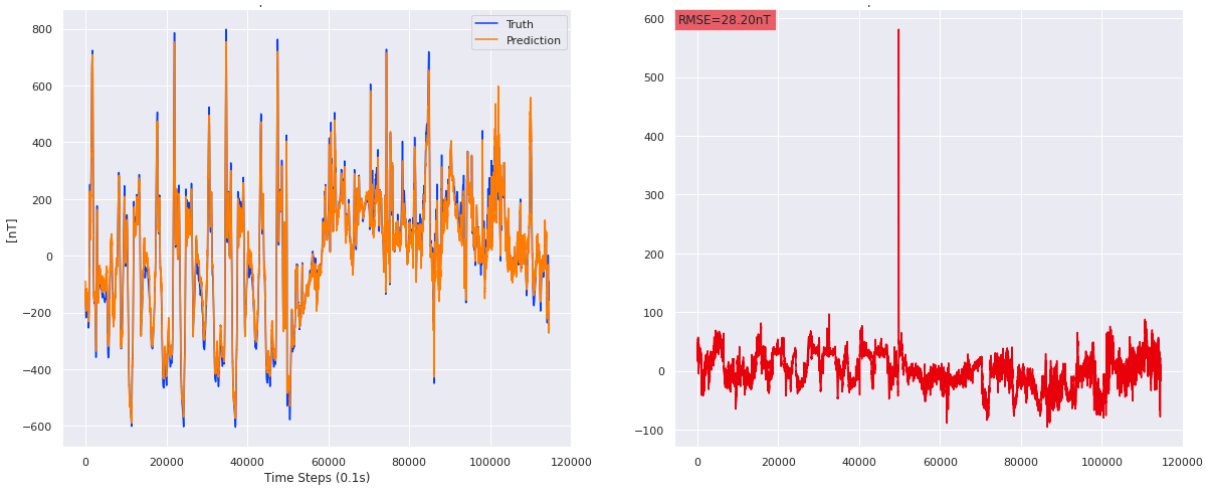


Figure 14 - Prediction and error on MLP Flight 1007.

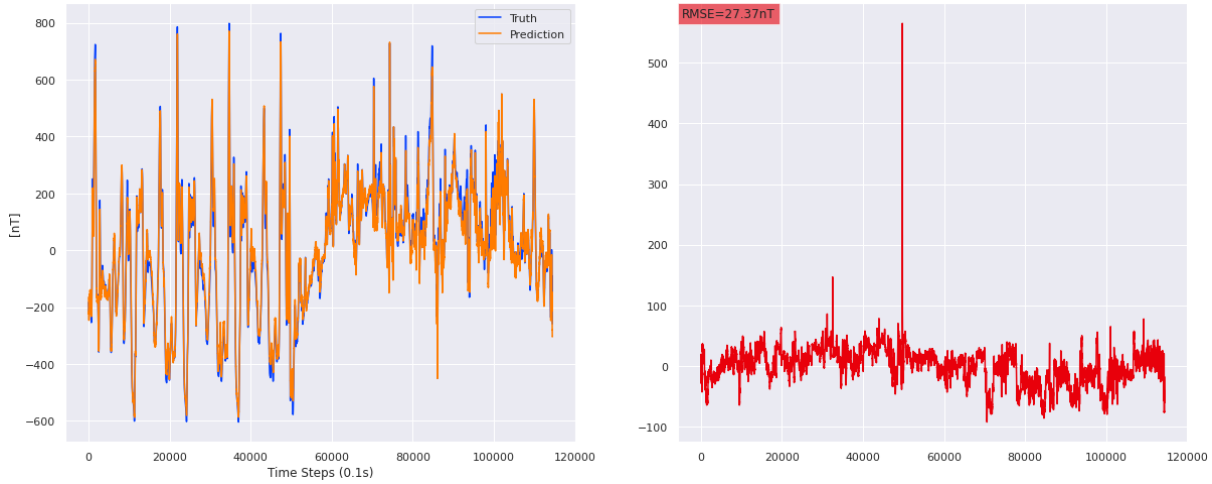


Figure 15 - Prediction and error on CNN Flight 1007.

Figure 16 - Prediction and error on LSTM Flight 1007.

Focusing on Figure 14, the MLP appears to have error jumps, where the CNN, in Figure 15, appears to be more constant. The CNN seems to be the model with the least error variation. There is a peak near the 50,000th time step on all error curves. This peak corresponds to a

measurement error of the magnetometers used. This has no impact on the performance of the models and applies only to flight 1007.

At first glance, the MLP seems more interesting than the CNN because it has a higher average error on both flights. However, when looking at the error per flight section for each flight, it is possible to see that the MLP models complex sections less well than the CNN and is very good at modeling easier sections such as straight lines. This has the effect of reducing the total RMSE on the flight but in reality, the performance on the complex sections of the flights is worse. Here is a section-by-section comparison of the two models:

Line number	RMSE [nT]
1007.01	26.38
1007.02	25.51
1007.03	28.47
1007.05	21.92
1007.06	25.46
1007.07	25.54

Line number	RMSE [nT]
1007.01	19.54
1007.02	21.63
1007.03	29.70
1007.05	19.90
1007.06	23.00
1007.07	19.30

Figure 17 - Sectional performance of MLP (left) and CNN (right) for Flight 1007.

Figure 17 shows that for the more complex sections 1007.02 and 1007.06, the CNN is systematically better. The difference is however not very big but still noticeable. The CNN thus seems to be the model of choice for our data as well as for the problem at hand.

In order to have more information on the predictions of the model, it is also possible to see at which position (longitude, latitude) the model has the most difficulty in modelling the disturbances thanks to figures 18 and 19:

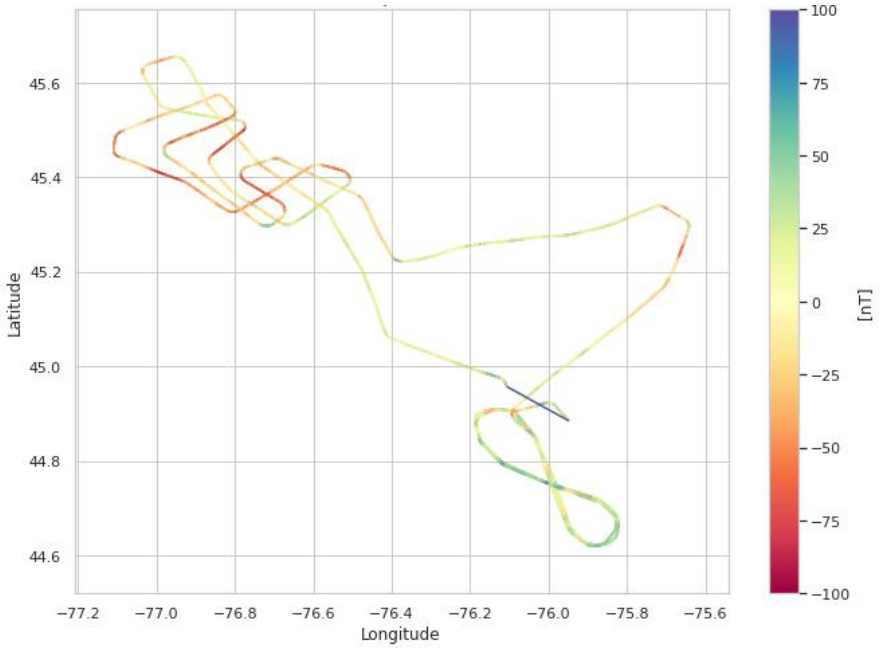


Figure 18 - 2D view of position-dependent compensation errors for the CNN.

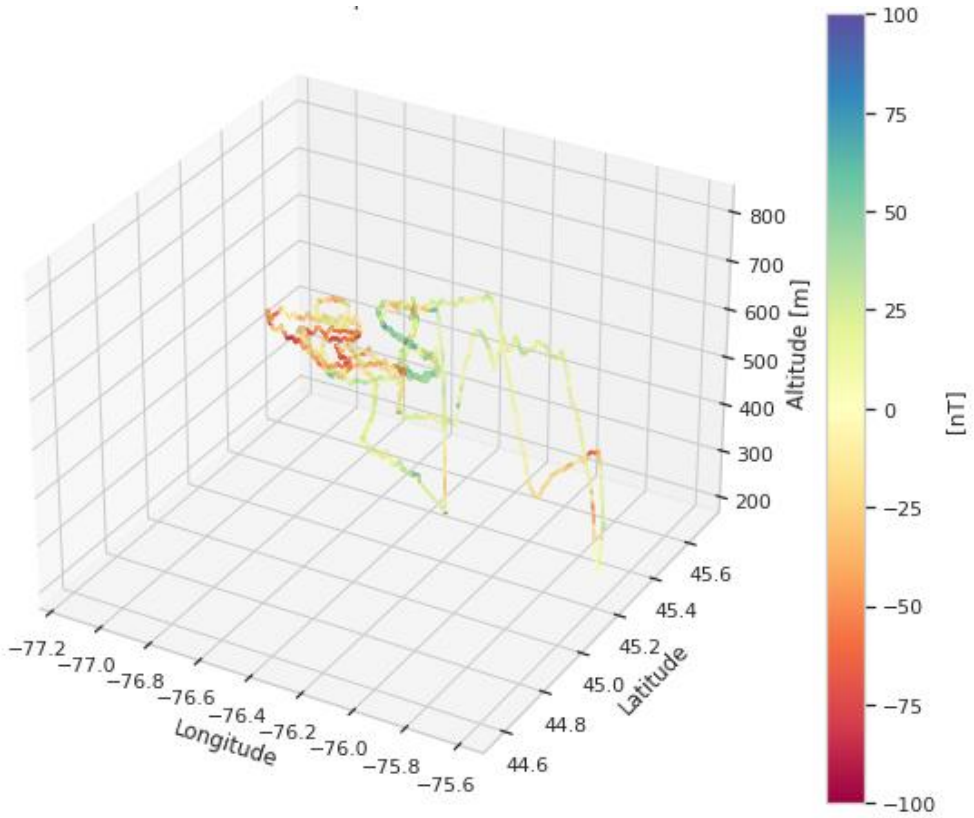


Figure 19 - 3D view of position-dependent compensation errors for the CNN.

The sections where the error is highest correspond to areas where the altitude is high and varies. Overall, altitude variation seems to have an impact on the denoising of the measurements. Since most measurements are made at constant altitude, it is possible that the

model does not see enough data corresponding to altitude variations. Certain types of maneuvers also seem to have an impact, such as turns. The main hypothesis is that when the aircraft moves, the magnetometers become noisier and therefore harder to compensate. Straight lines seem to be easier for the model to debug.

In the case of the CNN, it is possible to see the activation of its outputs according to the input data. To do this, it is possible to use the *SHAP* library which will use the gradients of the model to highlight the importance of the input data. This allows to see which data are the most useful for the prediction but also to see, on figure 20, which part of the sequence is the most used for the prediction:

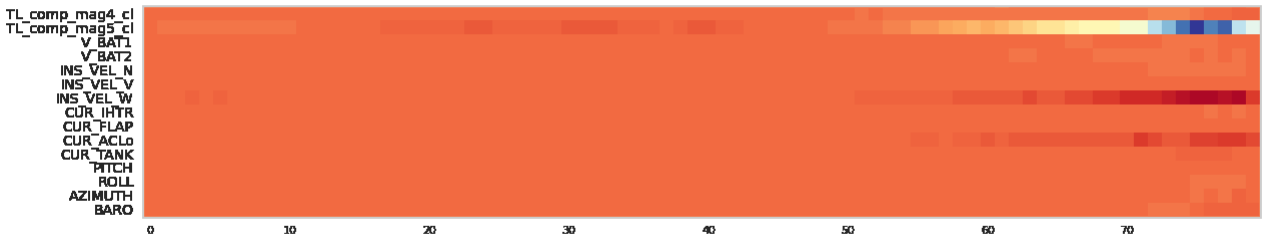


Figure 20 - Importance of training data for model prediction.

The sequence displayed corresponds to any instant of flight 1007. The model seems to rely mainly on magnetometer 5, which is the least noisy. This is consistent with what was seen earlier in the data selection section where magnetometer 5 was the most correlated magnetometer with magnetometer 1. The velocity of the aircraft also seems to be well used by the model for its prediction. More globally, it is the last 20 time steps that are mainly used for the model prediction

6- Prospects for improvement

This part will talk about the different possible points of improvement on artificial intelligence-based compensation methods.

6-1. Data

On the whole, the data set is sufficient for a qualitative study of the problem. However, to go further in the study of this problem, it is necessary to vary the data. Ideally, the trajectories of different aircraft should be spaced sufficiently far apart so that the magnetic field of the core

changes significantly. In the data set used, the same plane is used. Moreover, this plane is designed to make magnetic measurements. All flights are located west of Ottawa, it would be more interesting if the flights were located in different parts of Canada or even North America. It would also be possible to have data with more altitude variations.

This would allow us to see if artificial intelligence models can be applied to different types of aircraft in different areas.

Another possible approach would be to transform the aircraft's electrical measurements into teslas. This would put the data used to train the models under the same unit and potentially improve the way the models use this data.

6-2. Models

On the model side, there is still some optimization to be done. A combination of a CNN and an RNN could be considered. The data would first go through the CNN which would produce a map of the local relationships of the data. The data would then be passed through an RNN which would search for sequential relationships in the map produced by the CNN. This is a potential avenue that has already been used for time series prediction [14].

There is also a possibility of using ResNet-type models for CNNs adapted to time series. However, these models are more expensive to train and to use.

It is also possible to use the idea of [2], which is to use a neural network to modify the Tolles-Lawson coefficients. However, instead of using an MLP, it may be interesting to use a CNN or an RNN. It would therefore be potentially possible to find other relationships in the data than with an MLP for the prediction of Tolles-Lawson coefficients.

7- Conclusion

Magnetic navigation is a very promising way to obtain an accurate position even when GPS is unavailable. Only to be able to use it with more constrained systems such as fighter aircraft, one must be able to use the magnetic measurements made from inside the aircraft. Unfortunately, there is no current method to sufficiently compensate for the effects of the aircraft on the magnetometer measurements. The Tolles-Lawson method attempts to solve this problem but is very insufficient in the case where the magnetometers are too noisy mainly due

to the fact that this method has difficulty in modelling the non-linear magnetic disturbances of the aircraft.

New methods based on artificial intelligence seem promising. These methods allow to overcome the problem of Tolles-Lawson nonlinearity modeling. They allow to obtain up to 50% performance gain compared to the Tolles-Lawson method with a CNN. Moreover, these methods can be used in real time, which is an advantage for navigation.

However, even if the performance of these new methods is better, they are not sufficient to obtain an accuracy of the order of ten meters for navigation. Moreover, the data come from an aircraft where the magnetic disturbances are reduced. It would be necessary to see the performances on more traditional aircraft. It would also be interesting to see the performance of the models on areas very far from the training data to see if the models generalize over large areas.

Bibliography

[1] HOWARD, Karen and LUDWIGSON, Jon. Defense Navigation CapabilitiesDOD is Developing Positioning, Navigation, and Timing Technologies to Complement GPS. GOVERNMENT ACCOUNTABILITY OFFICE WASHINGTON DC, 2021. https://www.gao.gov/assets/gao-21-320sp.pdf

[2] GNADT, Albert Reuben. Advanced Aeromagnetic Compensation Models for Airborne Magnetic Anomaly Navigation. 2022. Ph.D. thesis. Massachusetts Institute of Technology. https://www.dropbox.com/s/forort8b6ilpi8j/Gnadt_Thesis.pdf

[3] MORTON, Y. Jade, VAN DIGGELEN, Frank, SPILKER JR, James J., et al. (ed.). Position, navigation, and timing technologies in the 21st century: Integrated satellite navigation, sensor systems, and civil applications, volume 1. John Wiley & Sons, 2021.

[4] TOLLES, Walter E. Magnetic field compensation system. U.S. Patent No. 2,692,970, Oct. 26, 1954. https://patents.google.com/patent/US2692970A

[5] TOLLES, Walter E. Compensation of aircraft magnetic fields. U.S. Patent No. 2,706,801, April 19, 1955. https://patents.google.com/patent/US2706801A

[6] TOLLES, W. E. and LAWSON, J. D. Magnetic compensation of MAD equipped aircraft. Airborne Instruments Lab. Mineola, NY, Rept, 1950, p. 201-1.

[7] LELIAK, Paul. Identification and evaluation of magnetic-field sources of magnetic airborne detector equipped aircraft. IRE Transactions on Aerospace and Navigational Electronics, 1961, No. 3, pp. 95-105. https://ieeexplore.ieee.org/document/4201799

[8] HAN, Qi, DOU, Zhenjia, TONG, Xiaojun, et al. A modified Tolles-Lawson model robust to the errors of the three-axis strapdown magnetometer. IEEE Geoscience and Remote Sensing Letters, 2017, vol. 14, no. 3, pp. 334-338. https://www.researchgate.net/publication/312482356_A_Modified_Tolles-Lawson_Model_Robust_to_the_Errors_of_the_Three-Axis_Strapdown_Magnetometer

[9] GNADT, Albert R., BELARGE, Joseph, CANCIANI, Aaron, et al. Signal enhancement for magnetic navigation challenge problem. arXiv preprint arXiv:2007.12158, 2020.
https://arxiv.org/abs/2007.12158

[10] HEZEL, Mitchell C. Improving aeromagnetic calibration using artificial neural networks. 2020.
https://scholar.afit.edu/etd/3589/

[11] EMERY, Kyle A. Modeling Aircraft Disturbance Fields for Magnetic Navigation Using Dense ANNs and the Novel MANNTL Architecture. 2021.
https://scholar.afit.edu/etd/4894/

[12] WILLIAMS, Peter M. Aeromagnetic compensation using neural networks. Neural Computing & Applications, 1993, vol. 1, no. 3, pp. 207-214.
https://link.springer.com/article/10.1007/BF01414949

[13] BATTAGLIA, Peter W., HAMRICK, Jessica B., BAPST, Victor, et al. Relational inductive biases, deep learning, and graph networks. arXiv preprint arXiv:1806.01261, 2018.
https://arxiv.org/pdf/1806.01261.pdf

[14] KIM, Tae-Young and CHO, Sung-Bae. Predicting residential energy consumption using CNN-LSTM neural networks. Energy, 2019, vol. 182, p. 72-81.
https://www.sciencedirect.com/science/article/abs/pii/S0360544219311223

[15] GOODFELLOW, Ian, BENGIO, Yoshua, and COURVILLE, Aaron. Deep learning. MIT press, 2016.

Netography

[w1] https://fr.wikipedia.org/wiki/Information_mutuelle, visited on 02/09/2022

[w2] https://fr.wikipedia.org/wiki/Fléau_de_la_dimension, visited on 02/09/2022

[w3] https://en.wikipedia.org/wiki/Convolutional_neural_network,visité on 03/08/2022

[w4] https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks, visited on 03/08/2022

Appendix A - Tolles-Lawson Equations and Resolution

Let's start from equation 2.3 by rewriting the Earth's magnetic sources as a single field:

$$\vec{B}_m = \vec{B}_t + \vec{B}_a \tag{A.1}$$

With $\vec{B}_m$ the total field strength measured by the magnetometer, $\vec{B}_t$ the total magnetic field of the Earth and $\vec{B}_a$ the magnetic field of the aircraft. The notations have been deliberately simplified to improve readability. A vector magnetometer measures $\vec{B}_m$ which is the vector of the measured magnetic field. A scalar magnetometer measures $|\vec{B}_m|$the intensity of $\vec{B}_m$. In the case of magnetic navigation, $|\vec{B}_t|$ which is the intensity of $\vec{B}_t$is required. By relating these terms according to equation (4.2) :

$$|\vec{B}_t|^2 = \vec{B}_t \cdot \vec{B}_t = (\vec{B}_m - \vec{B}_a) \cdot (\vec{B}_m - \vec{B}_a) \tag{A.2}$$

$$|\vec{B}_t|^2 = \vec{B}_m \cdot \vec{B}_m - 2\vec{B}_m \cdot \vec{B}_a + \vec{B}_a \cdot \vec{B}_a \tag{A.3}$$

$$|\vec{B}_t| = \sqrt{|\vec{B}_m|^2 - 2\vec{B}_m \cdot \vec{B}_a + |\vec{B}_a|^2} \tag{A.4}$$

$$|\vec{B}_t| = |\vec{B}_m| \sqrt{1 - 2\frac{\vec{B}_m \cdot \vec{B}_a}{|\vec{B}_m|^2} + \frac{|\vec{B}_a|^2}{|\vec{B}_m|^2}} \tag{A.5}$$

In the Tolles-Lawson model, it is assumed that the Earth's magnetic field is much larger than the plane's magnetic field. This implies that $\frac{|\vec{B}_a|^2}{|\vec{B}_t|^2}$ is negligible. Since the magnetic field measurement contains mainly the Earth's field, $|\vec{B}_m| \approx |\vec{B}_t|$ i.e., $\frac{|\vec{B}_a|}{|\vec{B}_m|} \approx \frac{|\vec{B}_a|}{|\vec{B}_t|}$. The previous expression can thus be simplified :

$$\left|\vec{B}_t\right| \approx \left|\vec{B}_m\right| \sqrt{1 - 2\frac{\vec{B}_m \cdot \vec{B}_a}{\left|\vec{B}_m\right|^2}}$$

(A.6)

We know from the integer series development that :

$$\sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} - \frac{x^4}{32} + \cdots$$

(A.7)

So by expanding our expression to the first order we get :

$$\left|\vec{B}_t\right| \approx \left|\vec{B}_m\right| - \frac{\vec{B}_m \cdot \vec{B}_a}{\left|\vec{B}_m\right|}$$

(A.8)

Vector magnetometers are used to calculate the director cosines, which gives us the following expression:

$$\hat{B}_m = \frac{\vec{B}_m}{\left|\vec{B}_m\right|}$$

(A.9)

Equation (4.9) can therefore be rewritten as :

$$\left|\vec{B}_t\right| \approx \left|\vec{B}_m\right| - \vec{B}_a \cdot \hat{B}_m$$

(A.10)

The term $\vec{B}_a \cdot \hat{B}_m$ is a corruption term. It is the effect of the projection of the plane's magnetic field into the Earth's field:

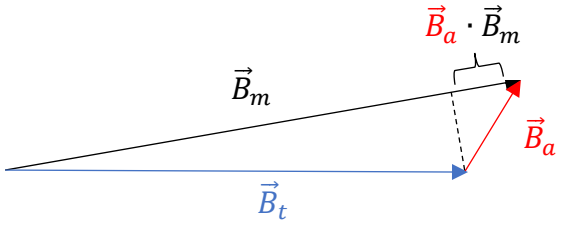


Figure 21 - Measured magnetic field vectors of the Earth and the aircraft.

Tolles-Lawson also assumes that the magnetic field of the aircraft is composed of three main effects which are:

- Permanent Effects - magnetic field created by elements with permanent magnetization. This field is static with respect to the aircraft but interacts differently with respect to the direction of the aircraft in the Earth's magnetic field.
- Remanent effects - magnetic field created by ferromagnetic elements placed in the earth's magnetic field. These effects are not static with respect to the aircraft and depend on the orientation of the aircraft. One of the elements that create the most remanent effects is usually the aircraft engine. They are also called *induced effects*.
- Eddy current effects - a magnetic field created in a conductive mass by electric currents or by the movement of that mass in a magnetic field. They are also called *eddy current effects*.

This gives us in the form of an equation:

$$\vec{B}_a = \vec{B}_{perm} + \vec{B}_{ind} + \vec{B}_{eddy} \tag{A.11}$$

The English notation will be used for consistency with other papers published on the subject. Tolles-Lawson also assumed that equation (4.12) can be written as follows:

$$\vec{B}_a = \boldsymbol{a} + \boldsymbol{B}\vec{B}_m + \boldsymbol{C}\dot{\vec{B}}_m \tag{A.12}$$

Where $\boldsymbol{a}$ is a vector and $\boldsymbol{B}$ and $\boldsymbol{C}$ are matrices of unknown coefficients. So here are the terms due to permanent effects:

$$\vec{B}_{perm} = \boldsymbol{a} = [a_1 \quad a_2 \quad a_3]^T$$

(A.13)

These three terms represent the permanent effects that are due to the aircraft components and are almost constant. These terms only change if the composition of the aircraft components is changed.

The following are also the terms of the residual effects:

$$\vec{B}_{ind} = \boldsymbol{B}\vec{B}_m = \left|\vec{B}_m\right| \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \hat{B}_m$$

(A.14)

These 9 terms correspond to residual effects. Since most of the aircraft is structurally made of non-magnetic elements, the main source of remanent effects is the engine.

Finally, the effects of eddy currents :

$$\vec{B}_{eddy} = \boldsymbol{C}\dot{\vec{B}}_m = \left|\vec{B}_m\right| \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \dot{\hat{B}}_m$$

(A.15)

These 9 terms correspond to the effects of eddy currents. These effects depend on the rate of change of the Earth's magnetic field. This phenomenon is similar to the way current is produced in a coil rotating in a uniform magnetic field.

Using expressions (4.14), (4.15) and (4.16):

$$\vec{B}_a \cdot \hat{B}_m = \left(\boldsymbol{a} + \left|\vec{B}_m\right|\boldsymbol{B}\hat{B}_m + \left|\vec{B}_m\right|\boldsymbol{C}\dot{\hat{B}}_m\right) \cdot \hat{B}_t$$

(A.16)

There are therefore a total of 21 coefficients in the Tolles-Lawson model. Due to the symmetry in the b of the after-effects, there are 18 unknown coefficients. Thus, using equation (4.11):

$$\left|\vec{B}_t\right| \approx \left|\vec{B}_m\right| - \left(\hat{B}_m^T \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \left|\vec{B}_m\right| \hat{B}_m^T \begin{bmatrix} \alpha_4 & \alpha 5 & \alpha_6 \\ \cdot & \alpha_7 & \alpha_8 \\ \cdot & \cdot & \alpha_9 \end{bmatrix} \hat{B}_m + \left|\vec{B}_m\right| \hat{B}_m^T \begin{bmatrix} \alpha_{10} & \alpha_{11} & \alpha_{12} \\ \alpha_{13} & \alpha_{14} & \alpha_{15} \\ \alpha_{16} & \alpha_{17} & \alpha_{18} \end{bmatrix} \dot{\hat{B}}_m \right)$$

(A.17)

This is the standard Tolles-Lawson model. Unfortunately, it is not possible to solve this equation without the knowledge of the coefficients α as well as $\left|\vec{B}_t\right|$. In a first step, starting from (4.18) it is possible to write :

$$\vec{\delta} = \begin{bmatrix} \hat{B}_x \\ \hat{B}_y \\ \hat{B}_z \\ \left|\vec{B}_{vec}\right| \hat{B}_x \hat{B}_x \\ \left|\vec{B}_{vec}\right| \hat{B}_x \hat{B}_y \\ \left|\vec{B}_{vec}\right| \hat{B}_x \hat{B}_z \\ \left|\vec{B}_{vec}\right| \hat{B}_y \hat{B}_y \\ \left|\vec{B}_{vec}\right| \hat{B}_y \hat{B}_z \\ \left|\vec{B}_{vec}\right| \hat{B}_z \hat{B}_z \\ \left|\vec{B}_{vec}\right| \hat{B}_x \dot{\hat{B}}_x \\ \left|\vec{B}_{vec}\right| \hat{B}_x \dot{\hat{B}}_y \\ \left|\vec{B}_{vec}\right| \hat{B}_x \dot{\hat{B}}_z \\ \left|\vec{B}_{vec}\right| \hat{B}_y \dot{\hat{B}}_x \\ \left|\vec{B}_{vec}\right| \hat{B}_y \dot{\hat{B}}_y \\ \left|\vec{B}_{vec}\right| \hat{B}_y \dot{\hat{B}}_z \\ \left|\vec{B}_{vec}\right| \hat{B}_z \dot{\hat{B}}_x \\ \left|\vec{B}_{vec}\right| \hat{B}_z \dot{\hat{B}}_y \\ \left|\vec{B}_{vec}\right| \hat{B}_z \dot{\hat{B}}_z \end{bmatrix}^T$$

(A.18)

Where $\hat{B}_x$, $\hat{B}_x$ and $\hat{B}_x$ are the director cosines calculated by the vector magnetometer magnetic field measurement and $\left|\vec{B}_{vec}\right|$ the measurement of the total magnetic field strength of the Earth by the vector magnetometer. In the case where more than one measurement is available :

$$\boldsymbol{A} = \begin{bmatrix} \vec{\delta}_1 \\ \vec{\delta}_2 \\ \vdots \\ \vec{\delta}_N \end{bmatrix}$$

(A.19)

Where A is a matrix $N \times 18$ with N each time step of measurement made by the magnetometer. With the rewritings (4.19) and (4.20) it is possible to reformulate (4.18) to obtain :

$$\left|\vec{B}_m\right| - \left|\vec{B}_t\right| = \boldsymbol{A}\alpha$$

(A.20)

With $\left|\vec{B}_t\right|$ the intensity of the Earth's magnetic field still unknown, α the 18 Tolles-Lawson coefficients still unknown and $\boldsymbol{A}$ the measurement matrix of the director cosines from the vector magnetometer.

The trick to get rid of $\left|\vec{B}_t\right|$ and thus to have only the Tolles-Lawson coefficients as an unknown is to use a finite impulse response (fpb) bandpass filter:

$$fpb\left(\left|\vec{B}_m\right| - \left|\vec{B}_t\right|\right) = fpb(\boldsymbol{A}\alpha)$$

(A.21)

A frequency range of 0.1-0.9 Hz is chosen, which has been shown to give good performance because the effects of the aircraft dominate over those of the Earth in this frequency range. However, this filtering is not perfect and therefore some effects of the Earth will not be filtered but also some effects of the aircraft will be filtered. It is therefore at this point that the data from the compensation maneuvers are used. They allow us to inject information on the aircraft's disturbances in the previously chosen frequency range. As the bandpass filter is supposed to filter out the effects of the Earth :

$$fpb\left(\left|\vec{B}_t\right|\right) \approx 0$$

(A.22)

This gives us :

$$fpb\left(\left|\vec{B}_m\right|\right) = fpb(\boldsymbol{A})\alpha$$

(A.23)

We know $|\vec{B}_m|$ which comes from the measurement of the scalar magnetometer and $\boldsymbol{A}$ which comes from the vector magnetometer measurement. We are left with the Tolles-Lawson coefficients as an unknown. The ordinary least squares method can be used for the solution,

$$\alpha = \left(\boldsymbol{A}_f^T A_f\right)^{-1} \boldsymbol{A}_f^T \left|\vec{B}_m\right|_f \tag{A.24}$$

or with regularization,

$$\alpha = \left(\boldsymbol{A}_f^T \boldsymbol{A}_f + \lambda \boldsymbol{I}\right)^{-1} \boldsymbol{A}_f^T \left|\vec{B}_m\right|_f \tag{A.25}$$

With $fpb\left(|\vec{B}_m|\right) = \left|\vec{B}_m\right|_f$ the filtered scalar magnetic field measurement, $fpb(\boldsymbol{A}) = \boldsymbol{A}_f$ the matrix containing the magnetometer measurements and λ the regularization parameter. Applying regularization is useful when $\boldsymbol{A}_f^T \boldsymbol{A}_f$ is ill-conditioned. There are different methods of determining this parameter, in this work the parameter used is based on [2].

Appendix B - List of data in the dataset

The following table lists all the available data. The vector magnetometer A is however voluntarily removed from the list because it is not present in all flights and therefore has not been studied. The original version of the table can be found in Appendix B of [9].

Name	Unit	Description
LINE	-	Line number XXXX.YY; XXXX=line and YY=section
FLT	-	Flight number
TIME	s	Seconds after midnight UTC
UTM_X	m	X coordinate, WGS-84 UTM ZONE 18N
UTM_Y	m	Y coordinate, WGS-84 UTM ZONE 18N
UTM_Z	m	Z coordinate, GPS elevation (above the WGS-84 ellipsoid)
MSL_Z	m	Z coordinate, GPS elevation (above the EGM2008 geoid)
LAT	deg	Latitude, WGS-84
LONG	deg	Longitude, WGS-84
BARO	m	Barometric altitude
RADAR	m	Filtered radar altimeter
TOPO	m	Radar topography (over the WGS-84 ellipsoid)
DEM	m	Digital Elevation Model from SRTM (over WGS-84 ellipsoid)
DRAPE	m	Planned drape of the study (over the WGS-84 ellipsoid)
PITCH	deg	Aircraft pitch calculated by INS
ROLL	deg	Aircraft roll calculated by INS
AZIMUTH	deg	Azimuth of the aircraft calculated by the INS
DIURNAL	nT	Measurement of daytime effects
COMPMAG1	nT	Mag 1: Compensated airborne magnetic field
LAGMAG1	nT	Mag 1: Offset-corrected airborne magnetic field
DCMAG1	nT	Mag 1: Airborne magnetic field corrected for diurnal effects
IGRFMAG1	nT	Mag 1: Airborne magnetic field corrected by the IGRF model and diurnal effects
UNCOMPMAG1	nT	Mag 1: Uncompensated airborne magnetic field
UNCOMPMAG2	nT	Mag 2: Uncompensated airborne magnetic field
UNCOMPMAG3	nT	Mag 3: Uncompensated airborne magnetic field
UNCOMPMAG4	nT	Mag 4: Uncompensated airborne magnetic field
UNCOMPMAG5	nT	Mag 5: Uncompensated airborne magnetic field

FLUXB_X	nT	Flux B: Fluxgate X axis
FLUXB_Y	nT	Flux B: Fluxgate Y axis
FLUXB_Z	nT	Flux B: Fluxgate Z axis
FLUXB_TOT	nT	Flow B: Total Fluxgate
FLUXC_X	nT	Flux C: Fluxgate X axis
FLUXC_Y	nT	Flux C: Fluxgate Y axis
FLUXC_Z	nT	Flux C: Fluxgate Z axis
FLUXC_TOT	nT	Flow C: Total Fluxgate
FLUXD_X	nT	Flux D: Fluxgate X axis
FLUXD_Y	nT	Flux D: Fluxgate Y axis
FLUXD_Z	nT	Flux D: Fluxgate Z axis
FLUXD_TOT	nT	Flux D: Total Fluxgate
OGS_MAG	nT	OGS measurements corrected for daytime effects
OGS_HGT	m	GPS elevation of OGS measurements (above the WGS-84 ellipsoid)
INS_ACC_X	m/s².	INS X acceleration
INS_ACC_Y	m/s².	INS Y-acceleration
INS_ACC_Z	m/s².	INS Z-acceleration
INS_WANDER	rad	Uncertainty agnel calculated by the NSI (counter-clockwise from North)
INS_LAT	rad	Latitude calculated by INS
INS_LON	rad	Longitude calculated by INS
INS_HGT	rad	Altitude calculated by INS (above the WGS-84 ellipsoid)
INS_VEL_N	m/s	Speed with respect to North calculated by INS
INS_VEL_W	m/s	Speed to the West calculated by INS
INS_VEL_V	m/s	Vertical speed calculated by INS
PITCHRT	deg/s	Pitch angle speed calculated by the avionics
ROLLRT	deg/s	Roll rate calculated by the avionics
YAWRT	deg/s	Yaw rate calculated by the avionics
LONG_ACC	m/s².	Longitudinal acceleration calculated by the avionics
LAT_ACC	m/s².	Lateral acceleration calculated by the avionics
NORM_ACC	m/s².	Normal (vertical) acceleration calculated by the avionics
TRUE_AS	m/s	Aircraft speed relative to the air mass calculated by the avionics
PITOT_P	kPa	Pitot pressure calculated by the avionics
STATIC_P	kPa	Static pressure calculated by the avionics
TOT_P	kPa	Total pressure calculated by the avionics
CUR_COM1	A	Current sensor : Aircraft COM1 radio
CUR_ACHi	A	Current sensor : Air conditioner fan in high position
CUR_ACLo	A	Current sensor : Air conditioner fan in low position
CUR_TANK	A	Current sensor : Cabine fuel pump
CUR_FLAP	A	Current sensor : Aircraft flap motor
CUR_STRB	A	Current sensor : Strobe lights

CUR_SRVO_O	A	Current sensor : INS external servo
CUR_SRVO_M	A	Current sensor : INS intermediate servo
CUR_SRVO_I	A	Current sensor : INS inner servo
CUR_IHTR	A	Current sensor : INS heating
CUR_ACPWR	A	Current sensor : Electrical power of the aircraft
CUR_OUTPWR	A	Current sensor : Electrical power output of the system
CUR_BAT1	A	Current sensor : Battery 1
CUR_BAT2	A	Current sensor : Battery 2
V_ACPWR	V	Voltage sensor: Aircraft electrical power
V_OUTPWR	V	Voltage sensor: Electrical power output of the system
V_BAT1	V	Voltage sensor: Battery 1
V_BAT2	V	Voltage sensor: Battery 2
V_RESp	V	Voltage sensor: Resolver Board +.
V_RESn	V	Voltage sensor: Resolver Board -
V_BACKp	V	Voltage sensor: Motherboard +.
V_BACKn	V	Voltage sensor: Motherboard -
V_GYRO1	V	Voltage sensor: Gyro 1
V_GYRO2	V	Voltage sensor: Gyro 2
V_ACCp	V	Voltage sensor: Accelerometer + from INS
V_ACCn	V	Voltage sensor: Accelerometer - from INS
V_BLOCK	V	Voltage sensor: Block
V_BACK	V	Voltage sensor: Backplane
V_SERVO	V	Voltage sensor: Servo motors
V_CABT	V	Voltage sensor: Electrical cabinet
V_FAN	V	Voltage sensor: Cooling fan

Appendix C - Neural Networks

This appendix aims to introduce the basics in order to understand the different neural networks used in this work.

C.1 - Multilayer Perceptron

Multilayer perceptron (MLP), also called *feedforward neural networks* or dense network is one of the most widespread models in deep learning, they are the basis of many other models like CNN or RNN that we will see later. The goal of an MLP is to approximate an f function. A concrete example could be a classification function $y = f(x)$ that associates x to a category y. The MLP will then define a function f^* such that $y = f^*(x\,;\theta)$ and learn the parameters θ that give the best approximation of f [14]. MLPs are known to be universal function approximators.

These networks are called *feedforward* because the information is propagated through layers of neurons, i.e., from x à y through the function f^* determined by the network. They are multi-layer networks because they have several layers, e.g., $f^*(x) = f^{*(3)}\left(f^{*(2)}\left(f^{*(1)}(x)\right)\right)$. This type of chain structure is globally used for neural networks. So, we have $f^{*(1)}$ the first layer of the network then $f^{*(2)}$ the second layer of the network etc. The last layer is called the *output layer* which produces y. The intermediate layers are called *hidden layers* and are layers defined before the training step where the learning algorithm modifies the parameters to obtain the best approximation of f. They are hidden because we usually do not look at the output of these layers but at the output layer.

C.2 - A neuron

The layers of an MLP are made up of neurons, also called *nodes* in the English literature. The neurons of a network are strongly inspired by those of our brain. The neurons have a weight and a bias and it is these parameters that constitute f^* that we are looking for to approximate as well as possible f. A neuron in an MLP can be modeled according to the following equation:

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

(C.1)

With z_j the output of the neuron j, w_{ji} the weight number i of the neuron j, x_i the input i in the neuron j, n the number of weights in the neuron and b_j the bais of the neuron j. In MLPs, also called dense networks, all neurons are connected to each other. So, each neuron receives as input the output of all the neurons of the previous layer except for the input layer where it is only our input variables.

C.2.1 - Activation function

The output of the neuron is then passed into an activation function. This function is essential for neural networks because it allows to add non linearity. Some functions are widely used like ReLU, tanh or sigmoid. Figure 22 is a graph of these different functions and their equations:

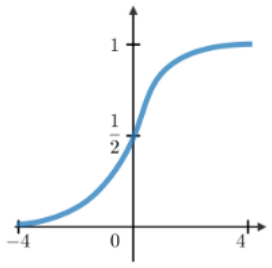
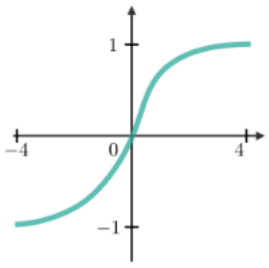
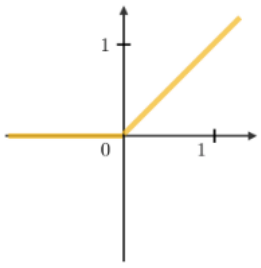
Sigmoid	Tanh	RELU
$g(z) = \frac{1}{1+e^{-z}}$	$g(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$	$g(z) = \max(0, z)$
		

Figure 22 - Different activation functions for a neural network.

These functions are generally designed so that the gradient is as relevant as possible because it is thanks to it that the network can learn its weights and biases. So, we have the output after the activation function:

$$h_j = g_j(z_j)$$

(C.2)

However, the sigmoid and tanh functions are more generally used as activation functions for the output layer because they give a result in the form of a normalized score. They saturate very quickly when the input is very large or very small and the inputs are only relevant when they

are close to 0. They are therefore not very suitable for the intermediate layers of the network, i.e. the *hidden layers*, because they make gradient learning more difficult [14].

C.2.2 - Learning

Neural networks need to learn the weights w and biases b of f^* that best approximate f. For this they use a process called *backpropagation* (BP). This algorithm is used for learning *feedforward neural networks*, another type of algorithm is used for recurrent networks (BPTT, Backpropagation Through time).

The algorithm calculates the gradient of the cost function with respect to the network weights using the *chain rule*. The cost function calculates the performance of the model by telling it how far it is from the objective it should reach. The *cross-entropy is an* example of a cost function widely used for classification. In this work, we perform a regression task and the cost function used will be the RMSE (*Root Mean Squared Error*). Figure 23 shows other cost functions that can be used for regression:

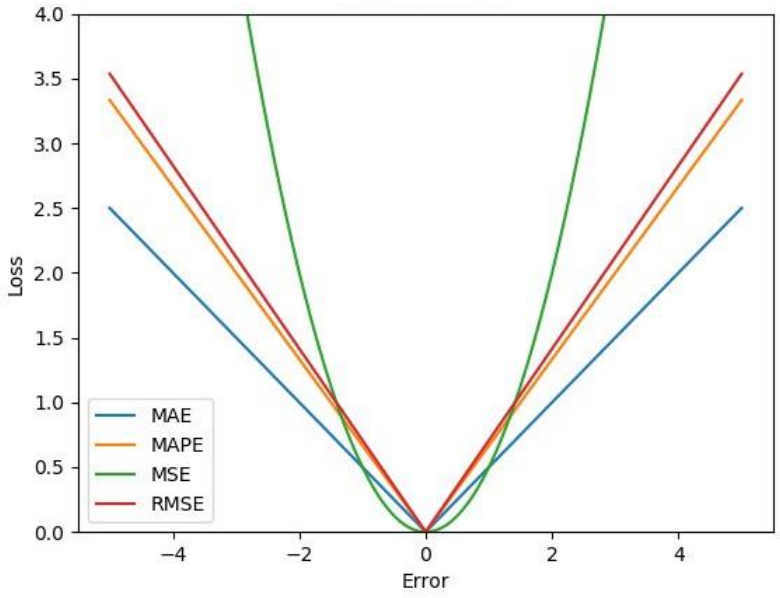


Figure 23 - Cost functions for a regression task.

The algorithm therefore calculates the gradient of the cost function with respect to the weights layer by layer and then the weights are updated by going in the opposite direction to the gradient. The gradient is multiplied by a learning rate that regulates how much the weights and biases are updated by the gradient. A low learning rate leads to long computation times and the possibility of getting stuck. Conversely, a learning rate that is too high leads to instability

in learning. We have the following equations for the update of the weights and biases by the *backpropagation*:

$$w \coloneqq w - \epsilon \frac{\partial C}{\partial w}$$
$$b \coloneqq b - \epsilon \frac{\partial C}{\partial b} \quad \text{(C.3)}$$

We have w and b the weights and biases of the network which are randomly chosen during the initialization, ϵ the learning rate which defines the influence of the gradient on the weights and biases and $\frac{\partial C}{\partial w}$, $\frac{\partial C}{\partial b}$ which are respectively the partial derivative of the cost function with respect to the weights and with respect to the biases. This is how the model learns the parameters of the function f^* to best fit the function f.

C.3 - Convolutional neural networks

Convolutional *neural* networks are networks specialized in the processing of data with a grid topology. For example, time series are one-dimensional grids and images are two-dimensional grids. Convolutional networks are mainly known for their efficiency on image/video data. These networks are also called *space invariant artificial neural networks* because they share within the same architecture the weights of convolution filters (more details in the next section) that move on the input data, which provides equivariant responses in translation called *feature maps* [w3]. In general, convolutional networks input data into convolutional layers, and then the output of these convolutional layers is then passed into an MLP which outputs the desired value such as a classification for example. Since MLPs have already been presented in section 5.1, only the components of convolutional layers are presented here.

C.3.1 - Convolutional layer

The convolutional layers of the network are simply convolution operations between our input data and the *kernel*. The kernel of a convolutional layer corresponds to a filter that is used to extract features from the input data. It is a matrix generally of size 3×3 by default for

computer vision tasks. Figure 24 shows an example of applying a convolutional layer filter to input data:

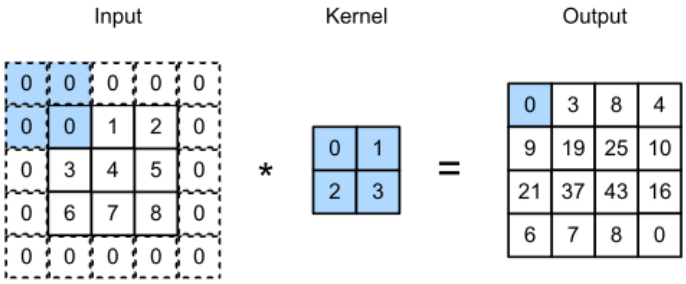


Figure 24 - Application of a convolutional layer filter.

In machine learning, we work mainly with discretized data so we use the discrete convolution operation. In the case of a two-dimensional image I and a two-dimensional kernel N we have the following convolution operation [14]:

$$S(i,j) = (I * N)(i,j) = \sum_m \sum_n I(m,n)N(i-m, j-n) \quad \text{(C.4)}$$

In reality, in most machine learning libraries it is the *cross-correlation* that is used because it avoids returning the kernel :

$$S(i,j) = (I * N)(i,j) = \sum_m \sum_n I(i+m, j+n)N(m,n) \quad \text{(C.5)}$$

The learning algorithm will therefore learn the values of the kernel appropriate to the problem. Following the convolution layer, an activation function is applied, usually ReLU.

C.3.2 - Pooling layer

After applying the convolutional layer and the activation function, we apply a *pooling* operation. The *pooling is* similar to a sub-sampling function that allows to reduce the size of the data while keeping the main information. This reduces the amount of parameters in the network and therefore improves the computation time. The most common example is *max pooling* 2×2 which amounts to compressing the data by a factor of 4 by the maximum value. Figure 25 is an example of *max pooling*:

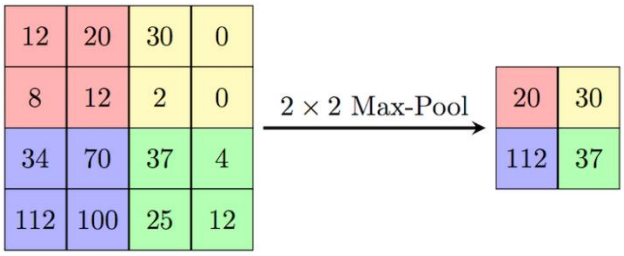


Figure 25 - Max-Pooling 2×2.

There are other operations such as *average pooling* or *L2-norm pooling* which each allow to highlight different characteristics in the data. As a general rule, *max pooling is the* most used because of its efficiency and the fact that it highlights strong network activations. The result is then sent to one or more dense layers for regression or classification tasks.

C.4 - Recurrent neural networks

Recurrent *neural networks* (RNN) are a family of neural networks specialized in the processing of sequential data.

These networks are mainly used with time series. They are the models mainly used for speech or handwriting recognition. They process the data one by one for each sequence. Their difference with a classical neural network is their internal memory updated at each time step. This allows them to keep in memory the data previously seen in the data sequence. That's why this type of model is widely used with sequential data.

RNNs can be seen as a sequence of classical neural networks where the output of one network corresponds to the input of the next by sharing the weights.

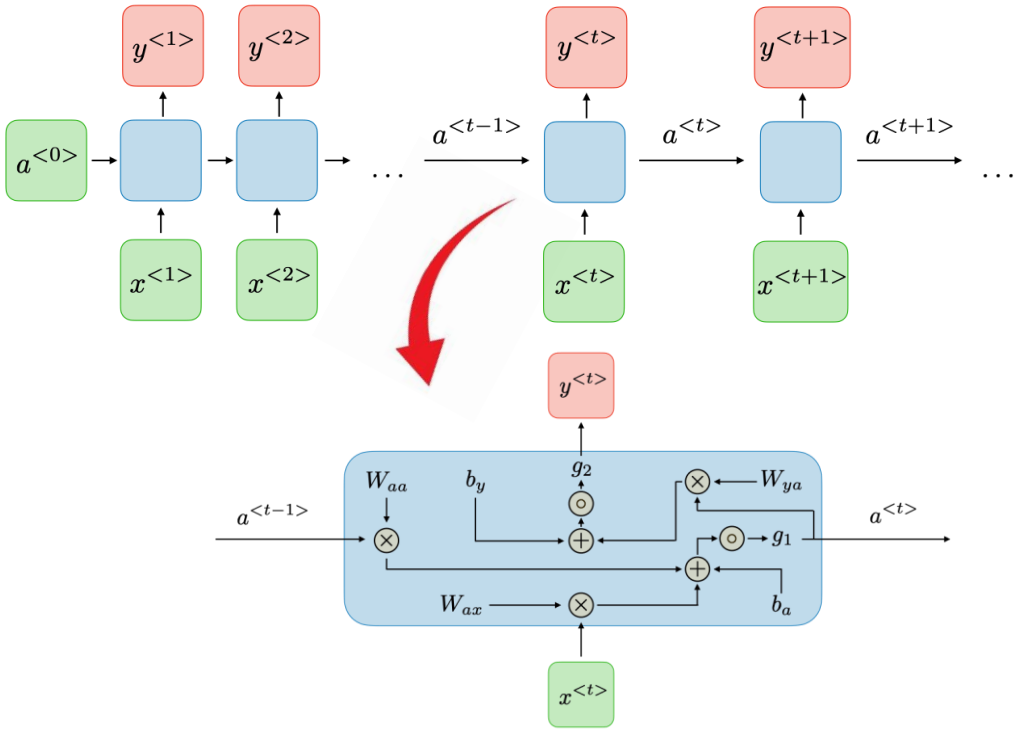


Figure 26 - Architecture of an RNN. Source: stanford.edu

If we look at the network equations [w4] :

$$a^{<t>} = g_1(W_{aa}a^{<t-1>} + W_{ax}x^{<t>} + b_a$$
$$y^{<t>} = g_2\left(W_{ya}a^{<t>} + b_y\right)$$

(C.6)

With W_{ax}, W_{aa}, W_{ya}, b_a and b_y the weights and biases of the network shared in time and g_1 and g_2 the activation functions. The equations highlight the recurrent aspect of these networks, the computation at time t depends on information of the instant $t-1$ which itself depends on the instant $t-2$ and so on, where the output of classical neural networks depends only on the input values.

Only these networks are subject to the *vanishing gradient* problem. This means that the value of the gradient becomes so small that no more updating of the weights is done. It is difficult to train these networks to keep the long-term dependency because of the number of layers which exponentially increases the number of multiplications of the gradient and thus amplifies the effect of *vanishing gradient*. To overcome this problem, new architectures have been introduced such as LSTM or GRU which allow a better control of the information conservation and thus a better propagation of the gradient.