Predictive Modelling of Sea Debris Around Maltese Coastal Waters

Mark Dingli

Supervisor: Dr Kristian Guillaumier

May 2024

Submitted in partial fulfilment of the requirements for the degree of B.Sc. IT Artificial Intelligence.

L-Università ta' Malta
Faculty of Information & Communication Technology

Abstract

The accumulation of sea surface debris around the coastal waters of Malta, presents numerous ecological and environmental challenges that negatively affect both marine ecosystems and human activities. This is exacerbated by the absence of an effective system that can predict their movement, making it more challenging to address and mitigate this issue effectively.

The primary objective of this project was to develop a system that can predict dispersion patterns of sea surface debris around Malta's coast. To achieve this, we developed a comprehensive machine learning and physics-based pipeline. This pipeline uses historical sea surface currents (SSC) data to predict future conditions, while also having the ability to visualise the movement of debris.

Central to this system is the integration of LSTM and GRU models, trained to predict the next 24 hours of SSC within a specific area. These predictions were subsequently utilised by a Lagrangian model to visualise the movement of surface debris, offering insights into future dispersion patterns.

A comparative evaluation was conducted for both models, examining the accuracy of their predictions and the quality of the simulations generated by the Lagrangian model, based on these predictions. The results indicated that the LSTM model outperformed the GRU model. This was evident in consistently lower error metrics such as mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE), alongside narrower interquartile ranges (IQR) in the results, thereby providing a more reliable basis for the subsequent simulation of debris dispersal patterns.

Overall, this project presents an approach that is specifically tailored to address the challenges of sea surface debris around Malta. By harnessing the power of machine learning in tandem with a physics-based Lagrangian model, we have established a framework that not only predicts SSC with high accuracy, but also visualises the movement of sea surface marine debris, allowing us to make more informed decisions about our environment and our effect on it.

Acknowledgements

I would like to extend my deepest gratitude to several individuals whose support and guidance were invaluable in the completion of this project.

Foremost, I express my heartfelt appreciation to my supervisor, Dr Kristian Guillaumier, for his unwavering guidance, encouragement, and insightful critiques throughout the whole project. His expertise and dedication were instrumental in navigating the complexities of this project and in pushing the boundaries of my academic capabilities.

My sincere thanks also go to Dr Adam Gauci, who not only provided the essential data for this project but also offered his expertise and support. His contributions have been pivotal in enriching the quality of this work.

Additionally, I am deeply grateful to my family, especially my parents and sister. Their unwavering support and love have been the cornerstone of my resilience throughout this journey. Their belief in my potential and constant encouragement has been instrumental in empowering me to pursue my goals.

Lastly, I must express profound appreciation to my girlfriend, Ilenia, and my close friends. Their understanding, patience, and encouragement have provided me with the motivation needed to persevere through the challenges of this project.

To all mentioned, and to those who contributed, I am eternally thankful. Your roles in this academic endeavour have left an indelible mark on both the project and my personal growth.

Contents

List of Figures

List of Tables

List of Abbreviations

AI Artificial Intelligence.

ANN Artificial neural networks.

CSV Comma-separated Variable.

GRU Gated recurrent unit.

IQR Interquartile Range.

LSTM Long short-term memory.

MAE Mean absolute error.

MSE Mean squared error.

NaN Not a Number.

NetCDF Network Common Data Form.

RMSE Root mean squared error.

RNN Recurrent neural network.

SSC Sea surface currents.

1 Introduction

This project is an integration of machine learning techniques with a physics based Lagrangian model [1] to address environmental issues of sea surface debris. At the core of this project is a pipeline that uses historical data to predict the next 24 hours of sea surface currents (SSC) velocities. These predictions serve as inputs to a Lagrangian model [1] enabling it to simulate the movement of surface marine debris. Finally, a comparative evaluation of both LSTM and GRU models is conducted, focusing on their predictive accuracy and the quality of the visualisations. This project presents an approach of merging machine learning with a physics-based model, providing valuable insights to marine conservation efforts and enhancing decision-making processes for the management of marine debris around the Maltese Islands.

1.1 Problem Definition

Sea surface debris around the coastal waters of Malta presents a significant environmental issue. Predominantly composed of plastics, which constitute 82% of all man-made floating materials encountered in the Mediterranean sea [2], this debris endangers marine life, disrupts ecological balances, and compromises the ecological integrity of coastal areas [3]. This problem is further aggravated by the lack of an effective system that can predict the movement of this surface debris since, as of writing, there exists no system that addresses this challenge specifically for the coastal areas around Malta. This underscores the need for a system that can accurately predict and visualise the dispersion patterns of sea surface debris.

1.2 Motivation

The geological characteristics of the Mediterranean Sea makes it difficult for surface debris to escape the area naturally, resulting in the accumulation of surface debris [4]. The current absence of a predictive system tailored to the coastal regions of Malta impedes effective interventions to mitigate environmental harm. This gap opens an opportunity for the implementation of a system that, through the application of machine learning and physics-based modelling, aims to address an urgent ecological issue, which is widely recognised as a global crisis [5]. By fulfilling this need, the project aims to provide accurate predictions that can guide effective cleanup operations and inform strategies for long-term marine conservation around the coast of Malta.

1.3 Aims and Objectives

The aim of this project is to create a system that simulates and predicts the movement of marine debris in the coastal waters of Malta, thereby supporting marine conservation efforts. To achieve this, the following objectives have been identified:

01. Data integration: To preprocess and integrate the SSC velocity datasets, ensuring compatibility and consistency for input into both models.

02. Lagrangian model development: To develop a physics-based Lagrangian model for simulating the movement of surface marine debris, utilising historical data to ensure accurate simulations.

03. AI model development: To develop both LSTM and GRU models to predict future SSC velocities. These models will serve as a crucial component of the forecasting system, leveraging their respective strengths in time series data processing to ensure accurate predictions.

04. Integrating the AI models with the Lagrangian model: To integrate the model's predictions into the Lagrangian model. This integration aims to create future simulations and visualisations of marine debris movement, enhancing the project's predictive capabilities for marine conservation.

05. Comparison of AI models: To conduct a comparative evaluation of both LSTM and GRU models, focusing on their predictive accuracy and the quality of the final visualisations.

1.4 Proposed Solution

The process begins with the preprocessing of the SSC velocities datasets that will be used as input for the subsequent models. A Lagrangian model will be developed to visualise the debris movement. This approach is designed to clarify both the expected input from the AI models and the expected nature of the ensuing visualisations. The core of the solution involves developing and fine-tuning two types of machine learning models namely, LSTMs and GRUs. These models will undergo extensive testing to determine the optimal architecture and hyper-parameters, aiming to accurately predict SSC velocities for a future 24-hour period. Upon establishing the predictive models, the pipeline integrates these predictions into a Lagrangian model, transforming the predicted data into dynamic visualisations of future debris movement. The project culminates in a comparative analysis of the LSTM and GRU models, evaluating their effectiveness through various metrics, including their predictive accuracy and the quality of the generated visualisations. By analysing the results and visualisations, this project aims to provide actionable insights for effective cleanup operations and strategies for long-term marine conservation around the coastal waters of Malta.

1.5 Summary of Results

This project achieved successful outcomes by developing an integrated system that combines machine learning models with a physics-based Lagrangian framework to forecast the movement and dispersion of sea surface debris. The results demonstrate the system's capability to make accurate 24-hour predictions and dynamically visualise the trajectory of marine debris. Through a comparative evaluation, it was determined that the LSTM model outperforms the GRU model in predicting SSC velocities, evidenced by better performance in error metrics such as mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE). These findings validate the effectiveness of the integrated approach and demonstrate its potential to enhance marine conservation efforts.

1.6 Document Structure

The remainder of this document is organised into the following chapters:

Background and Literature Review: This section outlines the foundational elements of the project, including a description of the dataset and the functionalities of the Lagrangian and AI models. The literature review explores existing research on marine debris, Lagrangian models, and the use of AI in environmental forecasting, establishing the scientific grounding for the project's methodologies.

Methodology: This chapter details the processes undertaken in implementing this project. It includes the steps for data preprocessing, the development of the Lagrangian and AI models, the pipeline for merging both models, and the evaluation strategy.

Evaluation: A comprehensive outline of the strategies used to test and evaluate the effectiveness and reliability of the implementation is presented in this chapter. This will be followed by the presentation and discussion of our results.

Conclusion: The document is concluded by summarising our work, revisiting the aims and objectives, discussing any limitations, highlighting the results, and finally suggesting any proposals for future work.

2 Background and Literature Review

In this chapter, we provide a comprehensive background of the main technologies related to this project. This is followed by a detailed overview of pertinent academic papers and literature to support the project's scientific foundations.

2.1 Background

This section is divided into five sections, each offering detailed information regarding this project. First, we begin with an exploration of the impacts of marine debris on ecosystems and also delve into the specific datasets used in this project. Then, we discuss the intricacies of the physics-based Lagrangian Model, provide an explanation of time series modelling, and finally discuss deep learning models, specifically LSTMs and GRUs.

2.1.1 The Impact of Marine Debris on Ecosystems

The environmental and ecological impact of marine debris, particularly in coastal and marine ecosystems, has been extensively researched, as shown in [6] and [7]. Several studies in this area reveal significant negative effects, ranging from harm to marine wildlife due to ingestion and entanglement [8], to the disruption of natural habitats [9]. The impact on coastal ecosystems extends beyond the environment, affecting economic sectors reliant on marine health, such as tourism and fishing [9]. Further research covers the long-term ecological consequences, highlighting the urgent need for effective management and mitigation strategies as discussed in [10]. These studies collectively emphasise the critical nature of addressing marine debris for ecosystem sustainability and conservation.

2.1.2 The Dataset

The dataset forms the backbone of this project, with its selection and preprocessing being crucial for creating our models. In this project, we utilise a dataset provided by the Department of Geosciences at the University of Malta. This dataset consists of SSC velocities data, recorded in hourly increments across four years, spanning from January 2020 to December 2023. These data points are derived from a model generated by high-frequency (HF) radar systems [11], located on the northern regions of the Maltese islands and southern Sicily. The locations of these radar systems, depicted in Figure 2.1 and identified from [12], provide a temporal snapshot of the SSC movements.

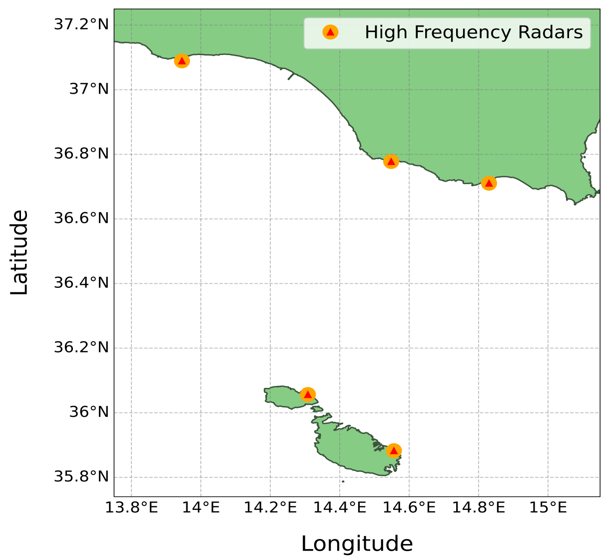

Figure 2.1 – High frequency radar locations.

The data is composed of several variables including longitude, latitude, time, and SSC velocities which are denoted as u and v. The variable u represents the east-west component, indicating the horizontal velocity of the SSC. Positive values of u signify an eastward movement, while negative values indicate a westward motion. On the other hand, v represents the north-south velocity, with positive values denoting northward movement and negative values representing southward motion. Together, these variables form a vector that represents the directional movement of SSC. The data's geographical scope is defined within the boundaries of 14.15° to 14.81° longitude and 35.79° to 36.30° latitude. This coverage translates into a grid of 52 latitude points by 43 longitude points, for a total of 180 data points, as shown in Figure 2.2. The dataset is in Network Common Data Form (NetCDF) format [13], a commonly used format for climate and meteorological data, ensuring compatibility with the Lagrangian Model.

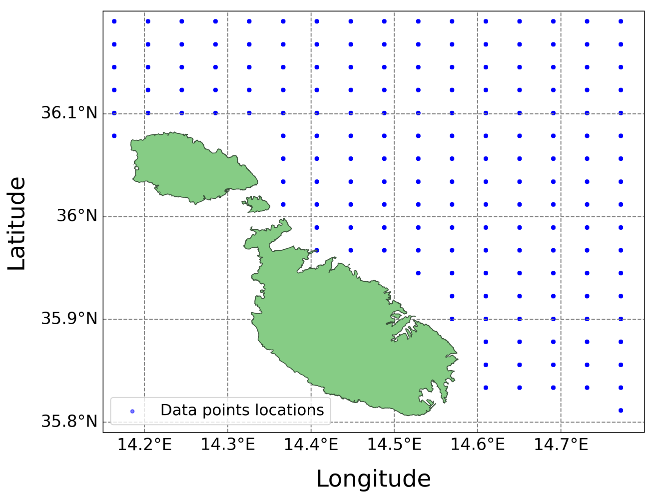

Figure 2.2 – Locations of radar data points.

2.1.3 Physics-Based Lagrangian Model

The practice of tracking ocean surface movements in a Lagrangian framework dates back to the earliest days of oceanography. Early methods involved observing the drift of ships or the paths of specially designed floats to document sea current movement, as outlined by [14]. The physics-based Lagrangian model [1] plays a pivotal role in environmental simulations. By offering a dynamic method to trace individual particle trajectories within fluid mediums, the model ensures precise tracking of the particle's spatio-temporal movement. Its broad applicability spans from localised studies, to global-scale systems. This is evident in its varied applications, such as tracking oil spills diffusion [15], mapping floating plastic debris [16], simulating jellyfish migrations [17], and smoke dispersion [18].

The Lagrangian model [1] operates by representing particles within a fluid medium, tracking their position and properties as they move with the fluid's flow. The model calculates the trajectory of each particle by integrating the velocity field of the fluid, which may vary in time and space. This approach enables the simulation of dispersal patterns of particles, such as marine debris, by accounting for both advection and diffusion processes. Advection represents the movement of particles by the flow of a fluid [19]. Diffusion, on the other hand, models the dispersion of particles through random motion [19]. This is done by applying techniques such as random walks or Gaussian distributions. This inclusion of randomness enhances the realism of the simulation.

To facilitate these Lagrangian simulations, several Python toolkits like OceanParcels [20], PyGnome [21], and Flexpart [22] have been developed. These toolkits enable the customisation and execution of particle tracking simulations, leveraging data on ocean currents, wind fields, and other environmental phenomena. OceanParcels is distinguished by several features that make it suitable for our project. One of its notable capabilities are custom kernels. These are user-defined functions that allow for tailored simulation scenarios at each time step. Through custom kernels, users can implement complex behaviours and interactions of particles within the fluid, such as particle reflection or response to environmental variables like temperature and wind. Another significant feature is particle initialisation. This feature enables the creation of particles at specific locations, times, and with distinct properties, allowing for more detailed and accurate simulations.

All these attributes make OceanParcels an optimal choice for this project. By integrating these features, this toolkit facilitates the development of comprehensive simulations. This is crucial for understanding and predicting the movement of marine debris, thereby enhancing our strategies for marine conservation and debris management.

2.1.4 Time Series Modelling

Time series modelling is a technique used to predict future data points by analysing the trends, cycles, and patterns in a series of data points collected over an interval of time [23]. The main focus is on analysing historical data to uncover the underlying structure of the data, which can then be used to forecast future trends. This method is particularly powerful for its ability to incorporate the sequence and time dependence within the dataset. By examining how values are interconnected over time, time series models can forecast future values based on the inherent temporal dynamics present in the historical data [24]. This form of predictive modelling assumes that past patterns shape future behaviours, making it an indispensable tool in a variety of fields ranging from weather forecasting [25] to stock market predictions [26].

While time series modelling is a powerful tool for forecasting future data, it also possesses some limitations. Time series data often exhibit seasonality and trends, which can complicate the forecasting process [27]. Outliers, missing sequences of data, and anomalies can also significantly impact the accuracy of forecasting models, requiring careful identification, and handling. The capacity of these models to integrate external influential factors and variables is also somewhat limited, often necessitating the integration of additional features for enhanced predictive accuracy [28]. Additionally, time series models require significantly more data for training, which can be cumbersome in situations where data is limited [28]. These challenges highlight the importance of adopting a methodical approach to time series modelling, emphasising the need to carefully consider the specific context and characteristics of the data being analysed when utilising time series models for effective forecasting.

In the context of this project, we utilise time series modelling to predict SSC velocities. Accurate predictions require a detailed analyses of the data sequences to discern patterns that could forecast future predictions. The historical hourly data of SSC form a time series, which is inherently continuous but sampled at discrete intervals. To address this, deep learning models, a subset of Artificial neural networks (ANNs), are leveraged due to their proficiency in handling vast amounts of sequential data and their capacity to learn complex temporal patterns [29]. Through training on past SSC data, these models are equipped to predict future values.

2.1.5 Deep Learning Models

Deep learning is a subset of machine learning that uses the power of ANNs to interpret and predict data through multiple layers [30]. Deep learning is useful because of its capacity to detect intricate patterns in various types of data [30].

LSTM networks are a specialised type of Recurrent neural networks (RNNs).

They are designed to address the challenge of learning long-term dependencies and overcoming the limitations faced by traditional RNNs, notably the vanishing gradient problem [31]. This challenge inhibits RNNs from effectively learning and retaining information over long sequences. LSTMs employ a unique architecture, characterised by a system of gates, namely the input, forget, and output gates. These gates collectively decide which information should be stored, discarded, or passed through, based on the relevance to the task at hand [32]. Memory cells within LSTMs retain information over long intervals, making them adept at managing sequences where understanding past context is crucial for future predictions [32]. This capability is pivotal for predicting SSC, as demonstrated in this project. Their ability to remember previous information for extended durations without degradation makes them ideal for capturing the underlying patterns in historical data of SSC, which is crucial for accurate prediction and subsequent debris dispersion simulations.

GRU networks are another variant of RNNs that aim to solve the vanishing gradient problem [31] with a more simplified structure than LSTMs. GRUs simplify the LSTM model by combining the input and forget gates into a single update gate and merging the cell state and hidden state [32, 33]. This reduction in complexity leads to a model that is faster to train, without significantly compromising the model's ability to capture dependencies in a sequence [32]. In the context of this project, GRUs are deployed alongside LSTMs to forecast SSC. Their efficiency and effectiveness in handling time series data render them adapt at predicting the movements of marine debris, offering a comparative perspective to the LSTM's performance.

LSTMs and GRUs distinguish themselves primarily through their structure and information processing; LSTMs offer a more detailed gating mechanism that excels in managing long-term dependencies, while GRUs provide a streamlined architecture that enables quicker training without significantly sacrificing performance [33]. Their inherent capabilities make them exceptionally suited for time series modelling, where understanding and predicting sequential data patterns is crucial [29], thereby making them highly applicable to the objectives of this project.

2.2 Literature Review

The literature review is divided into three subsections, each focusing on a critical aspect of marine debris dispersion and the methodologies used to predict and simulate it. The first subsection delves into studies that forecast the movement and accumulation of marine debris. The second subsection highlights research that applies machine learning techniques to predict SSC. Finally we explore the integration between AI models predictions and physics-based models.

2.2.1 Prediction of Marine Debris Dispersal

Researchers have explored multiple methodologies to understand and forecast the movement and accumulation zones of debris in marine environments. The variation in these approaches reflects the complexity of the problem, encompassing various methods that aim to capture the dynamic nature of marine debris movement. Through the implementation of numerical simulations [34, 35], deep learning techniques [36], and advanced simulation tools [37, 38], the field continues to evolve, seeking more accurate and efficient ways to predict debris dispersal patterns.

Numerical Simulations

E. van Sebille et al. [34] focus on the crucial role of numerical simulations in predicting and understanding the dispersion of marine debris. This study utilises a number of numerical simulations that leverage various physical oceanographic phenomena to model the movement of floating marine debris. Central to their approach is the use of extensive datasets, capturing various environmental factors such as the velocity and direction of ocean currents, wind patterns, and wave dynamics. These variables are crucial for determining the dispersal patterns of marine debris.

E. van Sebille et al. implement both Eulerian and Lagrangian frameworks in their research. The Eulerian approach models plastics as tracers within a grid, focusing on the interaction between fluid and particle phases, incorporating turbulence through diffusivity parameterisation. The Lagrangian framework, preferred for its three-dimensional transport analysis, traces virtual particles using pre-computed velocity data, integrating stochastic terms to reflect the impact of turbulence on dispersion patterns. Both these methods highlight the significant influence of environmental phenomena have on debris movement, especially in nearshore processes. However, accurately simulating coastal dynamics and beaching patterns remains challenging.

Similar to the research conducted by B. D. Hardesty et al. [35], E. van Sebille et al. also highlight the need for enhanced models that better capture surface interactions. Experiments conducted by these authors include the deployment of drifters and buoys equipped with GPS tracking, enabling the researchers to validate their simulation results against real-world data. These findings highlight the importance of integrating numerical simulations with empirical data to refine model accuracy and forecast reliability. Such efforts demonstrate the versatility and efficiency of numerical simulations and ultimately contribute to more effective mitigation and management strategies for marine pollution.

Deep Learning Techniques

Computer vision has also emerged as powerful tool in addressing environmental challenges, notably in the management and mitigation of marine debris dispersal, as demonstrated by the research of W. R. Winans et al. [36]. These techniques offer approaches to interpret large datasets, enabling more precise and effective solutions to combat marine pollution. The authors utilise deep learning and object detection methodologies combined with remote sensing, to automate the identification and classification of marine debris across extensive coastal areas. Specifically, their research targets a stretch of 1900km along the Hawaiian coastline.

W. R. Winans et al. carry out an evaluation of three distinct object detection models, demonstrating an analysis of each model's ability to tackle the complexities involved in detecting marine debris. The research is based on an extensive dataset comprising of 1587 image chips, which together contain 10,703 individual debris labels across various categories. The inclusion of data augmentation techniques further enhances the quality and reliability of the analysis.

Among the key findings, the Single Shot MultiBox Detector (SSD) [39], paired with a MobileNet-v2 feature extractor (SS-MN), stands out for its performance by achieving an average precision rate of 72%. This metric reinforces the practical viability and efficiency of leveraging deep learning for environmental monitoring.

The research also acknowledges some challenges and limitations, notably in maximising recall rates to ensure minimal oversight of debris objects. Despite this, the work of W. R. Winans et al. stands out as an effective demonstration of how deep learning and computer vision can be strategically deployed to tackle the issue of marine debris dispersal.

Advanced Simulation Tools

In our review of methodologies for modelling and simulating surface marine debris dispersal, we have come across numerous studies that utilise the OceanParcels toolkit [20] to tackle the issue of simulating marine debris. M. Yuniarti et al.[37] provide an insightful analysis by using OceanParcels [20] to simulate the distribution patterns of microplastics originating from the Seto Inland Sea and extending throughout Japanese waters. A similar approach is adopted by S. Aijaz [38], where OceanParcels [20] and the Regional Ocean Modelling System (ROMS) activated with its built-in Lagrangian model, facilitated the tracking of river plume dispersal, highlighting the toolkit's versatility across different marine environments.

By utilising OceanParcels [20], the authors in [37] and [38] delve into the trajectories of particles within marine environments, offering insights into how these different particles navigate through varied marine environments throughout the year.

The area simulated by M. Yuniarti et al. [37], spanning latitudes 28° to 55°N and longitudes 120° to 160°E, was strategically selected to optimise the study's focus, similar to the approach of S. Aijaz [38], where the authors identified a specific polygon area on the northeast coast of Australia.

M. Yuniarti et al. used data from the Hybrid Coordinate Ocean Model (HYCOM) to provide the necessary current velocities for their simulations, while statistical evaluations leveraged RMSE validated this data against in-situ observations. The validation process confirmed the suitability of the data for simulation inputs, with RMSE values indicating a close match to observed data, thus improving the accuracy of the simulation results. S. Aijaz used ROMS with a built-in Lagrangian model, incorporating wind fields from global models and recorded river volume discharges, akin to how M. Yuniarti et al. utilised HYCOM data to provide current velocities for the simulations.

The findings by M. Yuniarti et al. reveal that microplastics dispersion exhibits significant seasonal variations, with distinct pathways and accumulation zones becoming apparent in different seasons. The observed distribution patterns align with those documented within the review of previous studies by M. Yuniarti et al., affirming the reliability of the simulation approach utilised. Furthermore, this enables visualisations that illustrate the dispersal patterns of microplastics, enhancing the spatial and temporal dynamics of marine debris movement.

The authors in [37] and [38] address the challenges associated with tracking large numbers of particles across vast marine areas. These challenges highlight the need for advanced computational resources and methodologies to accurately simulate marine dispersal patterns. The insights from these studies are pivotal in demonstrating the importance of such approaches, paving the way for the development of effective mitigation strategies against marine pollution. This not only advances our understanding of debris trajectories but also establishes a standard for the application of advanced simulation tools like OceanParcels [20].

2.2.2 Machine Learning Models for Predicting SSC

SSC are a fundamental phenomenon within ocean hydrodynamics, having a significant influence on various marine processes as demonstrated in [40]. Numerous studies have turned to machine learning to unravel the intricacies of SSC, crucial for understanding marine debris dispersion. By leveraging different algorithms, these studies offer new perspectives on marine environmental monitoring, demonstrating the potential of machine learning to provide accurate predictions of SSC.

Predicting Ocean Currents at Multiple Depths (Single Location)

Dauji et al. [41] use ANNs for the task of predicting ocean currents across multiple depths, not just the sea surface. Ali et al. [42] implement a similar approach of predicting ocean currents across multiple depths by using LSTM networks. These studies propose time series models to overcome the constraints inherent in numerical models, which necessitate extensive external information, substantial computational resources, and often struggle with noise and gaps in data.

Both studies highlight the challenge of accurately forecasting ocean currents in different regions. Dauji et al. focus on two locations within the North Atlantic and North Pacific oceans. This dataset comprises of hourly records of current velocity and direction. These measurements were taken at depths of 18.3m and 460m, representing shallow and deep-water situations. On the other hand, Ali et al. [42] conduct their study in the Gulf of Mexico. The dataset includes measurements at 50 different depth levels, reaching down to 3000m below the surface, and spans horizontally from 88.5°W to 85°W and 24.65°N to 27°N.

In addressing the challenges posed by the data, Dauji et al. set up a feed-forward back-propagation ANN architecture, which according to the authors, is recognised for its efficiency. The consideration of additional inputs, specifically currents from lower depths, was explored but ultimately showed no significant improvement to the model's prediction accuracy. Ali et al. [42] use LSTM networks, chosen for their ability to handle long-term dependencies in data. Both studies explored the optimum length of past data segments for input, emphasising the temporal dynamics of sea currents. Both studies also encountered and addressed several limitations. One limitation was the initial under-prediction of extreme values. Dauji et al. tackled this issue by introducing methods for scaling target extreme values during training. Moreover, due to the high cost and complexity of collecting SSC data, both studies faced limitations in the availability of long-term observations.

Dauji et al. evaluated the performance of the ANN models quantitatively and qualitatively. The ruslts showed high correlation coefficients and low RMSE and MAE error metrics across various testing durations and prediction intervals. The study also compared the ANN model performance with past works and a random walk model. Notably, the models maintained high performance for currents at both shallow and deep-water layers and were effective across different forecasting durations. The ANN models outperformed traditional forecasting methods, marking a significant improvement in predictive accuracy. The performance of the LSTM models was evaluated using similar error metrics, including RMSE, Peak Signal to Noise Ratio (PSNR), and Structural Similarity (SSIM).

Both studies validate deep learning models as powerful tools for the real-time prediction of SSC, demonstrating their success in exceeding the accuracy of traditional forecasting methods.

Predicting SSC (Single Location)

I. I. Zulfa et al. [43] investigate the potential of LSTM networks for predicting the velocity and direction of SSC in Labuan Bajo, Indonesia. Given Labuan Bajo's significance as a pivotal point for trade and tourism, the study aimed to improve maritime navigation and safety through precise forecasts of SSC.

To conduct this study, the authors utilised a dataset consisting of hourly SSC velocities collected by the Perak Maritime Meteorology Station II. This dataset is comprised of 24 data points and captures the SSC velocities at a single geographical point. Before applying any predictive modelling, the data underwent preliminary preprocessing, which included normalisation using the *min-max* method. This step was crucial for adjusting the data values to a common scale, thereby facilitating the subsequent training of the predictive model.

The choice of LSTM as the predictive model was driven by its proven effectiveness in handling time-series data, making it particularly suited for forecasting tasks such as predicting SSC velocities. I. I. Zulfa et al. faced certain limitations, particularly the challenge of applying LSTM to short-term datasets. These models typically excel with long-term data, benefiting from extensive datasets to learn underlying patterns effectively.

In evaluating the performance of the LSTM model, the mean absolute percentage error (MAPE) metric was utilised. MAPE measures the accuracy of predicted values compared to actual values. The study achieved notably accurate predictions for the u and v components of SSC, with MAPE values of 14.15% and 8.43%, respectively, using a LSTM model configured with 50 hidden layers, a batch size of 32, and a learning rate drop period of 150.

The research concluded that using LSTM networks with specific parameter configurations, serves as a reliable tool for predicting the velocity and direction of SSC. However,the authors also suggest that further exploration into methods more suited to short-term data or the inclusion of seasonal variations and tidal factors could enhance predictive accuracy.

Bayindir [40] has a similar approach to I. I. Zulfa et al. [43], where the focus is also on using LSTMs to predict SSC. This choice is motivated by the LSTM's capability to capture long-term dependencies in sequential data, a common characteristic of SSC. The study uses a dataset collected by the National Oceanic and Atmospheric Administration (NOAA) in Massachusetts Bay, covering the period from November 2002 to February 2003, with measurements taken at a depth of 23.5m and recorded at intervals every 3 minutes and 44 seconds. This dataset, consisting of the current speed in two directions (u and v), undergoes preprocessing to standardise the data, ensuring zero mean and unit variance.

The methodology section stands out by providing a clear and concise explanation of how LSTM networks operate, including their sequence-to-sequence regression capability, which is central to predicting future states of SSC. Bayindir evaluates the LSTM model's performance by leveraging the RMSE error metric. This metric offers a direct comparison between the predicted and actual current velocities.

The results by Bayindir [40] demonstrate the LSTM model's ability to make accurate predictions, with significant improvements observed when the model incorporates real-time data updates. Initially, even without these updates, the LSTM model shows a strong capacity for predicting SSC, suggesting it can make reliable forecasts within a few future time steps. This is highlighted by the model's predictions exhibiting a higher frequency peak compared to the actual observed data, indicating a solid baseline accuracy. However, the research further reveals that when the model is refined with observed values, essentially updating it with real data, the accuracy of predictions markedly increases. This aspect underscores a common hurdle in machine learning and deep learning applications, where the quantity and quality of historical data can significantly impact the accuracy of the predicted values.

Predicting Sea Surface Temperatures (Multiple Locations)

H.-M. Choi et al. [44], develop LSTM networks to predict sea surface temperatures (SSTs) near the Korean Peninsula. The aim is to mitigate the impacts of rising SSTs due to global warming on marine ecosystems and aquaculture. The LSTM models demonstrate promising results in predicting SSTs and identifying high water temperature events with high accuracy for short-term forecasts.

They acknowledge limitations such as decreased prediction accuracy for longer-term forecasts and a reliance solely on SST data without considering other environmental factors. The evaluation of the model's performance was done through metrics like R^2, RMSE, MAPE, and F1 score. These metrics collectively assess the model's accuracy and its capability in classifying high water temperature events. The results indicate promising accuracy, particularly for short-term predictions up to four days in advance. However, the model's accuracy decreases for longer prediction windows, highlighting a critical area for improvement.

In H.-M. Choi et al. [44], the goal of predicting SSTs across a grid of 1519 data points near the Korean Peninsula closely mirrors the challenge we face in forecasting SSC velocities for various data point locations, which will subsequently be merged and utilized as inputs into a Lagrangian model. By training a predictive model on a 12-year dataset for all 1519 data points and then making predictions for subsequent days, this approach demonstrates a structured methodology for accurate environmental predictions. This process can be effectively applied to predict conditions across multiple locations in a marine area, paving the way for other applications.

2.2.3 Model Integration with Physics-Based Lagrangian Model

In [45], J. Mansui et al. set out to explore the dispersal of floating macro litter across the Mediterranean Sea by implementing a two-stage modelling approach to achieve their objective. The authors utilised the NEMO Oceanic General Circulation Model, which is designed to simulate global and regional ocean dynamics and marine ecosystem changes. This model was specifically configured for the Mediterranean basin, to simulate the sea state and velocity fields necessary for the drift simulations. This model configuration allows for a fine-scale representation of the region's oceanic conditions, which is crucial for accurate simulation of ocean currents and phenomena. The velocity outputs from this model are taken as daily averages, providing a consistent long-term description of the surface conditions.

Following the generation of these velocity fields, the second stage involves Lagrangian simulations. These simulations also use the data from the NEMO Oceanic General Circulation Model to simulate the movement of virtual particles that mimic the behaviour of floating macro litter at the sea surface. To specifically observe the surface transport pathways of the floating macro litter, the method focuses solely on surface movements without considering vertical dynamics or windage. The integration of these two models enables a comprehensive simulation of debris movement and accumulation.

The research by J. Mansui et al. [45] produced significant results, demonstrating seasonal and regional variations in floating macro litter distribution across the Mediterranean Sea. These findings were visualised to illustrate accumulation zones, offering a dynamic view of how marine debris disperses over time. The results aligned well with empirical data from previous studies, reinforcing the model's validity and effectiveness. J. Mansui et al. conclude that the integration of Lagrangian simulations with the OGCM offers a powerful framework for predicting marine litter distribution, highlighting its reliability.

In the report by the United Nations Environment Programme (UNEP) [46], the primary focus is to evaluate floating marine litter within the Northwest Pacific region. This is done by integrating different models with a physics-based Lagrangian model, aiming to enhance the understanding and management of marine litter trajectories.

This methodology is noteworthy for the integration of Eulerian models with Lagrangian particle tracking to predict and analyse the behaviour and dispersion patterns of marine litter. Eulerian models provide crucial data on ocean currents and winds by solving fluid dynamics equations on a fixed grid, establishing the environmental values that will be later utilised in the pipeline. Building upon this, the Lagrangian model simulates the trajectories of individual particles as they navigate through the ocean's dynamic conditions like currents and winds, which are determined by the Eulerian outputs. Through various case studies conducted within the Northwest Pacific region, the authors demonstrate the effectiveness of this combined methodology in predicting not only the movement but also the deposition areas of marine litter, providing valuable insights into effective management and mitigation strategies.

The report also addresses some challenges inherent in these models. A significant challenge highlighted is that the integration of wind-induced leeway drift poses discrepancies between observed and modelled trajectories, particularly under conditions of strong winds. Overall, The results derived from the applied models are largely successful, providing visual maps and simulations that depict litter trajectories and accumulation zones. These visualisations serve as crucial tools for understanding the impact of physical factors like currents and winds on the distribution of debris.

2.3 Summary

In This chapter we provided a detailed overview for understanding the complex dynamics of marine debris, its ecological impacts, and the methodologies used to forecast its movements. The background section offered an in-depth understanding of the datasets and outlined the essential components of the project, which included the environmental implications of marine debris and the modelling techniques utilised. The literature review further expanded on this by examining a number of studies, highlighting advancements and efforts in marine dispersal prediction. This section also aligned the project with various literature, reinforcing the project's contributions to predictive modelling and the integration of AI models with physics-based models.

3 Methodology

In this chapter we propose our approach to achieve the project's objectives, highlighting the reasoning behind each decision, and provide insights into the practical implementation of these objectives.

3.1 Data Integration and preprocessing

The preprocessing of datasets is a crucial first step in any project. In this project, we dealt with raw historical NetCDF [13] data spanning a total of four years, from January 2020 to December 2023 in hourly increments. The data was split into multiple folders and sub-folders for each day, necessitating a robust method to merge and preprocess the data without interfering with its temporal and spatial dimensionality.

To address this, we developed a framework that allows us to specify the start and end dates for the merging of the SSC data. The framework then merges the individual files along the time dimension, creating a single comprehensive dataset that encompasses all relevant data across the specified interval. This merged dataset is not only more manageable but also streamlined for any subsequent processes. A key feature of our framework is the preservation of the geographical boundaries and temporal aspects of the data. The dataset maintains the latitude and longitude ranges, ensuring that the spatial integrity is uncompromised. Similarly, checks were performed to ensure the time remained consistent, preserving the temporal integrity of the data.

Upon analysing the data, a substantial number of missing values, represented as Not a Number (NaN), were discovered. These NaNs are likely due to the proximity of the data to the coast, where high-frequency radars often struggle to capture all the data accurately. We decided not to address these NaN values at this stage. Each objective requires specific handling of missing data, which will be explored in the next sections. This preprocessing framework was utilised in every part of the project, from the Lagrangian simulations and AI models training, to the project's evaluation.

3.2 Lagrangian Model Development

For this part of the project, the OceanParcels toolkit [20] was utilised. The first step before proceeding with the simulation was to open the seven-day preprocessed SSC dataset. Following this, the *shapefile* of Malta, a digital vector storage format for storing geometric location and associated attribute information, was loaded and used to create the land-sea mask. This mask, illustrated in Figure 3.1, was produced by rasterising [47] the coastline *shapefile*. This mask effectively differentiates land from sea, ensuring accurate particle. The mask was saved as a NetCDF [13] file and added within the grid boundaries to match with the boundaries of the dataset. These coastal boundaries are crucial for defining the simulation area and facilitating land-sea interactions.

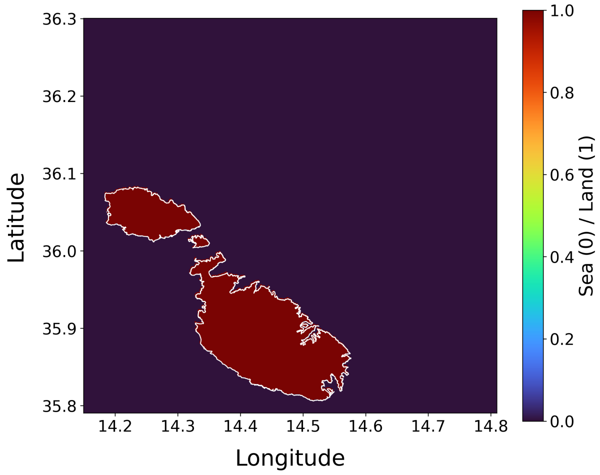

Figure 3.1 – Land-sea mask of Malta.

Subsequently, a *FieldSet* was created from the SSC dataset. This serves as the simulation environment, defining the velocity fields that drive particle movement. Additionally, the land-sea mask was integrated into the *FieldSet*, providing necessary data for handling particles upon reaching land. As depicted in Figure 3.2, simulation particles were initialised near a specific geographic coordinate (36.0475°N, 14.5417°E), with random offsets to simulate a dispersed release. The particles represent the objects of interest, such as sea surface debris, whose movements are to be simulated. Initially, the strategy involved simulating numerous randomly placed particles across the entire area; however, to enhance realism, we placed a cluster of 50 particles in close proximity. This configuration was selected to accurately represent how debris navigates marine environments, with each particle representing a cluster of debris.

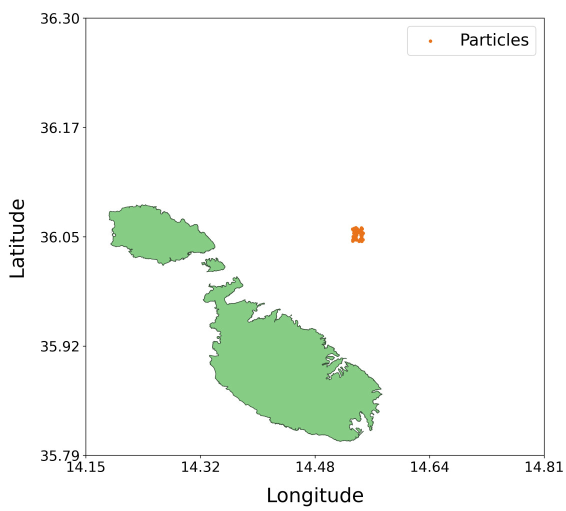

Figure 3.2 – Location of initial particles showing a cluster of debris.

The development and implementation of custom kernels was a critical component of the simulation. Custom kernels allow us to introduce specific behaviours into the simulation, modelling realistic scenarios that particles may encounter. The behaviours we implemented include:

- *CheckOutOfBounds:* deletes particles from the simulation if they move beyond the defined boundaries. This is necessary because no data is available outside the boundary, causing particles to get stuck.

- *CheckError:* deletes particles encountering computational errors. This ensures the simulation proceeds without disrupted or incorrect particle data.

- *UpdateElapsedTime:* shows how long a particle has been in the simulation. This tracks the duration of the particle within the environment.

- *UpdatePreviousPosition:* captures the position of particles before they move. This is useful as it allows us to save all the previous positions of the particles.

- *ReflectOnLand:* applies a reflection behaviour when particles encounter land, as defined by the land-sea mask. It also introduces a probabilistic component where there is a 15% chance that particles will 'beach' and be removed from the simulation, while the remaining 85% chance allows particles to be reflected back into the sea. This probabilistic distribution is informed by the geographic characteristics of Malta, where the predominance of rocky coastlines over sandy beaches increases the likelihood of debris being deflected back into the sea rather than beached.

The simulation was executed, and the resulting particle movements and dispersion patterns were visualised. These visualisations provide valuable insights into the trajectories of particles and their interactions with the environment. The time-step for the Lagrangian simulation is set at every 10 minutes, capturing the continuous dynamics of particle dispersion. This interval provides a good balance between computational efficiency and accuracy. The results are saved as an animated GIF file, offering a dynamic and easily interpretable visual representation of the simulated particle dispersion over time.

Some challenges emerged during this section of the project. Initially, simulations revealed that particles were getting stuck at the border boundaries. This issue was traced back to the dataset, which lacked data at the borders, rendering the particles unresponsive to environmental variables in these areas. To address this, the boundary of the simulation area was slightly reduced by 0.1°. Initially, our intention was to run the simulation for three years. However, it became apparent that such a lengthy period was unreasonable and did not align with the project's goals. Therefore, this was adjusted to a more useful seven-day simulation period. To achieve this, the preprocessing steps discussed in the previous section 3.1 were utilised to merge the data from January 1st to January 7th 2023. Originally, the goal was to incorporate both wind and current data into the simulation, but we decided to only use the SSC data. This decision was influenced by two main factors. Firstly, research like [48] indicates that sea surface debris is predominantly influenced by SSC rather than wind. M. Erikson et al. [48] also suggest that while wind does play a role, it is less significant than SSC in determining the spatial distribution of microplastics. Secondly, the complexity of integrating wind data and building custom behaviour kernels for wind interactions proved too challenging within the project's time frame. Despite this, we recognise that including wind data may potentially enhance the accuracy of the simulations in depicting real-life scenarios and propose future work.

To deal with missing data, we first attempted to interpolate it to fill in missing values. The visualisation results were noticeably different from those produced using the raw data which included NaNs. Despite experimenting with both linear and spline interpolation, the outcomes of the interpolated simulations remained consistent across different time frames, suggesting that the interpolation was homogenising the data excessively. This uniformity introduced by interpolation was misleading, as it failed to represent the true variability and dynamics of the SSC, compromising the actual behaviour and movement patterns of particles in the sea. Consequently, we decided not to remove NaN values from the data used in the Lagrangian simulations. This decision was based on the understanding that removing or interpolating these values could lead to simulations that do not accurately reflect real-world conditions. By preserving the integrity of the original dataset, including its inherent gaps, the simulations are more likely to represent the actual conditions and variations that marine debris would encounter in the sea. Examples of the final visualisations are presented in Figures 3.9 and 4.8.

The objective of initially implementing the Lagrangian model was to enhance our understanding and facilitate more informed decision-making for the subsequent sections. This exploratory phase was crucial in setting the stage for integrating the AI models, ensuring that we had a solid foundation with clear expectations.

3.3 AI Models Development

Predicting the dynamic and complex patterns of SSC velocities is a crucial objective of this project. Therefore, a comprehensive pipeline was established for this purpose. The initial phase involved selecting appropriate models, with LSTM and GRU architectures identified as the optimal choices. This decision was informed by their demonstrated effectiveness in processing time-series data, rendering them particularly suitable for this task, shown in studies by A. M. Ali et al. [42] and I. I. Zulfa et al. [43]. By utilising these models, this project aims to accurately predict the dispersion of marine debris around Malta's coastal waters, addressing both the temporal dynamics and spatial complexities inherent in SSC movements.

3.3.1 Data Preprocessing and Geospatial Filtering

This pipeline, illustrated in Figure 3.8, plays an essential role in facilitating future predictions. The process begins with the preprocessing of the data. Our initial plan was to train a model on a year's worth of data to forecast SSC for the next month. However, the complexity and four-dimensional nature of the data led to sub-optimal predictions. Consequently, the strategy was revised to extend the dataset used for training, which now spanned from February 25th 2020, to August 1st 2023. Given the substantial volume of data points and the need for geospatial filtering, a decision was made to concentrate on a smaller area of interest along the northern coast of Malta, as illustrated in Figure 3.3. The methodology involves predicting the u and v components individually for each longitude and latitude pair within the defined area. These individual predictions are subsequently merged into a unified file, which supports the execution of the subsequent Lagrangian simulation. This targeted approach, delineated by a specific polygon, ensures a more effective modelling process.

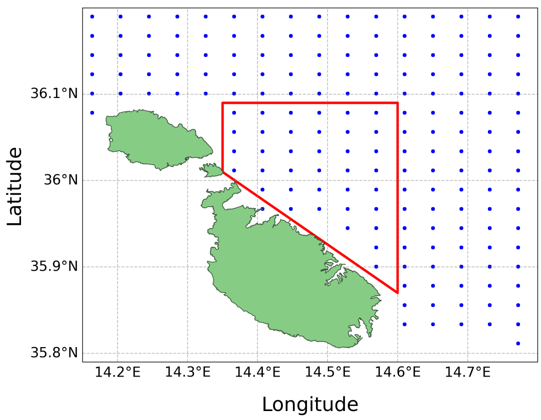

Figure 3.3 – All data points and selected area of interest.

The dataset was geospatially filtered to only include data points within the designated area of interest, as shown in Figure 3.4. This filtering resulted in a focused dataset consisting of 37 data points. The data did not require any normalisation as it had already been scaled between -1 and 1 during the data collection phase. Further preprocessing involved the removal of extraneous columns to streamline the dataset. Each coordinate pair was then processed into individual Comma-separated Variable

(CSV) files. These files were systematically named and organised according to the corresponding latitude and longitude coordinate points. These CSV files served as the basis for training the AI models for every individual data point location.

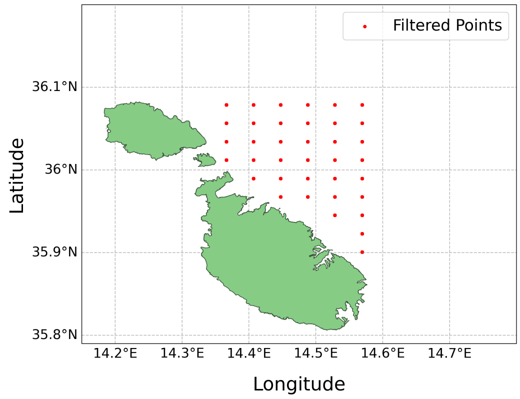

Figure 3.4 – Filtered points within area of interest.

In addressing missing data, we observed that areas closer to the coast exhibited fewer data points, as illustrated in the heat map in Figure 3.5. This lack of data near coastal regions is likely attributed to several factors. These include radar interference from nearby land or structures, obstruction of radar beams by coastal terrain or buildings, and refraction of radar waves at the coast, all contributing to distorted data collection. Efforts to solve this issue included experiments with data interpolation and filling missing values with the mean. However, these methods yielded worse results compared to those obtained by dropping the NaN values. Due to this, the most effective strategy proved to be the removal of NaN values, a decision informed by testing and also aligning with the methodologies applied for the Lagrangian model as discussed in section 3.2.

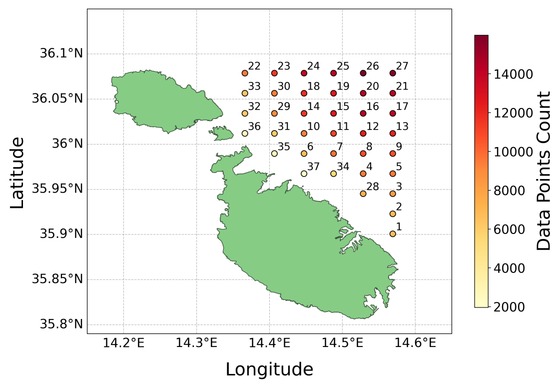

Figure 3.5 – Amount of data points per coordinate pair.

Prior to creating the pipeline, preliminary testing was conducted on a single model to determine the most effective features and targets. Experimentation involved the integration of both u and v as features, revealing marginally improved outcomes compared to using each as a single feature, prompting a focus on using both features for the predictions. Notably, the model yielded good results when predicting a single target, but the accuracy of predictions noticeably diminished upon testing the model to predict both (u and v) as targets. This observation led to the decision to develop separate models for each target variable to maximise the accuracy of the results. Therefore, we implemented a series of 37 models to predict the u component, and replicated this process for the v component, ensuring precise and reliable predictions.

3.3.2 The Main Loop

Inspired by the approach of H.-M. Choi et al. [44], who developed a model for every individual coordinate pair, we established a comprehensive pipeline that iterates through each pair of coordinates in the dataset and trains a dedicated model for each individual pair. This approach enables us to make predictions across the entire area of interest, which are later utilised for the Lagrangian simulations. The process begins by organising the CSV files according to their index. For each file, the u and v columns are extracted as input features. The dataset is then divided into training, validation, and testing segments in a 70-15-15 split respectively, a common split for ensuring a balance between adequate model training and thorough evaluation. Given the time series nature of the data, it was necessary to sequence the data appropriately; this was achieved using the *TimeseriesGenerator* library. Iterative testing of different parameter combinations revealed that a window size of 72 hours, a batch size of 64, and a sampling rate of 1 yielded the best overall results. This means that the data is sequenced into continuous blocks of 72 hours of data as input and paired with the value immediately following these 72 hours as the target output. This crucial step allows the model to essentially predict the next hour based on the preceding 72 hours of data, a crucial step for accurate forecasting.

Various architectures and hyperparameters were also tested to find which ones gave the best results. To ensure a fair comparison, we applied identical layer configurations and hyperparameters across both the LSTM and GRU architectures. The most effective architecture involved ten hidden layers, composed of four LSTM/GRU layers, three dropout layers, and two dense layers activated by *ReLU*, followed by an additional dropout layer after each dense layer. The models were set up with a learning rate of 0.001, *Adam* optimiser, and the MSE loss function. Importantly, the model was reinitialised in each iteration of the loop, ensuring that each dataset was trained on a fresh instance without any residual weights from previous iterations. This approach is crucial when dealing with multiple datasets to avoid any data leakage or influence from previously trained models (clean slate training). It also helps maintain the integrity of the learning process for each distinct dataset. Early stopping with a patience of 8 epochs was implemented to halt training and prevent overfitting. Model checkpoints were utilised to save the best-performing epoch automatically. After each training epoch, plots comparing training versus validation loss were generated to monitor the performance of each model. To ensure that each model was trained adequately, predictions were made on the test set and subsequently visualised through a graph comparing actual versus predicted values, as illustrated in Figure 3.6. Finally, adopting a similar evaluation approach to Adhikari et al. [23], MAE, MSE and RMSE error metrics were computed and displayed to evaluate the model's performance on the test set.

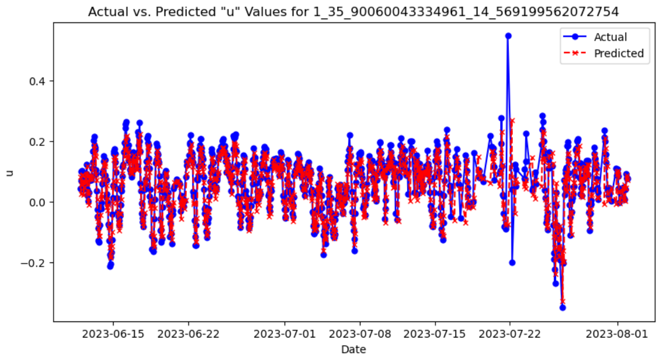

Figure 3.6 – Actual vs predicted values on test set.

3.3.3 Making Real World Predictions

In the final phase of the AI model pipeline, we undertook a simulation mirroring a real-world scenario by feeding historical data spanning 72 hours to predict the subsequent 24 hours. This 24-hour prediction window is chosen as it provides a balance between short-term accuracy and computational feasibility, which is often considered optimal for time series forecasting in dynamic environments like SSC. We decided to use data from August 1st to August 3rd 2023, as input, aiming to predict conditions for August 4th 2023. This setup allowed us to compare the predictions with actual historical data from our dataset. The process began with a loop to systematically extract SSC data across three days for all 37 individual coordinate pairs. Subsequently, actual data for the following 24-hour period on August 4th was extracted for comparative purposes, and both sets were saved as CSV files. Given the requirement for 72 consecutive hours of data to be able to make predictions and 24 hours for comparison, spline interpolation was utilised to address any present NaN values, ensuring the dataset's completeness. Using the rolling forecasting method as illustrated in Figure 3.7, predictions were generated for the subsequent 24-hour period for the individual targets, either u or v. Given the use of rolling predictions and the inclusion of interpolated values, it is important to recognise that the accuracy of the predictions may be affected. This is because predictions based on interpolated data serve as inputs for subsequent forecasts, potentially diminishing their precision. This effect is particularly noticeable in longer-term predictions, where accuracy tends to decrease as the forecast horizon extends, expectedly showing a decline in prediction accuracy the further the prediction extends into the future.

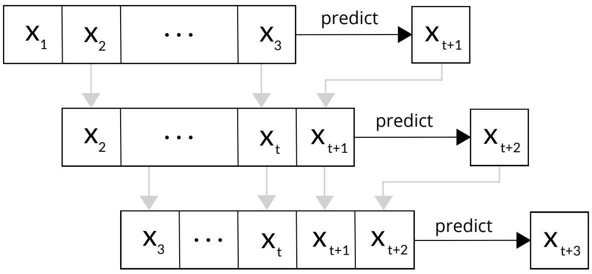

Figure 3.7 – The process of a rolling forecast.

This pipeline is repeated for all 37 data points and the predictions are converted into NetCDF format to be subsequently used for the Lagrangian model. Finally, the same error metrics are calculated and printed, to be used later for the evaluation. This pipeline was replicated four times in total, encompassing two LSTM models and two GRU models, one for each of the u and v components respectively. The architecture of the AI models pipeline is illustrated in Figure 3.8 below:

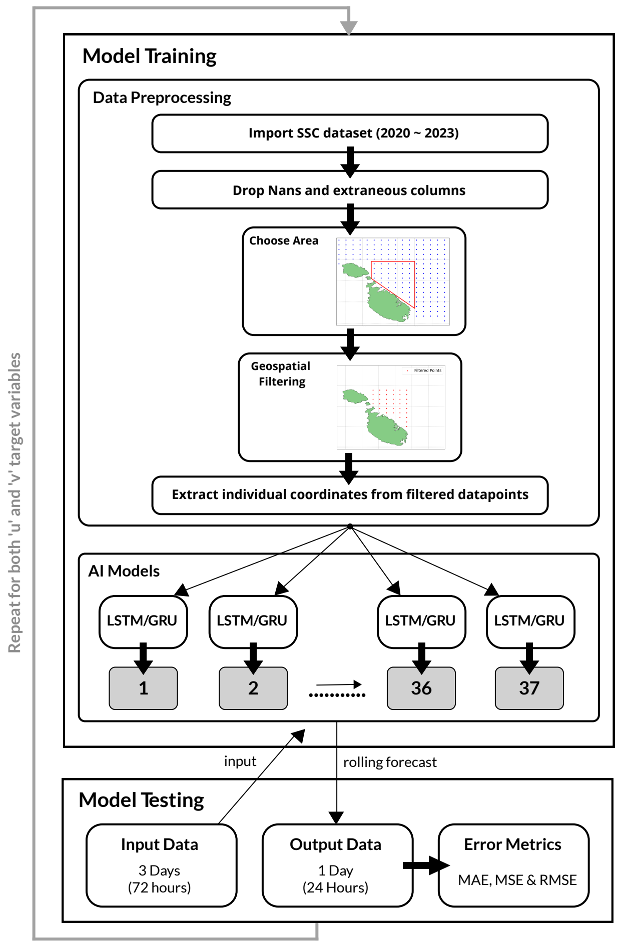

Figure 3.8 – Overview of entire AI model pipeline.

3.4 Integrating AI models with Lagrangian Model

The final stage of the pipeline involves the integration of the predictions generated by the AI models with the physics-based Lagrangian model to produce a 24-hour forecast simulation of sea surface debris dispersion.

This process was conducted separately for both the LSTM and GRU model predictions. The initial and most critical step involved the preprocessing and merging of the predicted u and v values. Multiple checks were implemented to ensure that the merging process was done correctly. The merged dataset was then converted into NetCDF format. Following this, the procedures outlined in Section 3.2 were implemented once again to set up the Lagrangian simulation framework. This involved the configuration of the land-sea mask, *FieldSet*, number of particles, kernels, and timestep. The only difference lies in the particle initialisation phase. Given that the AI model predictions are specific to the area of interest (as depicted in Figure 3.4), we decided to set the centroid of the polygon as the starting position. More specifically, the coordinates are at a latitude of 35.9895° and a longitude of 14.4944°. This approach is advantageous as it allows for an unbiased observation of dispersion patterns. The reason being, that since the centroid is equidistant from all edges of the polygon, it provides a neutral starting point that does not inherently favour any flow direction. In terms of the offset of the initial particles, the same random seed was used for both LSTM and GRU simulations. This ensures a fair comparison as the initial locations of the particles are identical for both models.

Finally, the Lagrangian simulations for both the LSTM and GRU models were executed and stored. The resulting particle movements and dispersion patterns were visualised and saved as *GIFs*. These visualisations (Figures 3.9 and 4.8), allow us to observe the surface debris movement predictions. They also facilitate the evaluation of the results produced by the LSTM and GRU models respectively.

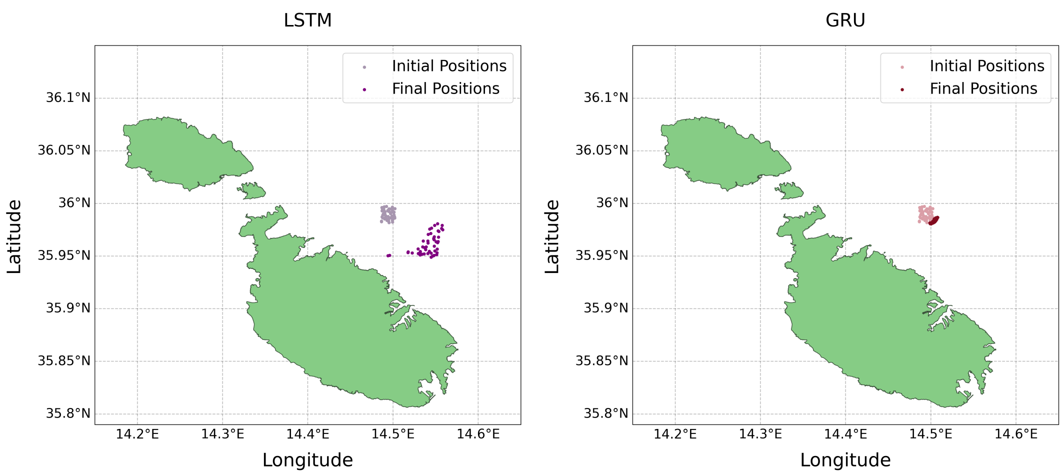

Figure 3.9 – LSTM and GRU initial vs final debris movement (4th August).

3.5 Evaluation Strategy

The last phase of the project focuses on a comparative evaluation between the LSTM and GRU predictions. We needed a way to test the performance of the models fairly, facilitating the comparison of the predicted results and determining which model performs best.

Originally we wanted to evaluate the Lagrangian framework by comparing dispersion patterns with drifter data, as undertaken in previous research by van Sebille et al. [34] and Aijaz et al. [38]. However, due to the coastal proximity of our area of interest and the general practice of deploying drifters in open waters to avoid beaching, the drifter data was not available for our specific area. While it would be the optimal approach to assess the Lagrangian simulations, we were unable to do so due to the unavailability of the necessary data. Instead, we decided to shift focus solely on the comparative analysis of LSTM and GRU model predictions. The accuracy of these predictions was evaluated using various error metrics to establish a comparative baseline. Furthermore, a spatial evaluation was incorporated to examine the similarity between the predicted dispersion patterns generated by each model.

To enhance the evaluation process, we opted to re-run the pipeline to predict SSC for an alternate time period. Specifically, we inputted interpolated data spanning from November 1st to November 3rd 2023, into the pipeline and utilising the same trained models as before, generated predictions for a 24-hour period for November 4th 2023. This allowed us to have two evaluation frameworks for two separate dates, thereby facilitating a more comprehensive evaluation of the results.

3.5.1 Error Metrics Evaluation

Adopting an approach similar similar to that of H. Yadav et al. [49], we developed a concise pipeline to assess the predictive accuracy of the LSTM and GRU models. We utilised error metrics to compare the actual historical values versus the predicted values. Specifically, we used the MAE, MSE, and RMSE:

$$\textbf{MAE} = \left(\frac{1}{N}\right) \sum |y_i - \hat{y}|$$

$$\textbf{MSE} = \left(\frac{1}{N}\right) \sum (y_i - \hat{y})^2$$

$$\textbf{RMSE} = \sqrt{\text{MSE}} = \sqrt{\left(\frac{1}{N}\right) \sum (y_i - \hat{y})^2}$$

Here, N is the number of observations or data points, y_i is the actual value for the i^{th} observation, and $\hat{y}$ is the predicted value for the i^{th} observation. Given that our analysis includes 37 distinct models for both the u and v components, we computed the average mean and standard deviation for each metric to facilitate a more comprehensive evaluation of the LSTM versus GRU results across two timeframes: August 4^{th} and November 4^{th}. During our analysis, we identified certain outliers within our results. To address this, we also calculated the average using the Interquartile Range (IQR), focusing on the differences between the 75^{th} and 25^{th} percentiles for each metric, thereby obtaining a more robust mean that excluded these outliers.

3.5.2 Geospatial Evaluation

In the second phase of our evaluation process, we introduced geospatial evaluation to enhance the comparative analysis of the LSTM and GRU models. This was motivated by the observation that each specific data point within our area of interest yielded varied results. Initially, we generated a heat map (Figure 3.5) to delineate which model corresponded to each location and to quantify the amount of data available for each model. Subsequently, to further our analysis for both the LSTM and GRU models, we visualised the MAE values for the u and v components using additional heat maps. The choice of MAE as a metric was deliberate, as it provides a straightforward and uniformly interpretable measure that treats all errors equivalently. These visualisations were instrumental in comparing the spatial accuracy of the models, highlighting areas where each model exhibited better or worse performance. These findings were then compared with the original heat map to discern patterns and discrepancies. This comparative analysis was executed for the predictions corresponding to both August 4^{th} and November 4^{th} 2023, facilitating a thorough assessment of each model across different temporal contexts.

In the final stage of our evaluation, we quantify the performance of the LSTM and GRU models across different regions by calculating their centroid, spread, and skewness for their respective Lagrangian simulation outputs. The specific measures include:

- **Mean, median, and standard deviation of centroids:** we computed the geographical centroids of the merged predictions from both the LSTM and GRU models to assess the proximity of the final debris movement predictions generated by the two models. The Euclidean distances from these centroids were then analysed, with the mean, median, and standard deviation calculated. Smaller values suggested a higher degree of concurrence between the models' predictions.

- **Spread of LSTM and GRU:** the spatial spread was determined by calculating the standard deviation of distances from each model's centroid. A lower standard deviation indicated a tighter clustering around the centroid, thus reflecting a more consistent model performance across the area.

- **Longitudinal and latitudinal skewness of LSTM and GRU:** to understand the directional tendencies of the models' predictions, we calculated the skewness for the distribution of the prediction points' longitude and latitude. A skewness close to zero indicates a symmetrical distribution of prediction errors, whereas a positive or negative skewness value points to a systematic bias in a particular direction.

By analysing these statistical measures, we aimed to determine not only which model had lower error values on average but also how those errors were distributed across the spatial domain. This approach offers insights into whether one model consistently outperformed the other across the entire study area, and also explores whether the models exhibited particular strengths or weaknesses in distinct regions. A comprehensive explanation and analysis of these evaluation results are presented in the subsequent Chapter 4.

3.6 Summary

In this chapter, we outlined the methodology implemented for this project, detailing the approaches undertaken for data integration, preprocessing, and model development. We began by establishing a framework for merging and preprocessing historical data, ensuring the preservation of its temporal and spatial integrity. This set the stage for the Lagrangian model simulations to simulate sea surface debris movements. Subsequently, we delved into the development of AI models, specifically LSTM and GRU architectures for predicting SSC velocities. This involved a detailed setup of data preprocessing, geospatial filtering, and a methodical iterative training process for each coordinate pair within our area of interest. The chapter concluded with the integration of these AI model predictions with the Lagrangian simulations to forecast sea surface debris dispersion, followed by an evaluation strategy utilising error metrics and geospatial analysis to assess the models' predictive accuracy and spatial distribution tendencies.

4 Evaluation

In this chapter, we present a detailed overview and a comprehensive explanation of the results obtained through the evaluation strategy outlined in Section 3.5. The primary objectives are to ascertain which model, LSTM or GRU, demonstrates superior performance, and to evaluate the similarity of the Lagrangian simulations generated using these models' predictions. As previously noted, the framework was executed on two specific dates, August 4th and November 4th 2023. This was done to gauge the models' consistency and reliability under varying seasonal conditions, offering a thorough analysis of their performance across different environmental dynamics. The chapter is structured into two main sections. Section 4.1 delves into the analysis of average error metrics to discern which model performed best. Section 4.2 focuses on a geospatial analysis to investigate whether the locations and amount of data influence the predictions. It also briefly compares the merged predictions with the Lagrangian simulations. This approach ensures a thorough evaluation of the models' effectiveness and their applicability in real-world scenarios.

4.1 LSTM vs GRU

In the initial experiment, we assessed the accuracy of the models in predicting SSC velocities by comparing the predicted results against actual historical values using three key error metrics:

- **MAE:** this metric computes the average absolute difference between actual and predicted values across the dataset. It quantifies the typical magnitude of the prediction errors without considering their direction, providing a clear measure of average error.

- **MSE:** this represents the average of the squared differences between the actual and predicted values. It accentuates larger errors more significantly than smaller ones by squaring the differences, highlighting impactful prediction discrepancies.

- **RMSE:** calculated as the square root of the MSE, it measures the standard deviation of residuals, offering a scale-sensitive accuracy measure. It provides an indication of the typical magnitude of prediction errors in the same units as the data.

4.1.1 Error Metrics Results

The average error metrics for the 24-hour rolling predictions from all 37 models on the 4th of August 2023 are presented in tables 4.1 to 4.4 below:

Table 4.1 LSTM u average error metrics (4th August).

Metric	Mean	Std Dev	IQR
MAE	0.141 26	0.22658	0.05830
MSE	0.116 93	0.51328	0.01012
RMSE	0.179 57	0.29100	0.05289

Table 4.2 LSTM v average error metrics (4th August).

Metric	Mean	Std Dev	IQR
MAE	0.143 70	0.13416	0.14072
MSE	0.064 05	0.10900	0.07281
RMSE	0.183 00	0.17483	0.21160

Table 4.3 GRU u average error metrics (4th August).

Metric	Mean	Std Dev	IQR
MAE	0.148 49	0.22171	0.06700
MSE	0.115 89	0.50339	0.01603
RMSE	0.186 93	0.28451	0.07046

Table 4.4 GRU v average error metrics (4th August).

Metric	Mean	Std Dev	IQR
MAE	0.144 72	0.13763	0.14908
MSE	0.065 58	0.11215	0.06614
RMSE	0.183 76	0.17936	0.20195

The analysis of these results reveals insightful differences in model performance. For the u component, LSTM models demonstrate slightly lower MAE and RMSE, indicating better average accuracy and consistency, although GRU models show a marginally lower MSE. Conversely, for the v component, both models perform similarly with minimal variations across all metrics, which suggests a near-equivalent capability in handling this type of prediction. Further examination of the variability through standard deviation and IQR metrics shows that LSTM models have a lower standard deviation in the v component predictions, suggesting more consistent performance relative to GRU. Additionally, the smaller IQR for LSTM in both components implies that its predictions are more tightly clustered around the median, indicating less variability and more reliability. While both models performed well, LSTM offered marginally better performance, particularly for the u component, establishing it as the preferable model.

On the other hand, the results for the 24-hour rolling predictions for all 37 models on the 4th of November are detailed in tables 4.5 to 4.8 below:

Table 4.5 LSTM u average error metrics (4th November).

Metric	Mean	Std Dev	IQR
MAE	1.031 41	2.11801	0.29923
MSE	14.765 09	40.75769	0.24335
RMSE	1.634 16	3.47773	0.39745

Table 4.6 LSTM v average error metrics (4th November).

Metric	Mean	Std Dev	IQR
MAE	2.621 64	5.50710	0.50617
MSE	97.857 86	253.42893	0.86051
RMSE	4.238 77	8.93816	0.84444

Table 4.7 GRU u average error metrics (4th November).

Metric	Mean	Std Dev	IQR
MAE	1.051 10	2.11189	0.56820
MSE	14.772 34	40.77340	0.46635
RMSE	1.651 05	3.47079	0.58636

Table 4.8 GRU v average error metrics (4th November).

Metric	Mean	Std Dev	IQR
MAE	2.653 16	5.50156	0.53727
MSE	97.980 28	253.62271	1.14379
RMSE	4.268 01	8.93109	0.99123

When it comes to the u component, both models display relatively high MAE, MSE, and RMSE, with LSTM showing lower metrics. The high standard deviations observed for both models suggest a significant presence of outliers, indicating some predictions were inaccurate. This is evident in the GRU u component where the IQR is higher, suggesting a broader spread compared to LSTM, pointing to more frequent outliers in the GRU model. In contrast, the v component shows considerably higher error values for both models, with GRU again having higher values across all metrics. The standard deviations and IQR values are significantly larger in the v component for both models, reinforcing the presence of outliers and indicating that predictions for the v component were generally less accurate and more variable. Overall, the LSTM model performs better than the GRU model, particularly in the v component, as evidenced by the lower error metrics and narrower IQR. Therefore, the more consistent performance of LSTM across both tests makes it the better model overall.

4.1.2 Error Metrics Discussion of Results

The analysis highlights that predictions for the u component (east-west velocity) were generally more accurate than for the v component (north-south velocity). This discrepancy can be attributed to the alignment of radar systems, which are aligned in a north-south orientation, as depicted in Figure 2.1. This alignment potentially impacts the accuracy of v component readings and inhibits the radar's ability to capture detailed north-south data, consequently leading to less accurate predictions for the v component. Moreover, the first experiment conducted on the 4th of August exhibited notably better results, characterised by lower error values and fewer outliers when compared to the subsequent November evaluation. This improvement is likely due to the August data being in close proximity to the final sequences of the test dataset, potentially leading to the models being better tuned to these conditions. Furthermore, the process of rolling forecasting, which bases predictions on preceding outputs, may lead to inaccuracies, particularly when initial predictions are derived from interpolated data. This method could inherently propagate errors, especially under conditions of notable missing data as discussed in Subsection 3.3.3.

Such findings emphasise the necessity of considering temporal proximity and data integrity when assessing model performance. These phenomena and their implications on model performance will be explored further in the next section.

4.2 Geospatial Analysis

In this section, we extended our analysis using the insights garnered from the heat map depicted in Figure 3.5. Our objectives were multifaceted. We aimed to determine whether an increased volume of data correlates with enhanced predictive accuracy, whether data points closer to the coast, typically characterised by less data, yield poorer performance, and whether the time of year and seasonality of the data impact the results. The primary goal was to assess whether the geographical location of data points influences the accuracy of the predictive models. Additionally, we sought to compare the final Lagrangian simulations generated by both the LSTM and GRU models to evaluate their similarities in modelling sea surface debris.

4.2.1 The Hypothesis

As highlighted in Subsections 3.3.1 and 3.5.2, the dataset contains a significant number of NaNs, with data points closer to the coast having more NaNs present, as evidenced in Figure 3.5. This is corroborated by Figure 4.1, which illustrates how some data points have significantly less data available. Figure 3.5 is instrumental in demonstrating the correlation between model performance and geographic location, thereby paving the way for a focused analysis on the impact of data availability at specific locations. Based on these observations, we formulated a hypothesis:

Hypothesis 4.1 *A plausible hypothesis could suggest that data points near the coast exhibit reduced data availability due to environmental and technical challenges that interfere with radar performance. This scarcity of data consequently impairs the accuracy of predictive models for SSC velocities near the coast. Specifically, it suggests that models predicting SSC at coastal locations perform less effectively compared to those further offshore, where radar data tends to be more complete.*

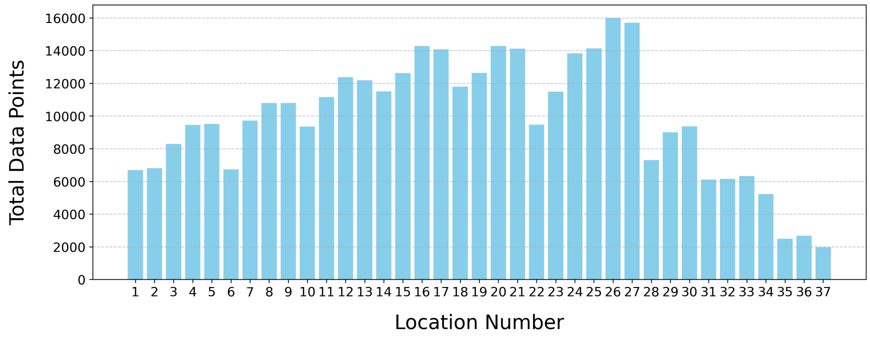

Figure 4.1 – Histogram of data points by location number.

In the following subsections, we will investigate hypothesis 4.1 by conducting a comparative analysis with the heat map in Figure 3.5. This analysis will help us validate or refute our assumption regarding the influence of geographical location on model accuracy.

4.2.2 Heat Maps Results and Analysis

As outlined in Subsection 3.5.2, we utilised heat maps to geospatially analyse the performance of our models, focusing on the MAE across all 37 models. These visualisations are crucial for testing the validity of hypothesis 4.1 regarding the impact of data availability on predictive accuracy near coastal areas. Heat maps for the MAE error metrics were produced for both the u and v components of both LSTM and GRU models, covering both the August 4th and November 4th 2023 predictions. To enhance the clarity of these visualisations and minimise the influence of outliers, we applied a clipping method at the 95th percentile of the data. This method effectively limited the range of data considered for colour scaling in the heat maps, allowing for more nuanced visual comparisons between most data points by excluding extreme outliers. The resultant heat maps for August 4th are displayed in figures 4.2 and 4.3 below:

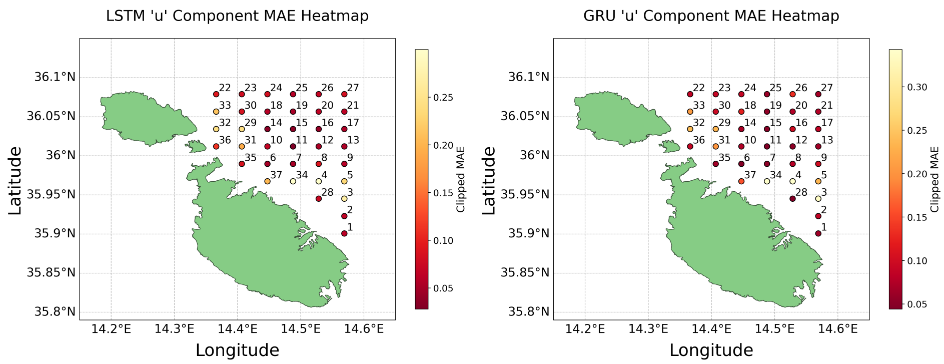

Figure 4.2 – u component MAE heat maps (4th August).

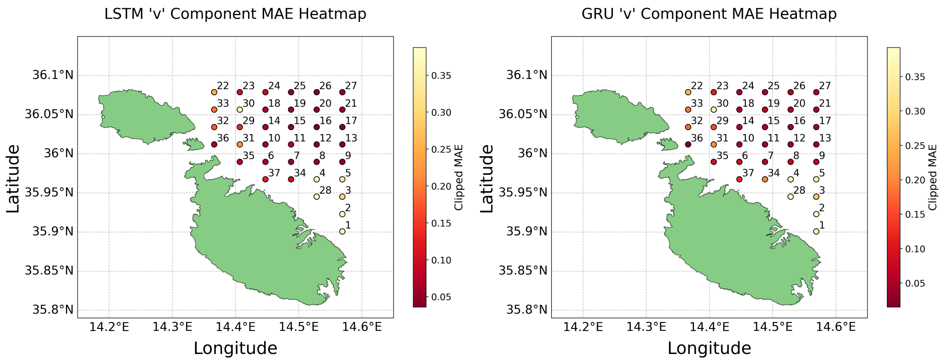

Figure 4.3 – v component MAE heat maps (4th August).

The heat maps for November 4th are are displayed in figures 4.4 and 4.5 below:

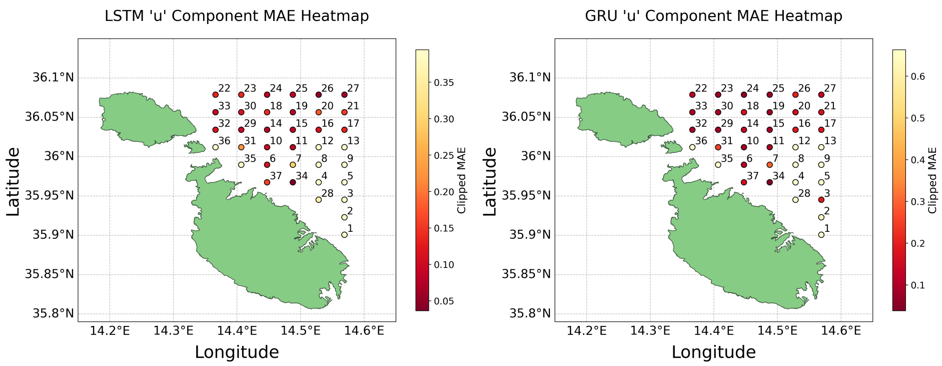

Figure 4.4 – u component MAE heat maps (4th November).

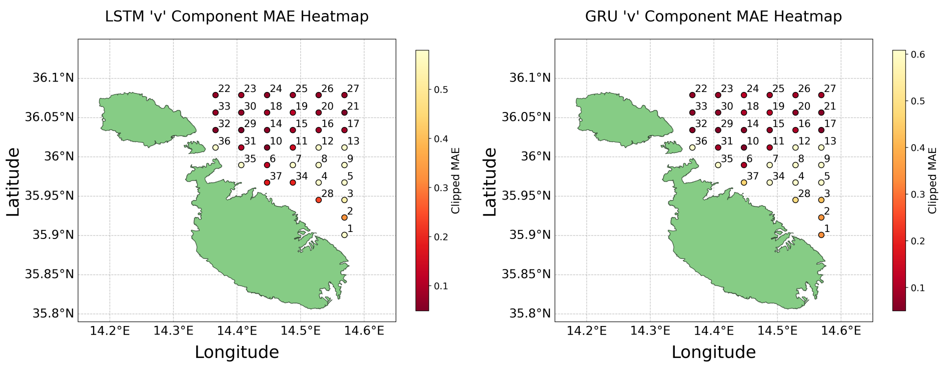

Figure 4.5 – v component MAE heat maps (4th November).

The hypothesis 4.1 is supported by the evidence presented in the MAE heat maps. These visualisations clarify that while models near the coast generally perform worse, presumably due to less data being available, there are notable exceptions where coastal predictions maintained good accuracy. This observation challenges the assumption that greater data volume directly correlates with higher predictive accuracy. Instead, it suggests that the models' capabilities to handle noise and extract meaningful patterns from sparse data can significantly impact their effectiveness.

Furthermore, the differential accuracy between the u and v components highlights the complexity of environmental factors and model sensitivities, which contribute to a diverse spectrum of outcomes that are not solely determined by the amount of data available.

Dividing the analysis between the August and November predictions reveals a stark contrast; predictions for August are markedly more accurate, underscoring the importance of the temporal proximity of training data to the predictive period and the impact of seasonal variations. The outliers observed on August 4^{th} are not consistent with those on November 4^{th}, highlighting the temporal variability of missing data and its non-uniform impact across different points and times. It becomes evident that no consistent pattern exists, thus challenging any definitive conclusions. While certain models excel in predicting the u component, their performance diminishes when applied to the v component. These observations attest to the intricate dynamics at play in predictive modelling, where factors like the models' ability to manage noise and the inherent directional properties of data lead to outcomes where less data does not always result in less accuracy. Therefore, these findings not only challenge but effectively refute hypothesis 4.1.

4.2.3 Comparison of Lagrangian Simulations

In the final component of the evaluation framework, we assessed the performance similarities between the LSTM and GRU models by analysing their centroids, spreads, and skewness within their respective Lagrangian simulation outputs. The findings for August 4^{th} are depicted in Figure 4.6.

```
Mean Centroid Distance: 0.019570° (2.17 km)
Median Centroid Distance: 0.018598° (2.06 km)
Std Dev of Centroid Distances: 0.013839° (1.54 km)
```
```
LSTM Spread: 0.016261° (1.81 km)
GRU Spread: 0.004905° (0.54 km)
```
```
LSTM Longitude Skewness: 0.248566
LSTM Latitude Skewness: -0.487350
GRU Longitude Skewness: -1.076138
GRU Latitude Skewness: 0.380908
```

Figure 4.6 – LSTM vs GRU comparison on Lagrangian simulations (4^{th} August).

For the 4^{th} of August, the average centroid distances for the LSTM and GRU models were around 2 km, suggesting both models achieved different geographical accuracy. The standard deviation of these distances was 1.54 km, indicating a moderate spread around the centroids. However, the GRU model demonstrated a more compact spread of 0.54 km compared to the LSTM's 1.81 km, suggesting GRU's predictions were more tightly grouped. The skewness metrics showed a mild eastward and southward bias in the LSTM predictions, whereas GRU exhibited a stronger westward bias and a slight northward tendency, highlighting directional tendencies in their prediction patterns. On the contrary, the results for the 4th of November are shown in 4.7.

```
Mean Centroid Distance: 0.014718° (1.63 km)
Median Centroid Distance: 0.017193° (1.91 km)
Std Dev of Centroid Distances: 0.006339° (0.70 km)

LSTM Spread: 0.014387° (1.60 km)
GRU Spread: 0.023217° (2.58 km)

LSTM Longitude Skewness: 0.562745
LSTM Latitude Skewness: -0.166972
GRU Longitude Skewness: 0.277801
GRU Latitude Skewness: 0.475496
```

Figure 4.7 – LSTM vs GRU comparison on Lagrangian simulations (4th November).

The results from the 4th of November showed improvements in clustering, with mean and median centroid distances reduced to approximately 1.63 km and 1.91 km respectively. This reduction, coupled with a decreased standard deviation of 0.70 km, suggests enhanced prediction accuracy and consistency for this period. Interestingly, LSTM showed a more consistent spread of 1.60 km, whereas GRU's predictions were more dispersed, with a spread of 2.58 km. Skewness values also shifted, indicating changes in predictive behaviour that might be influenced by different environmental conditions or model sensitivities to the input data at that time.

Overall, the analyses highlight that the LSTMs generally offer more consistent and reliable performance, marked by less variability in spread and skewness compared to GRUs. While GRUs seemed to adjust its performance based on different conditions, suggesting a possible sensitivity to seasonal or environmental changes, the LSTMs maintained steadiness across the evaluated metrics. This consistent performance makes LSTMs the more suitable model for this application.

4.2.4 Comparison of Final Lagrangian Visualisations

The visual comparisons between Figures 3.9 and 4.8 indicate notable differences in the outcomes of the LSTM and GRU simulations. Specifically, the August simulations (Figure 3.9) reveal distinct disparities in the final particle locations between the LSTM and GRU models. Conversely, the November simulations (Figure 4.8), display a high degree of similarity. This observation lends support to findings from our previous observations in subsection 4.2.3, where the November results demonstrated greater alignment between the models when compared to August. Such disparities point out the variable nature of these predictions and highlight the complex interplay of various factors that significantly impact the accuracy and consistency of the final outcomes. This variability is illustrative of the inherent challenges in modelling time series data, where slight variations in input or parameters can lead to markedly different predictions.

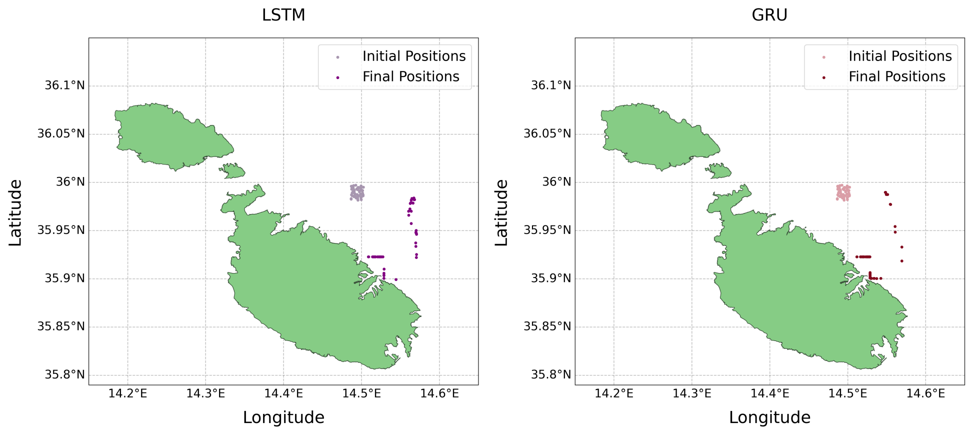

Figure 4.8 – LSTM and GRU initial vs final debris movement (4th November).

4.3 Summary

In this chapter we have explored the performance of LSTM and GRU models through a comprehensive evaluation strategy. The analysis was divided into an evaluation of error metrics and geospatial behaviour, which together provide a robust evaluation of the models' effectiveness and their practical applicability.

The findings indicate that LSTM provides more consistent and reliable predictions, establishing it as the preferred model when considering the performance of both components across diverse seasonal and environmental conditions. The geospatial analysis further corroborated these findings, showing that the LSTM generally maintains more consistent performance metrics, such as spread and skewness, when compared to the GRU. This consistency was evident despite the seasonal variations between the two dates, highlighting LSTM's robustness across different predictive scenarios.

Contrary to our hypothesis 4.1, the analysis does not consistently support the assertion that proximity to the coast and reduced data availability significantly degraded model performance. While coastal data points generally showed less accuracy, this was not universally the case. Some coastal predictions maintained good accuracy, suggesting that other factors, such as the models' capacity to handle sparse data and environmental noise, play a critical role in prediction outcomes. Therefore, the findings do not support the hypothesis that less data inherently results in poorer model performance.

5 Conclusion

This chapter reflects upon the objectives listed in Section 1.3, offering an assessment of the outcomes. It also addresses the challenges and limitations encountered during the study, presenting a transparent critique of the methodologies used. Finally, we explore potential improvements and directions for future research.

5.1 Revisiting our Aims and Objectives

The main aim of this project was to develop a predictive modelling system that leverages the strengths of AI models integrated with a physics-based Lagrangian model to predict the dispersion of sea surface debris within the coastal waters of Malta. To achieve this overarching goal, the project provided the following outcomes for our five objectives:

01. We preprocessed the SSC velocity datasets for a seamless integration into our models. This preprocessing was pivotal for maintaining consistency, laying the groundwork for the subsequent modelling phases.

02. A physics-based Lagrangian model was implemented to simulate the dispersion of marine debris on the sea surface withe the help of the OceanParcels toolkit.

03. LSTM and GRU models were successfully developed to predict SSC velocities through a methodical process involving iterative testing and hyperparameter tuning to optimise model performance, coupled with comprehensive training to ensure accurate and reliable predictions.

04. We merged the predictions from our AI models with the Lagrangian simulations, enabling us to create predictive visualisations of sea surface debris movement.

05. A comparative assessment of the AI models was conducted, evaluating their predictive accuracy and the reliability of the generated visualizations. This analysis revealed that the LSTM model outperforms the GRU model in predicting SSC velocities, as evidenced by superior performance in error metrics such as MAE, MSE, and RMSE.

5.2 Critique and Limitations

Throughout this project, we encountered several challenges and limitations that influenced the project's trajectory and outcomes. The primary challenge was the presence of missing data, particularly near coastal areas, which likely affected the precision of our predictive models. Originally, we planned to develop a web-page to visually demonstrate the model's predictions and offer our work as a tool to academics and as a public service. However, to better focus on implementing the AI and Lagrangian models, this component was deferred. Moreover, the geographical area of interest was limited, potentially constraining the broader applicability of our findings. The final limitation was the inability to empirically validate the Lagrangian model due to the absence of drifter data within the chosen area, preventing us from performing a direct evaluation of the model.

5.3 Future Work

This project lays the groundwork for various improvements. These include:

- Integrating additional weather phenomena like wind and wave height into the models could potentially enhance the predictions.
- The framework can be adapted to various applications, including predicting the movements of jellyfish and plankton, assisting in search and rescue operations, and simulating the dynamics of oil spills.
- Transitioning to ensemble learning methods may potentially enhance the accuracy of predictive models, as supported by the work of B. Naderalvojoud and T. Hernandez-Boussard [50]. Furthermore, the implementation of more sophisticated models, such as transformers, could also improve the predictions.
- Expanding the area of interest and consequently increasing the number of models could provide a more comprehensive understanding of marine debris dynamics. Furthermore, this would enable the evaluation of the Lagrangian model using historical drifter data.
- Developing a model specifically designed to predict and fill missing values within the datasets could enhance the accuracy of predictions and improve visualisations.
- As previously mentioned, developing a website to showcase future predictions with enhanced visualisations would make the research more accessible and informative. This would allow both academics and the general public to interact with the data in real-time, fostering greater engagement and understanding of the model's capabilities and environmental implications.

5.4 Final Remarks

In this project, we successfully integrated AI models with a physics based Lagrangian framework to predict and visualise the future 24-hour movement of sea surface debris around Malta's coastal waters. Through our evaluation process, we identified that the LSTM model outperformed the GRU model in terms of prediction accuracy. Our work is a comprehensive system that enhances marine conservation efforts by providing actionable insights for effective cleanup operations and simultaneously informing strategies for long-term marine conservation around the coast of Malta.

References

[1] C. Kehl, P. D. Nooteboom, M. L. A. Kaandorp, and E. van Sebille, "Efficiently simulating lagrangian particles in large-scale ocean flows — data structures and their impact on geophysical applications," *Computers and Geosciences*, vol. 175, p. 105 322, 2023. DOI: 10.1016/j.cageo.2023.105322. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0098300423000262.

[2] G. Suaria and S. Aliani, "Floating debris in the mediterranean sea," *Marine pollution bulletin*, vol. 86, no. 1, pp. 494–504, 2014. DOI: 10.1016/j.marpolbul.2014.06.025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0025326X14004056.

[3] M. Compa *et al.*, "Risk assessment of plastic pollution on marine diversity in the mediterranean sea," *Science of The Total Environment*, vol. 678, pp. 188–196, 2019. DOI: 10.1016/j.scitotenv.2019.04.355. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0048969719318984.

[4] J. Mansui, G. Darmon, T. Ballerini, O. van Canneyt, Y. Ourmieres, and C. Miaud, "Predicting marine litter accumulation patterns in the mediterranean basin: Spatio-temporal variability and comparison with empirical data," *Progress in Oceanography*, vol. 182, p. 102 268, 2020. DOI: 10.1016/j.pocean.2020.102268. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0079661120300069.

[5] P. G. Ryan, "A brief history of marine litter research," in (Marine Anthropogenic Litter), M. Bergmann, L. Gutow, and M. Klages, Eds., Marine Anthropogenic Litter. Cham: Springer International Publishing, 2015, pp. 1–25. DOI: 10.1007/978-3-319-16510-3_1. [Online]. Available: https://doi.org/10.1007/978-3-319-16510-3_1.

[6] P. R. Pawar, S. S. Shirgaonkar, and R. B. Patil, *Plastic marine debris: Sources, distribution and impacts on coastal and ocean biodiversity*, Jan. 2016.

[7] S. Katsanevakis, "Chapter 2 - marine debris, a growing problem: Sources, distribution, composition, and impacts," in New York: Nova Science Publishers, 2008, pp. 53–100.

[8] D. W. Laist, "Impacts of marine debris: Entanglement of marine life in marine debris including a comprehensive list of species with entanglement and ingestion records," in (Marine Debris: Sources, Impacts, and Solutions), J. M. Coe and D. B. Rogers, Eds., Marine Debris: Sources, Impacts, and Solutions. New York, NY: Springer New York, 1997, pp. 99–139. DOI: 10.1007/978-1-4613-8486-1_10. [Online]. Available: https://doi.org/10.1007/978-1-4613-8486-1_10.

[9] C. M. Rochman, M. A. Browne, A. J. Underwood, J. A. van Franeker, R. C. Thompson, and L. A. Amaral-Zettler, "The ecological impacts of marine debris: Unraveling the demonstrated evidence from what is perceived," eng, *Ecology*, vol. 97, no. 2, pp. 302–312, Feb. 2016. DOI: 10.1890/14-2070.1.

[10] P. Agamuthu, S. B. Mehran, A. Norkhairah, and A. Norkhairiyah, "Marine debris: A review of impacts and global initiatives," eng, *Waste management and research : the journal of the International Solid Wastes and Public Cleansing Association, ISWA*, vol. 37, no. 10, pp. 987–1002, Oct. 2019. DOI: 10.1177/0734242X19845041.

[11] J. Harlan *et al.*, "The integrated ocean observing system high-frequency radar network: Status and local, regional, and national applications," *Marine Technology Society journal*, vol. 44, no. 6, pp. 122–132, Nov. 2010. DOI: 10.4031/MTSJ.44.6.6. [Online]. Available: https://www.ingentaconnect.com/content/mts/mtsj/2010/00000044/00000006/art00017.

[12] *Portus 3.0*. [Online]. Available: https://portus.research.um.edu.mt/?p=13.833.

[13] *Unidata | netcdf*. [Online]. Available: https://www.unidata.ucar.edu/software/netcdf/.

[14] E. van Sebille *et al.*, "Lagrangian ocean analysis: Fundamentals and practices," *Ocean Modelling*, vol. 121, pp. 49–75, 2018. DOI: 10.1016/j.ocemod.2017.11.008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1463500317301853.

[15] S. A. Lonin, "Lagrangian model for oil spill diffusion at sea," *Spill Science and Technology Bulletin*, vol. 5, no. 5-6, pp. 331–336, 1999. [Online]. Available: http://inis.iaea.org/search/search.aspx?orig_q=RN:32026786.

[16] L. C. .-. Lebreton, S. D. Greer, and J. C. Borrero, "Numerical modelling of floating debris in the world's oceans," *Marine pollution bulletin*, vol. 64, no. 3, pp. 653–661, Mar. 2012. DOI: 10.1016/j.marpolbul.2011.10.027. [Online]. Available: https://dx.doi.org/10.1016/j.marpolbul.2011.10.027.

[17] M. N. Dawson, A. S. Gupta, and M. H. England, "Coupled biophysical global ocean model and molecular genetic analyses identify multiple introductions of cryptogenic species," *Proceedings of the National Academy of Sciences*, vol. 102, no. 34, pp. 11 968–11 973, 2005. DOI: 10.1073/pnas.0503811102. [Online]. Available: https://doi.org/10.1073/pnas.0503811102.

[18] D. Hertwig *et al.*, “Development and demonstration of a lagrangian dispersion modeling system for real-time prediction of smoke haze pollution from biomass burning in southeast asia,” *Journal of geophysical research. Atmospheres*, vol. 120, no. 24, pp. 12 605–12 630, Dec. 2015. DOI: `10.1002/2015JD023422`. [Online]. Available: `https://onlinelibrary.wiley.com/doi/abs/10.1002/2015JD023422`.

[19] R. G. Williams and M. J. Follows, *Ocean Dynamics and the Carbon Cycle: Principles and Mechanisms*. Cambridge: Cambridge University Press, 2011. DOI: `10.1017/CBO9780511977817`. [Online]. Available: `https://www.cambridge.org/core/product/31EF28FEF48A172FF746B3E654F9455A`.

[20] *Oceanparcels*. [Online]. Available: `https://oceanparcels.org`.

[21] *Pygnome*. [Online]. Available: `https://gnome.orr.noaa.gov/doc/pygnome/index.html`.

[22] I. Pisso *et al.*, “The lagrangian particle dispersion model flexpart version 10.4,” 2019. DOI: `10.5194/gmd-12-4955-2019`. [Online]. Available: `http://hdl.handle.net/11250/2634384`.

[23] R. Adhikari and R. K. Agrawal, “An introductory study on time series modeling and forecasting,” *ArXiv*, vol. abs/1302.6613, 2013. [Online]. Available: `https://api.semanticscholar.org/CorpusID:17070340`.

[24] T. Raicharoen, C. Lursinsap, and P. Sanguanbhokai, “Application of critical support vector machine to time series prediction,” vol. 5, 2003, p. V. DOI: `10.1109/ISCAS.2003.1206419`.

[25] S. Raksha, J. S. Graceline, J. Anbarasi, M. Prasanna, and S. Kamaleshkumar, “Weather forecasting framework for time series data using intelligent learning models,” 2021, pp. 783–787. DOI: `10.1109/ICEECCOT52851.2021.9707971`.

[26] A. Chatterjee, H. Bhowmick, and J. Sen, “Stock price prediction using time series, econometric, machine learning, and deep learning models,” 2021, pp. 289–296. DOI: `10.1109/MysuruCon52639.2021.9641610`.

[27] P. Wang, S. H. Gurmani, Z. Tao, J. Liu, and H. Chen, “Interval time series forecasting: A systematic literature review,” *Journal of forecasting*, vol. 43, no. 2, pp. 249–285, Mar. 2024. DOI: `10.1002/for.3024`. [Online]. Available: `https://onlinelibrary.wiley.com/doi/abs/10.1002/for.3024`.

[28] S. Jadon, J. Milczek, and A. Patankar, “Challenges and approaches to time-series forecasting in data center telemetry: A survey,” Cornell University Library, arXiv.org, Tech. Rep., Feb. 2021. DOI: `10.48550/arxiv.2101.04224`. [Online]. Available: `https://search.proquest.com/docview/2477387048`.

[29] A. Alsharef, Sonia, K. Kumar, and C. Iwendi, “Time series data modeling using advanced machine learning and automl,” *Sustainability*, vol. 14, no. 22, 2022. DOI: 10.3390/su142215292.

[30] I. H. Sarker, “Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions,” *SN Computer Science*, vol. 2, no. 6, p. 420, 2021. DOI: 10.1007/s42979-021-00815-1. [Online]. Available: https://doi.org/10.1007/s42979-021-00815-1.

[31] S. Hochreiter, “The vanishing gradient problem during learning recurrent neural nets and problem solutions,” *International journal of uncertainty, fuzziness, and knowledge-based systems*, vol. 6, no. 2, pp. 107–116, Apr. 1998. DOI: 10.1142/S0218488598000094. [Online]. Available: http://www.worldscientific.com/doi/abs/10.1142/S0218488598000094.

[32] M. J. Hamayel and A. Y. Owda, “A novel cryptocurrency price prediction model using gru, lstm and bi-lstm machine learning algorithms,” *AI*, vol. 2, no. 4, p. 496, 2021. DOI: 10.3390/ai2040030.

[33] P. T. Yamak, L. Yujian, and P. K. Gadosey, *A comparison between arima, lstm, and gru for time series forecasting*, 2020. DOI: 10.1145/3377713.3377722. [Online]. Available: https://doi.org/10.1145/3377713.3377722%20http://dx.doi.org/10.1145/3377713.3377722.

[34] E. van Sebille *et al.*, “The physical oceanography of the transport of floating marine debris,” *Environmental research letters*, vol. 15, no. 2, pp. 23 003–32, Feb. 2020. DOI: 10.1088/1748-9326/ab6d7d. [Online]. Available: https://iopscience.iop.org/article/10.1088/1748-9326/ab6d7d.

[35] B. D. Hardesty *et al.*, “Using numerical model simulations to improve the understanding of micro-plastic distribution and pathways in the marine environment,” *Frontiers in Marine Science*, vol. 4, Mar. 2017. DOI: 10.3389/fmars.2017.00030. [Online]. Available: https://search.proquest.com/docview/2307448993.

[36] W. R. Winans, Q. Chen, Y. Qiang, and E. C. Franklin, “Large-area automatic detection of shoreline stranded marine debris using deep learning,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 124, p. 103 515, 2023. DOI: 10.1016/j.jag.2023.103515. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1569843223003394.

[37] M. Yuniarti, G. M. Faid, Subiyanto, and M. R. Ismail, “Trajectory mapping of microplastics originating from the seto inland sea, japan,” *AACL Bioflux*, vol. 16, no. 6, pp. 3138–3149, 2023. [Online]. Available: http://www.bioflux.com.ro/aacl.

[38] S. Aijaz, F. Colberg, and G. B. Brassington, "Lagrangian and eulerian modelling of river plumes in the great barrier reef system, australia," *Ocean Modelling*, vol. 188, p. 102 310, 2024. DOI: 10.1016/j.ocemod.2023.102310. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1463500323001506.

[39] W. Liu *et al.*, "Ssd: Single shot multibox detector," in *Computer Vision – ECCV 20*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., Cham: Springer International Publishing, Sep. 2016, pp. 21–37.

[40] C. Bayindir, "Predicting the ocean currents using deep learning," *TWMS Journal of Applied and Engineering Mathematics*, vol. 13, no. 1, pp. 373–385, 2023. [Online]. Available: https://jaem.isikun.edu.tr/web/images/articles/vol.13.no.1/34.pdf.

[41] S. Dauji, M. C. Deo, and K. Bhargava, "Prediction of ocean currents with artificial neural networks," *ISH Journal of Hydraulic Engineering*, vol. 21, no. 1, pp. 14–27, 2015. DOI: 10.1080/09715010.2014.938133. [Online]. Available: https://doi.org/10.1080/09715010.2014.938133.

[42] A. M. Ali, H. Zhuang, J. VanZwieten, A. K. Ibrahim, and L. Chérubin, "A deep learning model for forecasting velocity structures of the loop current system in the gulf of mexico," *Forecasting*, vol. 3, no. 4, p. 953, 2021. DOI: 10.3390/forecast3040056.

[43] I. I. Zulfa, D. C. R. Novitasari, F. Setiawan, A. Fanani, and M. Hafiyusholeh, "Prediction of sea surface current velocity and direction using lstm," *IJEIS (Indonesian Journal of Electronics and Instrumentation Systems) (Online)*, vol. 11, no. 1, pp. 93–102, Apr. 2021. DOI: 10.22146/ijeis.63669. [Online]. Available: https://doi.org/10.22146/ijeis.63669.

[44] H.-M. Choi, M.-K. Kim, and H. Yang, "Deep-learning model for sea surface temperature prediction near the korean peninsula," *Deep Sea Research Part II: Topical Studies in Oceanography*, vol. 208, p. 105 262, 2023. DOI: 10.1016/j.dsr2.2023.105262. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0967064523000127.

[45] J. Mansui, G. Darmon, T. Ballerini, O. van Canneyt, Y. Ourmieres, and C. Miaud, "Predicting marine litter accumulation patterns in the mediterranean basin: Spatio-temporal variability and comparison with empirical data," *Progress in Oceanography*, vol. 182, p. 102 268, 2020. DOI: 10.1016/j.pocean.2020.102268. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0079661120300069.

[46] E. P. U. Nations, *Review and analysis of floating marine litter prediction models in the nowpap region - technical report no. 36*, 2018. [Online]. Available: `https://wedocs.unep.org/20.500.11822/26237`.

[47] A. Bhunia, P. Chowdhury, Y. Yang, T. M. Hospedales, T. Xiang, and Y. Song, "Vectorization and rasterization: Self-supervised learning for sketch and handwriting," 2021, pp. 5668–5677. DOI: `10.1109/CVPR46437.2021.00562`. [Online]. Available: `http://doi.ieeecomputersociety.org/10.1109/CVPR46437.2021.00562`.

[48] M. Eriksen *et al.*, "Plastic pollution in the world's oceans: More than 5 trillion plastic pieces weighing over 250,000 tons afloat at sea," eng, *PloS one*, vol. 9, no. 12, e111913, Dec. 2014. DOI: `10.1371/journal.pone.0111913`.

[49] H. Yadav and A. Thakkar, "Noa-lstm: An efficient lstm cell architecture for time series forecasting," *Expert Systems with Applications*, vol. 238, p. 122 333, 2024. DOI: `10.1016/j.eswa.2023.122333`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S095741742302835X`.

[50] B. Naderalvojoud and T. Hernandez-Boussard, "Improving machine learning with ensemble learning on observational healthcare data," *AMIA Annu Symp Proc*, vol. 2023, pp. 521–529, 2024.

Appendix A Pipeline Pseudocode

```
# Initial processing
main_data_frame = load_data()  # 1
main_data_frame = remove_nans(main_data_frame)  # 2
visualize_points(main_data_frame)  # 3

# Polygon filtering
polygon_filtered_data = polygon_filter(main_data_frame)  # 4
save_to_csv(polygon_filtered_data, "filtered_data.csv")  # 5

# Loop to process and save individual data frames
for coords in polygon_filtered_data:
    individual_data_frame = create_data_frame(coords)  # 6
    save_data_frame(individual_data_frame, f"folder/{coords}.csv")

# Loop through folder to create sequences
for data_frame in load_data_frames("folder"):  # 7
    sequences = create_sequences(data_frame)
    train, val, test = split_data(sequences)  # 8

    # Train on AI model
    for model in [LSTM()]:  # 9
        model.train(train, val)
        save_model_results(model, f"results/{data_frame}")  # 10

# Processing coordinate points
for i in range(37):  # 11
    input_data_72_hrs, output_data_24_hrs = create_input_output(i)
    store_data(input_data_72_hrs, f"storage/input_{i}.csv")
    store_data(output_data_24_hrs, f"storage/output_{i}.csv")

# Rolling forecast predictions
for input_data_72_hrs in load_data_frames("storage"):  # 12
    predictions = make_forecast(input_data_72_hrs, 24)
    for coord in input_data_72_hrs.coords:
        error_metrics = calculate_errors(predictions, output_data_24_hrs)
        print_metrics(error_metrics)  # 13

    merge_predictions(predictions, "merged_predictions.csv")  # 14

# Save and merge final data
save_data("merged_predictions.csv", "final_storage/")  # 15

# Repeat for other target variable 'v'
repeat_process(variable) # 16
```

```

# Final merge and simulation
merged_data = merge_target_variables('u', 'v')  # 17
simulate_lagrangian_model(merged_data)  # 18

# Repeat all same steps for GRU model
repeat_for_model(GRU()) # 19
```

Listing A.1 – Pseudocode for main Pipeline