



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Escola Tècnica Superior d'Enginyeria
de Telecomunicació de Barcelona

Design of algorithms for improving resilience of sensors in space exploration

MASTER THESIS

SUBMITTED TO THE FACULTY OF THE ESCOLA TÈCNICA D' ENGIYERIA DE
TELECOMUNICACIÓ DE BARCELONA

UNIVERSITAT POLITÈCNICA DE CATALUNYA

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE IN
TELECOMMUNICATIONS ENGINEERING

WRITTEN BY

Yi Sheng

UNDER THE DIRECTION OF

Univ. Prof. Dr. Dominguez Pumar, Manuel M. Univ.

Prof. Dr. Sayrol Cloles, Elisa

May 2021

Abstract

The main goal of this thesis is to design and implement software, based on artificial neural networks, capable of predicting data values for failures in the wind sensor aboard NASA's rovers. To achieve this objective, it is necessary to understand the context and operation of the wind sensors, designed by the UPC-ETSETB. Distinct conditions and situations are described to characterize the variables and neural networks proposed. This project demonstrates that it is possible to obtain variables, which entails that the behavior of a particular sensor can be predicted by the conduct of the rest.

In the process of the project, different machine learning models are constructed according to different data sets and target results. Then, the performance of the model is evaluated by such indicators as error rate. The result of data recovery is measured by root mean square error. At last, the model results are compared with the data interpolation results, and the conclusion is drawn. The result of model prediction is better than that of data interpolation.

These models show promising results with different input data and might provide robustness to the measurements of wind sensors. At the same time, the more input data, the higher the accuracy of output results.

Acknowledgments

After nearly three years of master's study in ETSETB-UPC, it is close to the end. Looking back on this experience of learning, too much emotion is hard to express. Although I have made a lot of hard efforts, I deeply feel that all the efforts are worth it.

First of all, I would like to express my heartfelt gratitude to my tutors, Prof. Manuel M. Dominguez Pumar, and Prof. Elisa Sayrol Clols. As the whole thesis is completed in the case of remote, so the two teachers and I have a regular meeting every week. In the meeting, we need to discuss the current experimental direction of the project, whether the experimental progress is in line with expectations and other practical issues. In addition, the main communication need to be completed by email. Teleworking is also a very energy consuming thing. Thank you very much for the teacher's weekly guidance, quick email reply and the ideas provided to me in the face of problems.

At the same time, I would like to express my heartfelt gratitude to all the teachers who have taught, cared for, and helped me in the three years of study at ETSETB-UPC.

I'd like to thank my parents for their kindness in raising me. They silently provide me with all possible help in my life and study, so that I can enjoy the warmth of my family. Your empathy, comfort and moral support have been fundamental to achieve whatever I proposed to. Your philosophy of effort, your spirit of perseverance, and your good manners are reflected in who I am now.

To my friends and university colleagues, thank you for your unconditional support, patience and advice during these years. Your energy, attitude, and desire to learn have inspired me from day one. Without a shadow of a doubt, I would not be here if it were not in part for your company and friendship.

The following pages below symbolise the fulfilment of years of effort and hard work that have opened me a passionate world of knowledge and opportunities.

Thank you all for making this possible.

Revision history and approval record

Revision	Date	Purpose
0	20/04/2021	Document creation
1	07/05/2021	Document revision
2	14/05/2021	Document revision

Written by:		Reviewed and approved by:	
Date	12/05/2021	Date	18/05/2021
Name	Yi Sheng	Name	Manuel Dominguez-Pumar Elisa Sayrol Cloles
Position	Project Author	Position	Project Supervisor

Contents

Abstract	2
Acknowledgements	3
Revision history and approval record	4
Contents	5
List of Figures	7
List of Tables	9
1. Introduction	10
1.1 Statement of Purpose	11
1.2 Requirement	13
1.3 Methodology and Procedure	13
1.4 Work Plan	14
1.5 Deviations and Modifications	15
2. Context	16
2.1 Mars Exploration	16
2.2 Insight Mission	18
2.2.1 Science	20
2.2.2 Temperature and Wind for Insight	22
3. State of the Art	24
3.1 Artificial Intelligence	24
3.2 Regression Prediction	25
3.2.1MLP	26
3.2.2 LSTM	28
3.3 DSSM	29
3.4 Other Methods	29
4. Project Development	29
4.1 Data Set	29
4.1.1Input/Output Variables	29
4.2 Development of the Neural Network Architecture	31
4.3 Prediction from Data for Single Point Model	33
4.3.1 Prediction from Two Booms for Single Point Model	33
4.3.2 Prediction from One Boom for Single Point Model	35
4.4 Prediction from Data for Two Consecutive Points Model	37
4.4.1 Prediction from Two Booms for Two Consecutive Points Model ---	37

4.4.2	Prediction from One Boom for Two Consecutive Points Model	----38
4.5	Data Interpolation	-----39
5.	Results	-----39
5.1	Prediction from Data for Single Point	-----39
5.1.1	Prediction from Two Booms Data for Single Point	-----39
5.1.2	Prediction from One Booms Data for Single Point	-----43
5.2	Prediction from Data for Two Consecutive Points	-----47
5.2.1	Prediction from Two Booms Data for Two Consecutive Points	----47
5.2.2	Prediction from One Boom Data for Two Consecutive Points	----53
6.	Budget	-----60
7.	Conclusion	-----61

List of Figures

Figure 1: One InSight's TWIN Boom. NASA/JPL-Caltech

Figure 2: Artist rendition of the InSight lander. NASA/JPL-Caltech

Figure 3: Work plan diagram

Figure 4: Another artist's concept showing NASA's Perseverance rover on the surface of Mars. NASA/JPL-Caltech

Figure 5: NASA's Curiosity Mars rover used two cameras to create this selfie in front of Mont Marcou, a rock outcrop that stands 20 feet (6 meters) tall. The panorama is made up of 60 images from the MAHLI camera on the rover's robotic arm along with 11 images from the Mastcam on the mast, or "head," of the rover. NASA/JPL-Caltech/MSSS

Figure 6: An artist's rendition of the InSight lander operating on the surface of Mars.

Figure 7: Earth picture

Figure 8: Mars picture

Figure 9: TWINS Booms ready to be installed. CAB/INTA-CSIC

Figure 10: TWINS Block diagram (Boom Y(-Y))

Figure 11: Four dice array orientation. [20]

Figure 12: Proposed model to predict with data from 2 Booms

Figure 13: Proposed model to predict with data from one Boom.

Figure 14: Proposed model to predict with data from two Boom

Figure 15: Proposed model to predict with data from one Boom

Figure 16: Training process

Figure 17: Time consumption for every step

Figure 18: Normalized data plot

Figure 19: Unnormalized data plot BLONG

Figure 20: Unnormalized data plot BTRANS

Figure 21: Unnormalized data plot BLONG expand the time

Figure 22: Unnormalized data plot BTRANS expand the time

Figure 23: Calculate the RMS for two parameters

Figure 24: Training process for single-point prediction from one boom

Figure 25: Data plotting for normalized point

Figure 26: Data plotting for normalized point BTRANS

Figure 27: Data plotting for normalized point BLONG

Figure 28: Data plotting for normalized point BTRANS time expand

Figure 29: Data plotting for normalized point BLONG time expand

Figure 30: RMS for two parameters

Figure 31: Training process for two consecutive points from two booms

Figure 32 33: Normalized data plot for two consecutive points from two booms

Figure 34 35: Unnormalized data plot for two consecutive points from two booms BTRANS

Figure 36 37: Unnormalized data plot for two consecutive points from two booms BLONG

Figure 38 39: Unnormalized data plot for two consecutive points from two booms BTRANS

Figure 40 41: Unnormalized data plot for two consecutive points from two booms BLONG

Figure 42: RMS for two points of each BLONG and BTRANS

Figure 43: Training process for two points from one boom data

Figure 44: Normalized data plot for two points from one boom data

Figure 45 46: Unnormalized data plot for two points from one boom data BTRANS

Figure 47 48: Unnormalized data plot for two points from one boom data BLONG

Figure 49 50: Unnormalized data plot for two points from one boom data BTRANS second point

Figure 51 52: Unnormalized data plot for two points from one boom data BTRANS second point

Figure 53: RMS for two points four parameters

List of Tables

Table 1: Comparison of Mars environment and earth environment

Table 2: Dataset split configuration

Table 3: Training process diagram

1.Introduction

“That’s one small step for a man, one giant leap for mankind”.

– Neil Armstrong

The need to search for answers diligently, as well as the instinct to explore and understand the things around us, has always been the driving force of human development and progress to today. On July 20, 1969, Apollo 11 effectively landed on the surface of the moon, which became a milestone in history and opened the limits of exploration and future missions. It is considered the end of the so-called space race of the 20th century.

Recently, an important trend of returning to the moon has become increasingly obvious, leading to accelerated exploration activities in the next few years, focusing on the development of technologies and systems to enable a series of human and robot missions to be extended to Mars. In May 2019, NASA officially announced the launch of the Artemis program, which will be divided into three phases. The first stage: in 2020, the Orion spacecraft will complete the unmanned flight around the moon and return to earth safely. The second stage: in 2022, Orion spacecraft will realize manned circumlunar flight. The third stage: by 2024, Orion will carry two astronauts to the moon.

Mars is a planet of the solar system. It is very similar to the earth. It has four seasons, atmosphere and weather changes, fixed forms, and ice and snow (polar crowns) at the poles. Therefore, Mars is called small earth. From 1946 to now, the United States has launched about 20 Mars probes. The main reason why Mars has been focused on by scientists is that we have the conditions to land on Mars. The spacesuit made by human beings now can land safely on Mars only after a little processing! In addition, Mars and the earth are relatively suitable in distance and the environment is closer to the earth. If we transform Mars, it will be easier to do, at least much easier than other planets, so scientists will pay close attention to Mars. The expansion of human beings across the solar system, while bringing new knowledge and opportunities to the earth, seems to be a science fiction idea. Today, however, we can learn about the formation and evolution of the planet, its history and geological process, the potential of the planet to carry life, and the future exploration of Mars by human beings through robot exploration of certain areas of

Mars. Over the years, artificial intelligence research has become interdisciplinary. In addition to the dominance of computer vision research, including image and video processing, neural computing, robotics, signal processing, and language disciplines have also conducted a lot of research. The emergence of machines is accompanied by the industrial revolution, to a certain extent, it plays a strong role in promoting the development of society. With the development and progress of science and technology, based on the remarkable development of the machine field, improving its intelligence has become the main direction of machine development at this stage. Artificial intelligence has emerged in the vision of human society and has become a typical representative of the new industrial revolution era. Through the simulation of human thinking, highlights the use-value and becomes a new research discipline. These incredible achievements are the result of rigorous technological progress and human ability to develop and adopt new technologies in current research or business, as well as the ability to introduce these systems into the daily lives of each of us. Artificial intelligence (AI) has become a new technology trend with the improvement of advanced algorithms, computing power, and storage capacity. More precisely, this master thesis is included in the concept of deep learning (DL) due to the proposed strategies, procedures, and architecture.

1.1 Statement of Purpose

The high risks and compromises faced by space missions are well known, so it needs the efforts of many researchers to make these missions perfect in all aspects. In this process, researchers also need to consider some common situations. Like any other electronic equipment, the sensors and experimental tools on these detectors also have their life cycles. In the process of data detection, storage, and transmission, some external factors will cause the data not to be detected or stored in a certain period. There is no way to ignore these situations.

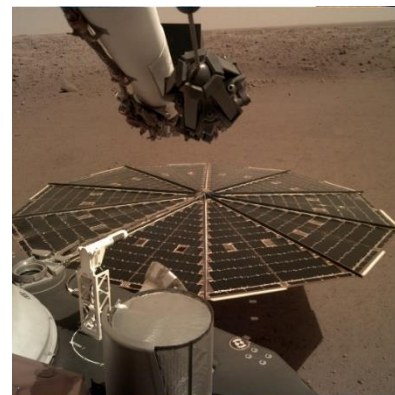


Figure 1: One InSight's TWIN Boom.

NASA/JPL-Caltech

On November 26th of 2018, NASA fulfills the ambitious mission of landing the InSight lander (Interior Exploration using Seismic Investigations, Geodesy, and Heat Transport) on Martian soil. This project focuses on the Temperature and Wind for InSight (TWINS), two finger-like Booms on the lander deck of the InSight. They monitor measurements of wind speed and direction, pressure, and air temperature, essential features to characterize the climate and atmosphere of Mars.

The ATS consist of two thin rods of FR4 material, each attached to one of the REMS booms at the mast of the MSL Curiosity rover. Each rod has three bonded Minisens RTD Thermistors Pt1000 class A, which provide temperature readings with an accuracy of 5 K and a resolution of ± 0.1 K.[19] This project counts on an extensive dataset based on real air temperature measurements taken by InSight on Martian soil, which makes incalculable the award that means working with this type of information.

The main goal of this project is to find a reasonable model, using the existing data as input, to calculate the previous missing WS data. Like any other study, the first thoughts came as questions: ‘Can we estimate values from a sensor, which is not capturing data, based on other equal sensors differently placed around? Therefore, the challenging objectives are the following: 1) characterize the behavior of temperature variables over time, 2) implement a Deep Learning neural network to predict wind data values for an out of service wind sensor, and 3) validate the results obtained and determine the performance of the neural network.

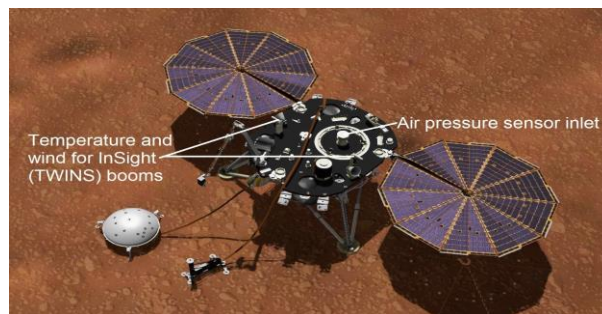


Figure 2: Artist rendition of the InSight lander. NASA/JPL-Caltech

1.2 Requirement

The requirements that have been proposed and that are necessary to achieve the project objectives are the following ones:

- 1.Design an input data loader capable of inserting specific data into the network in an appropriate configuration.
- 2.Implement software that can estimate values from the WS in a specific.
- 3.The resulting estimated values must be compared with the real values of the dataset, to obtain an error or a qualitative difference between these samples.
4. Using data insertion, the obtained results are compared with the real results to check the relevant errors.
- 5.Complete the prediction of a single point and continuous point.
- 6.The judgment of the final result, the difference of the results in different ways.

The stipulated technical specifications are as follows:

1. Deep Learning architecture software.
- 2.Programming language (potential libraries): Python (Scikit-learn, Keras, TensorFlow, PyTorch, etc).
- 3.GPU server

1.3 Methodology and Procedure

This project is a new starting point, is for wind sensor data processing, and based on these data to propose a better data prediction method.

The first part focuses on understanding the background, the nature of the problem, and the working principle. To understand the characteristics of the data, we need to understand the working principle of the sensor. This will help us use the data better. The second step is to find a suitable solution for the project and try as much as possible. The third step is to compare the results of the attempt with the interpolation results to judge the superiority of the solution. Finally, project prediction is made in different states.

1.4 Work Plan

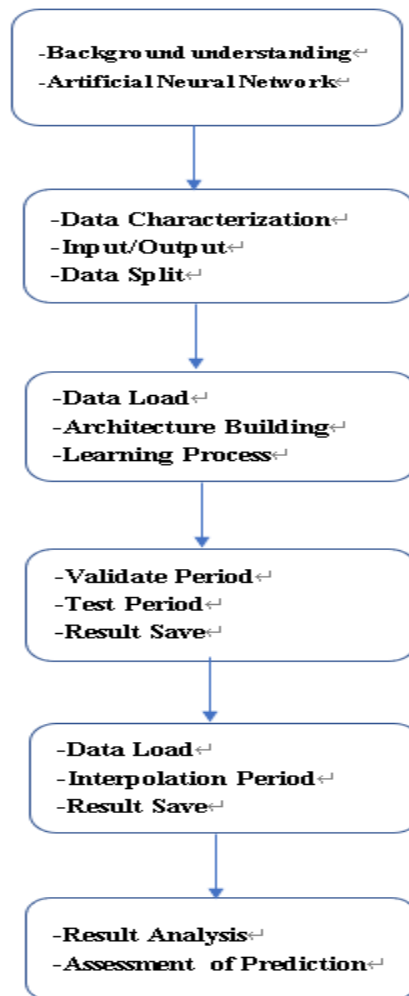


Figure 3: Work plan diagram

The milestones defined at the beginning of the project are the following:

1. Understand and find a suitable neural network architecture for a regression prediction problem.
2. Characterise and configure the global dataset to fit correctly as an input of the neural network software.
3. Implement a software that learns from correlated input data and predict desired output values.
4. Compare velocity-angle Temperature parameter between the estimated and the ground truth value.
5. Get interpolation result and compare it with the ground truth value.
6. Change the conditions to predict and interpolate, and judge which way is better

1.5 Deviations and Modification

The most important point in this process is the need to clarify the input and output and how to manage our dataset. We can also look at this problem from another perspective. We think that a part of the data in the data set disappears and needs to be recovered by relevant data means. This means that we need to do different configuration and processing of data, and the proposed model needs to meet different conditions. At the beginning of the project, a lot of time is spent on how to deal with different conditions. It depends very much on which strategy we choose. At the same time, need some time to adapt to the programming environment, familiar with the programming language.

Finally, the value is converted into velocity and angle to obtain more information. This work is not done in this thesis. At the end of the project, the results are sent in the format given by the features for subsequent work.

2.Context

2.1 Mars Exploration

Mars exploration refers to the scientific exploration of Mars by launching a space probe to Mars. Humans began to use telescopes to observe Mars in the 1600s. Mariner 4, launched on December 28, 1964, is the first probe ever to successfully reach Mars and send back data. The year 1976 will be remembered as a milestone in deep space missions. Viking 1 and Viking 2, two identical spacecraft, each composed of an orbiter and a lander, were the source of the greatest amount of data returned, completing more than six years of their mission and sending over 50,000 pictures of Mars.[11] NASA has returned to the exploration of Mars aiming at determining the atmospheric and surface characteristics of the planet to survey possible habitability. In this sense, it has become essential to develop new instrumentation capable of detecting with higher resolution and precision the main variables that determine the atmospheric parameters on the planet's surface.[13]

Reliable researches[8] suggest that meteorite impact events are fundamental processes in planetary evolution, playing an important role in the origin of life on Earth. This gives rise to new studies that compare the evolution of Earth and Mars, how the Red Planet has changed over time, and determines whether our planet is following the same steps or even if it is possible to 'terraforming' Mars, a hypothetical process of changing the natural conditions of the planet. [5] As a result of previous NASA missions, it is well known that the average planet surface temperature on Mars is 220 K and varies widely over a Martian day, from 145 K during the polar night to 300 K on the equator at midday at the closest point in its orbit around the Sun, with diurnal variations of up to 80–100 K[14].

Insight is not the only spacecraft to studying Mars. Curiosity was designed to assess whether Mars ever had an environment able to support small life forms called microbes.[15]. The rover landed on Mars in 2012. In support of science, Curiosity has instruments to better examine the environment.

As of mid-2018, Curiosity is working on the surface along with another NASA rover called Opportunity, which has been roaming the surface since 2004.[16] The Opportunity was initially designed for a 90-day mission but remains active after more than 14 years on Mars[16].

Perseverance Rover landed on Mars in February 2021. In April 2021, the Perseverance Rover successfully converted part of the carbon dioxide in the Martian atmosphere into oxygen, setting a new record.



Figure 4: Another artist's concept showing NASA's Perseverance rover on the surface of Mars.

NASA/JPL-Caltech

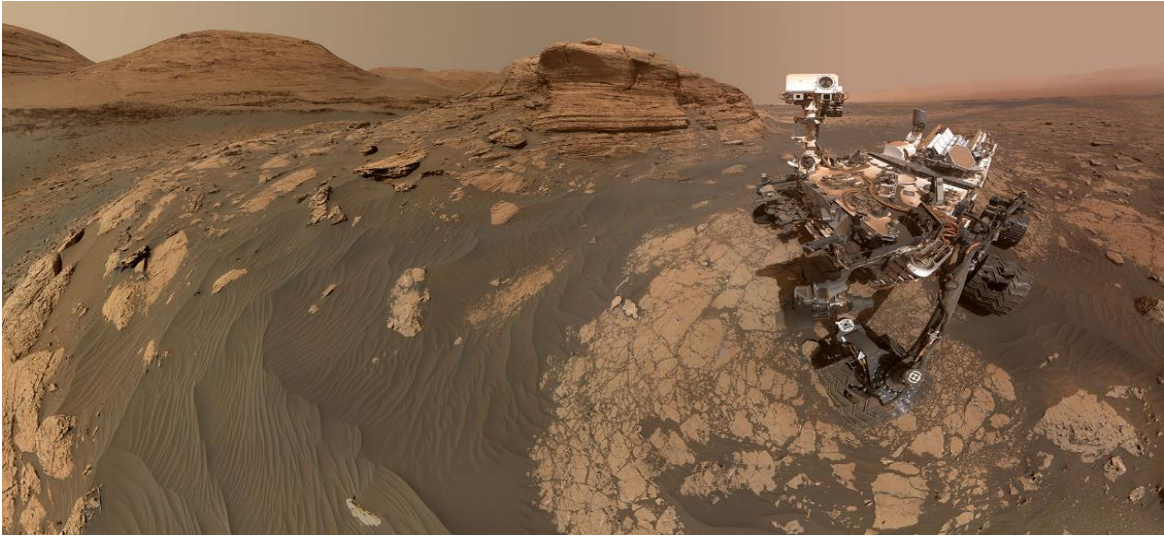


Figure 5: *NASA's Curiosity Mars rover used two cameras to create this selfie in front of Mont Marcou, a rock outcrop that stands 20 feet (6 meters) tall. The panorama is made up of 60 images from the MAHLI camera on the rover's robotic arm along with 11 images from the Mastcam on the mast, or "head," of the rover. NASA/JPL-Caltech/MSSS*

2.2 Insight Mission

To understand and explore Mars more comprehensively, NASA launched the "Insight lander" probe from Vandenberg Air Force Base in California. At that time, it detected the internal structure of the red planet by tracking seismic waves and measuring its underground temperature. And make a comprehensive analysis of the foundation structure of Mars, to collect first-hand information for future human colonization of Mars. On May 5th of 2018, the rocket Atlas V-401 was launched. Six months later, InSight landed on the Red Planet and began to communicate with Earth until our days. Due to the specific experiments and operations that this lander was born to do, and the safety and accurate considerations that a lander of these characteristics requires, the smooth plain called Elysium Planitia was chosen as the definitive landing site.[9] This flat surface is located on the western side near Mar's equator, exactly on 4.502°N latitude and 135.623°E longitude (was planned to land at 135.970°E), merely 600 km from Curiosity's landing site, Gale Crater. [7]According to the plan, insight will use two main

instruments: seismograph and heat detector to detect the planet thousands of kilometers away from the earth. According to the data of previous Mars probes, evidence shows that Mars has groundwater, which is very salty and full of perchlorate, but there are bacteria on earth that can survive in this condition. There are also lava tubes and caves on Mars, and there may be places with mild climates. Deep underground, it can be quite warm and humid, and there will be extremophiles living in this condition.

Formation & Evolution: Understand the formation and evolution of terrestrial planets through investigation of the interior structure and processes of Mars. [9]

Tectonic Activity: Determine the present level of tectonic activity and meteorite impact rate on Mars.[9]

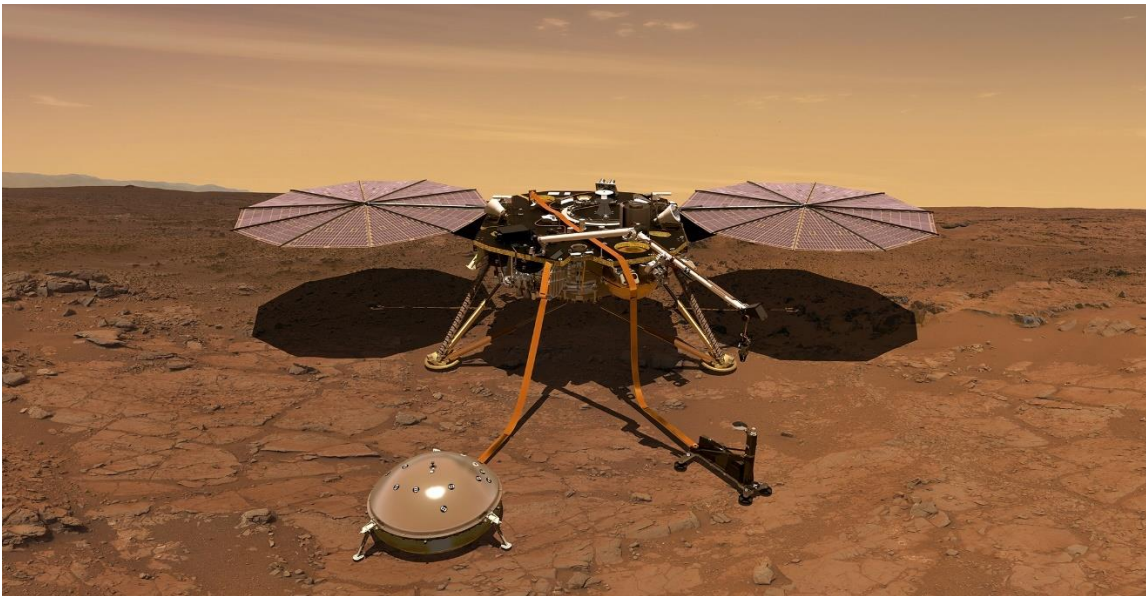


Figure 6: *An artist's rendition of the InSight lander operating on the surface of Mars.*

2.2.1 Science

Mars day refers to a "Sol" time on Mars, that is, a solar day on Mars. The average length of a stellar day on Mars is 24 hours 37 minutes 22.663 seconds, while the average length of a solar day on Mars is 88775.24409 seconds or 24 hours 39 minutes 35.24409 seconds. The average solar day on Mars is only 2.7% longer than that on earth. The international planetary organization declares August 27 as world Mars day. Mars, like the earth, has the same time difference, which can be calculated by the solar track of Mars. Because the orbit eccentricity of Mars is larger than that of the earth, the solar day length of Mars is quite variable. On Mars, the sun is sometimes 50 minutes slower and sometimes 40 minutes faster than the clock on Mars. The orbits of Mars and earth are elliptical, so they are closest when they are at the two closest points on the orbit.

Mars also has four seasons because of the inclination of its axis and the time of one rotation similar to that of the earth. However, one year of Mars is about two years of the earth's, and its orbit eccentricity is much greater than that of the earth. Therefore, the change of the four seasons of Mars will be much worse than that of the earth, and the data obtained by using a sundial to measure the time of Mars will become very inaccurate.

To deal with the problem that the Martian solar day is 39 minutes and 35 seconds longer than the earth's solar day, different space agencies have proposed many Martian clock designs, but so far they have not been accepted. One of the most common Martian clocks is 24 hours, 39 minutes, and 35 seconds.

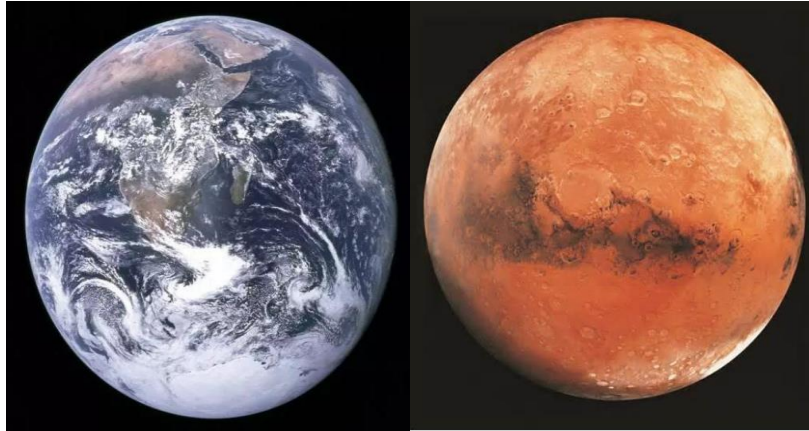


Figure 7: Earth picture

Figure 8: Mars picture

Planet	Earth	Mars
Solar Constant[w/m ²]	1373	591
Atmosphere air element composition[%]	9.8	3.7
Surface pressure[hPa]	N ₂ 78.08% Ar 0.93% O ₂ 20.9%	CO ₂ 95.22% N ₂ 2.7% Ar 1.6% O ₂ 0.13%
Surface density[kg/m ³]	1013hPa(1Ba)	6-8hPa(6-8mBar)
Kinematic viscosity[m ² /s]	1.2	0.0015
Average temperature[k]	150000	0.01
Tmerparatyre Variation[k]	300	220
average speed around the sun[m/s]	14.5	18.5

Table 1: Comparison of Mars environment and earth environment

The InSight Lander carries advanced equipment of seven principal science instruments, to provide key information on the internal structure and composition of Mars, as well as to carry out a series of surface geophysical investigations. This spacecraft carries three primary

investigations: Seismic Experiment for Interior Structure (SEIS); Heat Flow and Physical Properties Package (HP3); and Rotation and Interior Structure Experiment (RISE). SEIS is a seismometer, it measures the surface internal vibrations generated by “Marsquakes” or meteorite impacts to illuminate the properties of the crust, mantle, and core. The second instrument is used to generate pulses of heat through the inner layer of Mars, to reveal how much heat or vibrations flows out to the surface. RISE is an X-band radio instrument that measures the wobble of Mars' the North Pole as the sun pushes and pulls it in its orbit, providing clues on the size and composition of Mars' metallic core. Last but not least, these primary investigations are supported by a robotic arm, two cameras, and the set of sensors (wind, temperature, pressure, and magnetic field) called Auxiliary Payload Sensor Subsystem (APSS). [2][3]

2.2.2 Temperature and Wind for Insight

TWINS is arrayed with wind and temperature sensors is part of the complete InSight Lander weather station (APSS). The sensors that form it are very similar to those used in REMS (Curiosity), which were the first Spanish instruments that traveled to Mars. TWINS helps to understand the current weather and climate of the Red Planet.

The TWINS booms are located on top of InSight’s main deck. The two booms are disposed of horizontally, parallel to one another in diametrically opposite directions, facing away from the center of the deck and out over the solar panels. As a result, the two TWINS booms are facing outward on roughly opposite sides of the lander.



Figure 9: TWINS Booms ready to be installed.
CAB/INTA-CSIC

To understand this setting, it is easy to imagine a horizontal plane above the surface of the InSight’s platform, parallel to the ground, which is created by an axis (Y) aligned with booms and another (X) perpendicular to this, with its center in the middle of the lander. This deck placement is intended to minimize the effects of wind-flow perturbations induced by the other elements on the lander top deck by ensuring that at least one of the booms will be windward of the bulk of the lander body at all times and wind azimuths.

Each boom is composed of three principal components: Wind Sensors, which are three recording points based on a hot film anemometer, an Air Temperature Sensor (ATS), a small rod made of a low thermal conductivity that obliquely comes out of the bottom of each boom. This rod has three Platinum Resistance Thermometers (PRT) placed on different levels. ATS provides air temperature measurements, and their readings are combined and processed to provide clean wind data despite the thermal contamination from the boom. Finally, the Pressure Sensor (PS), a pressure transducer located in the lander body, where the effects of temperature that corrupt the pressure measurement are minimized. [10][4].

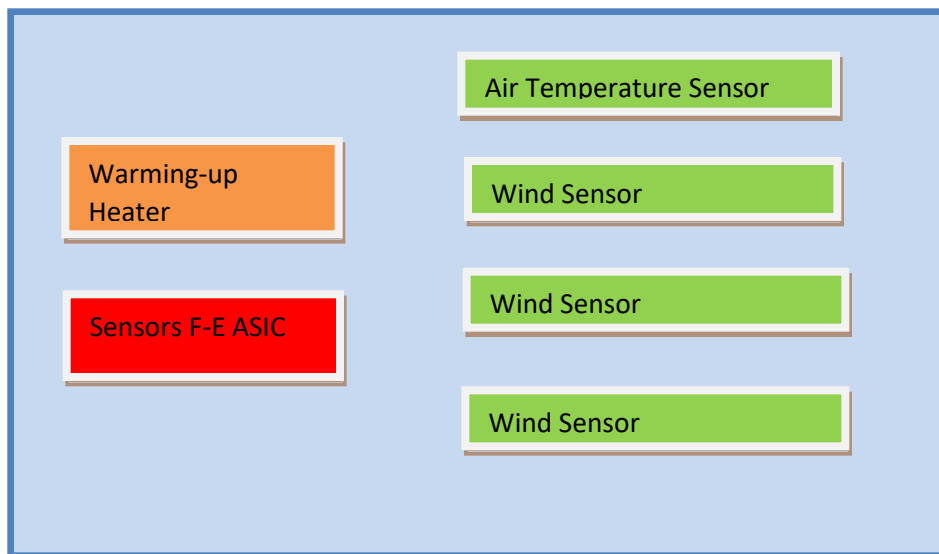


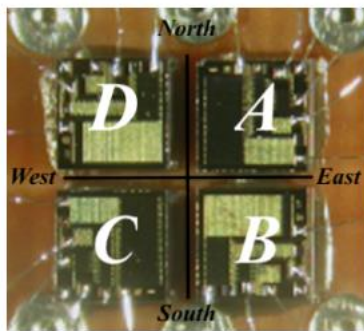
Figure 10: *TWINS Block diagram(Boom Y(-Y))*

The wind speed and direction sensor consists of three sensor boards (2-dimensional hot film anemometers) arrayed around the tip of each boom, 120° apart from one another. Each of these sensor boards consists of 4 hot dice and 1 cold dice that sense the local wind speed and direction in the plane of the sensor—those are the raw, direct measurements: acquisitions from all the dice are sent back to Earth by InSight. The wind sensor will be sampled at 1 Hz, matching its physical response time of roughly 1 s to wind perturbations. [18]

Each boom records the local 2D speed at three boards on the boom surface. Each board is oriented at 120° from the next, and the combined reading from those boards determine the wind speed and upwind direction.

Each boom has three boards (Boards 1 and 3 on the sides, and 2 in the lower part of the boom) identical. Each board has four hot dice and one cold die (the four hot dice are in the front of the board and the cold one is in the back).[20] In the hot dice, a resistor is used to heat it, another one is used as a sensor to measure the temperature, and the third one is used as a reference sensor in the measurement circuit. In the cold die, the only resistor used is the reference resistor.[20] The control loop compares the temperature of the hot dice with that of the cold die, to control the power injected to keep a constant predefined temperature difference (delta temperature) between them. Each board has an additional thermistor on its inside face to monitor the board's temperature and evaluate the conductive thermal losses of the dice.[20] With the ambient temperature and the pressure, it is possible to compute the fluid properties establishing the relation the wind speed at each recording point.

For a further understanding of variables behavior over time, it is interesting to introduce two relevant variables. The four dice from a single board can be defined as follows:



$$\text{LONG} = \text{North} - \text{South} = (A + D) - (B + C)$$

$$\text{TRANS} = \text{East} - \text{West} = (A + B) - (C + D)$$

Figure 11: Four dice array orientation. [20]

In the lander, each boom has three boards and each board has four dices. According to the definition of long and trans, each board can get one trans and one long through calculation. So it is easy to calculate the value of input variables, it is six pairs from two booms.

3 State of the Art

3.1 Artificial Intelligence

Artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think like humans and mimic their actions. It is a new technical science to

research and develop the theory, method, technology, and application system for simulating, extending, and expanding human intelligence. In 1956, it was first put forward by John McCarthy. At that time, it was defined as "the science and engineering of manufacturing intelligent machines". The purpose of artificial intelligence is to make machines think like human beings and have intelligence. Today, the connotation of artificial intelligence has been greatly expanded, is interdisciplinary. The ideal characteristic of artificial intelligence is its ability to rationalize and take actions that have the best chance of achieving a specific goal. A subset of artificial intelligence is machine learning, which refers to the concept that computer programs can automatically learn from and adapt to new data without being assisted by humans.

The information revolution represented by the Internet has brought about a qualitative change in the way people connect and in the efficiency of the combination of different elements of society. People do not have to memorize a lot of useless knowledge every day, and their brains begin to be liberated rapidly. In the era of the industrial revolution, we make machines by thinking. In the era of AI, we make machines that can think. The change of artificial intelligence to the future is the automation of the process of forming knowledge bit by bit. It's the automation of our process of replacing manpower with machines.

Due to the nature of the topic presented in this project, a Supervised Learning paradigm defines the basic concepts of the neural network. It is called supervised learning when the algorithm learns by making predictions on the training data and being corrected by the ground truth (correct answer) of the dataset.[6] Therefore, a mapping function for the known input (x) and output (y) variables is created and the network tries to generate close predictions to unseen input data. During the learning process (epochs/iterations), the fundamental idea is based on the comparison of the obtained and target output to keep on modifying the neural network weights to minimize the error and generalize the model. Furthermore, the ANN needed has to identify the better variables for a correct and accurate prediction.

3.2 Regression Prediction

As the name itself shows, regression is a technique to find correlations or strong mathematical relationships between a group of variables. The word regression was first

invented by Darwin's cousin Francis Galton. This statistic method gets conclusions and characteristics of variables over a specific period, which makes it possible to predict “dependent” variables identified as output (\hat{y}). Nowadays, many regression procedures are being applied in classification problems, where algorithms try to predict a discrete class label output for a specific input image. Nevertheless, the concept of this project does not correspond to this type of practice, the intention is to predict a continuous quantity of real values for each output. In this case, there is more than one input to find similarities in the behavior of the data, as well as, more than one variable to predict in the same period, so the system and strategy needed are based on a multiple regression prediction. The techniques used in this project are briefly presented hereunder.

3.2.1 MLP

Perceptron is a concept of ANN, which was first introduced by Frank Rosenblatt in the 1950s. Single-layer perceptron is the simplest ANN. It includes the input layer and output layer, and the input layer and output layer are directly connected. Single-layer perceptron can only deal with linear problems, not nonlinear problems.

MLP multilayer perceptron is a kind of ANN neural network with a forward structure. MLP can deal with nonlinear and separable problems. MLP multi-layer perceptron is a forward structure artificial neural network ANN, which maps a set of input vectors to a set of output vectors. MLP can be regarded as a directed graph, which is composed of multiple node layers, and each layer is fully connected to the next layer. Except for the input nodes, each node is a neuron with a nonlinear activation function. The supervised learning method of the BP backpropagation algorithm is used to train MLP. MLP is a generalization of perceptron, which overcomes the weakness that perceptron can't recognize linear indivisible data.

Compared with the single-layer perceptron, the output of MLP multi-layer perceptron is changed from one to many; There is not only one layer between the input and output, but now there are two layers: the output layer and the hidden layer. Backpropagation learning is a typical feedforward network. Its information-processing direction is from the input layer to the hidden layer and then to the output layer. The hidden layer realizes the nonlinear mapping of the input space, the output layer realizes the linear classification, and the nonlinear mapping mode and the linear discriminant function can be learned at the same time.

MLP training process is as follows

- 1) The weights of all edges are assigned randomly;
- 2) Forward propagation: using the input characteristics of all samples in the training set as the input layer, the artificial neural network is activated for all inputs in the training data set, and then the output value is obtained through forwarding propagation.
- 3) Backpropagation: use the output value and sample value to calculate the total error, and then use backpropagation to update the weight.
- 4) Repeat 2) ~ 3) until the output error is lower than the established standard.

After the above process, we get a learned MLP network, which is considered to accept new input.

Advantages of MLP:

- 1) Highly parallel processing;
- 2) Highly nonlinear global action;
- 3) Good fault tolerance;
- 4) It has an associative memory function;5) Very strong adaptive, self-learning function.

3.2.2 LSTM

A recurrent neural network (RNN) is a kind of neural network used to process sequence data. Compared with the general neural network, it can deal with the data of sequence change. Long short term memory (LSTM) is a special RNN, which is mainly used to solve the problem of gradient disappearance and gradient explosion in the process of long sequence training. In short, LSTM can perform better in longer sequences than ordinary RNN. In other words, the algorithm can learn long-term correlations in a sequence, which is useful for time sequence prediction problems. By the addition of a circuit of loops, LSTM is created by memory blocks that work as feedback to the next input layer. Consequently, the network

learns from the memory of recent temporal sequences, however, previous deep knowledge of the behavior of the data is needed.[12][1]

There are three stages in LSTM

1. Forgetting stage. This stage is mainly to selectively forget the input from the previous node. In short, it means "forgetting the unimportant and remembering the important".
2. . select the memory phase. This stage selectively "remembers" the input of this stage.
3. Output phase. This phase will determine which outputs will be taken as the current state.

One problem with LSTM is that it is computationally expensive because of the relative complexity of its internal structure.GRU is also designed to solve the gradient disappearance problem in standard RNN, which can be seen as a variant of LSTM.In fact, GRU performance is almost the same as LSTM in most cases (and sometimes LSTM is even better, and LSTM was proposed in 1997 and has been more historically tested), but GRU's biggest advantages are simplicity (because there are only two gates), low computational overhead, and more suitable for large data sets.

3.3 DSSM

DSSM (deep structured semantic models), also known as a deep semantic matching model, is the first article published by Microsoft to calculate semantic similarity in NLP.DSSM model uses two relatively independent complex networks to construct related features, so it is called double tower mode. From bottom to top, DSSM can be divided into three layers: input layer, presentation layer, matching layer.DSSM is most commonly used for semantic matching.

With the development of deep learning, CNN and LSTM have been proposed, and these structures of special diagnosis extraction have also been applied to DSSM. The main difference lies in replacing the fully connected structure of feature extraction layer with CNN or LSTM.

3.4 Other Methods

AutoEncoder is an unsupervised learning algorithm that uses a backpropagation algorithm to make the target equal to the input value. The AutoEncoder framework contains two major modules: the encoding process and the decoding process.The input sample is mapped to the

feature space through encoder, that is, the coding process; Then, the abstract features are mapped back to the original space through decoder to obtain the reconstructed samples, that is, the decoding process. The optimization goal is to optimize encoder and decoder simultaneously by minimizing the reconstruction error, so as to learn the abstract feature representation for the sample input. In the optimization process, the AutoEncoder does not need to use the label of the sample. In essence, the input of the sample is used as the input and output of the neural network at the same time. By minimizing the reconstruction error, the AutoEncoder hopes to learn the abstract feature representation of the sample. This unsupervised optimization greatly improves the generality of the model.

4 Project Development

4.1 Data Set

To carry out this research, it has been necessary to count on a huge amount of TWINS data. The dataset used has been provided and filtered by the Centro de Astrobiología (CAB) in Madrid, which has been in charge of the ATS tests and possesses the calibration tables for these. The first information of the dataset is dated from Earth day 334 of 2018 (Sol 559, Martian Year 34) at 19:09:38 UTC (Coordinated Universal Time), which means that the available data starts four days after InSight's landed, 30th November 2018. The last information of the dataset is from Earth day 263 of 2019 (Sol 177, Martian Year 35), 20th September 2019. Consequently, due to the InSight's landing site (North Hemisphere) and the duration of a Martian year, the available data characterizes the final stages of winter on Mars, the spring and early summer seasons.

Each document (.csv) supplied by CAB contains TWINS information of a specific day, with a specific hour interval and a second per sample. Therefore, the first thing that has been done is to assemble the entire collection of documents to create a single global and unique dataset. Through the python code implemented, the first data processing has been done. Moreover, it has been observed that the global dataset created has data (NAN) in line with time, so all the rows with no data values have been removed from the global dataset, succeeding to a final dataset of 6364733 rows in total.

4.1.1 Input/Output variables

For machine learning training, the number of inputs has a great impact on the results. In this project, the number of inputs is limited, and there is no way to achieve good results. On this basis, we use simple mathematical calculations to construct some new eigenvalues and expand the eigenvalues to help us get a better mode.

In the process of machine learning model development, we hope that the trained model can perform well on new and unprecedented data. To simulate new and unknown data, the available data is segmented. A common method of data segmentation is to divide the data into three parts (1) Training set, (2) verification set, and (3) test set. The training set is used to build the prediction model and evaluate the validation set. Based on the prediction, the model can be optimized (such as super parameter optimization), and the model with the best performance can be selected according to the results of the validation set. The test set does not participate in the establishment and preparation of any model. Therefore, the test set can act like new and unknown data. Next, the training set is used to establish the prediction model, and then the trained model is applied to the test set (that is, as new and unprecedented data) for prediction. The split percentage has been defined with the following proportion:

Subset	Split ratio	Number of rows
Train	64%	185.686
Validation	16%	46.422
Test	20%	58.027

Table 2: Dataset split configuration

Before data analysis, we usually need to normalize the data and use the standardized data for data analysis. Data standardization is the indexation of statistical data.

At present, there are many methods of data standardization, which can be divided into the linear method (such as extreme value method, standard deviation method), broken line method (such as three broken line method), curve method (such as semi-normal distribution). Different standardization methods have different effects on the evaluation results of the system.

From the experience, normalization is to make the features between different dimensions have a certain numerical comparison, which can greatly improve the accuracy of the classifier.

In the project, the data modification done has been standardization, which distributed data in a Gaussian form with zero mean and unit standard deviation or variance. Finally, the predicted values are inversely transformed, recovering the form and range of the original signal.

4.2 Development of the Neural Network Architecture

Due to the short duration of this project, only one neural network model has been established and proved. Nevertheless, two possible candidate models were initially assessed. The evaluated models were a multilayer perceptron network, which is the native deep learning network, and also the LSTM algorithm. LSTM needs long memory and is more suitable for dealing with problems with long time intervals. Because the amount of data we have is not very large, the time interval is short, and the characteristics are not obvious enough. In this case, LSTM is not a reasonable solution. Currently, the LSTM has been put aside perceiving it as a possible solution in future researches. Consequently, the DSSM network has been the convenience neural network model for this project.

The activation function used in the collection of neurons in the recent SELU (Scaled Exponential Linear Unit) and TANH(hyperbolic tangent). This function correctly treats positive values and permits higher control over negative values, which is critical in this case due to the standardization of the data. Dealing with regression in deep learning entails no activation function for the output layer. Tanh, is similar to sigmoid, which converts the input value to between - 1 and 1. The derivative of tan ranges from 0 to 1, which is better than 0 to 1 / 4 of sigmoid. To a certain extent, the problem of gradient vanishing is alleviated. The output and input of tanh can keep the nonlinear monotonic ascending and descending relation, which accords with the gradient solution of the BP (backpropagation) network, and has good fault tolerance and boundedness.

Training a neural network is an iterative process that in this case consists of two phases: forward and backpropagation. In the first one, all the training dataset (in batches and shuffled) is passed through the network at each epoch. Once each time step is estimated, a loss function is calculated by the difference between the prediction made and the ground truth. Moreover, a forward process of validation data quantifies the network behavior for unseen data. Afterward, in the second phase, the error calculated before is propagated backward, which is used by the model to upgrade the weights to minimize the loss function.

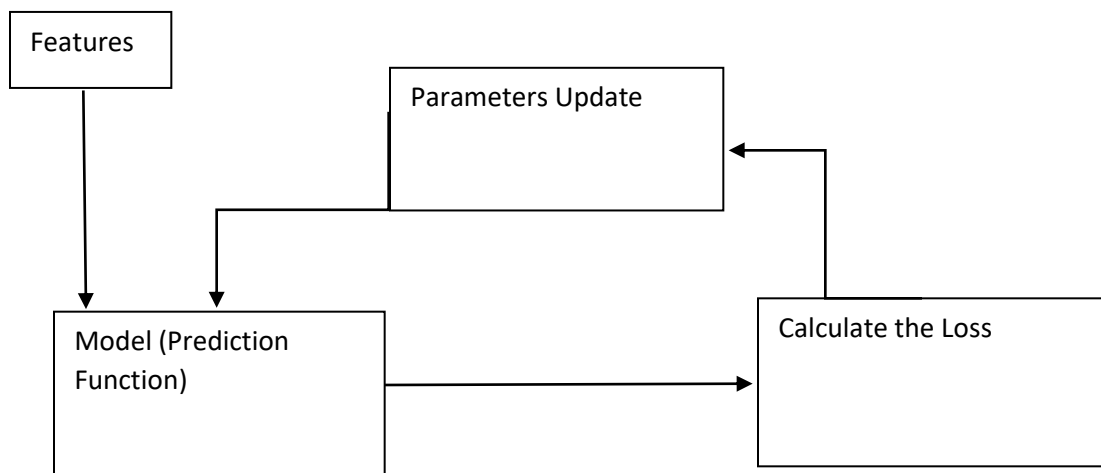


Table 3: *Training process diagram*

The regression metric employed as a loss function is the MSE (Mean Squared Error). The statistical parameter is the mean value of the square sum of the errors of the corresponding points of the predicted data and the original data. MSE can evaluate the change degree of the data. The smaller the MSE value is, the better the accuracy of the prediction model to describe the experimental data is.

To optimize the training process and efficiently minimize the loss function, the algorithm used as the optimizer has been the SGD (Stochastic Gradient Descent). SGD was proposed in 1847. Select a mini-batch instead of all samples at a time, and use gradient descent to update the model parameters. As long as the noise is not particularly large, SGD can converge well, and the training speed is very fast when large data sets are used.

The hyperparameters used have been tested for different situations and model architectures.

Learning rate is one of the most important hyperparameters affecting performance. If we can only adjust one hyperparameters, then the best choice is it. Compared with other hyperparameters, the learning rate controls the effective capacity of the model in a more complex way. When the learning rate is optimal, the effective capacity of the model is the largest. The lower the learning rate, the slower the loss function changes. Although using a low learning rate can ensure that we will not miss any local minima, it also means that we will spend more time converging. In different prediction processes, we use different learning rates.

The size of the batch affects the time required to complete each epoch and the smoothness of the gradient between each iteration in the process of deep learning training. Large batch size reduces the training time and improves the stability, but the generalization ability of the model decreases. The randomness introduced by small batch size will be greater, sometimes it can have a better effect, but the convergence speed is slow. Batch size is important in the project.

When a complete data set passes through the neural network once and returns once, this process is called an epoch[21]. In other words, all the training samples have a forward propagation and backpropagation in the neural network. The number of epoch for the fitting process is 350.

4.3 Prediction from Data for Single Point Model

4.3.1 Prediction from Two Booms for Single Point Model

This section contains the methods and procedures of the main attempts to predict the WS values of an entire two boards. In this first development, it is decided to include all the available information as input variables. Therefore, the model is trained by the data from both Booms BPY and BMY, to generate a prediction based on the most evidence and correlations possible. The values to be estimated are all those time instances that make up the test dataset.

Layer (type)	Output Shape	Param #	Connected to
input_2 (InputLayer)	(None, 160)	0	
dense_40 (Dense)	(None, 150)	24150	input_2[0][0]
dense_59 (Dense)	(None, 150)	24150	input_2[0][0]
dense_41 (Dense)	(None, 140)	21140	dense_40[0][0]
dense_60 (Dense)	(None, 140)	21140	dense_59[0][0]
dense_42 (Dense)	(None, 130)	18330	dense_41[0][0]
dense_61 (Dense)	(None, 130)	18330	dense_60[0][0]
dense_43 (Dense)	(None, 120)	15720	dense_42[0][0]
dense_62 (Dense)	(None, 120)	15720	dense_61[0][0]
dense_44 (Dense)	(None, 110)	13310	dense_43[0][0]
dense_63 (Dense)	(None, 110)	13310	dense_62[0][0]
dense_45 (Dense)	(None, 100)	11100	dense_44[0][0]
dense_64 (Dense)	(None, 100)	11100	dense_63[0][0]
dense_46 (Dense)	(None, 90)	9090	dense_45[0][0]
dense_65 (Dense)	(None, 90)	9090	dense_64[0][0]
dense_47 (Dense)	(None, 80)	7280	dense_46[0][0]
dense_66 (Dense)	(None, 80)	7280	dense_65[0][0]
dense_48 (Dense)	(None, 70)	5670	dense_47[0][0]
dense_67 (Dense)	(None, 70)	5670	dense_66[0][0]
dense_49 (Dense)	(None, 60)	4260	dense_48[0][0]
dense_68 (Dense)	(None, 60)	4260	dense_67[0][0]
dense_50 (Dense)	(None, 50)	3050	dense_49[0][0]
dense_69 (Dense)	(None, 50)	3050	dense_68[0][0]
dense_51 (Dense)	(None, 40)	2040	dense_50[0][0]
dense_70 (Dense)	(None, 40)	2040	dense_69[0][0]
dense_52 (Dense)	(None, 30)	1230	dense_51[0][0]
dense_71 (Dense)	(None, 30)	1230	dense_70[0][0]
dense_53 (Dense)	(None, 20)	620	dense_52[0][0]
dense_72 (Dense)	(None, 20)	620	dense_71[0][0]
dense_54 (Dense)	(None, 16)	336	dense_53[0][0]
dense_73 (Dense)	(None, 16)	336	dense_72[0][0]
dense_55 (Dense)	(None, 12)	204	dense_54[0][0]
dense_74 (Dense)	(None, 12)	204	dense_73[0][0]
dense_56 (Dense)	(None, 8)	104	dense_55[0][0]
dense_75 (Dense)	(None, 8)	104	dense_74[0][0]
dense_57 (Dense)	(None, 4)	36	dense_56[0][0]
dense_76 (Dense)	(None, 4)	36	dense_75[0][0]
dense_58 (Dense)	(None, 2)	10	dense_57[0][0]
dense_77 (Dense)	(None, 2)	10	dense_76[0][0]
concatenate_2 (Concatenate)	(None, 4)	0	dense_58[0][0] dense_77[0][0]
dense_78 (Dense)	(None, 2)	10	concatenate_2[0][0]
Total params: 275,370			
Trainable params: 275,370			
Non-trainable params: 0			

Figure 12: Proposed model to predict with data from 2 Booms.

In a project, assume that the first two parameters are not recorded for the sensor fails for some reasons. This figure 12 shows the model architecture proposed, which imitates the encoder concept. The purpose of the dense layer is to extract the features from the input and extract the correlation between these features after nonlinear changes in the dense layer, and finally map them to the output space. In theory, a dense layer is enough, but this is only in theory, because you don't know how many nodes you need in a dense layer, or how many times you need to train. Add more dense layers and you'll get faster convergence. Concatenate operation: a very important operation in network structure design, which is often used to combine features, to fuse features extracted by multiple convolutional feature extraction frameworks or to fuse information of the output layer. Concatenate is the combination of the number of channels, that is to say, the features describing the image itself increase while the information under each feature does not increase.

The total parameters are 275370. For the Input parameter, there are BMY_WS_BLONG_FINAL_2..3, BMY_WS_BTRANS_FINAL_2..3, BMY_TIP_ROD_TEMP, BPY_WS_BLONG_FINAL_1...3, BPY_WS_BTRANS_FINAL_1...3, BPY_TIP_ROD_TEMP, PRESSURE. The output is BMY_WS_BLONG_FINAL_1 and BMY_WS_BTRANS_FINAL_1 that is the missing values that we want to estimate.

4.3.2 Prediction from One Boom for Single Point Model

The next step in this research has been to minimize the available information as much as possible. Comparing with the prediction from the two booms, the one boom data is smaller. The complexity and parameters of the model will decrease, and the accuracy will also decrease. If the model can predict values it suggests that requires less computing time and complexity. The input data just includes the values from BMY boom, BMY_WS_BLONG_FINAL_2..3, BMY_WS_BTRANS_FINAL_2..3, BMY_TIP_ROD_TEMP, and PRESSURE. The output is the same as the previous.

The following figure shows the modified model that is suited in this context after trying different depth models. The architecture consists of the same preceding concept. The total parameters are 49718.

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	(None, 62)	0	
dense_1 (Dense)	(None, 60)	3780	input_1[0][0]
dense_22 (Dense)	(None, 60)	3780	input_1[0][0]
dense_2 (Dense)	(None, 58)	3538	dense_1[0][0]
dense_23 (Dense)	(None, 58)	3538	dense_22[0][0]
dense_3 (Dense)	(None, 54)	3186	dense_2[0][0]
dense_24 (Dense)	(None, 54)	3186	dense_23[0][0]
dense_4 (Dense)	(None, 50)	2750	dense_3[0][0]
dense_25 (Dense)	(None, 50)	2750	dense_24[0][0]
dense_5 (Dense)	(None, 48)	2346	dense_4[0][0]
dense_26 (Dense)	(None, 48)	2346	dense_25[0][0]
dense_6 (Dense)	(None, 42)	1974	dense_5[0][0]
dense_27 (Dense)	(None, 42)	1974	dense_26[0][0]
dense_7 (Dense)	(None, 38)	1634	dense_6[0][0]
dense_28 (Dense)	(None, 38)	1634	dense_27[0][0]
dense_8 (Dense)	(None, 34)	1326	dense_7[0][0]
dense_29 (Dense)	(None, 34)	1326	dense_28[0][0]
dense_9 (Dense)	(None, 30)	1050	dense_8[0][0]
dense_30 (Dense)	(None, 30)	1050	dense_29[0][0]
dense_10 (Dense)	(None, 26)	806	dense_9[0][0]
dense_31 (Dense)	(None, 26)	806	dense_30[0][0]
dense_11 (Dense)	(None, 22)	594	dense_10[0][0]
dense_32 (Dense)	(None, 22)	594	dense_31[0][0]
dense_12 (Dense)	(None, 20)	460	dense_11[0][0]
dense_33 (Dense)	(None, 20)	460	dense_32[0][0]
dense_13 (Dense)	(None, 18)	378	dense_12[0][0]
dense_34 (Dense)	(None, 18)	378	dense_33[0][0]
dense_14 (Dense)	(None, 16)	304	dense_13[0][0]
dense_35 (Dense)	(None, 16)	304	dense_34[0][0]
dense_15 (Dense)	(None, 14)	238	dense_14[0][0]
dense_36 (Dense)	(None, 14)	238	dense_35[0][0]
dense_16 (Dense)	(None, 12)	180	dense_15[0][0]
dense_37 (Dense)	(None, 12)	180	dense_36[0][0]
dense_17 (Dense)	(None, 10)	130	dense_16[0][0]
dense_38 (Dense)	(None, 10)	130	dense_37[0][0]
dense_18 (Dense)	(None, 8)	88	dense_17[0][0]
dense_39 (Dense)	(None, 8)	88	dense_38[0][0]
dense_19 (Dense)	(None, 6)	54	dense_18[0][0]
dense_40 (Dense)	(None, 6)	54	dense_39[0][0]
dense_20 (Dense)	(None, 4)	28	dense_19[0][0]
dense_41 (Dense)	(None, 4)	28	dense_40[0][0]
dense_21 (Dense)	(None, 2)	10	dense_20[0][0]
dense_42 (Dense)	(None, 2)	10	dense_41[0][0]
concatenate_1 (Concatenate)	(None, 4)	0	dense_21[0][0] dense_42[0][0]
dense_43 (Dense)	(None, 2)	10	concatenate_1[0][0]

Total params: 49,718
Trainable params: 49,718
Non-trainable params: 0

Figure 13: Proposed model to predict with data from one Boom.

4.4 Prediction from Data for Two Consecutive Points Model

Based on the single-point prediction, we continue to build a model that can predict two points at the same time.

4.4.1 Prediction from Two Booms for Two Consecutive Points Model

Once the two consecutive points are expected, this model will be used. The input parameters are the same as the previous prediction for one point from two booms. The output is different. It will export two consecutive points for `BMY_WS_BLONG_FINAL_1` and `BMY_WS_BTRANS_FINAL_1`. The total parameters are 243356 in the model.

Layer (type)	Output Shape	Param #	Connected to
input_2 (InputLayer)	(None, 160)	0	
dense_32 (Dense)	(None, 150)	24150	input_2[0][0]
dense_47 (Dense)	(None, 150)	24150	input_2[0][0]
dense_33 (Dense)	(None, 140)	21140	dense_32[0][0]
dense_48 (Dense)	(None, 140)	21140	dense_47[0][0]
dense_34 (Dense)	(None, 130)	18330	dense_33[0][0]
dense_49 (Dense)	(None, 130)	18330	dense_48[0][0]
dense_35 (Dense)	(None, 120)	15720	dense_34[0][0]
dense_50 (Dense)	(None, 120)	15720	dense_49[0][0]
dense_36 (Dense)	(None, 110)	13310	dense_35[0][0]
dense_51 (Dense)	(None, 110)	13310	dense_50[0][0]
dense_37 (Dense)	(None, 100)	11100	dense_36[0][0]
dense_52 (Dense)	(None, 100)	11100	dense_51[0][0]
dense_38 (Dense)	(None, 80)	8080	dense_37[0][0]
dense_53 (Dense)	(None, 80)	8080	dense_52[0][0]
dense_39 (Dense)	(None, 60)	4860	dense_38[0][0]
dense_54 (Dense)	(None, 60)	4860	dense_53[0][0]
dense_40 (Dense)	(None, 40)	2440	dense_39[0][0]
dense_55 (Dense)	(None, 40)	2440	dense_54[0][0]
dense_41 (Dense)	(None, 30)	1230	dense_40[0][0]
dense_56 (Dense)	(None, 30)	1230	dense_55[0][0]
dense_42 (Dense)	(None, 20)	620	dense_41[0][0]
dense_57 (Dense)	(None, 20)	620	dense_56[0][0]
dense_43 (Dense)	(None, 16)	336	dense_42[0][0]
dense_58 (Dense)	(None, 16)	336	dense_57[0][0]
dense_44 (Dense)	(None, 12)	204	dense_43[0][0]
dense_59 (Dense)	(None, 12)	204	dense_58[0][0]
dense_45 (Dense)	(None, 8)	104	dense_44[0][0]
dense_60 (Dense)	(None, 8)	104	dense_59[0][0]
dense_46 (Dense)	(None, 4)	36	dense_45[0][0]
dense_61 (Dense)	(None, 4)	36	dense_60[0][0]
concatenate_2 (Concatenate)	(None, 8)	0	dense_46[0][0] dense_61[0][0]
dense_62 (Dense)	(None, 4)	36	concatenate_2[0][0]

Total params: 243,356
Trainable params: 243,356
Non-trainable params: 0

Figure 14: Proposed model to predict with data from two Boom

4.4.2 Prediction from One Boom for Two Consecutive Points Model

The input parameters are the same as the previous prediction for one point from one boom. The output is different. It will export two consecutive points for `BMY_WS_BLONG_FINAL_1` and `BMY_WS_BTRANS_FINAL_1`. The total parameters are 41400 in the model.

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	(None, 60)	0	
dense_1 (Dense)	(None, 52)	3172	input_1[0][0]
dense_23 (Dense)	(None, 52)	3172	input_1[0][0]
dense_2 (Dense)	(None, 48)	2544	dense_1[0][0]
dense_24 (Dense)	(None, 48)	2544	dense_23[0][0]
dense_3 (Dense)	(None, 44)	2156	dense_2[0][0]
dense_25 (Dense)	(None, 44)	2156	dense_24[0][0]
dense_4 (Dense)	(None, 40)	1800	dense_3[0][0]
dense_26 (Dense)	(None, 40)	1800	dense_25[0][0]
dense_5 (Dense)	(None, 38)	1558	dense_4[0][0]
dense_27 (Dense)	(None, 38)	1558	dense_26[0][0]
dense_6 (Dense)	(None, 36)	1404	dense_5[0][0]
dense_28 (Dense)	(None, 36)	1404	dense_27[0][0]
dense_7 (Dense)	(None, 34)	1258	dense_6[0][0]
dense_29 (Dense)	(None, 34)	1258	dense_28[0][0]
dense_8 (Dense)	(None, 32)	1120	dense_7[0][0]
dense_30 (Dense)	(None, 32)	1120	dense_29[0][0]
dense_9 (Dense)	(None, 30)	990	dense_8[0][0]
dense_31 (Dense)	(None, 30)	990	dense_30[0][0]
dense_10 (Dense)	(None, 28)	868	dense_9[0][0]
dense_32 (Dense)	(None, 28)	868	dense_31[0][0]
dense_11 (Dense)	(None, 26)	754	dense_10[0][0]
dense_33 (Dense)	(None, 26)	754	dense_32[0][0]
dense_12 (Dense)	(None, 24)	648	dense_11[0][0]
dense_34 (Dense)	(None, 24)	648	dense_33[0][0]
dense_13 (Dense)	(None, 22)	550	dense_12[0][0]
dense_35 (Dense)	(None, 22)	550	dense_34[0][0]
dense_14 (Dense)	(None, 20)	460	dense_13[0][0]
dense_36 (Dense)	(None, 20)	460	dense_35[0][0]
dense_15 (Dense)	(None, 18)	378	dense_14[0][0]
dense_37 (Dense)	(None, 18)	378	dense_36[0][0]
dense_16 (Dense)	(None, 16)	304	dense_15[0][0]
dense_38 (Dense)	(None, 16)	304	dense_37[0][0]
dense_17 (Dense)	(None, 14)	238	dense_16[0][0]
dense_39 (Dense)	(None, 14)	238	dense_38[0][0]
dense_18 (Dense)	(None, 12)	180	dense_17[0][0]
dense_40 (Dense)	(None, 12)	180	dense_39[0][0]
dense_19 (Dense)	(None, 10)	130	dense_18[0][0]
dense_41 (Dense)	(None, 10)	130	dense_40[0][0]
dense_20 (Dense)	(None, 8)	88	dense_19[0][0]
dense_42 (Dense)	(None, 8)	88	dense_41[0][0]
dense_21 (Dense)	(None, 6)	54	dense_20[0][0]
dense_43 (Dense)	(None, 6)	54	dense_42[0][0]
dense_22 (Dense)	(None, 4)	28	dense_21[0][0]
dense_44 (Dense)	(None, 4)	28	dense_43[0][0]
concatenate_1 (Concatenate)	(None, 8)	0	dense_22[0][0] dense_44[0][0]
dense_45 (Dense)	(None, 4)	36	concatenate_1[0][0]

Total params: 41,400
Trainable params: 41,400
Non-trainable params: 0

Figure 15: Proposed model to predict with data from one Boom

4.5 Data Interpolation

Interpolation is a common way to deal with missing values. When for time series, if the data before and after are clear, then the simplest way is to get the data of the middle time point by calculating the average value of the two points before and after. In the project, this is simple approach for simplicity. In the following content, you will see the results of the data interpolation scheme and model prediction scheme.

5 Results

This section presents the global results of this project to evaluate the different algorithms and techniques proposed. In the different methodologies used, the obtained prediction has been plotted overtime above the ground truth values, that is, the real values., the prediction value from the model, and the interpretation. Therefore, the closer the predicted values are to the real one, the higher quality and accuracy are defined. To characterize the prediction precision, the loss is plotted in the picture and the RMS is calculated for each part.

Root mean square error is the arithmetic square root of mean square error. In the previous part, MSE is introduced.

5.1 Prediction from data for Single Points

5.1.1 Prediction from Two Booms Data for Single Point

The training and validation process is represented in the loss function below:

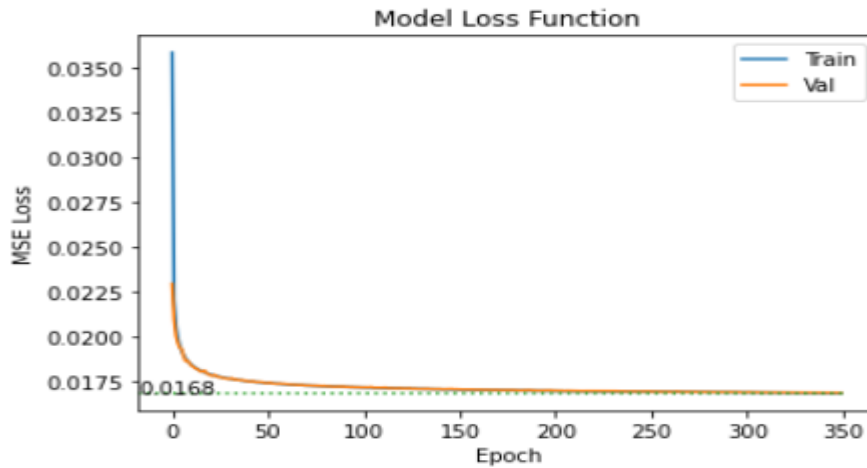


Figure 16: *Training process*

It can be observed how the training loss function decreases over the epochs as a result of the model's forward and backpropagation. Moreover, the validation loss function converges correctly, confirming that the model has also been trained for a similar behavior of unseen data.

The running time for each epoch and every step is listed here. During the process, the loss is decreasing quickly at first and then slowly.

```
Epoch 2/350
4073428/4073428 [=====] - 282s 69us/step - loss: 0.0220 - val_loss: 0.0209
Epoch 3/350
4073428/4073428 [=====] - 283s 69us/step - loss: 0.0206 - val_loss: 0.0200
Epoch 4/350
4073428/4073428 [=====] - 282s 69us/step - loss: 0.0199 - val_loss: 0.0196
Epoch 5/350
4073428/4073428 [=====] - 282s 69us/step - loss: 0.0195 - val_loss: 0.0193
Epoch 6/350
4073428/4073428 [=====] - 282s 69us/step - loss: 0.0192 - val_loss: 0.0193
Epoch 7/350
4073428/4073428 [=====] - 282s 69us/step - loss: 0.0189 - val_loss: 0.0189
Epoch 8/350
4073428/4073428 [=====] - 280s 69us/step - loss: 0.0188 - val_loss: 0.0186
Epoch 9/350
```

Figure 17: *Time consumption for every step*

Below are plotted the test real values (ground truth) in green, while the red cross represents the predicted data. It is the normalized value. From all the data points, 100 points are selected for display.

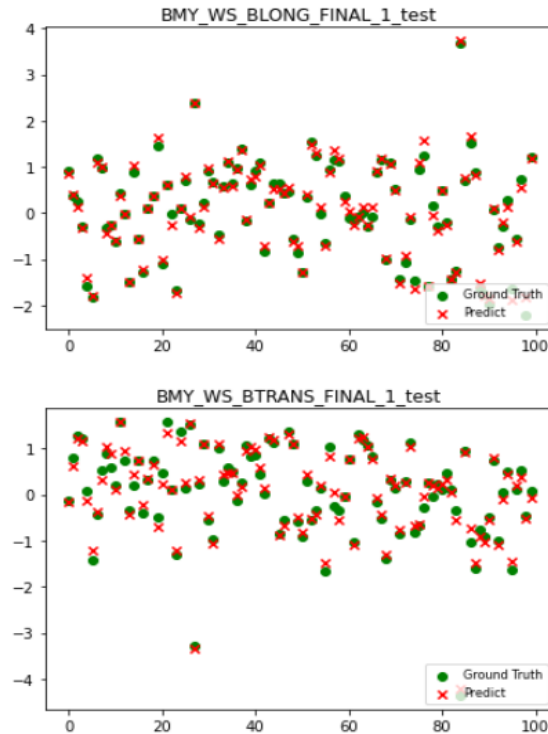


Figure 18: Normalized data plot

The whole effect of data display after standardization is very good. The green dots represent the real value, and the red dots represent the predicted value. The picture can clearly show the coincidence degree of each point.

Next, we compare the predicted results, interpolation results, and real values, hoping to find out which one is better. Before that, we need to de-standardize.

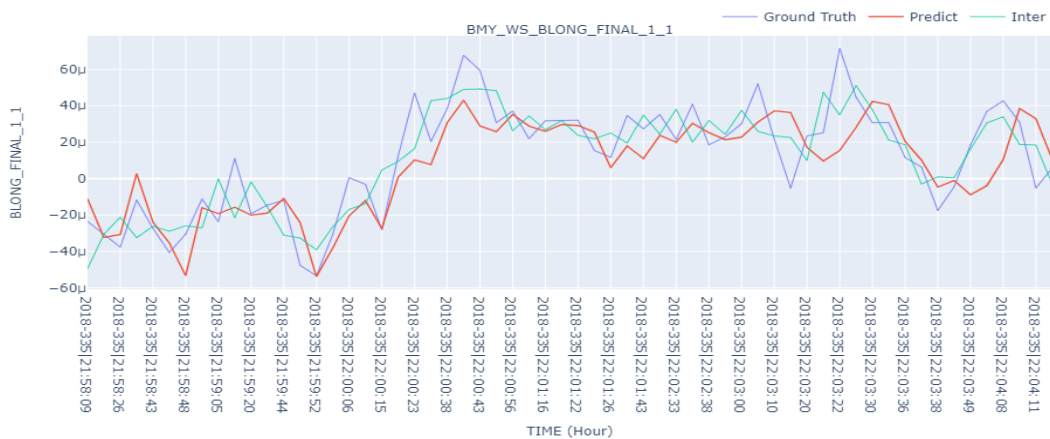


Figure 19: Unnormalized data plot BLONG

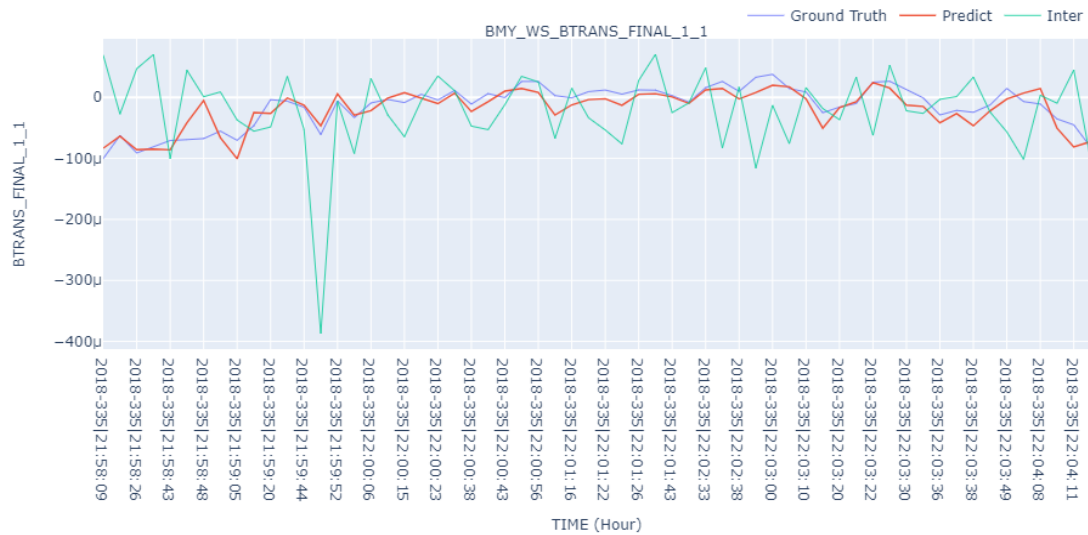


Figure 20: Unnormalized data plot BTRANS

It is the unnormalized data(Original Signal). The above two figures show the different results for `BMY_WS_BLONG_FINAL_1` and `BMY_WS_BTRANS_FINAL_1` of the interpolation scheme and the prediction scheme of the two variables in a relatively small time window(2018-335|21:58:09-2018-335|22:04:11). We can also continue to lengthen the time window and continue to observe our experimental results.

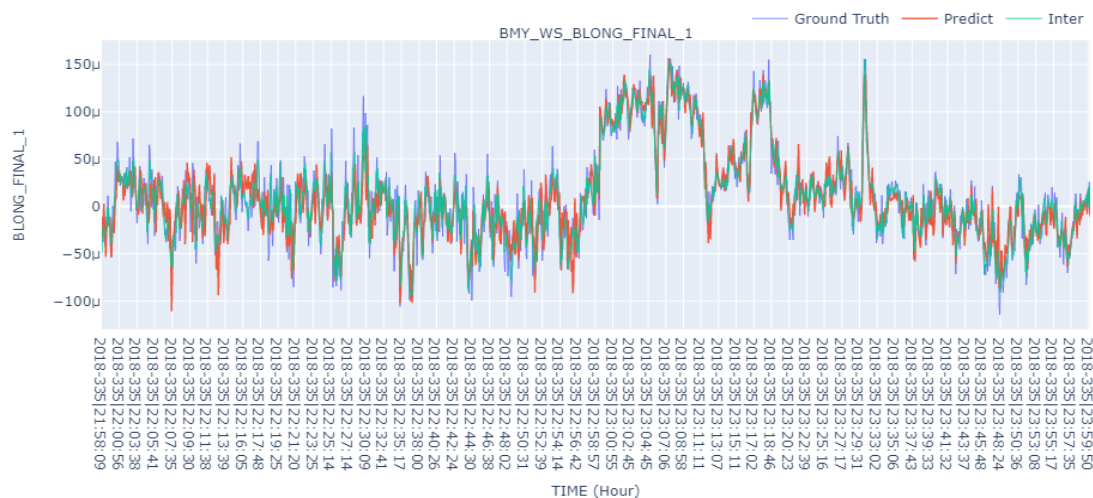


Figure 21: Unnormalized data plot BLONG expand the time

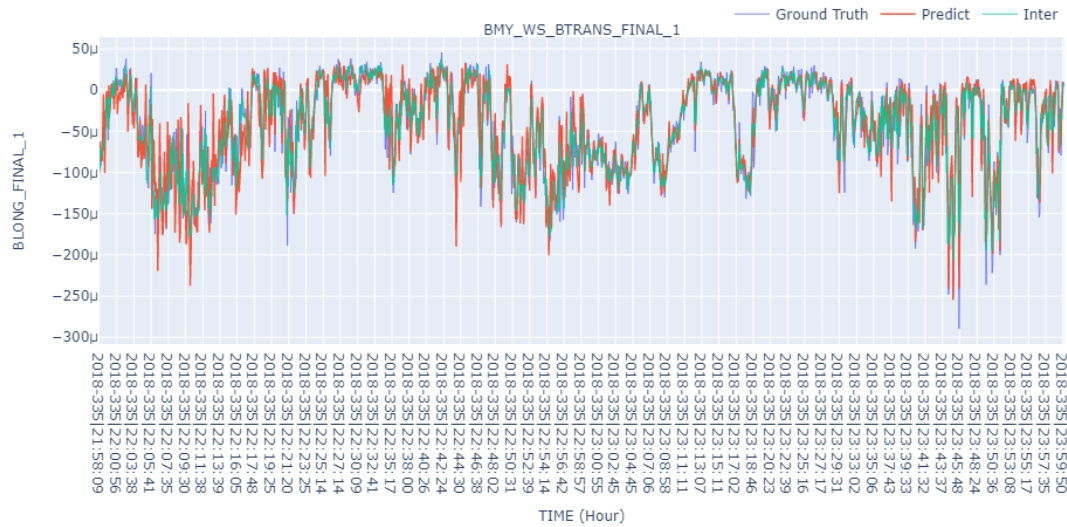


Figure 22: Unnormalized data plot BTRANS expand the time

Adjust the time window from 2018-335|21:58:09 to 2018-335|23:59:50 to see the performance of the two different schemes. From the picture information, the scheme using model prediction is better than the scheme using data interpolation. But further data proof is still needed, that is, by calculating RMS.

```
pred rms error is: 1.1414246026296635e-05
inter rms error is: 1.563542300698765e-05
```

Figure 23: Calculate the RMS for two parameters

The data shows that the total error predicted by the model is less than that of the data interpolation scheme.

5.1.2 Prediction from One Boom Data for Single Points

To make a direct comparison with the previous model, in this part, we use the data of a board to predict the model and observe whether we can get better results. Similarly, it is necessary to calculate the error of the model for intuitive comparison.

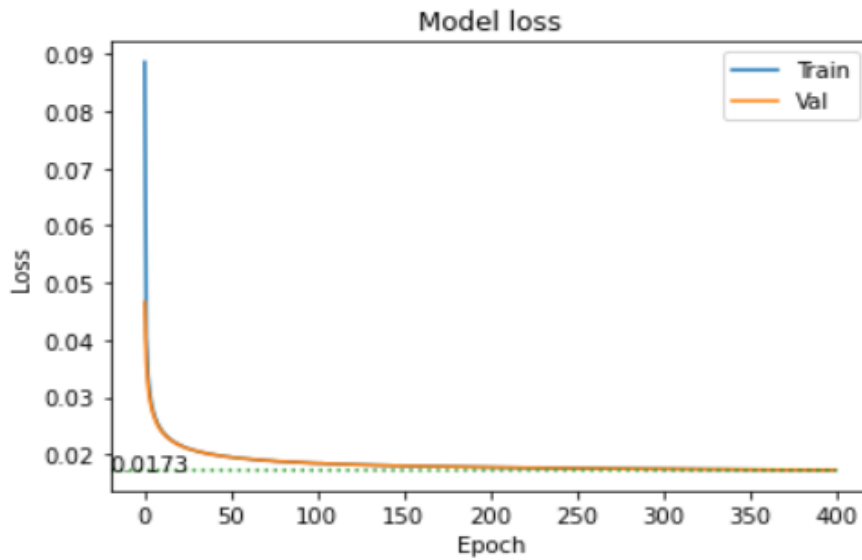


Figure 24: Training process for single-point prediction from one boom

Compared with the results of the two boards, the prediction accuracy is reduced by 0.005, but it is still a very good prediction result.

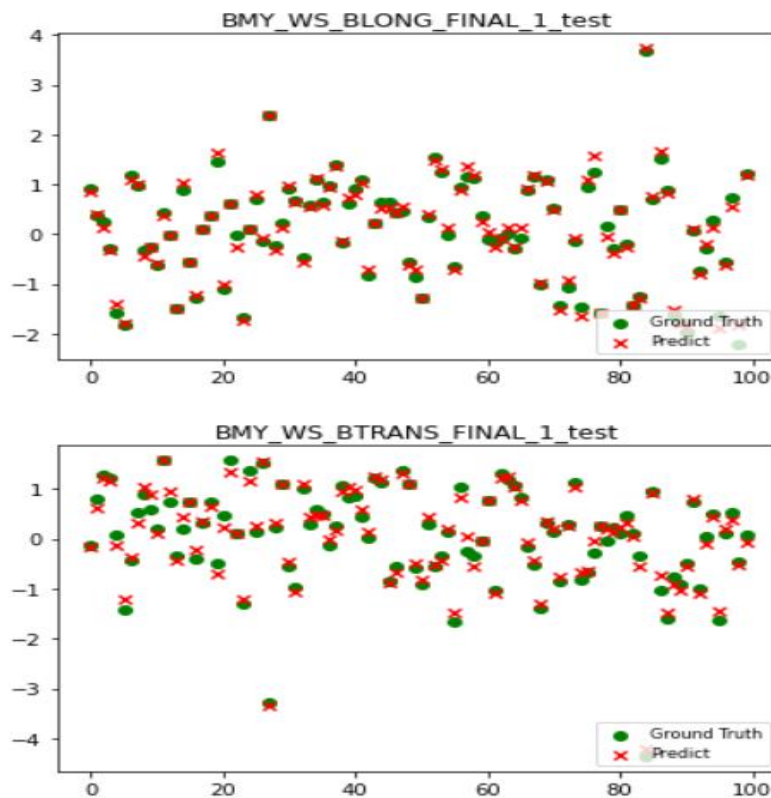


Figure 25: Data plotting for normalized point

It is still random sampling 100 data points to observe the data prediction results after standardization, to observe the prediction of the model.

Data results after de standardization. The comparison of model prediction results, data interpolation results, and real values.

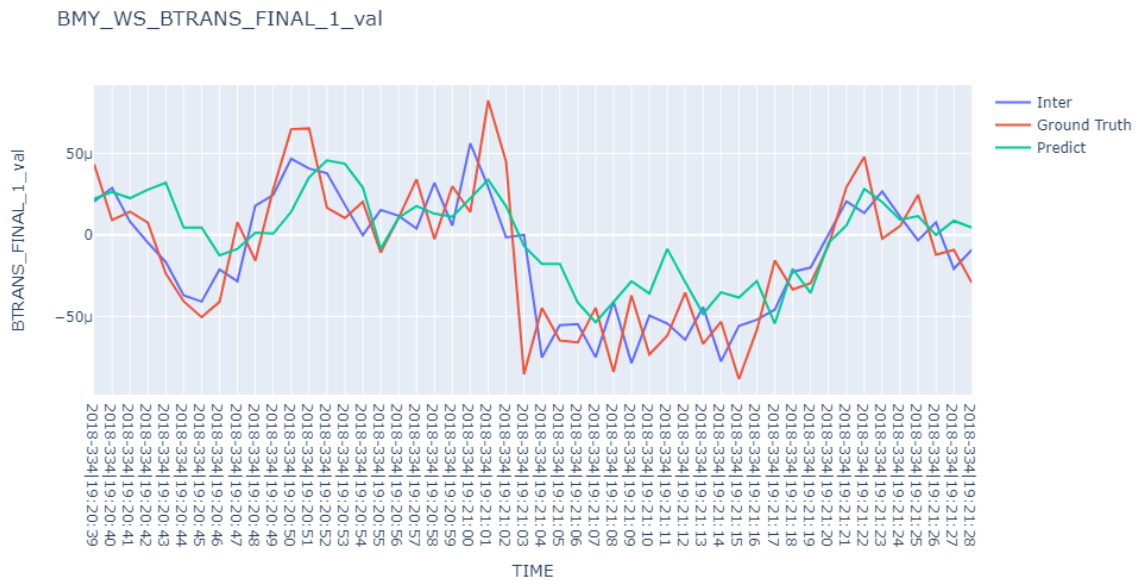


Figure 26: Data plotting for normalized point BTRANS

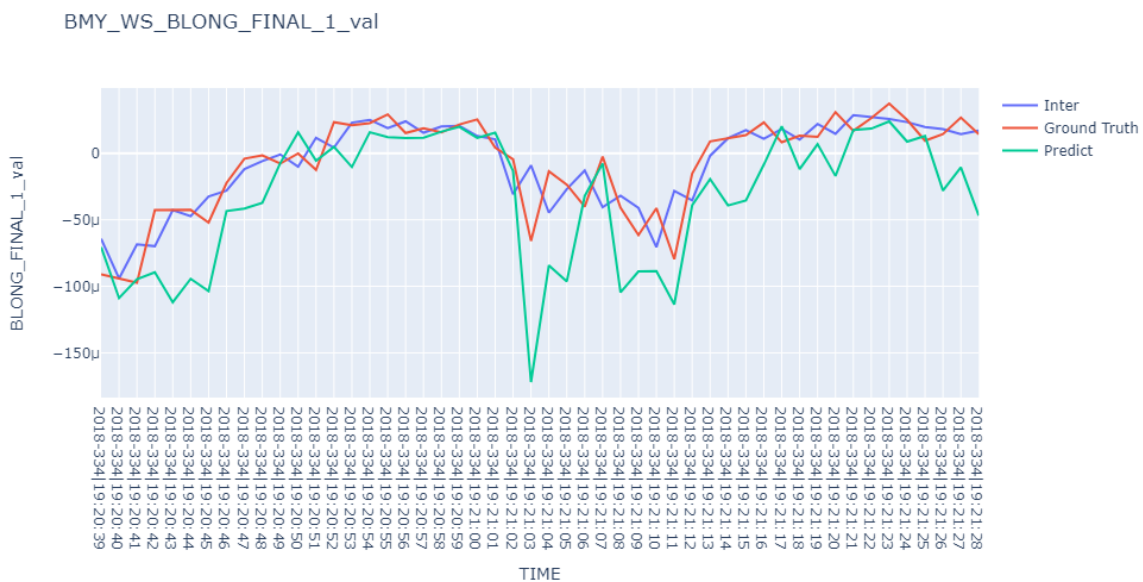


Figure 27: Data plotting for normalized point BLONG

Increasing the time window

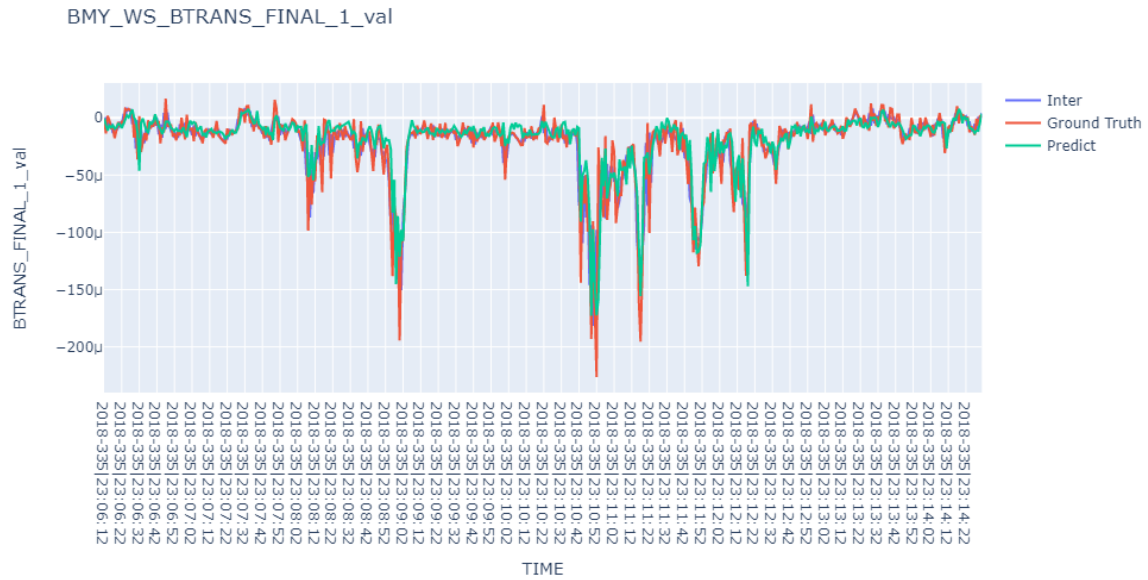


Figure 28: Data plotting for normalized point *BTRANS* time expand

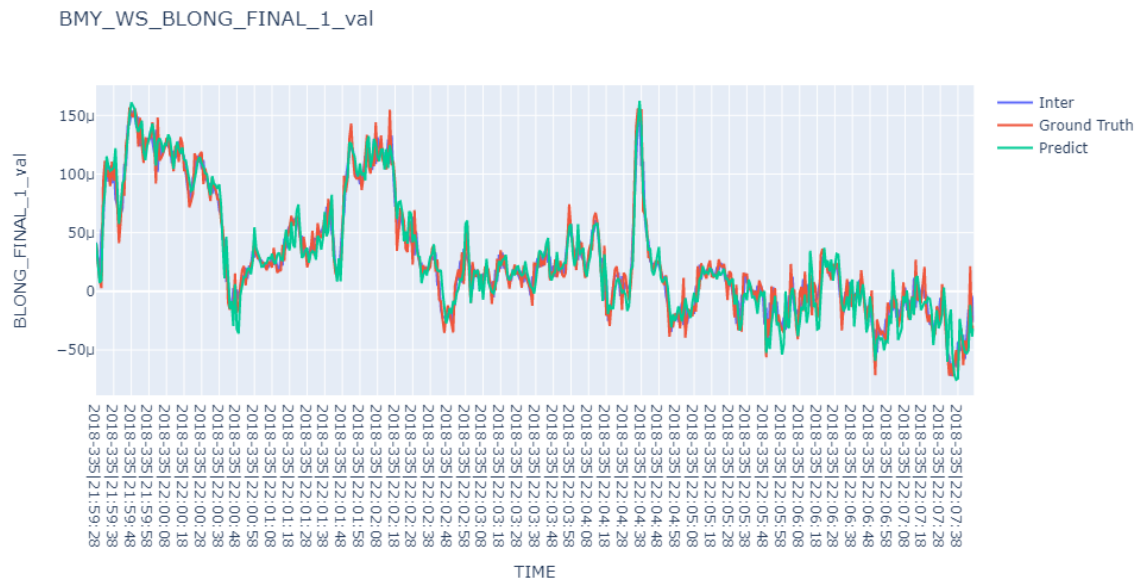


Figure 29: Data plotting for normalized point *BLONG* time expand

Calculate the error like the previous step.

```

pred rms error is: 1.1853697407899454e-05
inter rms error is: 1.563542300698765e-05

```

Figure 30: *RMS for two parameters*

Compared with the data interpolation scheme, the model prediction scheme has a better effect and lower error rate. At the same time, compared with the model prediction of two boards, the accuracy is decreased. To a certain extent, if more input can be given, the accuracy of output results will be improved. This experimental result is consistent with the previous experimental assumption.

5.2 Prediction from Data for Two Consecutive Points

5.2.1 Prediction from Two Booms Data for Two Consecutive Points

In this part, a new model is used to predict continuous time points

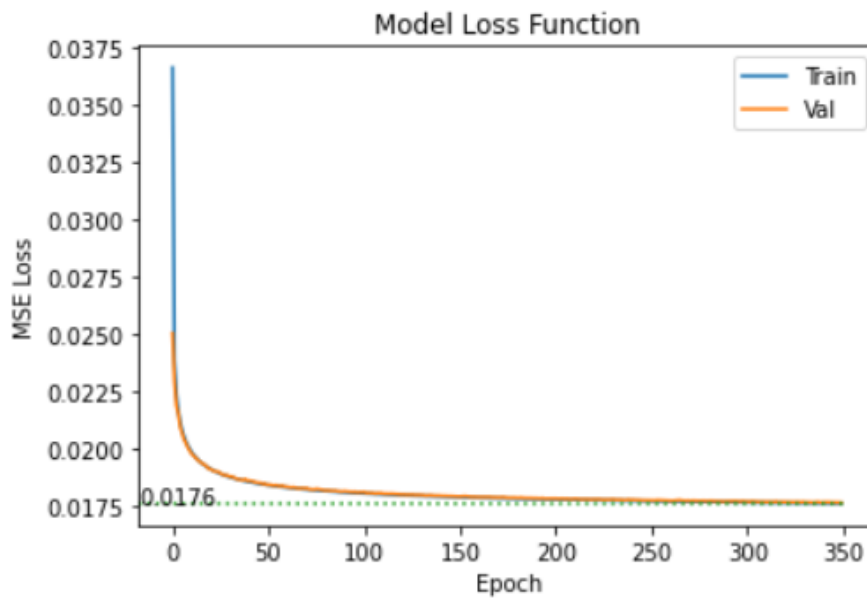


Figure 31: *Training process for two consecutive points from two booms*

Compared with the data using two boards to predict a point, the loss function of the whole model increases from 0.168 to 0.176. This means that the accuracy of the model has declined.

Similarly, let's observe the performance of every data at every moment, including standardized data and non-standardized data.

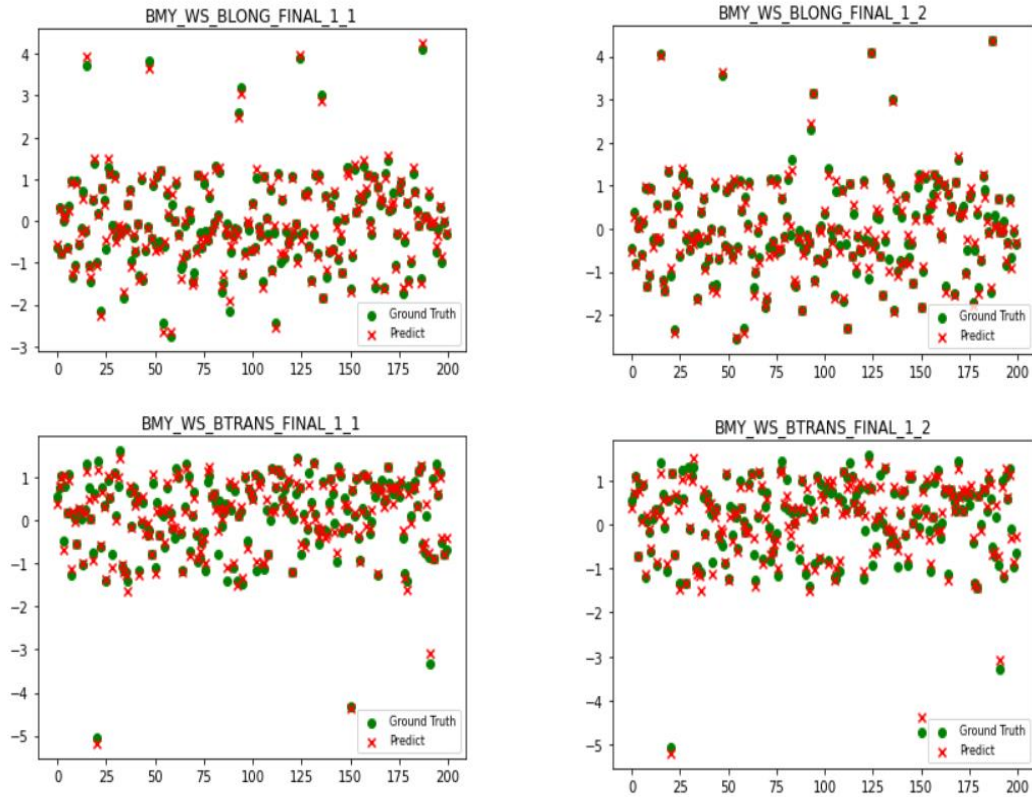


Figure 32 33: Normalized data plot for two consecutive points from two booms

It is also because the accuracy of the model is lower when two points are predicted continuously than when one point is predicted. It is obvious from the picture that the error of red and green points is obvious at some time.

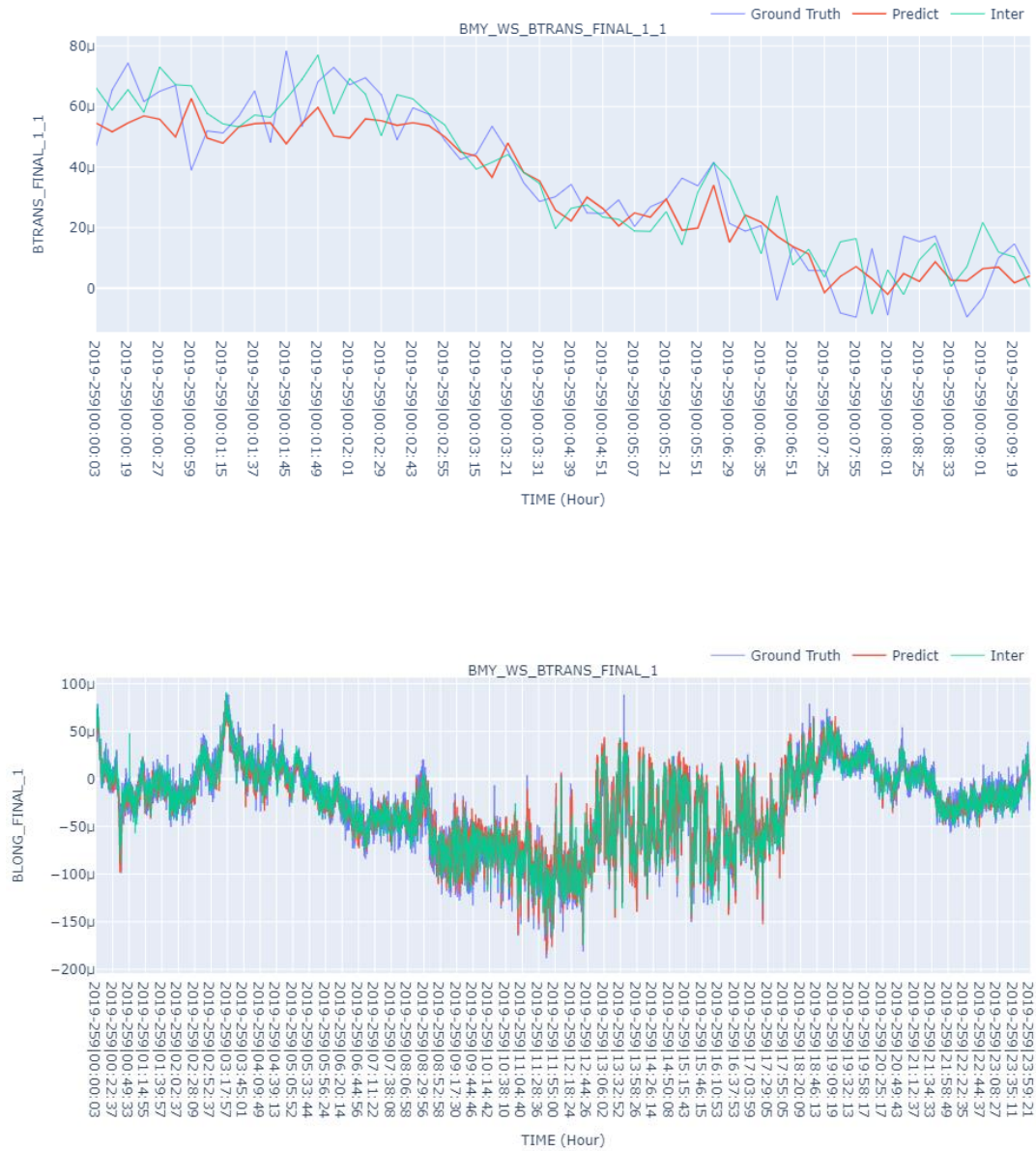


Figure 34 35: Unnormalized data plot for two consecutive points from two booms BTRANS

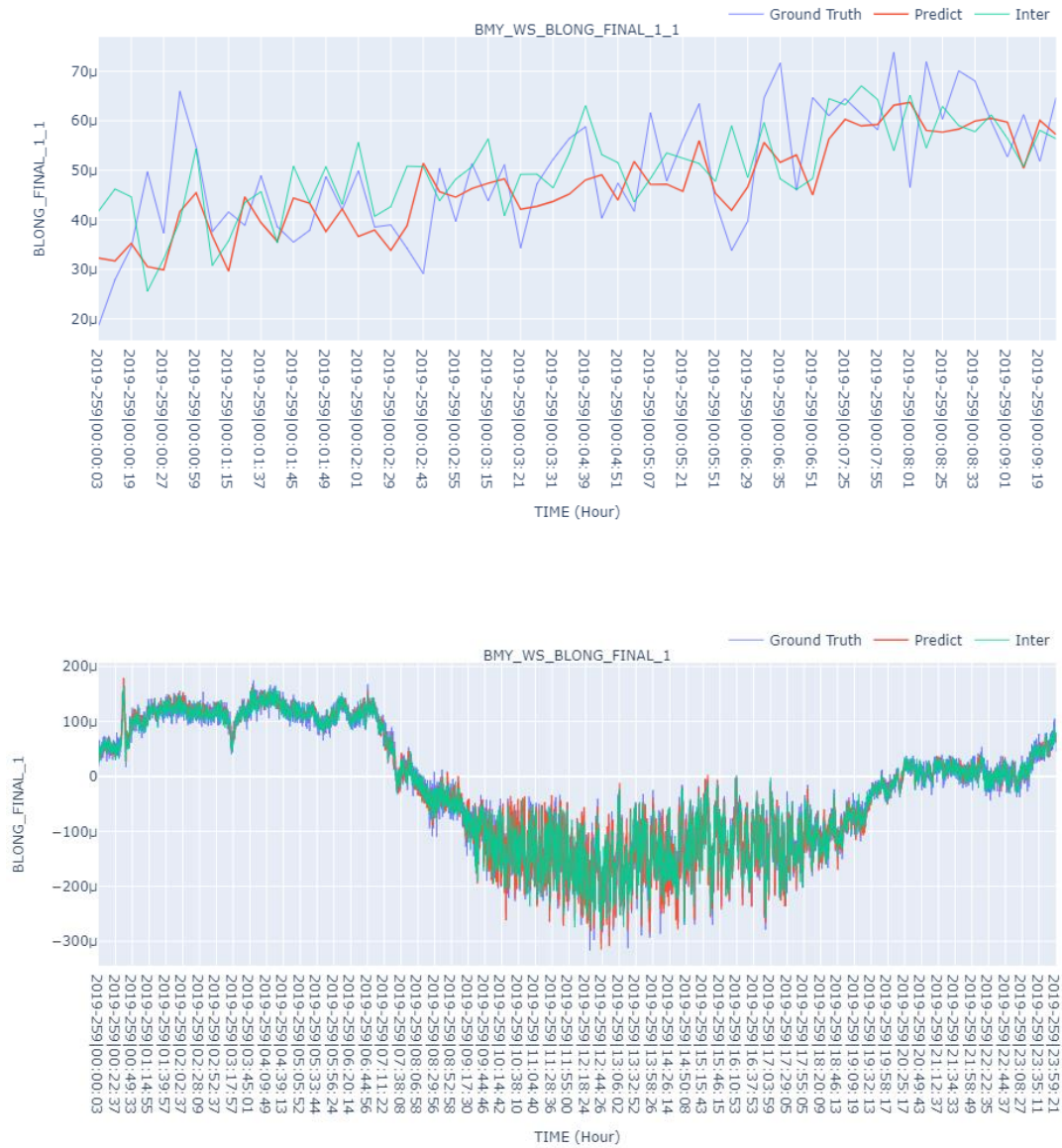


Figure 36 37: Unnormalized data plot for two consecutive points from two booms BLONG

The resulting plot for the second point.



Figure 38 39: Unnormalized data plot for two consecutive points from two booms BTRANS

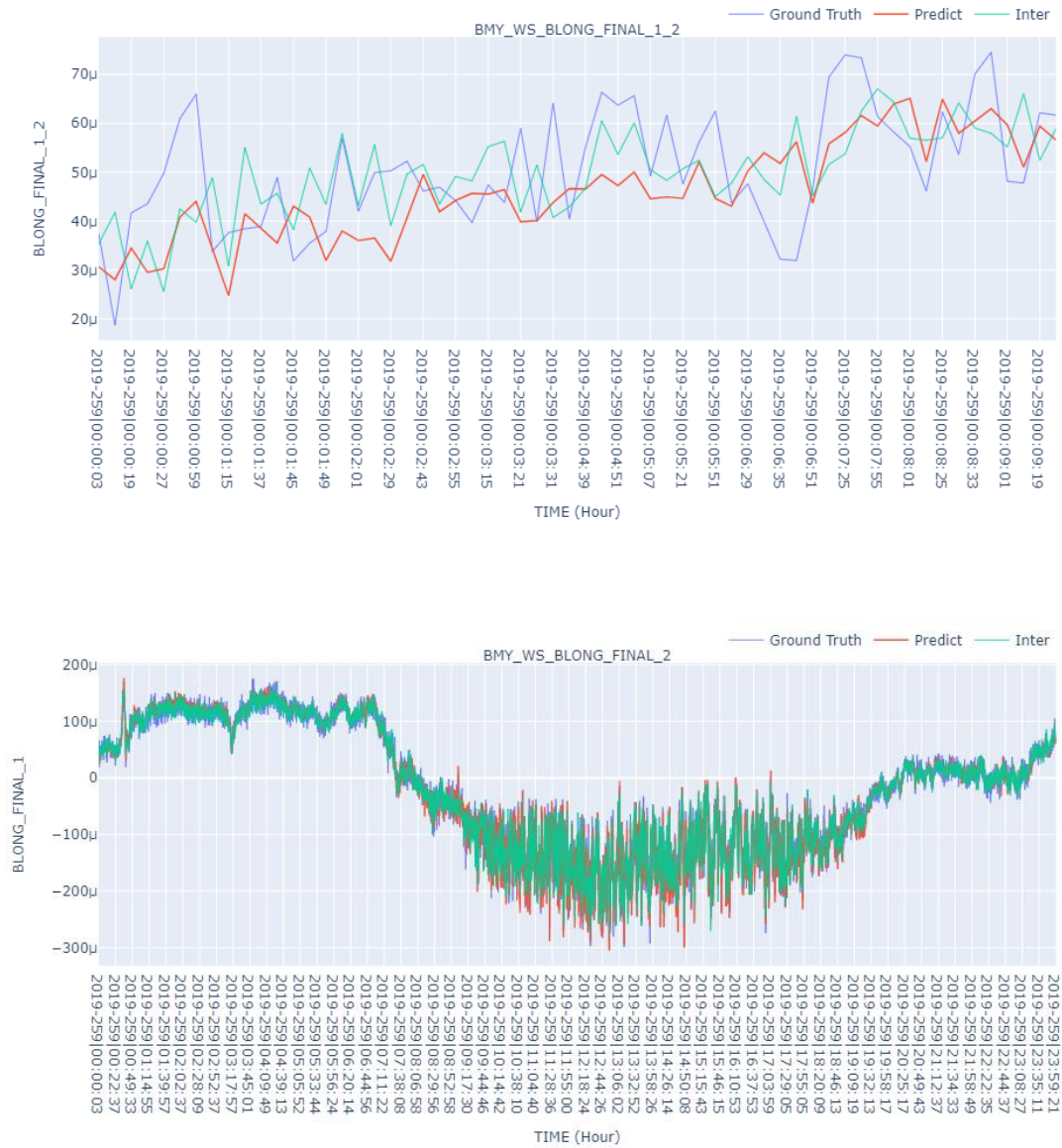


Figure 40 41: Unnormalized data plot for two consecutive points from two booms BLONG

```
BMV_WS_BLONG_FINAL_1_1
pred rms error is: 1.175947865372458e-05
inter rms error is: 1.8027987683818513e-05
BMV_WS_BTRANS_FINAL_1_1
pred rms error is: 1.031713983710088e-05
inter rms error is: 1.4385137477451336e-05
BMV_WS_BLONG_FINAL_1_2
pred rms error is: 1.1611266373972563e-05
inter rms error is: 1.8016747648503453e-05
BMV_WS_BTRANS_FINAL_1_2
pred rms error is: 1.0286968454872765e-05
inter rms error is: 1.4388838599814668e-05
```

Figure 42: *RMS for two points of each BLONG and BTRANS*

By calculating the error rate of each parameter using different schemes, it can be found that no matter which of the four parameters, the results of model prediction are better than data interpolation. Among them, the BLONG parameters of the two points are better than the BTRANS parameters.

In the previous experiment, when using the same data of two boards to predict a point, the overall error ratio is not significantly better than using the same input to predict two points. This shows that, to a certain extent, when the information received by the model is consistent, whether one point is predicted or two points are predicted at the same time, the accuracy is close.

5.2.2 Prediction from One Boom Data for Two Consecutive Points

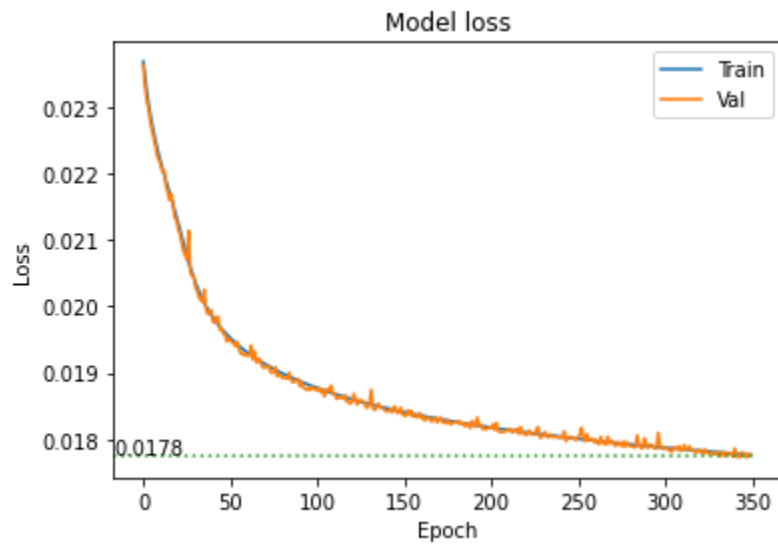


Figure 43: Training process for two points from one boom data

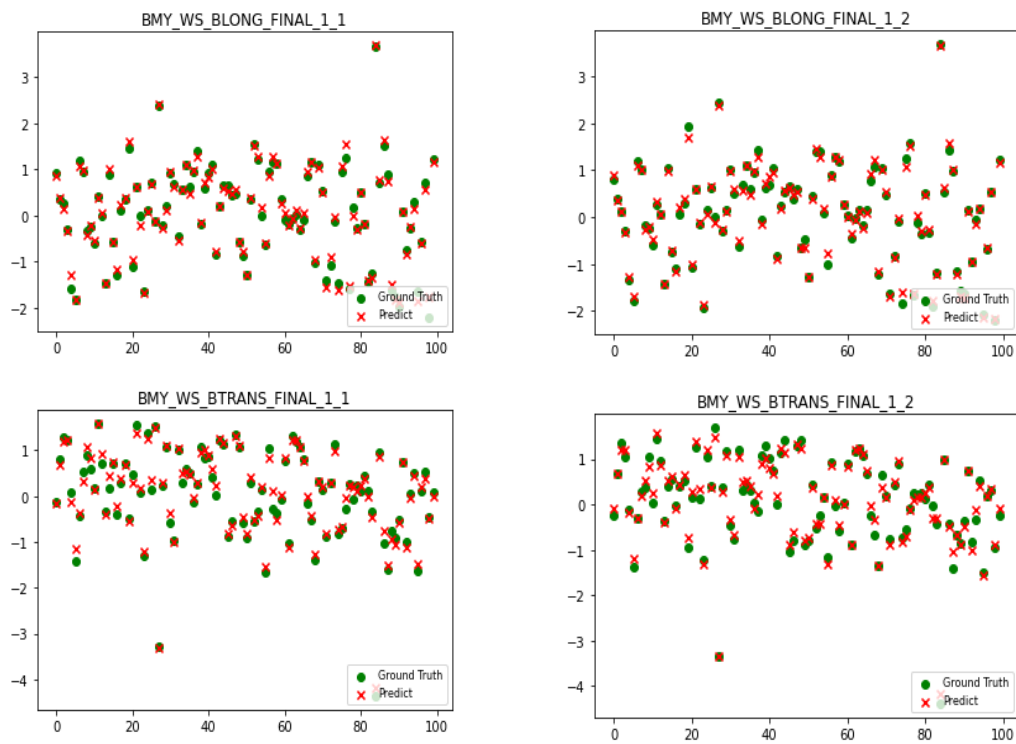


Figure 44: Normalized data plot for two points from one boom data

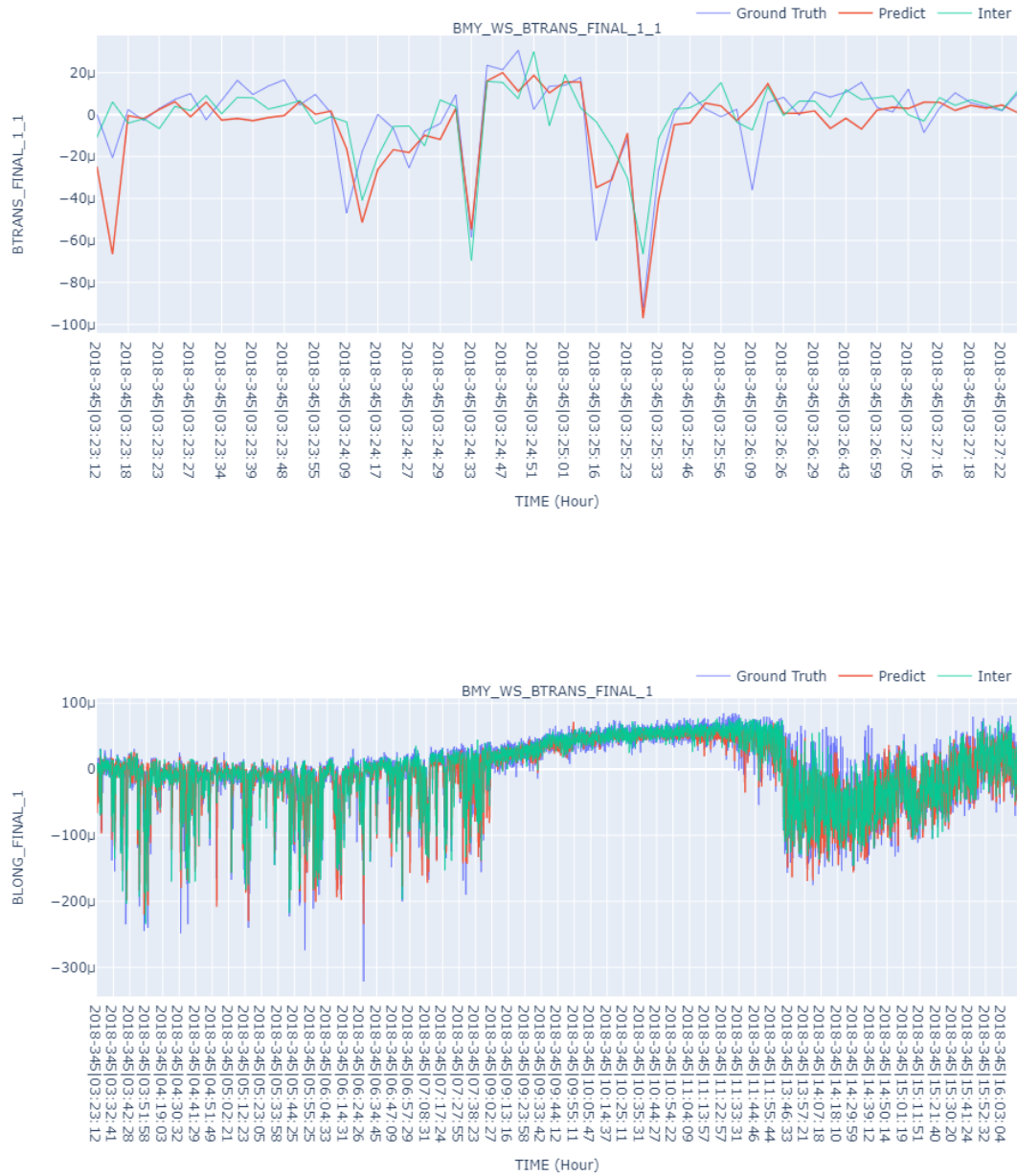


Figure 45 46: Unnormalized data plot for two points from one boom data BTRANS



Figure 47 48: Unnormalized data plot for two points from one boom data *BLONG*



Figure 49 50: Unnormalized data plot for two points from one boom data BTRANS second point

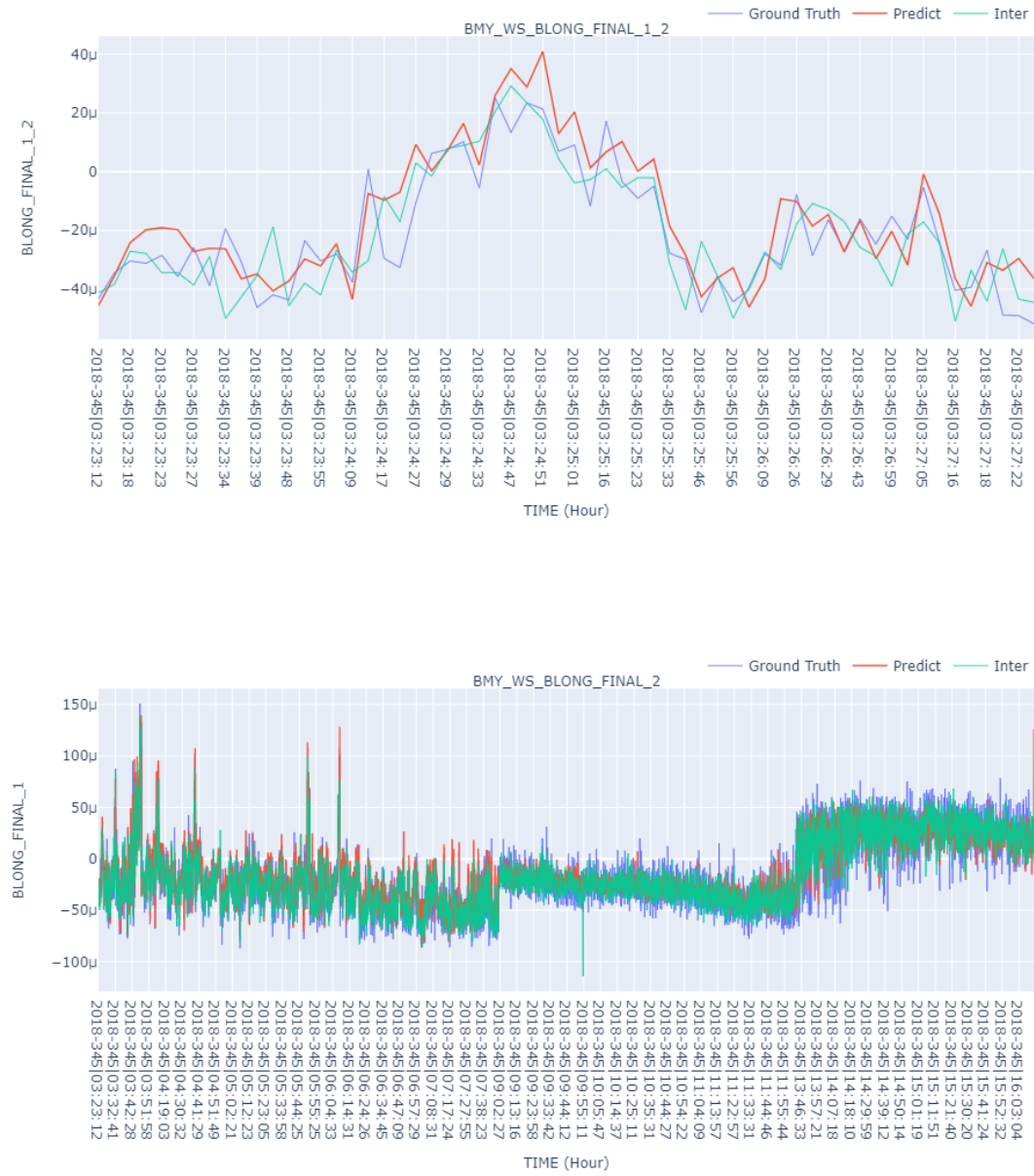


Figure 51 52: Unnormalized data plot for two points from one boom data BTRANS second point

```
BMV_WS_BLONG_FINAL_1_1
pred rms error is: 1.2150288284145119e-05
inter rms error is: 1.566778572331096e-05
BMV_WS_BTRANS_FINAL_1_1
pred rms error is: 1.0210870094379654e-05
inter rms error is: 1.3268091115401354e-05
BMV_WS_BLONG_FINAL_1_2
pred rms error is: 1.1914336966630896e-05
inter rms error is: 1.564600866901342e-05
BMV_WS_BTRANS_FINAL_1_2
pred rms error is: 1.0200065438286914e-05
inter rms error is: 1.326950890205582e-05
```

Figure 53: *RMS for two points four parameters*

For each parameter, the model prediction error is better than the interpolation calculation like the previous result. Therefore, model prediction is a better choice than the interpolation method.

The same result can be obtained by observing the calculated error data. When the input is consistent and only one board data is used, the accuracy of predicting two points at the same time is not worse than that of predicting only one point.

6. Budget

An overall approach of the direct and indirect costs of the project has been made based on the following premises:

- The duration of the project has been three months, 13 weeks, and a total of 25 hours per week.
- The student is considered a Junior Engineer, whose wage is 9 € per hour. Furthermore, the two professors responsible for the supervision of the project are defined as Senior Engineers with a salary of 25 € per hour, 2 hours per week are assumed.
- The computer costs 1500 € 32G RAM and GPU from AMD
- A computer, desk, chair, and other amortizable material have been acquired for a total price of 1.630,00 €.
- Due to the COVID-19 situation, this project has been carried out from home by telematic means. Therefore, the cost of the project assumes a part of the internet and home expenses, which go up to 45€ per month.

The total final cost ascends to 8.812,70 €.

7 Conclusion

The whole project has been carried out under the guidance of two teachers. It is not easy to find a model suitable for the project data itself from many machine learning models and get better results. Fortunately, the project was a success. The effect of the whole machine learning model is better than the previous expectation, and the prediction of loss data is well completed. Nevertheless we expected a higher difference with simple interpolation.

Through the understanding of the wind sensor, we can fully understand the characteristics of the sensor that can record the wind changing with time. It is feasible to predict unknown data from known data. Although in this process, we can find that the more data, the fewer points, the higher the accuracy. Of course, in future research, there is still room for improvement. As mentioned earlier, the project lacks part of the code that can convert angles and other parameters, so there is no way to interpret more information from the data. But by sending the existing data to the CAB, other researchers can carry out subsequent operations.

-
- [1] Anusri Pampari Abhishek Narwekar. Recurrent neural networks architectures. https://slazebni.cs.illinois.edu/spring17/lec20_rnn.pdf, 2016. [Online] Accessed: 15 December 2020.
- [2] W Bruce Banerdt, Suzanne E Smrekar, Don Banfield, Domenico Giardini, Matthew Golombek, Catherine L Johnson, Philippe Lognonné, Aymeric Spiga, Tilman Spohn, Clément Perrin, et al. Initial results from the insight mission on mars. *Nature Geoscience*, pages 1–7, 2020.
- [3] WB Banerdt, S Smrekar, L Alkalai, T Hoffman, R Warwick, K Hurst, W Folkner, P Lognonné, T Spohn, S Asmar, et al. Insight: an integrated exploration of the interior of mars. *LPI*, (1659):2838, 2012.
- [4] Luis Mora Sotomayor (CAB). Twins and ps data products sis. https://atmos.nmsu.edu/PDS/data/PDS4/InSight/twins_bundle/document/twinspsdp_sis_issue9.pdf, 2020. [On- line] Accessed: 18 November 2020.
- [5] Royal Observatory (Greenwich) Dhara Patel. How long is a day on mars? <https://www.rmg.co.uk/discover/explore/how-long-day-on-mars>, 2018. [Online] Accessed: 8 December 2020.
- [6] Bruce M Jakosky and Christopher S Edwards. Inventory of co2 available for terraforming mars. *Nature astronomy*, 2(8):634–639, 2018.
- [7] Lukasz Kowalski. *Contribution to advanced hot wire wind sensing*. PhD thesis, UPC Barcelona, 2016. Research in Electronic Engineering.
- [8] YQ Ni and M Li. Wind pressure data reconstruction using neural network techniques: A comparison between bpnn and grnn. *Measurement*, 88:468–476, 2016.
- [9] GR Osinski, CS Cockell, A Pontefract, and HM Sapers. The role of meteorite impacts in the origin of life. *Astrobiology*, 20(9):1121–1149, 2020.
- [10] Aymeric Spiga, Don Banfield, Nicholas A Teanby, François Forget, Antoine Lucas, Balthasar Kenda, Jose Antonio Rodriguez Manfredi, Rudolf Widmer-Schmidrig, Naomi Murdoch, Mark T Lemmon, et al. Atmospheric science with insight. *Space Science Reviews*, 214(7):109, 2018.

- [11] Li Tian, Guorui Li, and Cong Wang. A data reconstruction algorithm based on neural network for compressed sensing. In *2017 Fifth International Conference on Advanced Cloud and Big Data (CBD)*, pages 291–295. IEEE, 2017.
- [12] Colin F Wilson. *Measurement of wind on the surface of Mars*. PhD thesis, University of Oxford, 2003.
- [13] Sobrado J M , J. Martín-Soler, J. A. Martín-Gago. Mimicking Mars: A vacuum simulation chamber for testing environmental instrumentation for Mars exploration[J]. *Review of Scientific Instruments*, 2014, 85(3):035111-035111-8.
- [14] https://www.researchgate.net/publication/51872882_The_Rover_Environmental_Monitoring_Station_Ground_Temperature_Sensor_A_Pyrometer_for_Measuring_Ground_Temperature_on_Mars
- [15] [Curiosity Overview | NASA](#)
- [16] [Curiosity Rover: Facts and Information | Space](#)
- [17] [Overview | Mission – NASA's InSight Mars Lander](#)
- [18] Atmospheric Science with InSight[J]. *Space Science Reviews*, 2018, 214(7).
- [19] Ss A , Mpzb A , Mn A , et al. Wind retrieval from temperature measurements from the Rover Environmental Monitoring Station/Mars Science Laboratory[J]. *Icarus*, 346.
- [20] REMS: The Environmental Sensor Suite for the Mars Science Laboratory Rover[J]. *Space Science Reviews*, 2012, 170(1):583-640.
- [21] Liu G , Liu S , Wang S , et al. Research on Artificial Lateral Line Perception of Flow Field based on Pressure Difference Matrix[J]. *Journal of Bionic Engineering*, 2019, 16(6):1007-1018.

AI Artificial Intelligence.

ANN Artificial Neural Network.

APSS Auxiliary Payload Sensor Subsystem.

ATS Air Temperature Sensor.

Caltech California Institute of Technology.

CSIC Consejo Superior de Investigaciones Científicas.

DL Deep Learning.

ETSETB Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona.

HP3 Heat Flow and Physical Properties Package. 7

InSight Interior Exploration using Seismic Investigations, Geodesy and Heat Transport.

INTA Instituto Nacional de Técnica Aeroespacial.

LSTM Long Short-Term Memory.

MEDA Mars Environmental Dynamics Analyzer.

MLP Multilayer Perceptron.

MNT Micro and Nanotechnologies.

MSE Mean Squared Error.

NASA National Aeronautics and Space Administration.

PCB Printed Circuit Board.

PRT Platinum Resistance Thermometer.

PS Pressure Sensor

REMS Rover Environmental Monitoring Station.

RISE Rotation and Interior Structure Experiment

SEIS Seismic Experiment for Interior Structure

TWINS Temperature and Winds for InSight.

UPC Universitat Politècnica de Catalunya.

UTC Coordinated Universal Time

WS Wind Sensors

