

Marios Anastasopoulos

Optimization of satellite operations scheduling using an AI-based tool

School of Aerospace, Transport and Manufacturing

MSc in Astronautics and Space Engineering

MSc

Academic Year: 2022 - 2023

Supervisor: Dr. Nicola Garzaniti

September 2023



School of Aerospace, Transport and Manufacturing

MSc in Astronautics and Space Engineering

MSc

Academic Year 2022 - 2023

Marios Anastasopoulos

Optimization of satellite operations scheduling using an AI-based tool

Supervisor:  Dr. Nicola Garzaniti

September 2023

This thesis is submitted in partial fulfilment of the requirements for the degree of MSc in Astronautics and Space Engineering

© Cranfield University 2023. All rights reserved. No part of this publication may be reproduced without the written permission of the copyright owner.

Academic integrity declaration

I declare that:

* the thesis submitted has been written by me alone.
* the thesis submitted has not been previously submitted to this university or any other.
* that all content, including primary and/or secondary data, is true to the best of my knowledge.
* that all quotations and references have been duly acknowledged according to the requirements of academic research.

I understand that to knowingly submit work in violation of the above statement will be considered by examiners as academic misconduct.

Abstract

Satellite platforms need to be constantly monitored by satellite operators and action needs to be taken in order to keep the satellites’ performance nominal. Such activities can be scheduled in advance, because of the knowledge of the orbit of the satellites and the environment they cope with. However, the conditions they face are dynamic and constantly changing and as a result additional operational procedures need to be conducted out of schedule.

The aim of this thesis was to create a tool that can be incorporated in satellite operations and can assist in the evaluation and optimization of operations scheduling. Using machine learning, models were developed that can forecast the reaction wheel rotation rates and assess if the operation of wheel unloading will need to be performed. Then a ticket is automatically created to notify the operator that action needs to be taken out of the scheduled activities.

The research was conducted in collaboration with Hellas-Sat, a Greek-Cypriot satellite operator owned by Arabsat, and the tool was applied and developed on one of their platforms Hellas-Sat 3, that provides satellite communication services such as television and radio broadcasting, internet connectivity, and data transmission. Hellas-Sat operates GEO platforms that provide coverage over Europe, the Middle East and South Africa.

Keywords:

Machine Learning, Space, LSTM, Hyperpameter tuning, Timeseries forecasting, Reaction Wheels

Acknowledgements

I would like to thank my supervisor Dr. Nicola Garzaniti for the immense support he provided through my IRP. From the very first moment of conceptualizing the topic, to the very end of it, his insight and feedback has been extremely helpful and the guidance that he offered has been priceless.

Additionally, I would like to thank Hellas-Sat for making this project possible, by allowing and trusting me to have access to their data and information. All their engineers have provided me support and have been eager to help me with the project whenever asked, without second thoughts.

Finally, I would like to thank my family for supporting me all these years and for making it possible for me to study at Cranfield University and follow my dreams.

Table of Contents

[Academic integrity declaration i](#_Toc145611604)

[Abstract ii](#_Toc145611605)

[Acknowledgements iii](#_Toc145611606)

[List of Figures v](#_Toc145611607)

[List of Tables vii](#_Toc145611608)

[List of Equations viii](#_Toc145611609)

[List of Abbreviations ix](#_Toc145611610)

[1 Introduction 10](#_Toc145611611)

[1.1 Aim & Objectives 10](#_Toc145611612)

[1.2 Structure of Thesis 11](#_Toc145611613)

[2 Literature Review 13](#_Toc145611614)

[2.1 Time Series Forecasting 13](#_Toc145611615)

[2.2 Recurrent Neural Networks 14](#_Toc145611616)

[2.2.1 GRU 14](#_Toc145611617)

[2.2.2 LSTM 15](#_Toc145611618)

[2.2.3 Transformers 17](#_Toc145611619)

[2.3 Structuring Machine Learning Projects 19](#_Toc145611620)

[2.4 Hellas-Sat 3 21](#_Toc145611621)

[2.5 Tools 23](#_Toc145611622)

[3 Methodology 24](#_Toc145611623)

[3.1 Data Collection 24](#_Toc145611624)

[3.2 Data Pre-Processing 27](#_Toc145611625)

[3.3 Data Analysis and Preparation 29](#_Toc145611626)

[3.4 Hyperparameter Optimization 29](#_Toc145611627)

[3.5 Model Comparison 36](#_Toc145611628)

[3.6 Ticket Creation 42](#_Toc145611629)

[4 Results 46](#_Toc145611630)

[4.1 Model Selection 46](#_Toc145611631)

[4.1.1 Reaction Wheel 1 46](#_Toc145611632)

[4.1.2 Reaction Wheel 2 50](#_Toc145611633)

[4.1.3 Reaction Wheel 3 52](#_Toc145611634)

[4.1.4 Reaction Wheel 4 54](#_Toc145611635)

[4.2 Predictions 56](#_Toc145611636)

[4.3 Ticket 60](#_Toc145611637)

[4.4 Impact of the developed solution 61](#_Toc145611638)

[5 Conclusions 63](#_Toc145611639)

[5.1 Summary and Achievements 63](#_Toc145611640)

[5.2 Limitations 63](#_Toc145611641)

[5.3 Future work 64](#_Toc145611642)

[5.4 Lessons Learned 64](#_Toc145611643)

[References 65](#_Toc145611644)

[Appendices 68](#_Toc145611645)

# List of Figures

[Figure 2‑1 GRU cell structure, Source:[16] 13](#_Toc145527646)

[Figure 2‑2 Vanilla LSTM cell structure, Source:[16] 14](#_Toc145527647)

[Figure 2‑3 Architecture of Transformer forecasting model, Source:[17] 16](#_Toc145527648)

[Figure 2‑4 Multi-Layered LSTM model architecture 18](#_Toc145527649)

[Figure 3‑1 Process diagram 23](#_Toc145527650)

[Figure 3‑2 Reaction Wheel 1 rates Jan 2022 - Dec 2022 24](#_Toc145527651)

[Figure 3‑3 Reaction Wheel 2 rates Jan 2022 - Dec 2022 24](#_Toc145527652)

[Figure 3‑4 Reaction Wheel 3 rates Jan 2022 - Dec 2022 25](#_Toc145527653)

[Figure 3‑5 Reaction Wheel 4 rates Jan 2022 - Dec 2022 25](#_Toc145527654)

[Figure 3‑6 Non-normalized predictions for example LSTM model 27](#_Toc145527655)

[Figure 3‑7 Normalized predictions for example LSTM model 28](#_Toc145527656)

[Figure 3‑8 Ticket creation flow process 41](#_Toc145527657)

[Figure 3‑9 Rolling input window for predictions 42](#_Toc145527658)

[Figure 4‑1 RW1 RNN model structure 46](#_Toc145527659)

[Figure 4‑2 Code for RW1 model 46](#_Toc145527660)

[Figure 4‑3 RW1 Training and Validation Loss & RMSE at 100 epochs 47](#_Toc145527661)

[Figure 4‑4 RW1 Training dataset Predictions (top) and Actual Data (bottom) 48](#_Toc145527662)

[Figure 4‑5 RW1 Validation dataset Predictions (top) and Actual Data (bottom) 48](#_Toc145527663)

[Figure 4‑6 RW2 Training and Validation Loss & RMSE at 100 epochs 50](#_Toc145527664)

[Figure 4‑7 RW2 Training dataset Predictions (top) and Actual Data (bottom) 50](#_Toc145527665)

[Figure 4‑8 RW2 Validation dataset Predictions (top) and Actual Data (bottom) 51](#_Toc145527666)

[Figure 4‑9 RW3 Training and Validation Loss & RMSE at 100 epochs 52](#_Toc145527667)

[Figure 4‑10 RW3 Training dataset Predictions (top) and Actual Data (bottom) 52](#_Toc145527668)

[Figure 4‑11 RW3 Validation dataset Predictions (top) and Actual Data (bottom) 53](#_Toc145527669)

[Figure 4‑12 RW4 Training and Validation Loss & RMSE at 100 epochs 54](#_Toc145527670)

[Figure 4‑13 RW4 Training dataset Predictions (top) and Actual Data (bottom) 54](#_Toc145527671)

[Figure 4‑14 RW4 Validation dataset Predictions (top) and Actual Data (bottom) 55](#_Toc145527672)

[Figure 4‑15 RW1 Test dataset Predictions (top) and Actual Data (bottom) 56](#_Toc145527673)

[Figure 4‑16 RW2 Test dataset Predictions (top) and Actual Data (bottom) 56](#_Toc145527674)

[Figure 4‑17 RW3 Test dataset Predictions (top) and Actual Data (bottom) 57](#_Toc145527675)

[Figure 4‑18 RW4 Test dataset Predictions (top) and Actual Data (bottom) 57](#_Toc145527676)

[Figure 4‑19 Survey results for detected threshold violation instances a) Reaction Wheel 1, b) Reaction Wheel 2, c) Reaction Wheel 3, d) Reaction Wheel 4 58](#_Toc145527677)

[Figure B‑1 Imported Libraries 67](#_Toc145527678)

[Figure B‑2 Data Analyzer class, importing and concerting CSV file 67](#_Toc145527679)

[Figure B‑3 Data Split class, splitting data into train, validation and test sets 68](#_Toc145527680)

[Figure B‑4 Model class, pt1. model structure and training 69](#_Toc145527681)

[Figure B‑5 Model class, pt.2 rolling prediction window 69](#_Toc145527682)

[Figure B‑6 Model class, pt.3 plot predictions 70](#_Toc145527683)

[Figure B‑7 Threshold Scanner class, pt.1 scan predictions 71](#_Toc145527684)

[Figure B‑8 Threshold Scanner class, pt.2 scan results gathering 71](#_Toc145527685)

[Figure B‑9 Create Ticket class, pt.1 establish connection with database 72](#_Toc145527686)

[Figure B‑10 Create Ticket class, pt.2 insert critical information in database 73](#_Toc145527687)

[Figure B‑11 Create Ticket class, pt.3 insert procedure steps in database 73](#_Toc145527688)

[Figure B‑12 Hyperparameter Model class, pt.1 model building for optimization 74](#_Toc145527689)

[Figure B‑13 Hyperparameter Model class, pt.2 model building for optimization 74](#_Toc145527690)

[Figure B‑14 Hyperparameter Model class, pt3. hyperband search, model training and results reporting 75](#_Toc145527691)

List of Tables

[Table 3‑1 Optimized Hyperparamters 29](#_Toc145527692)

[Table 3‑2 Reaction Wheel 1 architecture results of hyperparameter optimization 31](#_Toc145527693)

[Table 3‑3 Reaction Wheel 2 architecture results of hyperparameter optimization 32](#_Toc145527694)

[Table 3‑4 Reaction Wheel 3 architecture results of hyperparameter optimization 33](#_Toc145527695)

[Table 3‑5 Reaction Wheel 4 architecture results sof hyperparameter optimization 34](#_Toc145527696)

[Table 3‑6 Reaction Wheel 1 Architectures Validation Loss and RMSE 37](#_Toc145527697)

[Table 3‑7 Reaction Wheel 2 Architectures Validation Loss and RMSE 38](#_Toc145527698)

[Table 3‑8 Reaction Wheel 3 Architectures Validation Loss and RMSE 39](#_Toc145527699)

[Table 3‑9 Reaction Wheel 4 Architectures Validation Loss and RMSE 40](#_Toc145527700)

[Table 4‑1 Selected model architecture for Reaction Wheel 1 45](#_Toc145527701)

[Table 4‑2 Selected model architecture for Reaction Wheel 2 49](#_Toc145527702)

[Table 4‑3 Selected model architecture for Reaction Wheel 3 51](#_Toc145527703)

[Table 4‑4 Selected model architecture for Reaction Wheel 4 53](#_Toc145527704)

[Table 4‑5 Database entry with critical operation information 59](#_Toc145527705)

[Table 4‑6 Database entry with procedure steps 59](#_Toc145527706)

List of Equations

[(1‑1) 14](#_Toc145527707)

[(1‑2) 14](#_Toc145527708)

[(1‑3) 14](#_Toc145527709)

[(1‑4) 14](#_Toc145527710)

[(1‑5) 15](#_Toc145527711)

[(1‑6) 15](#_Toc145527712)

[(1‑7) 15](#_Toc145527713)

[(1‑8) 15](#_Toc145527714)

[(1‑9) 15](#_Toc145527715)

[(1‑10) 15](#_Toc145527716)

[(3‑1) 30](#_Toc145527717)

List of Abbreviations

|  |  |
| --- | --- |
| ADAM | Adaptive Moment Estimation |
| ARIMA | Autoregressive Integrated Moving Average |
| CCW | Counter-Clockwise |
| CSV | Comma Separated Values |
| CW | Clockwise |
| GRU | Gated Recurrent Unit |
| LSTM | Long Short-Term Memory |
| ML | Machine Learning |
| MSE | Mean Squared Error |
| RMSE | Root Mean Squared Error |
| RNN | Recurrent Neural Network |
| RPM | Revolutions per Minute |
| RW | Reaction Wheel |

# Introduction

Ever since satellites have been deployed, our quality of life has been drastically improved. Most of the time we are using their services without even realizing it, such us telecommunication services or positioning services. However, what is often overlooked is the fact that all these spacecraft that are being used needs to be monitored and maintained operational and in good condition, through actions that are taken by satellite operators or automated procedures that are being executed on-board [1]–[4].

Additionally, in recent years there has been accelerated development of artificial intelligence and machine learning tools, that are being increasingly incorporated in everyday life. For instance, on a personal level many people use virtual home assistants that make use of AI, in order to control smart home devices or get notifications, and navigation applications use AI to predict traffic conditions and suggest a faster route. On a professional or industry level AI is used in finance to detect fraud or predict stock price evolution, in energy management to create smart grids and optimize energy production and distribution and in manufacturing to make robots more efficient and increase quality control[5], [6].

Unfortunately, artificial intelligence and machine learning has not yet been integrated fully or adopted partially in commercial space applications. An area of interest for machine learning or artificial intelligence applications is satellite operations scheduling. Although such operations can be scheduled in advance, because of the knowledge of the orbit of the satellites and the environment they cope with, the conditions they face are dynamic and constantly changing [7]–[9]. As a result, additional operational procedures need to be conducted out of scheduled and potential contingency situations need to be foreseen.

## Aim & Objectives

The aim of this thesis is to create a tool that can be integrated in current satellite operations infrastructure and can assist in the evaluation and optimization of operations scheduling.

**Objective 1:** Identification of the subsystem of interest, by studying the platform that will be used and its subsystems and assess in coordination with the operator company which subsystem is an area of interest that could benefit from the tool we want to create.

**Objective 2:** Development and validation of the tool using machine learning based on the data that will be provided and the operating structure of the company.

**Objective 3:** Integration and testing in the pipeline of processes of the operator company, so that in the future it can be used and considered a functional asset to them.

## Structure of Thesis

Chapter 1 is the introduction of the problem that we are challenged with, as well as the aim and objectives of the project.

Chapter 2 covers the literature review concerning time series forecasting, recurrent neural networks and its model variations that can be used, as well as their advantages and disadvantages. It also covers how a machine learning model is being built and the choices that were made for our model concerning the hyperparameters that were used. Moreover, this chapter includes the satellite that the tool will be applied on and the subsystem of interest and finally the tools that were used to create, validate and test the model.

Chapter 3 discusses the methodology that was followed. More specifically, how the data that were used were collected, analyzed and pre-processed. Moreover, it covers the hyperparameter optimization that was performed for the machine learning models, as well as comparison of the results. Finally, it includes the ticket creation process and the information that it will include.

Chapter 4 contains the results for the machine learning models that were chosen, a demonstration of usage of the tool that was created using never before seen data and the resulting created ticket, as part of the validation process and finally a comparison of the created tool with the current approach to the problem.

Chapter 5 is a summary of the work that has been done and the conclusions. It discusses the limitations of the tool that was created and possible future work.

# Literature Review

Satellites are highly complex systems that are constituted of many different subsystems, each responsible for specific operations. All subsystems are equipped with a wide range of sensors, such as temperature, pressure, voltage or current sensors, that continuously take measurements. Those measurements are the telemetry that is being transmitted to the ground for monitoring and detailed analysis by the satellite operators, in order to determine the state of health of each subsystem and eventually of the entire satellite. Since the amount of data that is being generated and needs to be analysed is massive, there has been identified the need for automation either on board the satellites or on the ground [10], [11]. Since almost all telemetry that is being produced has temporal dependency and the measurements that are being taken are in series, time series forecasting could be applied[12]–[14]. Additionally, many operations are time sensitive and the amount of data that needs to be processed for them is massive. Therefore a way is needed to analyze all the data fast and derive conclusions about the temporal evolution of the telemetry for those time critical procedures.

## Time Series Forecasting

Time series forecasting is an analytical technique used to predict future values of a variable based on its historical data points, where each observation is recorded at specific time intervals. This technique finds applications across numerous domains, including finance, economics, healthcare, weather prediction, supply chain management, and more. The ability to make accurate predictions about future trends and patterns in time-dependent data is invaluable for decision-making, resource allocation, risk management, and strategic planning.

In time series forecasting, the primary goal is to develop models that can capture underlying patterns, seasonality, trends, and irregular fluctuations present in the data. Accurate forecasts enable organizations and stakeholders to anticipate market trends, optimize resource allocation, and mitigate potential risks, leading to improved operational efficiency and better-informed decisions.

Traditional statistical methods for time series forecasting, such as ARIMA and exponential smoothing, often struggle to capture complex temporal dependencies and nonlinear relationships in data [15].

Other methods for time series forecasting include time series decompositioning, where the time series is broken down to its components such as the trend and seasonality, and time series regression models that uses multiple variables to predict future values. However, those methods have the disadvantages that they don’t work well for non-linear or complex data and require detailed parameter tuning and preprocessing to avoid overfitting and get good results.

Those limitations prompted the exploration of more sophisticated techniques, leading to the rise of deep learning and specifically, Recurrent Neural Networks (RNNs).

## Recurrent Neural Networks

RNNs are a class of neural networks uniquely designed to process sequences of data by maintaining a hidden state that captures information from previous time steps. However, vanilla RNNs have difficulties learning long-range dependencies due to the vanishing gradient problem. This limitation gave birth to a specialized type of RNN known as Long Short-Term Memory (LSTM).

Typical types of RNNs used for time series forecasting are GRUs, LSTM and transformers.

### GRU

Gated Recurrent Units (GRUs) are a type of gated RNN, similar in nature to LSTMs but with a simplified architecture. GRUs also incorporate gating mechanisms to control the flow of information, but they merge the cell state and hidden state, resulting in a more streamlined design compared to LSTMs [16]. GRUs tend to have fewer parameters, which can make them computationally efficient and sometimes better suited for smaller datasets. While they might not capture long-range dependencies as effectively as LSTMs, GRUs have shown competitive performance in various time series forecasting applications .

A diagram of a diagram

Description automatically generated

Figure ‑ GRU cell structure, Source:[16]

|  |  |
| --- | --- |
|  | (‑) |
|  | (‑) |
|  | (‑) |
|  | (‑) |

### LSTM

LSTM, as previously mentioned, is a specialized variant of Recurrent Neural Networks (RNNs) designed to overcome the challenges posed by vanishing gradients. LSTMs employ memory cells, gating mechanisms, and activation functions to effectively capture temporal dependencies in sequential data [5]. This makes them well-suited for time series forecasting tasks, where patterns and relationships can span across different time steps. LSTMs have been widely adopted in various domains, ranging from finance and energy to healthcare and beyond. They excel in capturing both short-term fluctuations and long-term trends in time series data.

A diagram of a tank

Description automatically generated

Figure ‑ Vanilla LSTM cell structure, Source:[16]

|  |  |
| --- | --- |
|  | (‑) |
|  | (‑) |
|  | (‑) |
|  | (‑) |
|  | (‑) |
|  | (‑) |

At each time step, the LSTM cell performs the following operations:

* The cell calculates how much of the new input Xt at the current time step should be added to the memory cell. This is controlled by the input gate, which is a sigmoid function that takes into account the current input Xt and the previous hidden state ht-1.
* The cell decides how much of the previous memory cell state Ct-1 should be forgotten or retained. This is determined by the forget gate, another sigmoid function that considers the previous hidden state and the current input.
* The cell updates its memory cell state by combining the information from the input gate it and the forget gate ft. This step involves element-wise multiplication and addition operations.
* The cell determines how much of the memory cell's current state should be output as the new hidden state ht. This is controlled by the output gate, which is one more sigmoid function based on the input Xt and the previous hidden state ht-1.
* The final hidden state ht is computed based on the output gate and the current memory cell state. This hidden state becomes the output of the current LSTM cell and is also passed as input to the next time step's LSTM cell in the same layer or to the corresponding LSTM cell in the next layer.

### Transformers

Transformers, originally designed for natural language processing tasks, have been adapted for time series forecasting with remarkable success. Unlike traditional RNNs, transformers process input data in parallel rather than sequentially, which can lead to faster training times. Attention mechanisms within transformers allow them to focus on relevant temporal relationships across various time steps, making them highly effective at capturing complex patterns in time series data.

In the context of time series forecasting, the "Transformer-XL" and "Temporal Fusion Transformers" are notable architectures that extend transformers to handle sequential data. These models take into account the order and temporal dependencies of the data, making them well-suited for tasks that involve both short-term and long-term forecasting.

A diagram of a decoder

Description automatically generated

Figure ‑ Architecture of Transformer forecasting model, Source:[17]

The architecture of a transformer based timeseries forecasting model includes the following [18]:

* Encoder and Decoder stacks: the encoder stack processes the historical time series data, and the decoder stack generates the future forecast.
* Encoder Input: time series data is first transformed into numerical representations that the Transformer can understand. Each data point in the time series is represented as a vector that includes relevant information, like the actual value, time step, and any other contextual features.
* Self-Attention: It allows the model to weigh the importance of each element in the input sequence with respect to all other elements. This captures both short-term and long-term dependencies in the sequence.
* Feed forward: After the self-attention layers, it includes feed forward neural networks for each position in the sequence. They transform the learned representations to capture more complex patterns.
* Multi Head Attention: The self-attention mechanism is often used in multiple "heads”, in order to capture different types of patterns and relationships. Each head focuses on a different aspect of the data, allowing the model to learn multiple dependencies simultaneously.
* Output layer: produces the final forecast for each future time step. The output can be a single scalar value or a vector representing multiple forecasted values.

## Structuring Machine Learning Projects

To create a model for timeseries forecasting using a Recurrent Neural network, there are several parameters that need to be taken into account and when combined result into many different architectures that have to be explored.

Those parameters are the number of layers within the model, as well as the type of layers and the number of neurons or units in each layer. In this project we chose to use the vanilla LSTM type of RNN for several reasons. First and foremost, LSTM cells have the ability to capture long-term dependencies in sequential data, like in our case, where we want to study the behavior of the reaction wheels of the satellite throughout a whole year. Secondly, LSTM have the advantage that they don’t need a complex model to be able to capture complex patterns within a large dataset. Additionally, when there are sufficient data for training, they tend to perform even better. Finally, by adding a number of layers of LSTM cells within the model both short- and long-term trends can be captured.

Another type of layer that was used is the Dense layer, which is the basic type of layer in machine learning.

The architecture that was chosen was a multi-layered LSTM or “stacked” LSTM. In a “stacked” LSTM network, like the example shown in Figure 2-4, all input data are inserted in the first LSTM layer to generate hidden states, which are then used as inputs in the next LSTM layer and so on, unlit we get to the output layer. By adding more hidden layers, we add greater model depth and can observe the problem in different time scales.



Figure ‑ Multi-Layered LSTM model architecture

In the beginning, the LSTM cells in the first layer receive the initial input sequence of length of the time-window that we are using. The input sequence is processed one time step at a time, moving from the beginning to the end of the sequence.

The output of each LSTM cell in a layer serves as input to the LSTM cells in the next layer. This layer-to-layer movement enables the network to capture increasingly complex temporal patterns and dependencies as the data progresses through the layers.

Other parameters that affect the performance of the model are the learning rate and the optimizer that is used and the input window.

The window size of the input data for an LSTM network refers to the number of time steps or observations that the LSTM considers as a single input sequence. In other words, it determines how much historical information the LSTM can utilize to make predictions or learn patterns in the data. The window size is a crucial hyperparameter in LSTM networks, and selecting an appropriate value depends on the nature of the data and the specific task at hand. A smaller window size may be suitable for tasks with rapid changes, while a larger window size might be more appropriate for tasks with long-term dependencies. In our case, we experimented with window sizes of 24, 48 and 72, which translates to using one, two or three days of data as input, since our data are hourly measurements.

The learning rate is a hyperparameter that controls the step size at which the optimizer adjusts the model's parameters during training. It determines how much the model should update its weights in response to the calculated gradients. A larger learning rate allows for faster convergence during training but may also lead to overshooting and instability. On the other hand, a smaller learning rate leads to slower convergence but provides more stable updates.

The optimizer that was used is ADAM (Adaptive Moment Estimation). It has the ability to adapt the learning rate for each parameter in the model. Instead of using a fixed learning rate, it maintains separate adaptive learning rates based on the historical gradients. In this way it allows the model to converge faster and more efficiently.

## Hellas-Sat 3

Hellas-Sat 3 is a telecommunications satellite operated by Hellas-Sat, that provides coverage to Europe, the Middle East and South Africa, by utilizing Ka and Ku bands and it provides Fixed Satellite Services as well as Broadband Satellite Services. The satellite is placed in a geostationary orbit at 39° East longitude and was launched on 28th of June 2017 [19].

Being a telecommunications satellite, Hellas-Sat 3 requires high pointing accuracy to be able to fulfil its mission and provide satisfactory services. Moreover, the pointing accuracy and the orbital position need to be maintained at all times.

Maintaining the orbital position, also known as stationkeeping, is achieved with a series of scheduled manoeuvres, that utilize the propulsion system of the spacecraft. It is composed by thrusters that are positioned all around the satellite with the aim to apply external force to it in all possible directions. The manoeuvres can be in a North, South, East and West direction.

The satellite is three-axis stabilized in order to achieve and maintain the high pointing accuracy. To achieve that there are four reaction wheels on board the satellite, that rotate to exchange angular momentum with the satellite body. In this way any external torque that is applied due to perturbations such as solar wind pressure or gravitational forces from the moon can be counter-balanced by spinning up or down the reaction wheels, with the result of applying the opposite torque and maintaining the attitude of the satellite.

The reaction wheels are placed in pyramidal configuration, so that they can contribute to the angular momentum of the spacecraft in all three axis. Additionally, by having 4 reaction wheels, there is redundancy aboard the spacecraft and in case of malfunction of one the wheels, the satellite can continue to provide services and fulfil its mission.

However, the momentum that can be stored in each reaction wheel is limited and defined by their properties, such as mass and moment of inertia. Therefore, there are limits to how fast they can spin before being in danger of breaking down. The operational limit for the reaction wheels aboard Hellas-Sat 3 is ±4500 rpm, since they can spin in both directions, clockwise and counterclockwise.

Hellas-Sat engineers and operators aim to keep the wheel speed well below the constructor’s limit, close to ±4000 rpm, so that there is a substantial margin in case of contingency.

When the wheel speed gets close or over the limit, they are considered to be saturated and they may be unable to provide the necessary torque to perform attitude adjustments. To avoid that, wheel unloading needs to take place frequently.

Wheel unloading is the process of reducing the accumulated angular momentum in reaction wheels on a satellite or spacecraft. By using the propulsion system of the spacecraft, thrusters are fired in specific directions to generate torques that counteract the wheel's angular momentum. These thrusters are used to unload momentum while conserving the satellite's limited propellant and preserving the attitude of the satellite.

It may be performed as a standalone procedure, but it can also be combined with stationkeeping manoeuvres, so that fuel is preserved. Such is the case on Hellas-Sat 3. During the North/South and the East/West manoeuvres, the reaction wheels take advantage of the thruster firings to offload their momentum. However, because the manoeuvres take place every two weeks, they is a need to perform the procedure on its own, between the scheduled manoeuvres.

## Tools

A variety of tools are available to process and visualize the data , as well as for building machine learning models , optimizing and testing them.

For this project we are going to use Python programming language because it is spread within the data science community, therefore easily accessible and most importantly there are any libraries that are compatible with it and expand its capabilities.

Also,,MySQ database will be created locally to perform testing, because the database for the tickets that Hellas Sat is using is of the same type and therefore this is where the results would have to be input.

Some of the python libraries that were used are the following:

* Pandas: To import the data that were provided in csv format, filter them and separate them into different sub-dataset that were useful for machine learning.
* Matplotlib: To visualize the data and plot the results.
* Tensorflow and Keras: To build machine learning models and test them. It is simple and user-friendly and requires no experience in machine learning for someone to build and test a model. Also, it allows to stack layers on top of each other, so that complex model architectures can be made.
* Kerastuner: to perform hyperparameter optimization on the built models. It is a library within tensorflow, that provides the tools to perform hyper parameter tuning on machine learning models that were built using keras. It has customizable search space to make it fit to one’s needs and includes multiple searching algorithms.
* Pymysql: to interact and import in a MySQL library the results for the created ticket. It has the capability not only to connect to a MySQL database, but also to execute MySQL queries from python.
* Numpy: to execute mathematical functions whenever necessary.

# Methodology

In this chapter we will discuss the dataset that was used and how they were processed in order to be fit to use for training the models. Afterwards the process flow of how the models were built and assessed will be demonstrated.

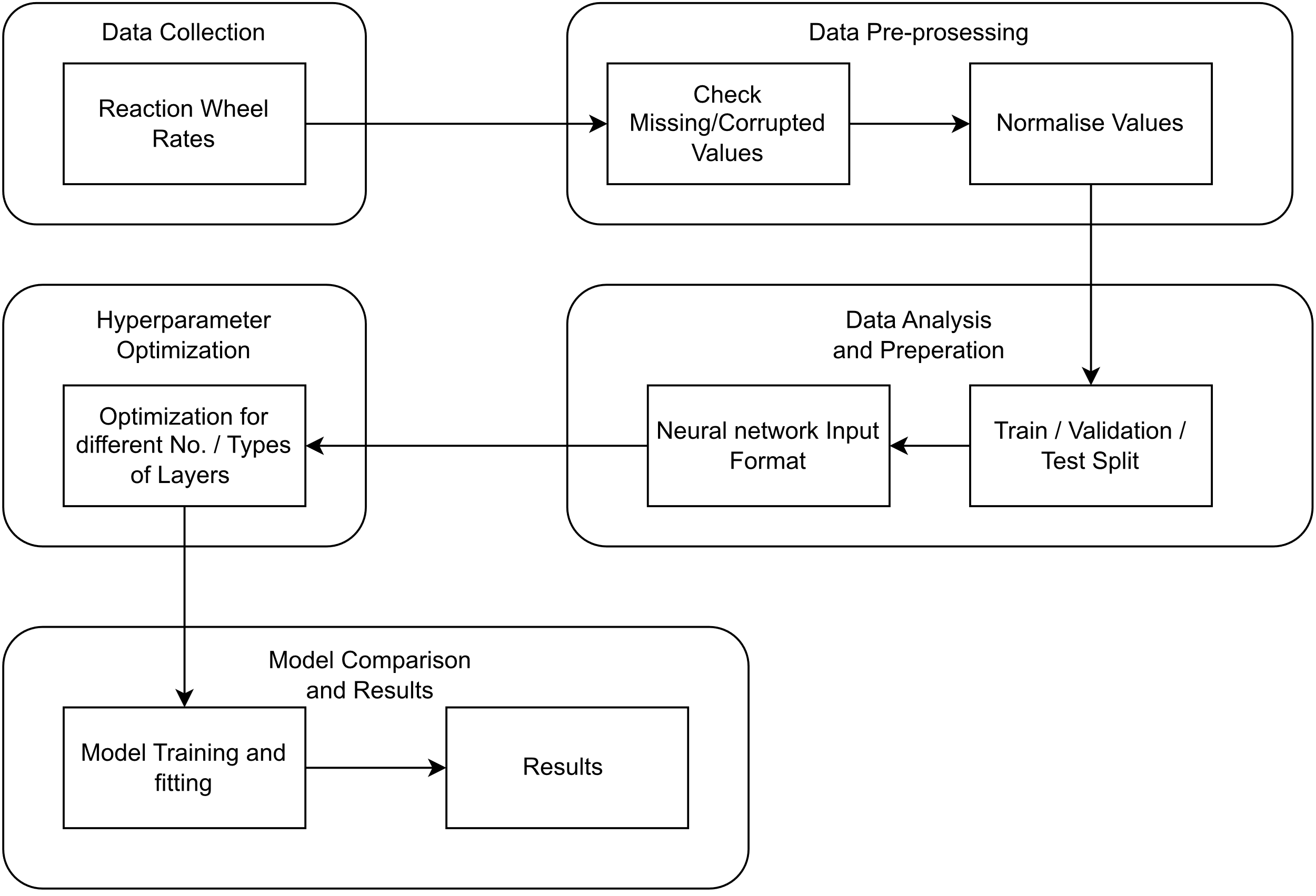


Figure ‑ Process diagram

## Data Collection

The data that were used were provided by Hellas-Sat and collected through their telemetry system for Hellas-Sat 3. The rates of all four reaction wheels were provided for a timespan of one year, from 1-1-2022 01:00 UTC to 31-12-2023 23:00 UTC. The measurements that were provided were the hourly rpm of each wheel, which was calculated by taking the mean of the minimum and the maximum value in a 60-minute window. For each wheel the dataset that was used contained 8749 measurements. Each reaction wheel has its own telemetry: AW1010R is RW 1, AW2010R is RW 2, AW3010R is RW3 and AW4010R is RW4.

A blue line graph with numbers and a white background

Description automatically generated

Figure ‑ Reaction Wheel 1 rates Jan 2022 - Dec 2022

A blue line graph with numbers and a white background

Description automatically generated

Figure ‑ Reaction Wheel 2 rates Jan 2022 - Dec 2022

A blue line graph with numbers and a white background

Description automatically generated

Figure ‑ Reaction Wheel 3 rates Jan 2022 - Dec 2022

A blue line graph with numbers and a white background

Description automatically generated

Figure ‑ Reaction Wheel 4 rates Jan 2022 - Dec 2022

Some noticeable features on the plots of the reaction wheel rates are the following:

* All reaction wheels follow a similar pattern across the year.
* The amplitude of the range in each season varies. The maximum value is smaller around April and September, while it is larger around January, July and November.
* In November all reaction wheels are being forced to rotate in the opposite direction from their prior direction. This is done on purpose by the operators for maintenance purposes.
* The sudden decrease in the rates of the wheels is either because a stationkeeping maneuver is performed, or because of a manual wheel unloading. When a stationkeeping maneuver is performed, automatically a wheel unloading is performed, so that the satellite can take advantage of the thrust and the torques that are applied on it and become more fuel efficient.

## Data Pre-Processing

The data were provided in CSV format and had to be converted into a dataframe, with the corresponding columns of date and time, telemetry code, and value in rpm. All data were scanned for missing or corrupted values and an additional step that had to be taken is that we took the absolute values of the measurements. This is related to the fact that the reaction wheels have the ability to rotate both clockwise and counterclockwise and at all times two of the wheels are spinning CW and two CCW. For purposes of maintenance, once per year the satellite operators make each wheel spin in the opposite direction and therefore we had to filter out of the dataset this transition, in order to avoid confusing the model during the training, because this is not a normal behavior of the wheels that the model needs to be able to predict.

Additionally, during the process of training, experimenting and building the models we came to the conclusion that all data needed to be normalized. We chose the method of min-max normalization, where all values in the time series are being divided with the maximum allowed value, 4300 in our case which is the maximum value that the rate of the wheel is allowed to reach as advised by the constructor and Hellas-Sat, and are consequently on a range from 0 to 1. The reason that such a normalization is beneficial for the model, is that as it was demonstrated in section 2.2.2, LSTM units use a sigmoid function in their gating mechanism and a tanh function in the output function internally. Both functions operate in a range where 1 is the maximum and therefore if the input data is on the same range, the model can be more efficient and provide better results.

The necessity for normalization can be demonstrated with the following example. The same model was used for training and validation two times. It included 1 LSTM layer with 64 units, 2 Dense layers with 128 units each and a Dense layer with 1 unit which is the output and a learning rate of 0.001 with Adam optimizer. On the first iteration the values were not normalized, while on the second they were. In the following figures, the results of the training on the validation can be seen. We notice that on the non-normalized predictions in Figure 3-6, the predictions are cut-off when they reach higher values, while in the normalized model in the Figure 3-7 we see patterns in the predictions in the higher values, that mimic those of the actual data, that are being plotted below them.

A picture containing text, screenshot, line, font

Description automatically generated

Figure ‑ Non-normalized predictions for example LSTM model

A picture containing text, screenshot, line, font

Description automatically generated

Figure ‑ Normalized predictions for example LSTM model

## Data Analysis and Preparation

Moreover, before the data was fed to the models, they had to be split into three subsets: Training, Validation and Testing. For training purposes 70% of the original dataset was used, for validation 20% and for testing 10%. The resulting subsets contained 6124 datapoints for training, 1750 for validation and 802 for testing.

Additionally, the data were reshaped in the appropriate way to be used as input in the models. LSTM models expect a 3D shaped input in the format of [batch size, time steps, features]. Batch size is the number of samples in each batch, time steps is the number of steps in each sequence and features is the number of measurable properties.

## Hyperparameter Optimization

In Table 3-1 all hyperparameters are listed, as well as the range of their values that were explored.

Table ‑ Optimized Hyperparamters

|  |  |
| --- | --- |
| Hyperparameter | Range |
| No. Layers | [1, 2, 3, 4, 5, 6] |
| Type of Layers | [Dense, LSTM] |
| No. LSTM units | [32, 64, 128, 192, 256] |
| No. Dense neurons | [32, 64, 128, 192, 256] |
| Learning Rate | [0.1, 0.01, 0.001] |
| Optimizer | [Adam] |
| Window Size | [24, 48, 72] |

For a more complete exploration of the possible architectures, a tool for hyperparameter optimization was created using kerastuner in python that utilized Hyperband optimization, so that our research would be more extensive and would include as many candidate architectures as possible.

Hyperband's key idea is to quickly identify and eliminate poorly performing configurations, focusing resources on the most promising ones. This approach allows Hyperband to explore a wide range of hyperparameter configurations more efficiently than traditional grid search.

It starts by sampling a large number of random hyperparameter configurations and evaluates them for a small number of epochs, which is referred to as the "initial" number of epochs. After each epoch, it discards the poorly performing configurations and retains the top-performing configurations. The retained configurations are then allocated more resources (epochs) and trained further.

The process continues, progressively reducing the number of configurations but increasing the number of epochs allocated to each configuration at each round. Eventually, only the best configuration remains and is trained for a large number of epochs.

In summary, Hyperband does not exhaustively test all possible combinations of parameters. Instead, it combines random sampling with early elimination of underperforming configurations to find the best hyperparameter configuration within a specified resource budget.

In Hyperband, the number of configurations tested at each round is determined by a formula:

|  |  |
| --- | --- |
|  | (‑) |

n : the number of configurations at the current round

ceil : function that rounds a number up to the nearest integer

R : the maximum number of epochs specified for the search

eta : the factor parameter

s : the round number (starting from 0)

The factor parameter in the Hyperband tuner controls the resource allocation and termination rate of the hyperparameter search. It determines the ratio between the number of configurations to evaluate at each round of successive halving. Its purpose is to control the aggressive elimination of poorly performing configurations and the allocation of more resources to promising configurations.

Although the reaction wheels are similar to their characteristics a model for each one had to be created, because they are positioned in different spots within the spacecraft and therefore they experience and are affected by the perturbations in a unique manner each. In the following tables we can see the results of hyperparameter optimization of different architectures for each reaction wheel. As we can see, the results vary among the reaction wheels for similar architectures.

After analyzing the data for the reaction wheel rates, it was observed that there are some patterns emerging with different frequencies. A daily, weekly, bi-weekly and monthly periodicity are observable as well as a fluctuation in the maximum value that is reached in each epoch throughout the year. Therefore, with the aim to capture these dependencies in the model, several “stacked” architectures were tested with different number of hidden layers. The results for all 4 reaction wheels from the hyperparameter optimization regarding the number of layers, units and trainable parameters are displayed in Tables 3-2 to 3-5 below.

Table ‑ Reaction Wheel 1 architecture results of hyperparameter optimization

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Network | Window Size | Layers | Output Shape | No. of Parameters | Total Parameters | Trainable Parameters |
| RNN | (24, 1) | LSTM | (None, 32) | 7296 | 7329 | 7329 |
|  |  | Dense | (None, 1) | 33 |  |  |
| RNN | (48, 1) | LSTM | (None, 256) | 312320 | 312577 | 312577 |
|  |  | Dense | (None, 1) | 257 |  |  |
| RNN | (48, 1) | LSTM | (None, 64) | 28928 | 45825 | 45825 |
|  |  | Dense | (None, 256) | 16640 |  |  |
|  |  | Dense | (None, 1) | 257 |  |  |
| RNN | (24, 1) | LSTM | (None, 64) | 16896 | 41857 | 41857 |
|  |  | Dense | (None, 128) | 8320 |  |  |
|  |  | Dense | (None, 128) | 16512 |  |  |
|  |  | Dense | (None, 1) | 129 |  |  |
| RNN | (48, 1) | LSTM | (None, None, 128) | 90624 | 119937 | 119937 |
|  |  | LSTM | (None, 32) | 20608 |  |  |
|  |  | Dense | (None, 256) | 8448 |  |  |
|  |  | Dense | (None, 1) | 257 |  |  |
| RNN | (72, 1) | LSTM | (None, None, 192) | 203520 | 273537 | 273537 |
|  |  | LSTM | (None, 64) | 65792 |  |  |
|  |  | Dense | (None, 64) | 4160 |  |  |
|  |  | Dense | (None, 1) | 65 |  |  |
| RNN | (72, 1) | LSTM | (None, None, 64) | 35072 | 245121 | 245121 |
|  |  | LSTM | (None, None, 32) | 12416 |  |  |
|  |  | LSTM | (None, 192) | 172800 |  |  |
|  |  | Dense | (None, 128) | 24704 |  |  |
|  |  | Dense | (None, 1) | 129 |  |  |
| RNN | (72, 1) | LSTM | (None, None, 64) | 35072 | 267585 | 267585 |
|  |  | LSTM | (None, None, 192) | 197376 |  |  |
|  |  | LSTM | (None, 32) | 28800 |  |  |
|  |  | Dense | (None, 64) | 2112 |  |  |
|  |  | Dense | (None, 64) | 4160 |  |  |
|  |  | Dense | (None, 1) | 65 |  |  |

Table ‑ Reaction Wheel 2 architecture results of hyperparameter optimization

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Network | Window Size | Layers | Output Shape | No. of Parameters | Total Parameters | Trainable Parameters |
| RNN | (24, 1) | LSTM | (None, 256) | 287744 | 288001 | 288001 |
|  |  | Dense | (None, 1) | 257 |  |  |
| RNN | (48, 1) | LSTM | (None, 64) | 28928 | 28993 | 28993 |
|  |  | Dense | (None, 1) | 65 |  |  |
| RNN | (48, 1) | LSTM | (None, 128) | 90624 | 94785 | 94785 |
|  |  | Dense | (None, 32) | 4128 |  |  |
|  |  | Dense | (None, 1) | 33 |  |  |
| RNN | (24, 1) | LTSM | (None, 64) | 16896 | 50049 | 50049 |
|  |  | Dense | (None, 256) | 16640 |  |  |
|  |  | Dense | (None, 64) | 16448 |  |  |
|  |  | Dense | (None, 1) | 65 |  |  |
| RNN | (48, 1) | LTSM | (None, None, 192) | 185088 | 382721 | 382721 |
|  |  | LTSM | (None, 128) | 164352 |  |  |
|  |  | Dense | (None, 256) | 33027 |  |  |
|  |  | Dense | (None, 1) | 257 |  |  |
| RNN | (72, 1) | LTSM | (None, None , 192) | 203520 | 233409 | 233409 |
|  |  | LTSM | (None, 32) | 28800 |  |  |
|  |  | Dense | (None, 32) | 1056 |  |  |
|  |  | Dense | (None, 1) | 33 |  |  |
| RNN | (72, 1) | LSTM | (None, None, 32) | 13440 | 217217 | 217217 |
|  |  | LSTM | (None, None, 192) | 172800 |  |  |
|  |  | LSTM | (None, 32) | 28800 |  |  |
|  |  | Dense | (None, 64) | 2112 |  |  |
|  |  | Dense | (None, 1) | 65 |  |  |
| RNN | (72, 1) | LSTM | (None, None, 128) | 102912 | 538625 | 538625 |
|  |  | LSTM | (None, None, 192) | 246528 |  |  |
|  |  | LSTM | (None, 128) | 164352 |  |  |
|  |  | Dense | (None, 128) | 16512 |  |  |
|  |  | Dense | (None, 64) | 8256 |  |  |
|  |  | Dense | (None, 1) | 65 |  |  |

Table ‑ Reaction Wheel 3 architecture results of hyperparameter optimization

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Network | Window Size | Layers | Output Shape | No. of Parameters | Total Parameters | Trainable Parameters |
| RNN | (24, 1) | LSTM | (None, 32) | 7296 | 7329 | 7329 |
|  |  | Dense | (None, 1) | 33 |  |  |
| RNN | (48, 1) | LSTM | (None, 256) | 312320 | 312577 | 312577 |
|  |  | Dense | (None, 1) | 257 |  |  |
| RNN | (48, 1) | LSTM | (None, 128) | 90624 | 107265 | 107265 |
|  |  | Dense | (None, 128) | 16512 |  |  |
|  |  | Dense | (None, 1) | 129 |  |  |
| RNN | (24, 1) | LSTM | (None, 128) | 66560 | 149121 | 149121 |
|  |  | Dense | (None, 256) | 33024 |  |  |
|  |  | Dense | (None, 192) | 49344 |  |  |
|  |  | Dense | (None, 1) | 193 |  |  |
| RNN | (48, 1) | LSTM | (None, None, 256) | 312320 | 706817 | 706817 |
|  |  | LSTM | (None, 192) | 344832 |  |  |
|  |  | Dense | (None, 256) | 49408 |  |  |
|  |  | Dense | (None, 1) | 257 |  |  |
| RNN | (72, 1) | LSTM | (None, None, 192) | 203520 | 536449 | 536449 |
|  |  | LSTM | (None, 192) | 295680 |  |  |
|  |  | Dense | (None, 192) | 37056 |  |  |
|  |  | Dense | (None, 1) | 193 |  |  |
| RNN | (72, 1) | LSTM | (None, None, 192) | 203520 | 417537 | 417537 |
|  |  | LSTM | (None, None, 32) | 28800 |  |  |
|  |  | LSTM | (None, 192) | 172800 |  |  |
|  |  | Dense | (None, 64) | 12352 |  |  |
|  |  | Dense | (None, 1) | 65 |  |  |
| RNN | (72, 1) | LSTM | (None, None, 64) | 35072 | 353825 | 353825 |
|  |  | LSTM | (None, None, 32) | 12416 |  |  |
|  |  | LSTM | (None, 256) | 295936 |  |  |
|  |  | Dense | (None, 32) | 8224 |  |  |
|  |  | Dense | (None, 64) | 2112 |  |  |
|  |  | Dense | (None, 1) | 65 |  |  |

Table ‑ Reaction Wheel 4 architecture results of hyperparameter optimization

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Network | Window Size | Layers | Output Shape | No. of Parameters | Total Parameters | Trainable Parameters |
| RNN | (24, 1) | LSTM | (None, 192) | 166656 | 166849 | 166849 |
|  |  | Dense | (None, 1) | 193 |  |  |
| RNN | (48, 1) | LSTM | (None, 32) | 10368 | 10401 | 10401 |
|  |  | Dense | (None, 1) | 33 |  |  |
| RNN | (48, 1) | LSTM | (None, 256) | 312320 | 328833 | 328833 |
|  |  | Dense | (None, 64) | 16448 |  |  |
|  |  | Dense | (None, 1) | 65 |  |  |
| RNN | (24, 1) | LSTM | (None, 128) | 66560 | 79393 | 79393 |
|  |  | Dense | (None, 32) | 4128 |  |  |
|  |  | Dense | (None, 256) | 8448 |  |  |
|  |  | Dense | (None,1) | 257 |  |  |
| RNN | (48, 1) | LSTM | (None, None, 128) | 90624 | 148481 | 148481 |
|  |  | LSTM | (None, 64) | 49408 |  |  |
|  |  | Dense | (None, 128) | 8320 |  |  |
|  |  | Dense | (None, 1) | 129 |  |  |
| RNN | (72, 1) | LSTM | (None, None, 192) | 203520 | 281985 | 281985 |
|  |  | LSTM | (None, 64) | 65792 |  |  |
|  |  | Dense | (None, 192) | 12480 |  |  |
|  |  | Dense | (None, 1) | 193 |  |  |
| RNN | (72, 1) | LSTM | (None, None, 256) | 336896 | 764289 | 764289 |
|  |  | LSTM | (None, None, 64) | 82176 |  |  |
|  |  | LSTM | (None, 256) | 328704 |  |  |
|  |  | Dense | (None, 64) | 16448 |  |  |
|  |  | Dense | (None, 1) | 65 |  |  |
| RNN | (72, 1) | LSTM | (None, None, 192) | 203520 | 396865 | 396865 |
|  |  | LSTM | (None, None, 128) | 164352 |  |  |
|  |  | LSTM | (None, 32) | 20608 |  |  |
|  |  | Dense | (None, 128) | 4224 |  |  |
|  |  | Dense | (None, 32) | 4128 |  |  |
|  |  | Dense | (None, 1) | 33 |  |  |

## Model Comparison

The next step from hyperparameter optimization was to assess the resulting models for each type of architecture. Assessing machine learning models is an important step in evaluating their performance and to determine whether they are suitable for our project. The metrics that were used for the evaluation are Root Mean Squared Error, Loss and Validation Loss.

Root Mean Squared Error (RMSE) is a metric used to measure the accuracy of a model’s predictions compared to the actual values. It calculates the square root of the average of squared differences between predicted values and actual values. A lower RMSE indicates better performance of the model.

Loss refers to the objective function used during the training process. In LSTM models it is usually Mean Squared Error. During training the model tries to minimize this loss, making the predictions closer to the actual values. Monitoring the loss during training allows you to understand how well the model is learning from the data.

Validation loss is computed using the same loss function as during training, but it is calculated on the validation dataset. This dataset is not used for training the model, and its purpose is to evaluate how well the model generalizes to unseen data. The validation loss helps you detect overfitting, when the model performs well on the training data but poorly on new data.

In the following Tables 3-6 to 3-9 are displayed the results from all the models that were trained regarding the validation loss and the Root Mean Squared Error. As we can see there is no absolute correlation between number of layers, window size and lower validation error and RMSE.

Generally having a greater window size, which means more data as input used to make the next prediction, results in better performance, but not always. There are cases where a 48-hour window has a better performance than a 72-hour window with the same model architecture.

Additionally, more layers don’t necessarily mean a better predicting performance for the models. This means that adding more complexity to the structure of the model is not mandatory to get good results. Simpler models with just two LSTM layers and two Dense layers are enough for some of the wheels’ models to be able to capture the complexity of the data in a satisfactory level.

Finally, it is important to note that not all reaction wheels models had the best performance with the same architecture. That is indicative of the fact that each reaction wheel is affected by environmental effects in a unique way and since each model is trained on dedicated data for each reaction wheel, it is expected to have different results.

Table ‑ Reaction Wheel 1 Architectures Validation Loss and RMSE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Network | Window Size | Layers | Val. Loss | RMSE |
| RNN | (24, 1) | LSTM | 0.010 | 0.100 |
|  |  | Dense |  |  |
| RNN | (48, 1) | LSTM | 0.015 | 0.122 |
|  |  | Dense |  |  |
| RNN | (48, 1) | LSTM | 0.00087 | 0.029 |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (24, 1) | LSTM | 0.0013 | 0.036 |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (48, 1) | LSTM | 0.00092 | 0.030 |
|  |  | LSTM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (72, 1) | LSTM | 0.00086 | 0.029 |
|  |  | LSTM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (72, 1) | LSTM | 0.00089 | 0.030 |
|  |  | LSTM |  |  |
|  |  | LSTM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (72, 1) | LSTM | 0.00083 | 0.029 |
|  |  | LSTM |  |  |
|  |  | LSTM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |

Table ‑ Reaction Wheel 2 Architectures Validation Loss and RMSE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Network | Window Size | Layers | Val. Loss | RMSE |
| RNN | (24, 1) | LSTM | 0.00158 | 0.040 |
|  |  | Dense |  |  |
| RNN | (48, 1) | LSTM | 0.00214 | 0.046 |
|  |  | Dense |  |  |
| RNN | (48, 1) | LSTM | 0.00131 | 0.036 |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (24, 1) | LTSM | 0.00099 | 0.031 |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (48, 1) | LTSM | 0.00147 | 0.038 |
|  |  | LTSM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (72, 1) | LTSM | 0.00150 | 0.039 |
|  |  | LTSM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (72, 1) | LSTM | 0.00167 | 0.041 |
|  |  | LSTM |  |  |
|  |  | LSTM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (72, 1) | LSTM | 0.00160 | 0.040 |
|  |  | LSTM |  |  |
|  |  | LSTM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |

Table ‑ Reaction Wheel 3 Architectures Validation Loss and RMSE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Network | Window Size | Layers | Val. Loss | RMSE |
| RNN | (24, 1) | LSTM | 0.00136 | 0.037 |
|  |  | Dense |  |  |
| RNN | (48, 1) | LSTM | 0.00178 | 0.042 |
|  |  | Dense |  |  |
| RNN | (48, 1) | LSTM | 0.00116 | 0.034 |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (24, 1) | LSTM | 0.00093 | 0.030 |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (48, 1) | LSTM | 0.00120 | 0.035 |
|  |  | LSTM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (72, 1) | LSTM | 0.00123 | 0.035 |
|  |  | LSTM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (72, 1) | LSTM | 0.00125 | 0.035 |
|  |  | LSTM |  |  |
|  |  | LSTM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (72, 1) | LSTM | 0.00127 | 0.036 |
|  |  | LSTM |  |  |
|  |  | LSTM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |

Table ‑ Reaction Wheel 4 Architectures Validation Loss and RMSE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Network | Window Size | Layers | Val. Loss | RMSE |
| RNN | (24, 1) | LSTM | 0.00129 | 0.036 |
|  |  | Dense |  |  |
| RNN | (48, 1) | LSTM | 0.00195 | 0.044 |
|  |  | Dense |  |  |
| RNN | (48, 1) | LSTM | 0.00111 | 0.033 |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (24, 1) | LSTM | 0.00083 | 0.029 |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (48, 1) | LSTM | 0.00123 | 0.035 |
|  |  | LSTM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (72, 1) | LSTM | 0.00120 | 0.035 |
|  |  | LSTM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (72, 1) | LSTM | 0.00127 | 0.036 |
|  |  | LSTM |  |  |
|  |  | LSTM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
| RNN | (72, 1) | LSTM | 0.00126 | 0.035 |
|  |  | LSTM |  |  |
|  |  | LSTM |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |
|  |  | Dense |  |  |

## Ticket Creation

Once the models for each reaction wheel are built and trained, we can make predictions for each one, to detect if their rate is going to exceed a specific threshold and then create a ticket in the Spacecraft Control Centre with the necessary information, so that the engineers, operators and flight dynamics team can act accordingly. The flow diagram below, Figure 3-8, explains the procedure in which the ticket will be created and the steps that need to take place.

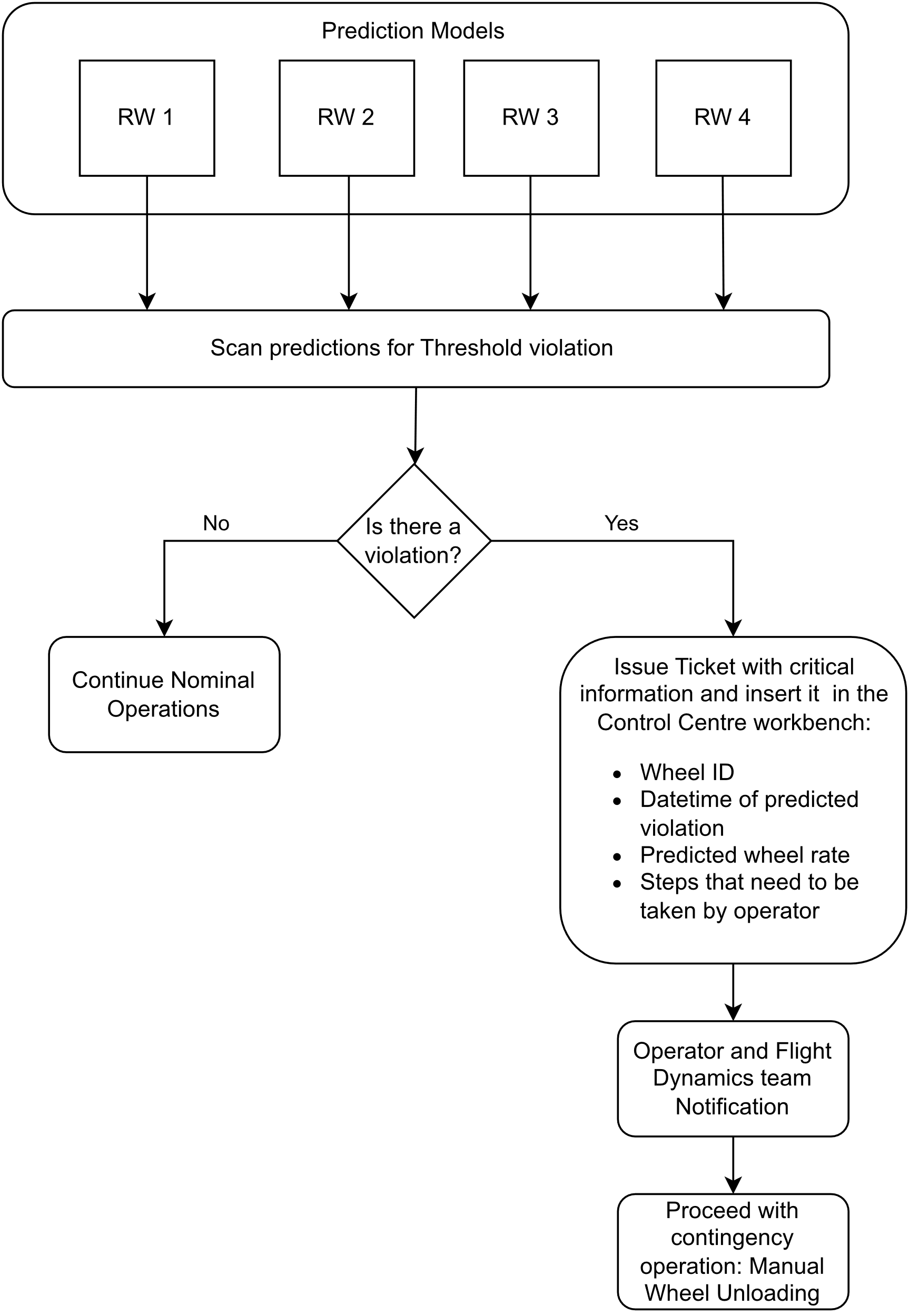


Figure ‑ Ticket creation flow process

Firstly, the predictions are created for each Reaction Wheel for a certain period. Then the results need to be examined, in order to assess if any of the predictions exceed a set threshold.

To make predictions a rolling input window system was created, as shown in Figure 3-9. To start making the predictions the hourly Reaction Wheel rates from the last 72 hours are needed. Then using these as input data a prediction is made for the next hour, which can be then used alongside with the last 71 hours, with the aim to predict the next hour. This procedure can be followed another 70 times, 71 in total, so that we can get a prediction in total for the next 71 hours, approximately 3 days. No further predictions should be made, because they would be based only on already predicted values and no real data.

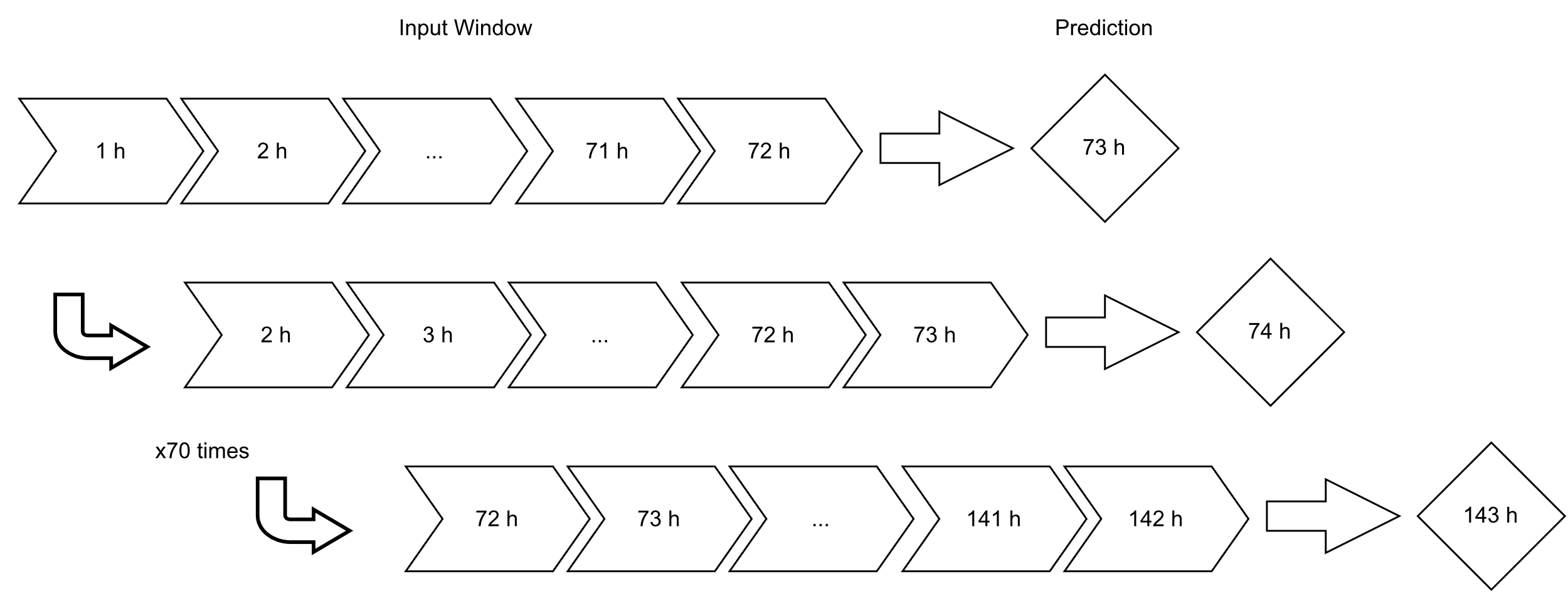


Figure ‑ Rolling input window for predictions

If none of the wheels exceed the set limit, then no action needs to be taken and operations can proceed nominally. However, if the predictions indicate that one of the wheels will surpass the limit at a specific time, a ticket needs to be created in the operation center workbench, so that the operators can be notified that a manual wheel unloading will need to take place.

To achieve that a process of scanning the prediction results was created, which identifies which wheel will exceed the limit and when, and then creates a ticket with the appropriate information on the workbench, by creating an entry in the right tables in its database.

A ticket is produced only for the most imminent threshold violation in one of the reaction wheels. This is due to the fact that when a wheel unloading takes place on the spacecraft, all four wheels are being unloaded at the same time. Therefore, if the procedure is executed for the most imminent limit violation chronologically, then afterwards the rates of all the wheels are being affected and the predictions that were made before become invalid, as well as any alarms that the threshold will be reached.

The information that is input is the date of the procedure, which is the date that the rate will pass the threshold, the predicted rate of the wheel, which wheel will it be and the steps of the procedure that will need to be followed.

The procedure that the operator will have to execute is 4000CA02 and includes 11 steps which are also included in the ticket that is produced and displayed on the workbench:

• Step 1: Initialize procedure 4000CA02.

• Step 2: Verify Observability. Verification of the housekeeping packet that is needed for the operation. This is done to ensure that all telemetry that is needed for the operation are downlinked. Need to activate the table DUMP NM\_DMP for the operation follow-up and in particular the duty cycle parameters checking.

• Step 3: Deactivate automatic wheel unloading. The automatic wheel unloading is disabled to avoid the random execution of the wheel unloading during the procedure.

• Step 4: Configure wheels unloading in the direction (East, West, North , South) that has been decided by the flight dynamics team. The maneuver direction shall be set to select the thruster set used during the wheel unloading. If it is not provided, the procedure will be rejected.

• Step 5: Activate wheels unloading. Wheel unloading starts as soon as telecommand 3002FHA is received on board. While wheel unloading has started, procedure execution has to be continued in order to monitor wheel unloading behavior.

• Step 6: Monitor wheels unloading. The wheel unloading is checked thanks to specific telemetry: spacecraft attitude, wheel angular momentum, on-control thruster valves status and temperature. If these telemetry are not correct during the wheel unloading the operator has to abort the procedure.

• Step 7: Verify end of wheels unloading. The thrusters’ history is recorded to know wheel unloading fuel consumption.

• Step 8: Reset boost parameters. The boost parameters must be reset after the boost to avoid performing any operational mistake.

• Step 9: Re-activate automatic wheels unloading. Set back on automatic wheel unloading, which was deactivated in the beginning of the procedure.

• Step 10: Deactivate observability. Come back to the default spacecraft telemetry plan by stopping NM\_DM table DUMP.

• Step 11: End procedure.

The ticket shall be marked as unverified, so that the flight dynamics team can process the information that is given and decide in what direction the maneuver for the unloading will be performed and if there is a more efficient time to execute the procedure than the one that is predicted. Also in consideration needs to be taken whether there are any other procedures that have to be executed on the same time and if the wheel unloading needs to be prioritized.

# Results

## Model Selection

According to the results from the training of the models that we mentioned in section 3.5, the following architectures were chosen for the predictions of each reaction wheel. In all cases, the input window that was chosen was 72 hours, even though in some cases a 48-hour input window resulted in slightly lower validation losses and RMSE. This choice was made, with the aim to make trustworthy predictions for the longest period of time possible, by using the rolling window method that was described in section 3.6.

### Reaction Wheel 1

A structure of three LSTM layers followed by 3 Dense layers was the most efficient, with validation loss 0.00083 and RMSE 0.029. The first LSTM layer is composed of 64 units, the second LSTM composed of 192 units, while the third composed of 32 units. The following first and second Dense layers have 64 neurons each and the final Dense layer is the output layer and has just 1 neuron.

A conceptual representation of the model that was built for reaction wheel 1 with the characteristics that were described and are mentioned in Table 4-1, can be seen in Figure 4-1.

Table ‑ Selected model architecture for Reaction Wheel 1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Network | Window Size | Layers | Output Shape | Val. Loss | RMSE |
| RNN | (72, 1) | LSTM | (None, None, 64) | 0.00083 | 0.029 |
|  |  | LSTM | (None, None, 192) |  |  |
|  |  | LSTM | (None, 32) |  |  |
|  |  | Dense | (None, 64) |  |  |
|  |  | Dense | (None, 64) |  |  |
|  |  | Dense | (None, 1) |  |  |

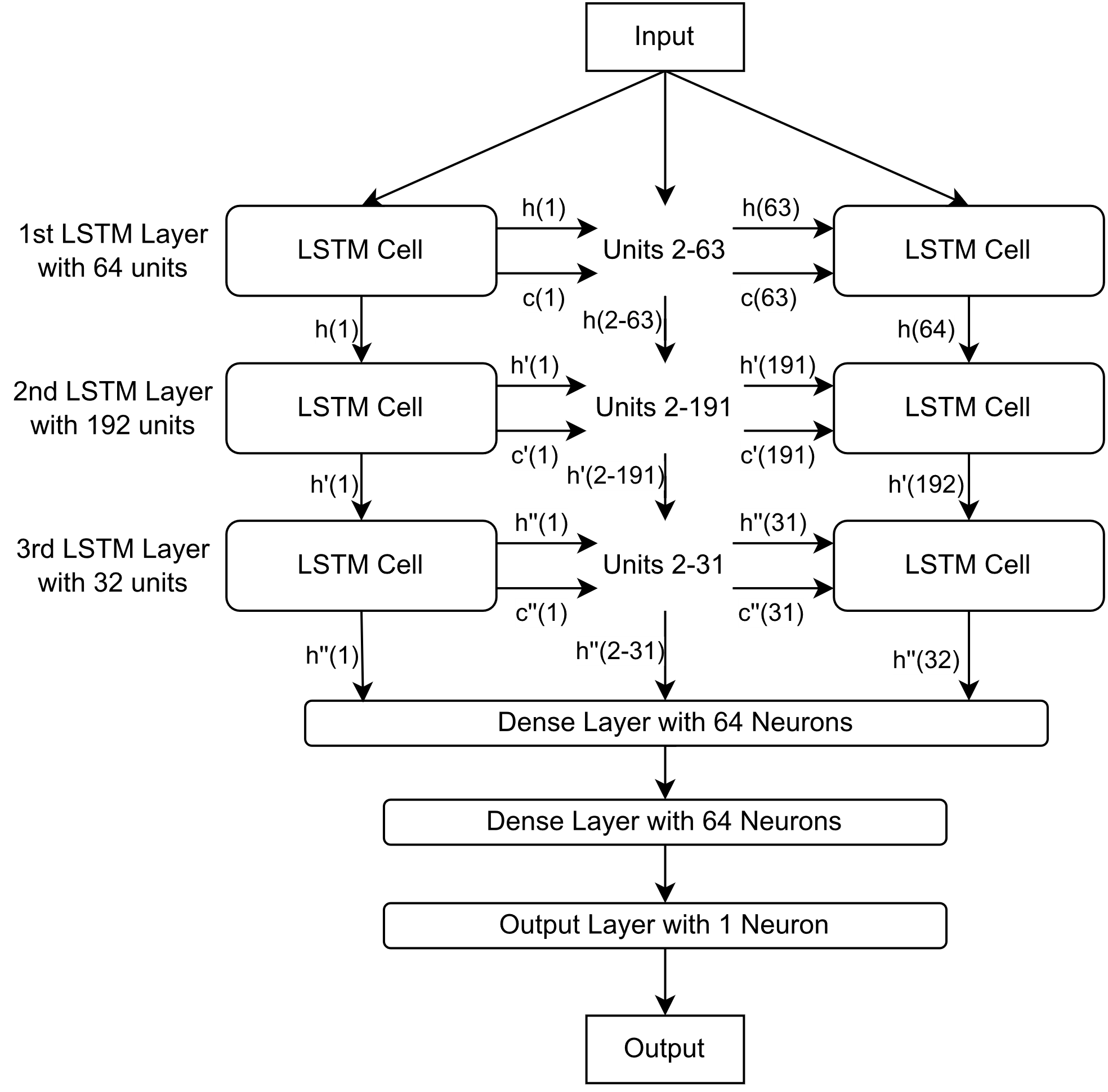


Figure ‑ RW1 RNN model structure

A computer screen shot of a program code

Description automatically generated

Figure ‑ Code for RW1 model

In Figure 4-2 is displayed the code that we wrote for the creation of the model. It was written in Python, using the Tensorflow and Keras libraries to import the LSTM and Dense layers to create the model. Also those libraries include the optimizer for the model and the metrics. The structure of the code is object oriented so that it can be used multiple times for the rest of the reaction wheels, just by changing the input variables, depending on the results of the hyperparameter tuning.

The model was trained for 100 epochs and the results for the training and validation data, as well as the losses and the RMSE, are displayed in Figures 4-3 to 4-5.

A graph of a training loss and validation loss

Description automatically generated

Figure ‑ RW1 Training and Validation Loss & RMSE at 100 epochs

A blue lines with white text

Description automatically generated with medium confidence

Figure ‑ RW1 Training dataset Predictions (top) and Actual Data (bottom)

A graph of blue lines

Description automatically generated with medium confidence

Figure ‑ RW1 Validation dataset Predictions (top) and Actual Data (bottom)

### Reaction Wheel 2

A structure of two LSTM layers followed by two Dense layers was the most efficient, with validation loss 0.0015 and RMSE 0.039. The first LSTM layer is composed of 192 units and the second LSTM composed of 32 units. The following Dense layer has 32 neurons and the final Dense layer is the output layer and has just 1 neuron.

The conceptual representation for the model of reaction wheel 2 is similar to the one displayed in Figure 4-1 for reaction wheel 1, but with different characteristics. The number of layers, the number of units in each layer as well as the number of neurons are only what changes, according to the hyperparameter optimization results that are shown in Table 4-2.

Table ‑ Selected model architecture for Reaction Wheel 2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Network | Window Size | Layers | Output Shape | Val. Loss | RMSE |
| RNN | (72, 1) | LTSM | (None, None , 192) | 0.0015 | 0.039 |
|  |  | LTSM | (None, 32) |  |  |
|  |  | Dense | (None, 32) |  |  |
|  |  | Dense | (None, 1) |  |  |

The model was trained for 100 epochs and the results for the training and validation data, as well as the losses and the RMSE, are displayed in Figures 4-6 to 4-8.

A graph of a training loss and validation loss

Description automatically generated

Figure ‑ RW2 Training and Validation Loss & RMSE at 100 epochs

A graph of blue lines

Description automatically generated with medium confidence

Figure ‑ RW2 Training dataset Predictions (top) and Actual Data (bottom)

A graph of blue lines

Description automatically generated with medium confidence

Figure ‑ RW2 Validation dataset Predictions (top) and Actual Data (bottom)

### Reaction Wheel 3

A structure of three LSTM layers followed by two Dense layers was the most efficient, with validation loss 0.00125 and RMSE 0.035. The first LSTM layer is composed of 192 units, the second LSTM composed of 32 units, while the third composed of 192 units. The Dense layers has 64 neurons and the final Dense layer is the output layer and has just 1 neuron.

The conceptual representation for the model of reaction wheel 3 is similar to the one displayed in Figure 4-1 for reaction wheel 1, but with different characteristics. The number of layers, the number of units in each layer as well as the number of neurons are only what changes, according to the hyperparameter optimization results that are shown in Table 4-3.

Table ‑ Selected model architecture for Reaction Wheel 3

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Network | Window Size | Layers | Output Shape | Val. Loss | RMSE |
| RNN | (72, 1) | LSTM | (None, None, 192) | 0.00125 | 0.035 |
|  |  | LSTM | (None, None, 32) |  |  |
|  |  | LSTM | (None, 192) |  |  |
|  |  | Dense | (None, 64) |  |  |
|  |  | Dense | (None, 1) |  |  |

The model was trained for 100 epochs and the results for the training and validation data, as well as the losses and the RMSE, , are displayed in Figures 4-9 to 4-11.

A graph of a training loss and validation loss

Description automatically generated with medium confidence

Figure ‑ RW3 Training and Validation Loss & RMSE at 100 epochs

A blue lines with white text

Description automatically generated with medium confidence

Figure ‑ RW3 Training dataset Predictions (top) and Actual Data (bottom)

A graph of blue lines

Description automatically generated with medium confidence

Figure ‑ RW3 Validation dataset Predictions (top) and Actual Data (bottom)

### Reaction Wheel 4

A structure of two LSTM layers followed by two Dense layers was the most efficient, with validation loss 0.0015 and RMSE 0.039. The first LSTM layer is composed of 192 units and the second LSTM composed of 32 units. The following Dense layer has 32 neurons and the final Dense layer is the output layer and has just 1 neuron.

The conceptual representation for the model of reaction wheel 4 is similar to the one displayed in Figure 4-1 for reaction wheel 1, but with different characteristics. The number of layers, the number of units in each layer as well as the number of neurons are only what changes, according to the hyperparameter optimization results that are shown in Table 4-4.

Table ‑ Selected model architecture for Reaction Wheel 4

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Network | Window Size | Layers | Output Shape | Val. Loss | RMSE |
| RNN | (72, 1) | LSTM | (None, None, 192) | 0.0012 | 0.035 |
|  |  | LSTM | (None, 64) |  |  |
|  |  | Dense | (None, 192) |  |  |
|  |  | Dense | (None, 1) |  |  |

The model was trained for 100 epochs and the results for the training and validation data, as well as the losses and the RMSE are displayed in Figures 4-12 to 4-14.

A graph of a training loss and validation loss

Description automatically generated with medium confidence

Figure ‑ RW4 Training and Validation Loss & RMSE at 100 epochs

A graph of blue lines

Description automatically generated with medium confidence

Figure ‑ RW4 Training dataset Predictions (top) and Actual Data (bottom)

A graph of blue lines

Description automatically generated with medium confidence

Figure ‑ RW4 Validation dataset Predictions (top) and Actual Data (bottom)

## Predictions

Using the abovementioned models for each reaction wheel, we can make predictions for their future speed rates, so that we can subsequently survey them to check if any of the wheels will need unloading.

To create predictions to showcase the functionality of the ticket creation and the whole pipeline, we used the test dataset, because it is constituted of data that any of the models have never interacted with. In this way we can also assess the performance of the models in never-before-seen data.

The threshold that was set is 4000 rpm. Figures 4-15, 4-16, 4-17 and 4-18 display the predictions that were made for each reaction wheel respectively by using their developed models, when their test dataset was used as an input.

A graph of blue lines

Description automatically generated with medium confidence

Figure ‑ RW1 Test dataset Predictions (top) and Actual Data (bottom)

A graph of blue lines

Description automatically generated with medium confidence

Figure ‑ RW2 Test dataset Predictions (top) and Actual Data (bottom)

A graph of blue lines

Description automatically generated with medium confidence

Figure ‑ RW3 Test dataset Predictions (top) and Actual Data (bottom)

A graph showing the results of a performance

Description automatically generated with medium confidence

Figure ‑ RW4 Test dataset Predictions (top) and Actual Data (bottom)

After surveying the predictions for each reaction wheel, the results thar are shown in Figure 4-19, show that all of them at some point will surpass the limit that was set and almost all of them more than a single time.

|  |  |
| --- | --- |
| a) | b) |
| c) | d) |

Figure ‑ Survey results for detected threshold violation instances a) Reaction Wheel 1, b) Reaction Wheel 2, c) Reaction Wheel 3, d) Reaction Wheel 4

However, only the first instance time-wise is important, because if a wheel unloading is performed at that point, all wheels will offload their momentum and reduce their rates. Therefore, any predictions that were made for after that time are obsolete. As we can see from the survey results, the first instance of limit violation is going to happen at reaction wheel 2, in 210 hours from the moment of creating the predictions or at the date that is calculated in J2000 format.

## Ticket

Having acquired the critical information of which reaction wheel is predicted to surpass the threshold and when, we create a ticket on the operator’s workbench to notify them about the situation and make the aware of the procedure they will have to execute. In the ticket are also included the steps of the procedure they will have to follow.

In Tables 4-5 and 4-6, we can see the inputs on the MySQL database that were created, with the critical information for the procedure and the steps that the operator has to take, to execute it successfully.

Table ‑ Database entry with critical operation information

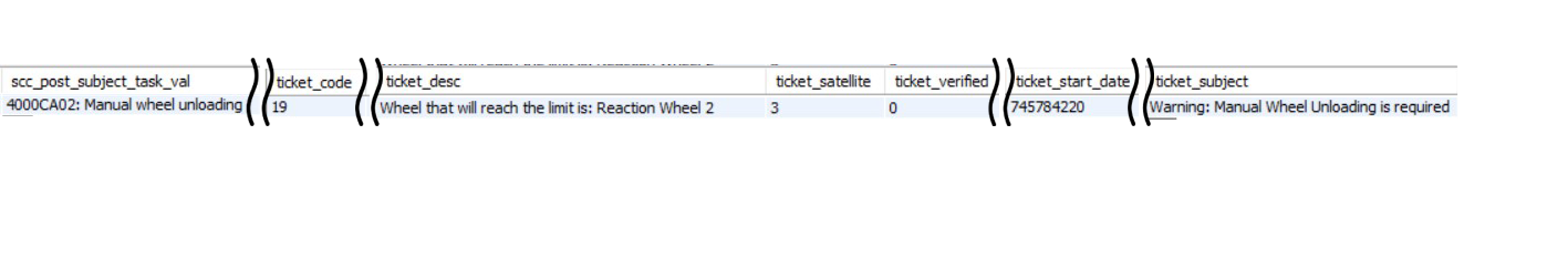


Table ‑ Database entry with procedure steps

A screenshot of a computer

Description automatically generated

## Impact of the developed solution

Currently there are no systems or solutions implemented in Satellite Control Center for Hellas-Sat 3 that can provide predictions regarding the rates of the reaction wheels, or any other kind of telemetry. With that being said, in case of a contingency or if a trend of increasing rates is observed by one of the operators, nothing can be done proactively except for waiting to see how the trend will evolve and take action afterwards, hoping that it is not too late.

By implementing our solution, accurate and trustworthy predictions can be made for the reaction wheels’ rates up to three days ahead and a ticket is created providing important information to the operators and notifying them about the situation. The models we developed have a Root Mean Squared Error of 0.029 for Reaction Wheel 1, 0.029 for Reaction Wheel 2 and 0.035 for Reaction Wheels 3 and 4. That is translated to predictions with an error of margin of ± 127 rpm for RW1, ± 150 rpm for RW2 and ± 168 rpm for RW3 and RW4. These small margins do not affect significantly the decision-making impact that the forecasts can have.

The three day prediction window that is being generated is a reasonable amount of time for engineers to analyse the situation and estimate the status of the spacecraft, in order to investigate if there are any contingencies and make the correct decisions. The system can be used to get a first assessment of the evolution of any alarming rates, but can also be used repeatedly as the critical moment comes closer, in order to get even more precise predictions.

During a demonstration of our solution to Hellas Sat satellite engineers, we received the following feedback: “*Being able to predict how the rates of the reaction wheels will evolve, in the next hour and even better for the following days is a very powerful tool. We recently had a contingency on one of the other satellites, where the rates of the reaction wheels had reached the operational limit that we have imposed. We could not take any action and just had to wait to see how the rates will evolve in the next hour and then plan to execute a contingency procedure if it would be necessary. Had we had the tool that you have developed, we would be able to have foreseen this situation and plan accordingly beforehand. Even if we hadn’t done that, we would be able to use the tool in the last minute, to make predictions for the following hours and execute a contingency procedure before it is too late. Thankfully the rates of the reaction wheels started decreasing and no further action needed to be taken.*”

# Conclusions

## Summary and Achievements

In this thesis we have achieved creating a tool that can be applied on satellite operations, by predicting when a procedure will need to be executed and automatically scheduling it.

The tool was developed for Hellas-Sat 3, a geostationary telecommunications satellite operated by Hellas-Sat, and the operation that it focused on was the reaction wheel unloading of the satellite.

The first half of the tool includes predicting when the reaction wheels will need unloading. To achieve that, machine learning models were built and trained, one for each reaction wheel, to perform time series forecasting for the rotation rates of the wheels. Those models had firstly their hyperparameters uniquely optimized based on the telemetry data of each reaction wheel and secondly were trained and validated based on the same data.

Once the models were finalized and we have credible forecasted values for every reaction wheel within a specified time window, the predictions are surveyed to check if there is a threshold violation. When such a case is detected, the critical information, such as the date of the violation and rate and ID of the reaction wheel, are extracted and automatically a workbench ticket is created to notify the engineering team and operators about the situation. The ticket also includes the steps that the operators have to take, in order to execute successfully the appropriate operation to mitigate the situation and ensure the nominal operability of the spacecraft.

## Limitations

Although the predictions that are made by the models are credible close to the limits that are important to the reaction wheel unloading operation, the forecasted values for the middle to low values in the operating range of the wheels sometimes are not of the same accuracy. However, they tend to correctly predict the trend that the values are going to follow.

Additionally, the tool can successfully be applied for predictions for up to the following three days. This limits the scheduling ahead ability that it could have, but does not render it useless. Since it predicts a contingency situation, any notification about it ahead of time can be useful to the engineering team to access the situation with time to spare and examine what is the correct plan of action. This limitation can be mitigated by applying the tool either daily or once at least every two days.

## Future work

To continue development of the tool further, future work could include creating new models that use multiple inputs to make the predictions, such as the temperature of the reaction wheels or their angular momentum. In this way the forecasts could be even more accurate and also one of the limitations of the tool would be addressed.

Moreover, to expand the utility of the tool, it could be applied on other spacecraft that have similar attitude control systems, just by training new models based on their dedicated telemetry.

Finally, having gain insight on how to automatically create the tickets for operations, schedule procedures and how to create models for time series forecasting, the same method could be applied to other critical subsystems of the spacecraft, to predict possible future contingency situations.

## Lessons Learned

Looking back at this thesis, we have been able to demonstrate that there is massive potential in combining Artificial Intelligence and Satellite Operations. By combining the huge amount of data that are being produced by spacecraft and the computing power and complexity of machine learning models, satellite operations can become more efficient and can reduce the contingency risks, by creating forecasts for critical subsystems of the spacecraft. Future common development in both Artificial Intelligence and Satellite Operations could potentially lead into highly automated spacecraft operations, overcoming current limitations.

References

[1] “Special Issue on Social Infrastructure that Guarantees Safety, Security, Fairness, and Efficiency Automation and Labor-Saving Technology for Satellite Operation GONAI Toshiya, INOMATA Sota.”

[2] Y. Li, Z. Jun, H. Chen, and J. Zhang, “Onboard autonomy for on-orbit GEO satellites,” in *2016 IEEE International Conference on Mechatronics and Automation, IEEE ICMA 2016*, Institute of Electrical and Electronics Engineers Inc., Sep. 2016, pp. 205–210. doi: 10.1109/ICMA.2016.7558561.

[3] J. Liang, Y. he Zhu, Y. zhong Luo, J. cheng Zhang, and H. Zhu, “A precedence-rule-based heuristic for satellite onboard activity planning,” *Acta Astronaut*, vol. 178, pp. 757–772, Jan. 2021, doi: 10.1016/j.actaastro.2020.10.020.

[4] G. Morelli, A. Mainguet, and M. Eustace, “Automated operations of large geo telecom satellites with digital transparent processors (Dtp): Challenges and lessons learned,” in *15th International Conference on Space Operations, 2018*, American Institute of Aeronautics and Astronautics Inc, AIAA, 2018. doi: 10.2514/6.2018-2666.

[5] K. Duan, R. Wang, S. Chen, and L. Ge, “Exploring the predictability of attention mechanism with LSTM: Evidence from EU carbon futures prices,” *Res Int Bus Finance*, vol. 66, p. 102020, Oct. 2023, doi: 10.1016/j.ribaf.2023.102020.

[6] S. A. Haider, M. Sajid, H. Sajid, E. Uddin, and Y. Ayaz, “Deep learning and statistical methods for short- and long-term solar irradiance forecasting for Islamabad,” *Renew Energy*, vol. 198, pp. 51–60, Oct. 2022, doi: 10.1016/j.renene.2022.07.136.

[7] B. Alamo and L. Moffelein, “WHO - Working Hours Operations for geostationary telecom satellites,” in *SpaceOps 2006 Conference*, 2006. doi: 10.2514/6.2006-5946.

[8] F. Xhafa and A. W. H. Ip, “Optimisation problems and resolution methods in satellite scheduling and space-craft operation: a survey,” *Enterp Inf Syst*, vol. 15, no. 8, pp. 1022–1045, 2021, doi: 10.1080/17517575.2019.1593508.

[9] H. Nagamatsu, “Development of automatic / remote satellite operation system,” in *13th International Conference on Space Operations, SpaceOps 2014*, American Institute of Aeronautics and Astronautics Inc., 2014. doi: 10.2514/6.2014-1855.

[10] M. K. Ben-Larbi *et al.*, “Towards the automated operations of large distributed satellite systems. Part 1: Review and paradigm shifts,” *Advances in Space Research*, vol. 67, no. 11, pp. 3598–3619, 2021, doi: https://doi.org/10.1016/j.asr.2020.08.009.

[11] M. K. Ben-Larbi *et al.*, “Towards the automated operations of large distributed satellite systems. Part 2: Classifications and tools,” *Advances in Space Research*, vol. 67, no. 11, pp. 3620–3637, 2021, doi: https://doi.org/10.1016/j.asr.2020.08.018.

[12] N. L. Crowley and V. Apodaca, “Analysis of satellite telemetry data,” in *1997 IEEE Aerospace Conference*, 1997, pp. 57–67 vol.4. doi: 10.1109/AERO.1997.577497.

[13] R. Stottler and J. Mao, “Automatic Satellite Telemetry Analysis for SSA using Artificial Intelligence Techniques,” 2017. [Online]. Available: http://www.af.mil/News/Article-Display/Article/719941/hyten-announces-space-enterprise-vision/

[14] J. Chen, D. Pi, Z. Wu, X. Zhao, Y. Pan, and Q. Zhang, “Imbalanced satellite telemetry data anomaly detection model based on Bayesian LSTM,” *Acta Astronaut*, vol. 180, pp. 232–242, 2021, doi: https://doi.org/10.1016/j.actaastro.2020.12.012.

[15] J. Tong, L. Xie, W. Yang, K. Zhang, and J. Zhao, “Enhancing time series forecasting: A hierarchical transformer with probabilistic decomposition representation,” *Inf Sci (N Y)*, vol. 647, p. 119410, Nov. 2023, doi: 10.1016/j.ins.2023.119410.

[16] R. Khaldi, A. El Afia, R. Chiheb, and S. Tabik, “What is the best RNN-cell structure to forecast each time series behavior?,” *Expert Syst Appl*, vol. 215, Apr. 2023, doi: 10.1016/j.eswa.2022.119140.

[17] N. Wu, B. Green, X. Ben, and S. O’Banion, “Deep Transformer Models for Time Series Forecasting: The Influenza Prevalence Case,” Jan. 2020, [Online]. Available: http://arxiv.org/abs/2001.08317

[18] A. Vaswani *et al.*, “Attention Is All You Need.”

[19] “Hellas Sat 3.” https://www.hellas-sat.net/hellas-sat-3 (accessed Sep. 05, 2023).

Appendices

Ethical approval letter

A letter with a blue circle and text

Description automatically generated with medium confidence

Python Code

In this appendix it is included the code that was developed in Python for the Thesis. The development of the code was object oriented so that the classes that were developed are smaller, more manageable and reusable.

All the code that was developed is available on the following repository on [GitHub](https://github.com/MariosAnastasopoulos/IRP/tree/main/Development).

A screen shot of a computer code

Description automatically generated

Figure ‑ Imported Libraries

A screen shot of a computer code

Description automatically generated

Figure ‑ Data Analyzer class, importing and concerting CSV file

A screen shot of a computer code

Description automatically generated

Figure ‑ Data Split class, splitting data into train, validation and test sets

A screen shot of a computer program

Description automatically generated

Figure ‑ Model class, pt1. model structure and training

A computer screen shot of a code

Description automatically generated

Figure ‑ Model class, pt.2 rolling prediction window

A computer screen shot of a program

Description automatically generated

Figure ‑ Model class, pt.3 plot predictions

A computer code on a dark background

Description automatically generated

Figure ‑ Threshold Scanner class, pt.1 scan predictions

A screen shot of a computer code

Description automatically generated

Figure ‑ Threshold Scanner class, pt.2 scan results gathering

A screen shot of a computer program

Description automatically generated

Figure ‑ Create Ticket class, pt.1 establish connection with database

A screen shot of a computer program

Description automatically generated

Figure ‑ Create Ticket class, pt.2 insert critical information in database

A computer screen shot of a program code

Description automatically generated

Figure ‑ Create Ticket class, pt.3 insert procedure steps in database

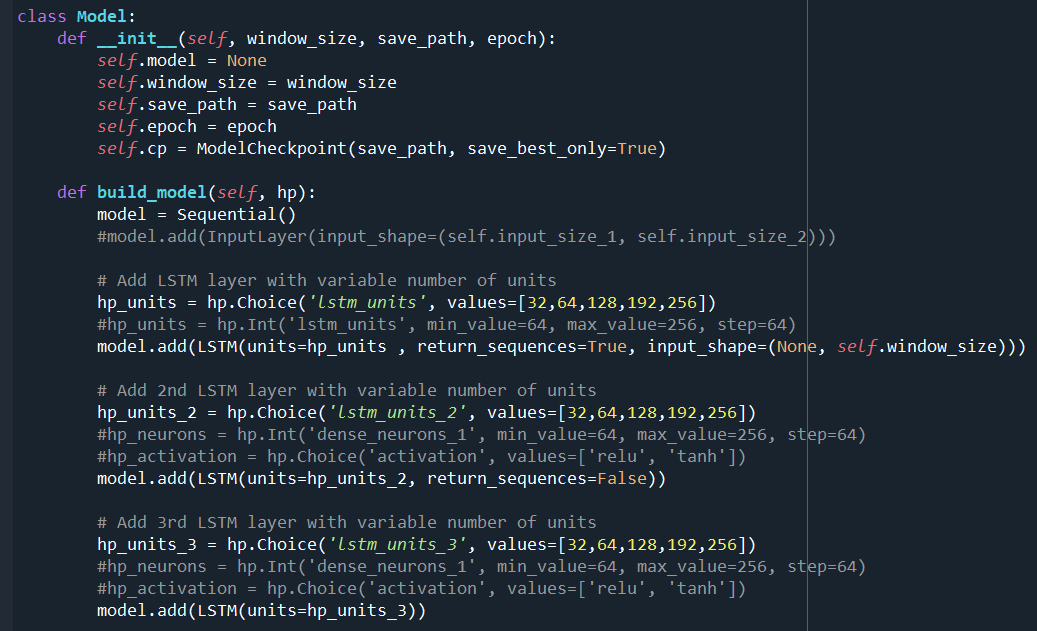


Figure ‑ Hyperparameter Model class, pt.1 model building for optimization

A computer screen shot of a program

Description automatically generated

Figure ‑ Hyperparameter Model class, pt.2 model building for optimization

A screen shot of a computer program

Description automatically generated

Figure ‑ Hyperparameter Model class, pt3. hyperband search, model training and results reporting