5-2024

Geomagnetic Substorms Prediction Model Using Combined Physics-Based and Deep Learning Modeling Techniques

Ruthba Yasmin

THE UNIVERSITY OF SOUTH ALABAMA
COLLEGE OF ENGINEERING

GEOMAGNETIC SUBSTORMS PREDICTION MODEL USING COMBINED PHYSICS-BASED AND DEEP LEARNING MODELING TECHNIQUES

BY

Ruthba Yasmin

A Thesis

Submitted to the Graduate Faculty of the
University of South Alabama
in partial fulfillment of the
requirements for the degree of

Master of Science

in

Electrical Engineering

May 2024

Approved: Date: 04/16/2024

Edmund Spencer
Edmund Spencer (Apr 17, 2024 13:22 CDT) Apr 17, 2024

Chair of Thesis Committee: Dr. Edmund Spencer

Purbi Adhya
Purbi Adhya (Apr 17, 2024 13:44 CDT) Apr 17, 2024

Committee Member: Dr. Purbi Adhya

Mohamed Shaban Apr 17, 2024

Committee Member: Dr. Mohamed Shaban

Silas Leavesley
Silas Leavesley (Apr 17, 2024 14:20 CDT) Apr 17, 2024

Committee Member: Dr. Silas Leavesley

M El Sharkh Apr 17, 2024

Chair of Department: Dr. Yusef El-Sharkh

C Woods Apr 17, 2024 BP
BP

Director of Graduate Studies: Dr. Clive Woods

Harold Pardue Apr 22, 2024

Dean of the Graduate School: Dr. J. Harold Pardue

GEOMAGNETIC SUBSTORMS PREDICTION MODEL USING COMBINED PHYSICS-BASED AND DEEP LEARNING MODELING TECHNIQUES

A Thesis

Submitted to the Graduate Faculty of the
University of South Alabama
in partial fulfillment of the
requirements for the degree of

Master of Science

in

Electrical Engineering

by
Ruthba Yasmin
B.A., Mount Holyoke College, 2014
M.S. United International University, 2021
May 2024

ACKNOWLEDGEMENTS

I would like to express my utmost gratitude to my research advisor Dr. Edmund Spencer for his endless support throughout this research. He has given me valuable feedback and knowledge at every turn of research. Being his direct student for the past year and a half, I gained tremendous knowledge in the field of Space Physics and Electromagnetism, which was pivotal to achieve the goals of this research. I would like to thank Dr. Yusef El-Sharkh, Chair of the Electrical and Computer Engineering department, for his relentless support. I am also indebted to all my faculty members and staff members at the Electrical and Computer Engineering department for their help and encouragement. I express my gratitude to my parents, my brother, family and friends who always motivated me to achieve my academic goals, work towards my highest potential and provided their consistent heartfelt support. I want to thank my friends and colleagues at the University of South Alabama for helping me during my graduate journey. Lastly, I want to thank Dr. Purbi Adhya for her substantial contributions and support throughout this graduate research.

TABLE OF CONTENTS

Page

LIST OF TABLES v

LIST OF FIGURES vi

LIST OF ABBREVIATIONS viii

ABSTRACT ix

CHAPTER I INTRODUCTION 1

1.1 Motivation 3

1.1.1 Incorporating Comprehensive Data Sources 4
1.1.2 Exploring Parameter Influences 5
1.1.3 Evaluating Substorm Onset Lists 5
1.1.4 Comparing Model Performances 5
1.1.5 Integration with Physics Models 5

1.2 Contributions and Challenges 6
1.3 Organization of the Thesis 6

CHAPTER II GEOMAGNETIC SUBSTORMS 7

2.1 External Energy Pathways 9
2.2 Reconnection and Magnetic Field Orientation 9
2.3 Energy Dynamic Summary 10
2.4 Geomagnetic Substorm Onset Phases 12

2.4.1 Substorm Growth Phase 13
2.4.2 Substorm Onset and Expansion Phase 13
2.4.3 Substorm Recovery Phase 14

CHAPTER III SUBSTORM ONSET DETECTION TECHNIQUES 16

3.1 Frey et al. (2004) and Mende et al. (2009) Onset Detection Technique.....17
3.2 Newell and Gjerloev (2011) Onset Detection Technique.....19
3.3 Forsyth et al. (2015) Onset Detection Technique.....21
3.4 Ohtani et al. (2020) Onset Detection Technique.....22
3.5 Methodology to Create an Aggregated List for Substorm Onset.....24

CHAPTER IV WINDMI AND NEURAL NETWORK MODELS.....28

4.1 WINDMI Model.....29
4.2 LSTM Architecture.....33
4.3 CNN Architecture.....36
4.4 Data Processing and Model Training.....39
4.5 Performance Analysis.....50

CHAPTER V MODEL PREDICTIONS AND RESULTS.....52

5.1 Model Training and Convergence.....53
5.2 LSTM Cases.....55
5.3 CNN Cases.....58
5.4 Model Performance for different lists.....61
5.5 Performance Analysis.....64

CHAPTER VI SUMMARY AND CONCLUSIONS.....70

REFERENCES.....73

APPENDICES.....78

Appendix A LSTM Results.....78
Appendix B CNN Results.....83

BIOGRAPHICAL SKETCH.....88

LIST OF TABLES

Appendix Table Page

A1. Performance Evaluation Metrics for Aggregated Substorm List using LSTM 78

A2. Performance Evaluation Metrics for Forsyth Substorm List using LSTM.....79

A3. Performance Evaluation Metrics for Newell Substorm List using LSTM.....80

A4. Performance Evaluation Metrics for Ohtani Substorm List using LSTM 81

A5. Performance Evaluation Metrics for Frey Substorm List using LSTM.....82

B1. Performance Evaluation Metrics for Aggregated Substorm List using CNN.....83

B2. Performance Evaluation Metrics for Forsyth Substorm List using CNN 84

B3. Performance Evaluation Metrics for Newell Substorm List using CNN.....85

B4. Performance Evaluation Metrics for Ohtani Substorm List using CNN.....86

B5. Performance Evaluation Metrics for Frey Substorm List using CNN 87

LIST OF FIGURES

Figure Page

2.1 Illustration of the interaction between the solar wind and Earth's magnetosphere (Tsurutani et al., 2022) 8

2.2 The Magnetosphere, a sketch from Kivelson and Russell (1995) 8

2.3 A schematic of the magnetospheric substorm 12

3.1 Map of North America showing the substorm onset locations in geographic coordinates regardless of the local time of onset 17

3.2 Substorm onsets criterion by Ohtani and Gjerloev (2020) 23

3.3 Illustration depicting the merging of substorm onset times obtained from different substorm lists sourced from the SuperMAG website 26

4.1 The WINDMI model describes the physical parameters of the magnetosphere-ionosphere system as defined by the 8 differential equations Spencer et al. (2007) 31

4.2 WINDMI model outputs with I (Geo-tail current), I_1 (Field-aligned current), θ (Trigger function) 32

4.3 WINDMI onset detection results 32

4.4 The Long Short-Term Memory (LSTM) cell, depicted here, can sequentially process data, and retain its hidden state across multiple time steps 33

4.5 A flowchart of processing pipeline for the LSTM Network, sequence input is the layer responsible for feeding all the input parameters (16 layers at once) and then passes it through the LSTM layer 35

4.6 Neurons depicted in blue belong to a convolutional layer, each connected to their respective receptive field highlighted in red 36

4.7 A flowchart of processing pipeline for the Convolution 2D Neural Network 38

4.8 An instance of the Solar Wind input data showing 9 parameters (case 13) segmented into 60 minutes training and 30 minutes prediction window.....39

4.9 A mind-map illustrating the 16 combinations cases adopted and sent into the two neural networks as sequence inputs of the time series data.....43

4.10 Description of Parameters and Physics-Incorporated Neural Network Model observational data and theoretical outputs from the WINDMI model, providing a comprehensive framework for substorm prediction and analysis.....48

5.1 Training Progress for one case using the LSTM Model.....53

5.2 A confusion matrix resulting from the evaluation of the LSTM model applied to the aggregated dataset for training case.....55

5.3 A confusion matrix resulting from the evaluation of the LSTM model applied to the aggregated dataset.....57

5.4 A confusion matrix resulting from the evaluation of the 2D-CNN model applied to the Newell dataset.....58

5.5 A confusion matrix resulting from the evaluation of the 2D-CNN model applied to the Newell dataset.....60

5.6 Heatmap depicting the comparison of LSTM and CNN model accuracies for the aggregated substorm list across 16 cases.....63

5.7 Performance heatmap illustrating the results of the LSTM network across 16 different cases for predicting substorm onset with Frey, Forsyth, Newell, Ohtani, and the aggregated substorm lists.....65

5.8 Performance heatmap illustrating the results of the CNN network across 16 different cases for predicting substorm onset with Frey, Forsyth, Newell, Ohtani, and the aggregated substorm lists.....67

LIST OF ABBREVIATIONS

ACE	Advanced Composition Explorer
AE	Auroral Electrojet
CNN	Convolutional Neural Network
IMF	Interplanetary Magnetic Field
LSTM	Long Short-Term Memory
PINN	Physics Incorporated Neural Network
SML	Super Magnetometer Lower Index
SYMH	Symmetric H-component of Ring Current

ABSTRACT

Ruthba Yasmin, M. S., University of South Alabama, May 2024. Geomagnetic Substorms prediction model using combined physics-based and deep learning modeling techniques. Chair of Committee: Edmund A. Spencer, Ph.D.

This thesis aims to develop a hybrid physics-incorporated neural network model (PINN) for classifying geomagnetic substorms in the Earth's Magnetosphere. The model is trained using a comprehensive list of substorm onsets, ground magnetometer data from a global network, and solar wind parameters from the Advanced Composition Explorer (ACE) satellite. Two different neural network architectures are used, and the physics model used for training is called WINDMI. The magnetic field components on the ground, which are a function of the ionospheric currents, are captured by the SML index. The methodology involves using 60-minute data segments preceding an event to train the hybrid PINN model to predict and classify substorms in the subsequent 30 minutes. A significant part of this research is predicting substorm occurrences based on substorm onsets from four distinct lists and an aggregated substorm list. The results, based on the performance of 16 adopted cases, show that the hybrid model performs significantly better than the WINDMI model alone. This finding underscores the importance of scrutinizing the physics model alongside the hybrid model.

CHAPTER I

INTRODUCTION

A Geomagnetic Substorm is a complex phenomenon that involves energy transfer from the magnetotail to the auroral ionosphere and is crucially linked to solar wind, magnetosphere-ionosphere parameters. Instrumentation, ranging from satellites to ground-based imagers and magnetometers, captures substorm onset metrics with distinct phases: growth, expansion, and recovery. Substorm onsets on the nightside manifest as auroral brightening and poleward expansion, crucially observed through all-sky imaging, simultaneously as westward auroral electrojet current (Newell et al., 2010).

Extensive research by experts Akasofu (1964), Frey et al. (2004), and McPherron (2020) among others has utilized space and ground-based instruments, as well as first-principles models, to understand substorms. Key determinants include the impact of solar wind and interplanetary magnetic field parameters, with IMF B_z being a significant factor affecting energy transfer.

The concept of substorms was first introduced in the 1960s to describe a sudden release of stored energy in the magnetosphere, resulting in dynamic changes in the magnetospheric and ionospheric systems (Akasofu, 1981). Substorms have been found to significantly alter the auroras in the polar region, as well as the magnetospheric magnetic field and plasma environments. Advancements in satellite observations, ground-based

measurements, and computer simulations have provided valuable insights into the underlying processes of substorm onset.

Debates exist on the triggers for substorm expansion, whether externally triggered by solar wind and IMF changes or dominated by internal triggers (Akasofu, 2004). The sequence of events during the growth phase is also a topic of discussion, questioning magnetosphere preconditioning and has contributed to our understanding of detailed measurements of magnetic field variations, particle populations, and plasma dynamics. In recent decades, satellite reliance has increased, and substorms pose threats to satellite operations and infrastructure. Prediction efforts, from early neural networks to recent support vector machines and binary classification, highlight the need for accurate forecasting.

The goal of this research is to analytically evaluate the substorm onset criteria based on ground-based magnetic indices, field-aligned currents, auroral observations, electrojet indices criteria to establish a relationship and lay the foundation for neural network training and modeling. This is a novel approach in this domain of Space Physics because neural network training to classification and tuning of the hyperparameters has never been attempted before across several substorm lists. The model correlated each set of parameters (input cases) and used different lists. This is the first time this methodology has been attempted which attests to the novelty of the framework. Substorm lists are key indicators that were used to enable the model to find a correlation between solar wind, ground measurements and the lists. The auroral electrojet and the ring current are compared to AL and Dst respectively. The motive of this thesis is to utilize a comprehensive substorm list along with pseudo-break up events identified by auroral

brightening signatures from ground-based imagers. The WINDMI model, being the Physics model incorporated into the bigger architecture is derived by assuming a magnetospheric configuration (Spencer et al., 2018). Numerically speaking, the hybrid model attains an accuracy of 77% compared to the 40% accuracy that the physics model gives.

1.1 Motivation

Understanding geomagnetic substorms and their impact on Earth's magnetosphere-ionosphere system is crucial for space weather forecasting and mitigating potential hazards to technological infrastructure and human activities in space. Despite significant advancements in observational techniques and theoretical models, there remain several gaps and challenges in accurately predicting substorm occurrences and their associated effects. This study is motivated by the pressing need to advance our understanding of geomagnetic substorms and develop more effective predictive models for space weather forecasting. Geomagnetic substorms can lead to disruptions in satellite communications, power grids, and GPS navigation systems, posing significant risks to modern technological infrastructure. By improving our understanding of substorm dynamics, we aim to enhance space weather forecasting capabilities, ultimately reducing the impact of substorms on critical systems. Substorms represent fundamental processes in the interaction between the solar wind and Earth's magnetosphere, offering valuable insights into plasma physics, magnetospheric dynamics, and space weather phenomena (Bittencourt, 2013). The integration of advanced machine learning techniques with observational data offers new opportunities for predictive modeling of substorm events.

Leveraging large datasets from ground-based magnetometers, space-based satellites, and theoretical models like WINDMI, we can develop more accurate and robust predictive models for substorm onset and evolution. Comparative analysis of different substorm onset lists and machine learning models allows us to assess the strengths and limitations of existing approaches, identifying optimal strategies for substorm prediction, and highlighting areas for improvement. Integrating data-driven machine learning models with physics-based models like WINDMI enables a synergistic approach to substorm prediction, combining insights from observational data with theoretical understanding to improve the accuracy and reliability of substorm forecasts, ultimately enhancing our ability to mitigate the impacts of space weather on society and technology.

Our research aims to enhance our understanding of geomagnetic substorms by integrating advanced data-driven techniques with first-principles physics models. We seek to investigate the dependencies between substorm occurrences and the current state of the magnetosphere-ionosphere system, leveraging insights from both observational data and theoretical models.

1.1.1 Incorporating Comprehensive Data Sources

Our first goal is to incorporate data from the WINDMI model, capturing the intricate dynamics of the magnetosphere, into our machine learning framework. Additionally, we aim to utilize SML and SYMH indices, which characterize the magnetosphere-ionosphere system before substorm occurrences, enhancing the predictive capabilities of our models.

1.1.2 Exploring Parameter Influences

We aim to explore the influence of various solar wind parameters and ground magnetometer observations on substorm occurrences. By systematically analyzing different combinations of these parameters, we seek to identify key factors driving substorm activity.

1.1.3 Exploring Parameter Influences

Our research involves highlighting the performances of four distinct substorm onset lists by Forsyth, Newell, Ohtani and Frey, along with an aggregated list. Through comparative analysis, we aim to understand the strengths and limitations of each list and assess their suitability for predictive modeling.

1.1.4 Comparing Model Parameters

We aim to compare the performances of different machine learning models, not necessarily for the purpose of direct comparison, but to understand their individual strengths and weaknesses. Additionally, we seek to analyze the best and worst-case scenarios to gain insights into model behavior under varying conditions.

1.1.5 Integration with Physics Models

Finally, we aim to compare the performances of our machine learning models with those of the WINDMI physics model. By juxtaposing data-driven and physics-based approaches, we aim to elucidate synergies and discrepancies, ultimately advancing our understanding of substorm dynamics.

1.2 Contributions and Challenges

This thesis contributes to the field by proposing an enhanced predictive model for substorm activity incorporating the identified initiation criteria for more accurate forecasts and two neural network models that can benchmark against the Physics model. Modeling the substorm onsets using IMF measurements from the ACE spacecraft as input into a system of 8 nonlinear ordinary differential equations where the state variables of the differential equations represent the energy stored in the geomagnetic tail, the central plasma sheet, ring current, and field-aligned currents. The output from the model is the current (I_1) which compares with the real-time measure of geomagnetic activities and disturbances (SML) (Spencer et al., 2007).

1.3 Organization of the Thesis

This thesis is comprised of five chapters. In Chapter I, the problem statement, research motivation, and contribution of this research are discussed. Chapter II addresses the background study of the challenges associated with substorm onset determination, the motivation behind the adoption of the combined physics-incorporated neural network, the criteria of the models, and the methodology that is used. Chapter III discusses the WINDMI model and the neural network. Chapter IV contains the performance evaluation of the WINDMI model with the neural network architectures used and has illustrations of all relevant outputs as well as the comparison of those results. Chapter V states the conclusion and discusses the future scope of the work of the thesis.

CHAPTER II

GEOMAGNETIC SUBSTORMS

The near-Earth space environment undergoes profound transformations through the interaction between the solar wind and Earth's magnetosphere. The solar wind, a continuous stream of charged particles emanating from the sun, shapes the dynamics of space around our planet. Composed primarily of protons and electrons, with traces of heavier ions like helium and oxygen, the solar wind carries the Sun's magnetic field, embedded within its magnetized plasma.

As the solar wind encounters Earth's magnetic field, a complex interplay ensues, illustrated in Figure 2.1. The red curved line delineates the bow shock, where the high-speed, low-density solar wind plasma encounters Earth's magnetic field and decelerates, creating a shockwave. Beyond the bow shock lies the magnetosheath, characterized by draped magnetic field lines and convected plasma. The magnetosphere, highlighted in green, forms the region where Earth's magnetic field dominates the solar wind, creating a protective shield around the planet.

Variations in solar wind speed and density, ranging from 300 km/s to 800 km/s and 5-10 cm3 respectively, significantly impact the interaction with Earth's magnetosphere. During heightened solar activity, speeds can surge to several thousand km/s, intensifying the interaction and potentially leading to geomagnetic storms.

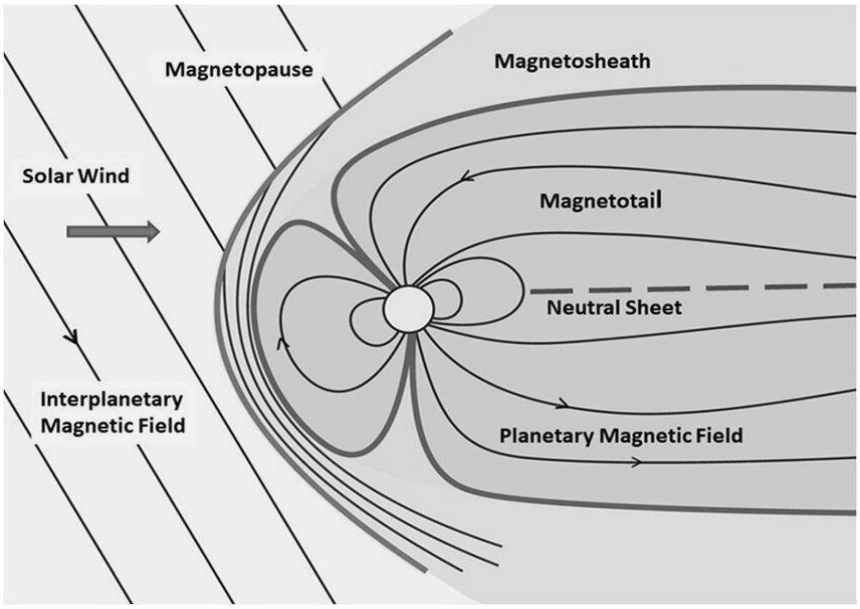

Figure 2.1. Illustration of the interaction between the solar wind and Earth's magnetosphere (Tsurutani et al., 2022).

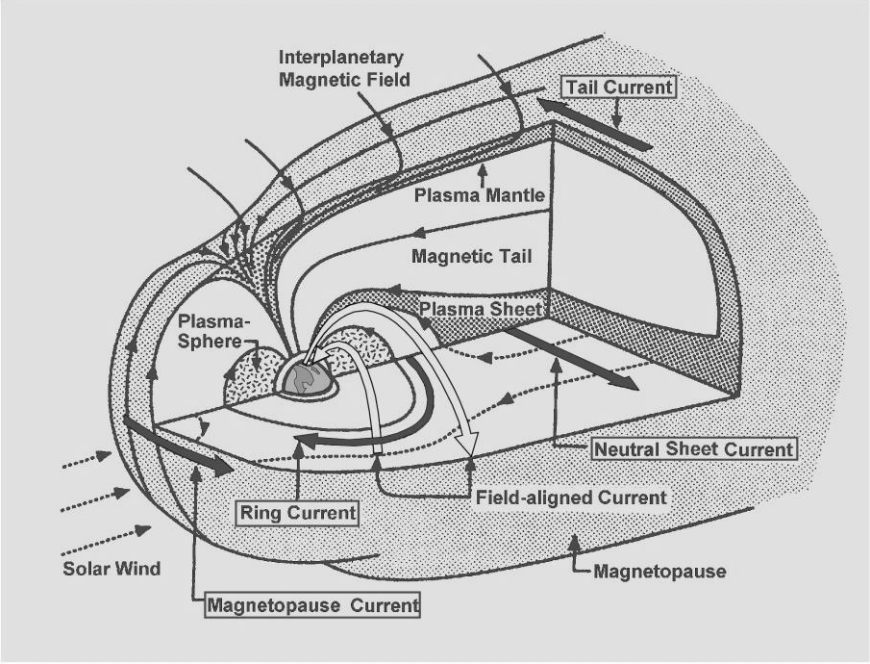

Figure 2.2. The Magnetosphere, a sketch from Kivelson and Russell (1995).

The magnetosphere (Figure 2.2) itself is a complex, dynamic structure with numerous energy reservoirs and transitional layers. The solar wind's electric field drives substantial power through various components of the magnetosphere, with a large fraction dissipated in the ionosphere and the ring current through charge exchange collisions with neutral atoms. Transient power influxes into the magnetosphere during storm times can reach levels up to 1012 W for periods of up to a day. During the main phases of storms most energy is lost from the closed region due to magnetopause erosion. However, most energy is gained through the lobe boundary. During stormy periods, the magnetosphere experiences dynamic energy exchanges:

2.1 External Energy Pathways

As the solar wind conditions change, energy enters the magnetosphere from the magnetosheath (the region between the magnetopause and the bow shock). This energy flows into the magnetospheric lobes and escapes through the closed filed region. The magnetic energy flux in this external pathway contributes to the overall energy dynamics. Previous work and theory have studied this process extensively.

2.2 Reconnection and Magnetic Field Orientation

The internal circulation pathway is controlled via the reconnected X-lines(s). Additionally, the interplanetary magnetic field orientation plays a crucial role in governing this energy flow.

2.3 Energy Dynamic Summary

During the main phases of storms, most energy is lost from the closed region due to magnetopause erosion and most energy is gained through the lobe boundary. The lobes act as an expandable reservoir, adjusting their energy context due to mismatches between incoming and outgoing energy flux. The findings enhance our understanding of how energy moves within Earth's magnetosphere, shedding light on processes with space weather impacts.

The plasma sheet, extending to the distant magnetotail, harbors hot, slow-moving particles and contributes to the overall energy dynamics of the magnetosphere. Energetically charged electrons and ions trapped in the Van Allen radiation belts and their interactions combined with the complex dynamics of the solar wind-magnetosphere coupling have profound implications for space weather and technological systems, highlighting the importance of understanding and monitoring these processes for space exploration and satellite operations. Magnetospheric substorms possess an elemental process involving the storage and explosive release of solar wind energy (Baker et al., 1985).

The magnetopause, magnetospheric boundary separating the solar wind and Earth's magnetic field, represents a critical interface in this interaction. Inside the magnetopause lies the magnetosphere, where Earth's magnetic field dominates, while outside lies the magnetosheath, characterized by mixed solar and terrestrial plasmas. The plasma sheet, extending to the distant magnetotail, harbors hot, slow-moving particles and contributes to the overall energy dynamics of the magnetosphere.

Additionally, energetic charged particles, predominantly protons and electrons, are trapped in the Van Allen radiation belts, executing intricate trajectories along magnetic field lines and drifting around Earth. These interactions, combined with the complex dynamics of the solar wind-magnetosphere coupling, have profound implications for space weather and technological systems, highlighting the importance of understanding and monitoring these processes for space exploration and satellite operations.

The Earth's magnetosphere is a dynamic system constantly interacting with the solar wind, resulting in a variety of phenomena crucial for understanding space weather. Among these phenomena, magnetospheric substorms stand out as elemental processes involving the storage and explosive release of solar wind energy. Substorms play a pivotal role in the dynamics of Earth's magnetosphere and contribute significantly to magnetic storms, making them a focal point of research in space physics and space weather applications.

In recent years, significant advancements have been made possible by modern multi-point space-based and ground-based platforms, providing comprehensive datasets that have revolutionized our understanding of substorms. These datasets have not only facilitated a system-wide perspective but have also propelled advancements in modeling approaches. This introduction reviews recent major advances enabled by these datasets, with a particular focus on two crucial areas: substorm onset timing and evidence for current sheet preconditioning, and fast flows and depolarizations within the magnetotail.

Substorm onset, marked by rapid brightening, breakup, and poleward expansion of the aurora, signifies the explosive development of dissipative processes starting from

localized parts of the plasma sheet. Understanding the sequence of key processes leading to substorm onset has been a long-standing and critical question in near-Earth space plasma physics. Although auroral observations have traditionally been considered the most accurate means of timing and locating onset signatures, recent advancements in observational capabilities have allowed for deeper insights into the underlying physics.

2.4 Geomagnetic Substorm Onset Phases

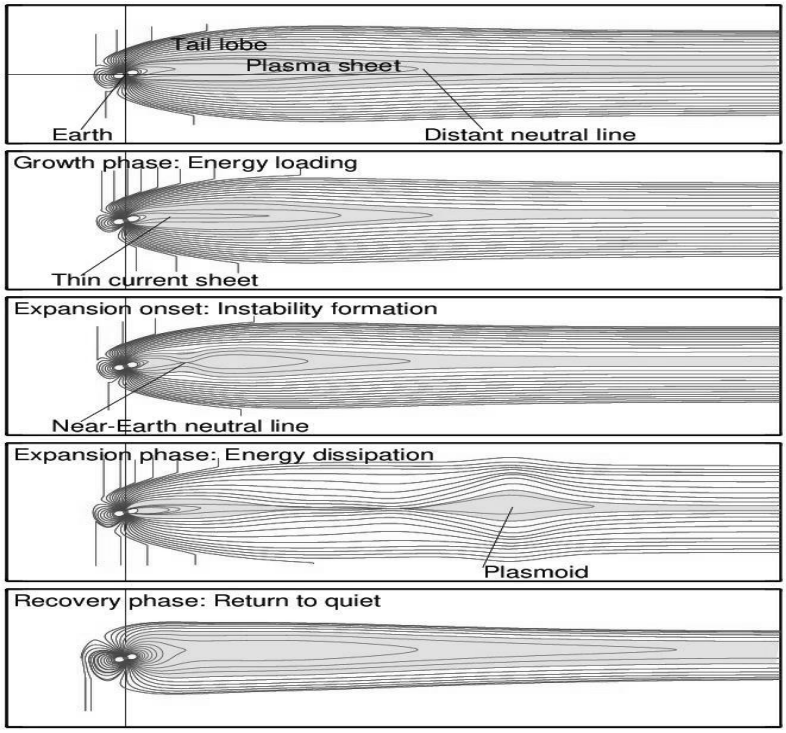

Figure 2.3. A schematic of the magnetospheric substorm. (a) Illustrates the onset of dayside reconnection (b) Growth phase: energy is loaded into the magnetotail which leads to the formation of a large-scale thin current sheet. (c) Expansion onset: this marks the instability formation and is followed by (d) Expansion phase: is where the energy dissipation occurs characterizing the ejection of the plasmoid at the tail end lastly (e) Recovery phase: this is where the substorm subsides and reconfigures itself to the quiet phase (Pulkkinen, 2007).

Geomagnetic substorms represent crucial mechanisms in the solar wind-magnetosphere interaction, facilitating the transfer, storage, and release of energy. Typically, a substorm comprises several events, including changes in aurora brightness, geomagnetic disturbances in polar regions, disruptions to the magnetotail current sheet, and injections of energetic particles. The phases of a substorm (Figure 2.3) are as follows:

2.4.1 Substorm Growth Phase

The growth phase initiates the substorm cycle, representing the stage where the magnetosphere accumulates and stores significant energy from the solar wind. Enhanced magnetospheric convection facilitates the penetration of solar wind energy into the magnetosphere, primarily concentrating on the nightside. During this phase, the stored energy gradually builds up, setting the stage for the subsequent explosive release phase. This phase is crucial for priming the magnetosphere and preparing it for the dynamic changes that will follow.

2.4.2 Substorm Onset and Expansion Phase

The expansion phase marks the climax of the substorm event, characterized by the sudden and explosive release of the stored energy within the magnetosphere. This phase is associated with intense activity, including disruptions to the magnetotail current sheet and the formation of field-aligned currents. These disruptions lead to enhancements in auroral brightness, particularly in polar regions, and manifest as geomagnetic disturbances across various locations. The injection of energetic particles into the polar ionosphere further intensifies the auroral displays, highlighting the dynamic and complex nature of the magnetospheric response during this phase.

2.4.3 Substorm Recovery Phase

Following the peak of activity during the expansion phase, the magnetosphere enters the recovery phase, also known as the quiet state. During this phase, the magnetosphere undergoes a process of reconfiguration, gradually returning to a stable and steady state. Disruptions such as enhanced auroras and geomagnetic disturbances weaken and eventually dissipate as the magnetosphere stabilizes. The recovery phase represents a period of restoration and normalization, where the magnetosphere adjusts back to pre-substorm conditions, completing the cycle of substorm activity.

Substorm triggering represents a complex and dynamic process involving the sudden release of stored magnetic energy within the Earth's magnetosphere. Accompanied by various phenomena such as the expansion and intensification of auroras, the formation of plasmoids, and the generation of energetic particles, substorm onset marks the beginning of an active phase within the magnetosphere.

The exact mechanisms triggering substorm onset remain a subject of active research, reflecting the intricate interplay between solar wind dynamics and magnetospheric responses. Despite significant progress in understanding substorms, the precise triggers for the release of stored magnetic energy and the associated energy transfer processes are still not fully understood. This uncertainty underscores the complexity of magnetospheric dynamics and the need for ongoing studies to refine existing models and explore new theories.

During the expansion phase following substorm onset, the released energy propagates through the magnetosphere, inducing changes in magnetic field configuration and plasma dynamics. These changes contribute to the dynamic evolution of the

magnetospheric system, leading to enhancements in auroral activity and geomagnetic disturbances.

Continued research efforts aim to enhance our understanding of the background and dynamics of substorm triggering, with the goal of improving our ability to predict and mitigate the impacts of space weather on technological systems and satellite communication. By unraveling the mechanisms driving substorm onset, scientists can develop more accurate models and forecasting tools, enabling better preparedness for space weather events and their potential impacts on Earth's technological infrastructure.

CHAPTER III

SUBSTORM ONSET DETECTION TECHNIQUES

Geomagnetic substorms are fundamental phenomena in space weather, characterized by sudden and intense releases of energy in Earth's magnetosphere-ionosphere system. These events play a significant role in shaping the dynamics of the near-Earth space environment, influencing auroral activity, magnetic disturbances, and plasma dynamics (Akasofu, 2013). Understanding the onset of substorms is crucial for space weather forecasting and mitigating potential impacts on technological infrastructure.

Scientists have developed various techniques to identify and characterize the onset of geomagnetic substorms. These techniques leverage observational data from ground-based magnetometers, satellite missions, and theoretical models to detect and analyze the precursor signals preceding substorm events. Each technique offers unique advantages and limitations, contributing to our comprehensive understanding of substorm onset dynamics.

In this chapter, we explore different substorm onset techniques, highlighting their principles, methodologies, and applications in space weather research. By examining the strengths and limitations of each technique, we aim to provide insights into the diverse approaches utilized in the study of geomagnetic substorms. Additionally, we will discuss

recent advancements and emerging trends in substorm onset detection, paving the way for future developments in space weather forecasting and predictive modeling.

3.1 Frey et al. (2004) and Mende et al. (2009) onset detection technique

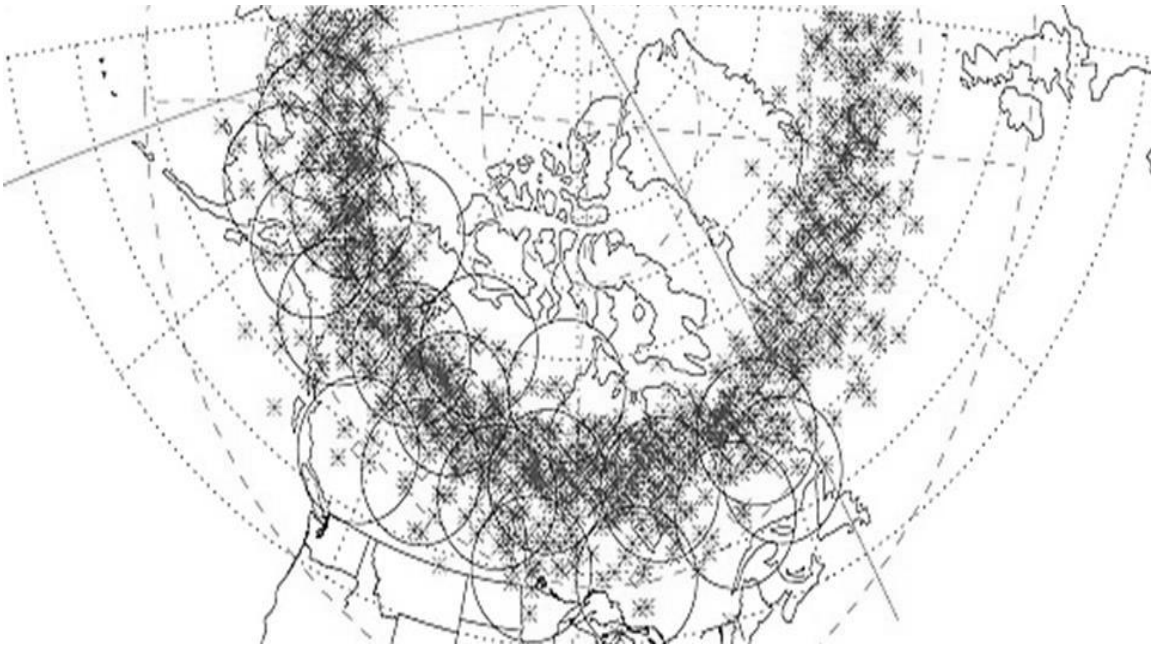

Figure 3.1. Map of North America showing the substorm onset locations in geographic coordinates regardless of the local time of onset. A geomagnetic grid is given in blue. The presently planned locations of THEMIS GBO (Pu et al., 2010) are indicated with the fields of view of their all-sky cameras. Ninety-eight percent of the substorms started within 600 km of the closest planned GBO. Each substorm onset location is given with a red asterisk. Those onsets outside of the field of view of the THEMIS all-sky cameras are marked with green asterisks (Frey et al., 2004).

Frey et al. (2004) devised a method for detecting substorm onsets using data from the Far Ultraviolet (FUV) instrument onboard the IMAGE spacecraft (Mende et al., 2009). Substorm onsets have been observed by IMAGE-FUV. For 2.5 years, they

identified more than 2400 substorm onsets in the Northern Hemisphere, aiming to publish this list to facilitate further investigation into auroral substorm phenomena.

Their detection technique involved analyzing FUV data, primarily utilizing Wideband Imaging Camera (WIC) images for their superior spatial resolution. In instances where WIC data was not optimal, they relied on SI-13 images. Substorms were identified based on specific criteria: (1) a discernible local brightening of the aurora, (2) expansion of the aurora to the poleward boundary and azimuthal spreading for at least 20 minutes, and (3) ensuring at least 30 minutes had passed since the previous onset to avoid closely spaced events.

The resulting list of substorm onsets includes detailed information such as date, time, instrument used for identification (WIC or SI-13), spacecraft geocentric distance, and characteristics of the brightest pixel within the onset surge. This dataset allows researchers to search for onsets based on various criteria, such as high magnetic latitude, late local time, proximity to specific ground stations, or closeness to the IMAGE spacecraft for improved spatial resolution (Kistler et al., 2002).

The analysis of the substorm onset locations (Figure 3.1) confirms previous findings regarding their average distributions in geomagnetic latitude and local time. Importantly, the dataset exhibits no bias towards specific geomagnetic longitude locations, as expected for observations from a satellite in a non-locked orbit.

The publication of this substorm onset list serves to facilitate further research in the field, inviting other researchers to utilize the data for their investigations. Additionally, the list was used to assess the probability of substorm onset observations

for the THEMIS Ground-Based Observatories (GBOs) (Mende et al., 2009), which are expected to capture a significant portion of substorms originating over North America.

Overall, Frey et al. (2004) work contributes valuable insights into the characteristics and distribution of auroral substorm onsets, providing a comprehensive dataset for future studies in magnetospheric and auroral physics.

3.2 Newell and Gjerloev (2011) Onset Detection Technique

Newell and Gjerloev (2011) developed a novel substorm onset detection technique utilizing magnetometer chains in collaboration with SuperMAG, a global network of magnetometers, to derive a generalized auroral electrojet index called SME (Substorm Onset Electrojet). This index represents the integrated nightside auroral power, primarily stemming from the diffuse aurora, with a remarkable correlation to the total nightside auroral power, reaching an impressive correlation coefficient of $r = 0.86$.

Their method involves utilizing data from more than 100 magnetometer sites, providing a more comprehensive view compared to the traditional AE (12) index, which only uses data from 12 stations. They demonstrated that SME, as well as its approximation AE, can predict a substantial portion of the variance in nightside auroral power, indicating a clear geophysical meaning for these indices.

The study also focused on improving the detection of substorm onsets. By using SME, which has a larger number of contributing stations, they found that their technique significantly improved the identification of substorm onsets compared to traditional methods. Specifically, their performance analysis showed that their technique identified about 30-40% more substorm onsets per year.

Furthermore, they showed that the onset times detected using SME were much closer to the epochs inferred from global auroral images, with a median time difference of about 4 minutes compared to 8 minutes using traditional indices. This suggests a higher accuracy and temporal resolution in detecting substorm onsets.

Newell and Gjerloev (2011) compiled a comprehensive list of substorms between January 1, 1997, and December 31, 2002, using their detection technique. This list, included as auxiliary material, provides detailed information about the detected substorms during that time interval. They conducted superposed epoch analyses using DMSP data to characterize the behavior of isolated and recurrent substorms, revealing distinct differences in their effects on auroral power.

Overall, Newell and Gjerloev (2011) substorm onset detection technique utilizing SME represents a significant advancement in understanding and characterizing substorm activity, offering improved accuracy and comprehensiveness compared to traditional methods.

The Substorm Onset Electrojet (SML) index, devised by Newell and Gjerloev (2011) exhibits robust correlations with both Polar UVI (Ultraviolet Imager) observations and DMSP (Defense Meteorological Satellite Program) data regarding substorm signatures. Comparisons between SML-identified onset times and those observed by Polar UVI reveal that SML tends to identify substorm onsets slightly after Polar UVI observations, with a significantly shorter median delay compared to traditional AL (12) indices. Furthermore, cumulative fraction analyses indicate that SML is substantially more likely to recognize substorm onsets within the first few minutes compared to AL (12). Superposed epoch analyses using DMSP data demonstrate significant increases in

both diffuse and broadband electron precipitation following SML-identified substorm onsets, aligning with expected substorm characteristics. Distinctions between isolated and recurrent substorms reveal that while isolated substorms exhibit a lower initial auroral power but a higher percentage increase, both isolated and recurrent substorms show comparable absolute changes in auroral power. These findings underscore the utility of the SML index in capturing geomagnetic disturbances associated with substorms and elucidating their subsequent auroral activity (Newell et al., 2010).

3.3 Forsyth et al. (2015) Onset Detection Technique

Forsyth et al. (2015) introduce a novel quantitative technique, termed Substorm Onsets and Phases from Indices of the Electrojet (SOPHIE), designed to identify the times and durations of substorm expansion, recovery, and possible growth phases based on percentiles of the rate of change of auroral electrojet indices. This approach allows for the determination of substorm phases using user-defined percentile values, enabling the identification of smaller or larger variations in auroral index data or any ground-based magnetometer time series.

SOPHIE leverages the SuperMAG AL (SML) index data, although it is adaptable to other auroral zone magnetic indices or ground magnetometer time series. Substorms are identified in a nonparametric manner based on the exceedance of a percentile threshold in the rate of change of SML. Negative changes in SML beyond a specified percentile level indicate expansion phases, while positive changes denote recovery phases. Additionally, SOPHIE considers possible growth phases between expansion and recovery phases, allowing for a comprehensive characterization of substorm dynamics.

The technique is applied in a three-stage process, involving low-pass filtering of data, calculation of percentiles of the rate of change of SML, and iterative modification of thresholds to ensure an equal number of expansion and recovery phases are identified. SOPHIE's effectiveness is demonstrated through comparisons with previous lists of substorm onsets, showing a high degree of agreement and providing valuable insights into substorm dynamics over a significant period. The supporting information includes detailed start times of possible growth phases, expansion phases, and recovery phases, along with flags indicating enhanced convection events.

Overall, SOPHIE represents an automated, quantitative approach that improves upon traditional visual identification methods and fixed threshold techniques for substorm onset detection.

3.4 Ohtani et al. (2020) Onset Detection Technique

Ohtani and Gjerloev (2020) present a technique for identifying the onset of isolated substorms utilizing the SuperMAG AL (SML) index. Unlike previous methods, which aimed at capturing all substorm onsets, this technique focuses specifically on isolating individual substorms with high confidence.

Drawing inspiration from prior work by Newell & Gjerloev (2011), the authors refine their approach to overcome limitations and enhance accuracy. The methodology unfolds in three distinct steps, as depicted in Figure 3.2. In the initial step, the SML index undergoes smoothing via a 9-minute-wide boxcar filter to minimize short-duration variations, laying the groundwork for subsequent analysis.

Step 2 comprises the identification of potential onsets based on six rigorous criteria, ensuring the isolation and gradual progression of substorm events. These criteria encompass aspects such as the magnitude and duration of SML variations, the slope of the growth and expansion phases, and the overall intensity and duration of the substorm. Each criterion is meticulously implemented to capture the distinct signatures associated with substorm onset while minimizing false positives.

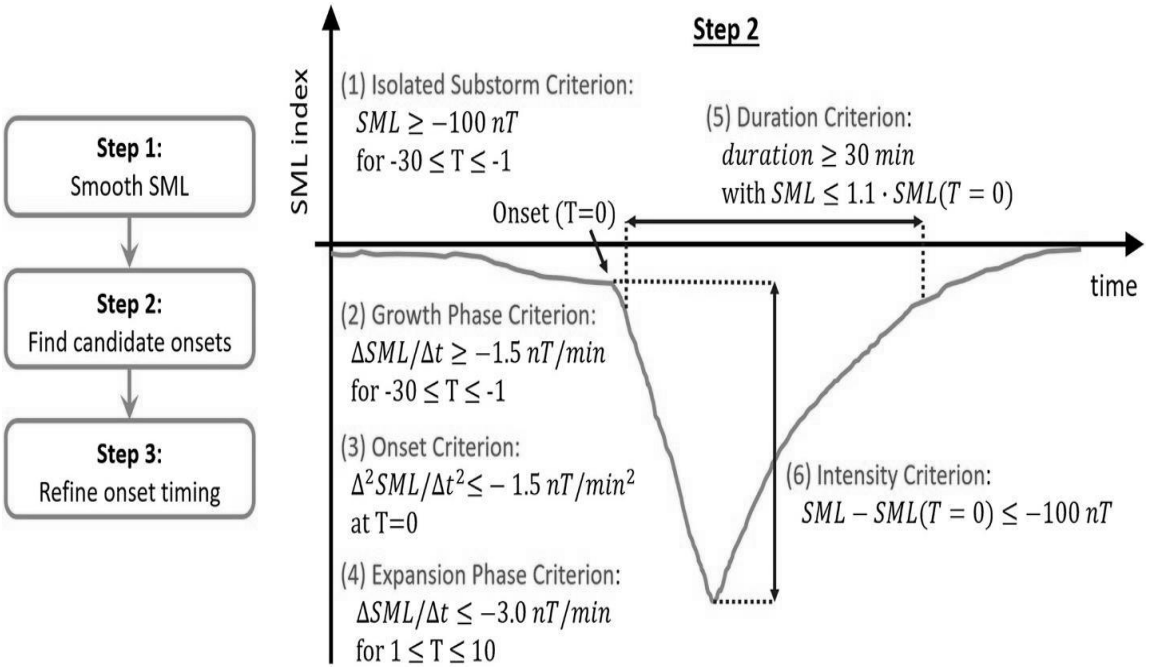

Figure 3.2. Substorm onsets criterion by Ohtani and Gjerloev (2020).

Step 3 further refines the onset selection process by considering the station location and the characteristics of the SML-contributing stations, particularly focusing on their placement within the midnight sector (20 MLT to 30 MLT). This ensures that the identified onset corresponds to a genuine substorm event and not a spurious artifact. The

technique's robustness is underscored by its validation through independent visual inspections by each author, affirming its ability to consistently identify isolated substorms with a high level of confidence.

Notably, the technique yields a typical annual count of 500–700 substorms, reflecting its selective nature and emphasis on clarity and precision. The resulting substorm list, containing detailed onset information, is made readily accessible through the supporting information and the SuperMAG website, ensuring its widespread availability for future research endeavors and data analysis.

3.5 Methodology to Create an Aggregated List for Substorm Onsets

To assess the consistency of WINDMI model outputs with substorm onset lists generated by Frey, Liou, Newell, Forsyth, and Ohtani, the triggering times of the WINDMI model were compared with a merged substorm onset list from these authors. A 15-minute window was considered, and if any substorm onset coincided with the WINDMI model triggering within this window, it was deemed a positive result. The SML index, an improvement over the AL index, was utilized to measure the near-Earth magnetic field in the polar region. The baseline elimination technique suggested by Gjerloev involved a three-step process for determining the baseline of a given station and component, addressing slowly varying offsets mainly attributed to the Earth's main field and diurnal components largely associated with the solar quiet current system. Substorm onset data from the SuperMAG website, obtained through methods devised by Frey, Liou, Newell, Forsyth, and Ohtani, were merged, averaging onset times occurring within

15 minutes of each other to consider them a single event. This resulted in over 15,000 substorm onsets over a decade.

Figure 3.3 demonstrates the merging of three substorm onsets by different methods into one. The SML data for January 7, 2000, from 09:00 to 10:30 UT, shows the index initially slightly negative around 09:00 UT, gradually decreasing starting around 09:20 UT, indicating the growth phase. Around 09:30 UT, the SML value started rapidly declining for about 30 minutes, indicating the expansion phase. After 10:00 UT, the SML index began to recover, signifying the substorm's recovery period. Substorm onset, typically at the beginning of the expansion phase, is detected differently by each method. For example, Newell's technique identifies an onset whenever the index shows a sharp change in value <-45 nT/sec, placing it just before the sharpest decline, while Ohtani's method positions the onset around the curve's inflection point, relying on a knee-like signature in the SML curve. Forsyth's approach, utilizing changes in percentage threshold, often detects substorm onsets before other methods.

In terms of data selection, eight years of data were taken (2000-2007) from the four authors as well as the aggregated list for substorm onsets. The selection criteria were determined by the following criteria:

- Frey’s list (2000-2005)
- Newell’s list (1970-2021)
- Forsyth’s list (1970-2021)
- Ohtani’s list (1970-2021)
- Aggregated list (2000-2007)

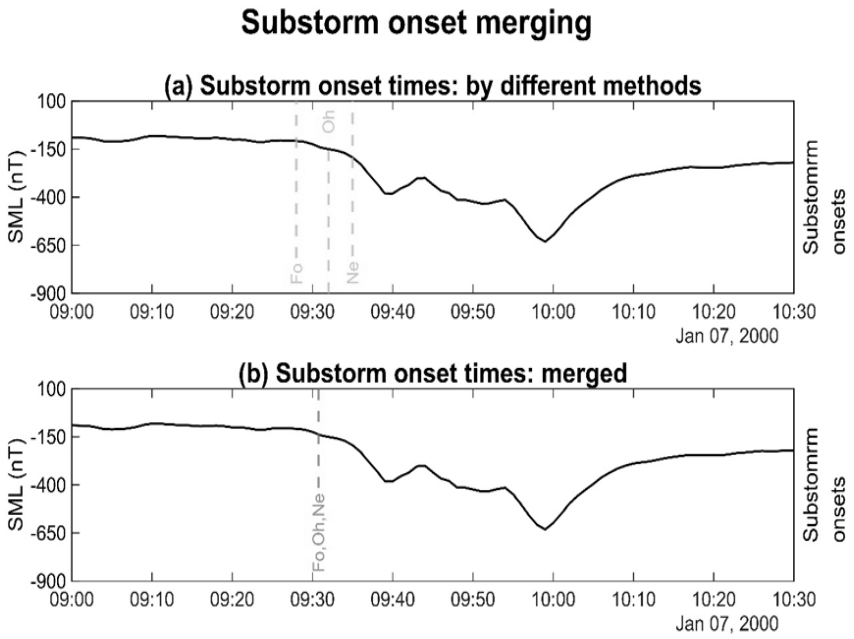

Figure 3.3. Illustration depicting the merging of substorm onset times obtained from different substorm lists sourced from the SuperMAG website. The figure showcases 1.5 hours, displaying the SML index in nT. Panel (a) exhibits three dotted yellow vertical lines, situated within approximately 5 minutes of each other, corresponding to substorm onset times identified by Forsyth, Ohtani, and Newell (from left to right). Panel (b) demonstrates the merged substorm onset as a single dotted vertical line positioned at the average of the three times depicted in Panel (a).

Since Frey et al. (2004) provided a list from 2000 to 2005, it was regarded as the pre-cursor for determining the duration for all the other substorm lists including the aggregated list, even though some of the lists date back to 1970s (Newell, Forsyth and Ohtani).

Negative control data incorporation was key in the Neural Network training, it refers to the sample or data points that do not belong to the target class. It is essential to incorporate negative control data in neural network training for the following reasons:

1. **Specificity Assessment:** By including negatives, we can assess the specificity of the model. A good model should find true positives when they are genuinely

present (sensitivity), but it should also correctly identify negatives (true negatives or true negative rate).

2. **Balanced Training:** Class-balanced data ensures that the model is exposed to an equal number of examples from each class (substorms and no-substorms). This prevents bias toward any specific

3. **Robustness:** Exposure to negative control data helps the model learn to handle cases where the target class is absent.

While negative control data is essential in experimental design, negative samples play a distinct role in neural network training. Both contribute to robust and accurate models, albeit in different contexts.

CHAPTER IV

WINDMI AND NEURAL NETWORK MODELS

In this chapter, we delve into the description and analysis of two essential components: the WINDMI model and two distinct neural network architectures, Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN).

The WINDMI provides a comprehensive framework for simulating the dynamics of the Earth's magnetosphere-ionosphere system. Developed through rigorous theoretical formulations and empirical data assimilation, the WINDMI model offers insights into the complex interplay between solar wind parameters and geomagnetic activity. By capturing the fundamental processes governing magnetospheric dynamics, the WINDMI model serves as a valuable tool for understanding substorms and their evolution.

Complementing the physics-based approach of the WINDMI model, neural network architectures offer powerful data-driven methodologies for predictive modeling and pattern recognition. Long Short-Term Memory (LSTM) networks, known for their ability to capture temporal dependencies and long-range dependencies in sequential data, have emerged as effective tools for time series forecasting in space weather research. Their recurrent nature allows for the retention of information over extended periods, making them well-suited for analyzing dynamic processes such as substorm onset.

In contrast, Convolutional Neural Networks (CNNs) excel in spatial pattern recognition tasks, leveraging hierarchical feature extraction to capture complex spatial relationships in input data. In the context of space weather, CNNs have demonstrated proficiency in analyzing satellite imagery, magnetic field distributions, and other spatially distributed datasets. Their ability to automatically learn spatial features makes CNNs valuable for identifying spatial patterns associated with substorm events. Since a 16 by 5 input sequence is trained by the neural network, the input is an image data incorporated for training. This approach is also comparable to the work contrasted in Maimaiti et al. (2019).

In this chapter, we provide a detailed exploration of the underlying principles, architectures, and applications of the WINDMI model, LSTM networks, and CNNs in the context of space weather research. Through comparative analysis and case studies, we aim to elucidate the strengths and limitations of each approach, paving the way for integrated modeling frameworks and advanced predictive capabilities in the study of geomagnetic substorms and other space weather phenomena.

4.1 WINDMI Model

The WINDMI model is a low-order nonlinear model of the nightside magnetosphere that can simulate substorm events based on solar wind input. The paper discusses the curated equations and parameters that have been derived by careful examination and research of energy transfer and dissipation between the magnetotail, plasma sheet, ring current and field-aligned currents (Horton & Doxas, 1996), (Horton & Doxas, 1998), (Spencer et al., 2007).

The model outputs the auroral electrojet and the ring current that are compared to the AL and Dst indices, respectively. The motive of this thesis is to utilize a comprehensive substorm list along with pseudo-breakup events identified by auroral brightening signatures from ground-based imagers. The model parameters are precisely tuned to match the onset times and the AL and SML indices as closely as possible. The WINDMI model predicts the occurrence of a substorm within a particular framework which is known as the energy loading and unloading paradigm (Kallio et al., 2000). In contrast with the recent work on substorm predictions presented by Maimaiti et al. (2019), this research suggests some possible improvements and future work for the WINDMI model (Spencer et al., 2007).

The quantities L, C, Σ, L_1, C_1, and Σ_1 represent the inductances, capacitances, and conductance of the magnetospheric and ionospheric components. These parameters play a crucial role in determining the behavior of the system. A_{eff} represents an effective aperture for particle injection into the ring current. The ring current is a flow of charged particles encircling the Earth in the magnetosphere.

The resistances in the partial ring current and region-2 current (I_2) regions are denoted as R_{prc} and R_{A2}, respectively whereas L_2 represents the inductance of the region-2 current. The Heat Flux Limiting Parameter is denoted by coefficient u_0 in equation (7).

The parameters t_E and t_{rc} refer to the confinement times for the central plasma sheet and ring current. The transition region magnetic field is given by B_{tr} and the effective and the width of the magnetosphere is L_y equation (3). The pressure gradient driven current is given by $I_{ps} = L_x(p/\mu_0)^{1/2}$, and L_x is the effective length of the magnetotail. The pressure unloading function $\Theta(u) = 1/2\ [1+ \tanh(u)]$ where $u = \frac{(I - I_c)}{\Delta I}$ in

equation (7) is given by an integral parameter known as critical current I_c and the interval ΔI for the transition to loss of plasma along newly opened magnetic field lines. This unloading function changes from zero to unity based on the current I compared to I_c (Figure 4.2) (Spencer et al., 2009, 2019). The unloading function follows from current gradient driven tearing modes or cross-field current instabilities, as described by Yoon and Lui (2004).

The primary objective of this thesis is to incorporate the substorm list and pseudo-breakup events that have been cross-referenced with the WINDMI model output to build the PINN model. The proposed neural network aims to enhance the precision and accuracy of predicting onset times as obtained by the WINDMI model in Figure 4.3.

The WINDMI model equations are given as follows:

$$L\frac{dI}{dt} = \boldsymbol{V_{SW}}(t) - V + M\frac{dI_1}{dt} \quad (1)$$

$$C\frac{dV}{dt} = I - I_1 - I_{ps} - \Sigma V \quad (2)$$

$$\frac{3}{2}\frac{dp}{dt} = \frac{\Sigma V^2}{\Omega_{cps}} - \mu_0 p K_{\parallel}^{\frac{1}{2}}\Theta(u) - \frac{pVA_{eff}}{\Omega_{cps}B_{tr}L_y} - \frac{3p}{2\tau_E} \quad (3)$$

$$\frac{dK_{\parallel}}{dt} = I_{ps}V - \frac{K_{\parallel}}{\tau_{\parallel}} \quad (4)$$

$$L_1\frac{dI_1}{dt} = V - V_1 + M\frac{dI}{dt} \quad (5)$$

$$C_1\frac{dV_1}{dt} = I_1 - I_2 - \Sigma_1 V_1 \quad (6)$$

$$L_2\frac{dI_2}{dt} = V_1 - (R_{prc} + R_{A2})I_2 \quad (7)$$

$$\frac{dW_{rc}}{dt} = R_{prc}I_2^2 + \frac{pVA_{eff}}{B_{tr}L_y} - \frac{W_{rc}}{\tau_{rc}} \quad (8)$$

Figure 4.1. The WINDMI model describes the physical parameters of the magnetosphere-ionosphere system as defined by the 8 differential equations Spencer et al. (2007).

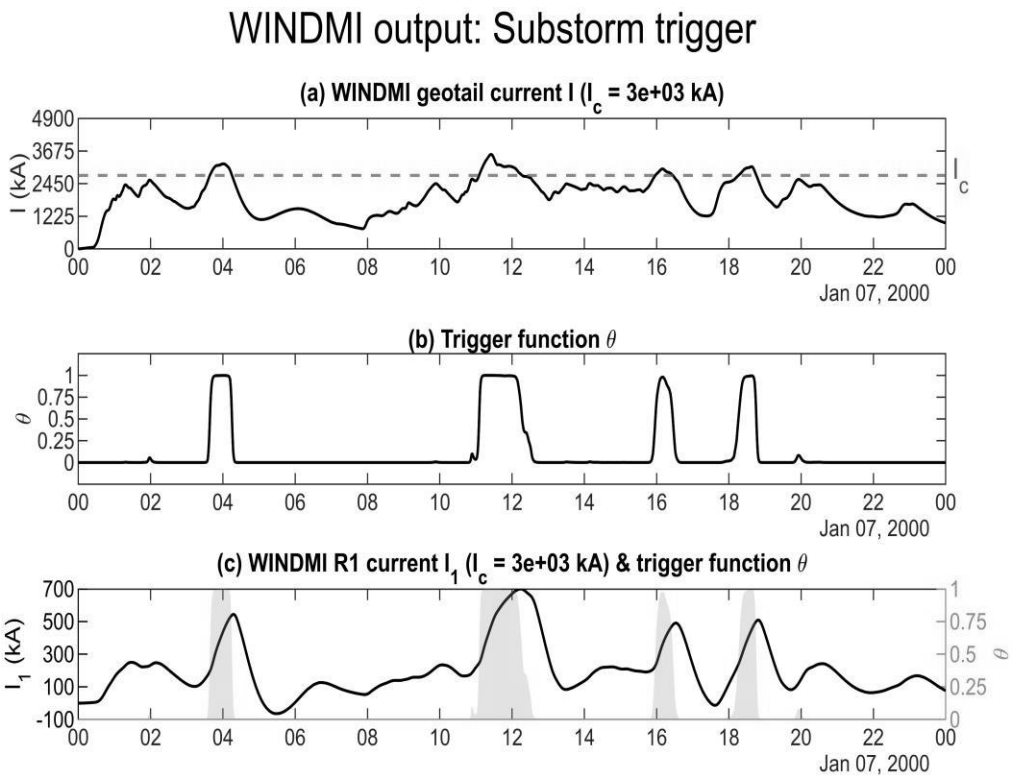

Figure 4.2. WINDMI model outputs with I (Geo-tail current), I_1 (Field-aligned current), θ (Trigger function). The figure demonstrates the geo-tail (I) exceeds a threshold (I_c, blue dashed line) when the trigger function θ is activated. I_1 shows enhancement.

Methods	WINDMI (detect)	WINDMI (no detect)	Total
Frey/Liou	896 (39%)	1416(61%)	2312
Newell	2109 (45%)	2558(55%)	4667
Forsyth	2230 (40%)	3369(60%)	5599
Ohtani	917 (36%)	1650(64%)	2567

Figure 4.3. WINDMI onset detection results. WINDMI performs 40% across all substorm lists. WINDMI model detected Frey/Liou's list 39%, Newell's list 45%, Forsyth's list 40% and Ohtani's list 36% (Adhya et al., 2024).

4.2 LSTM Architecture

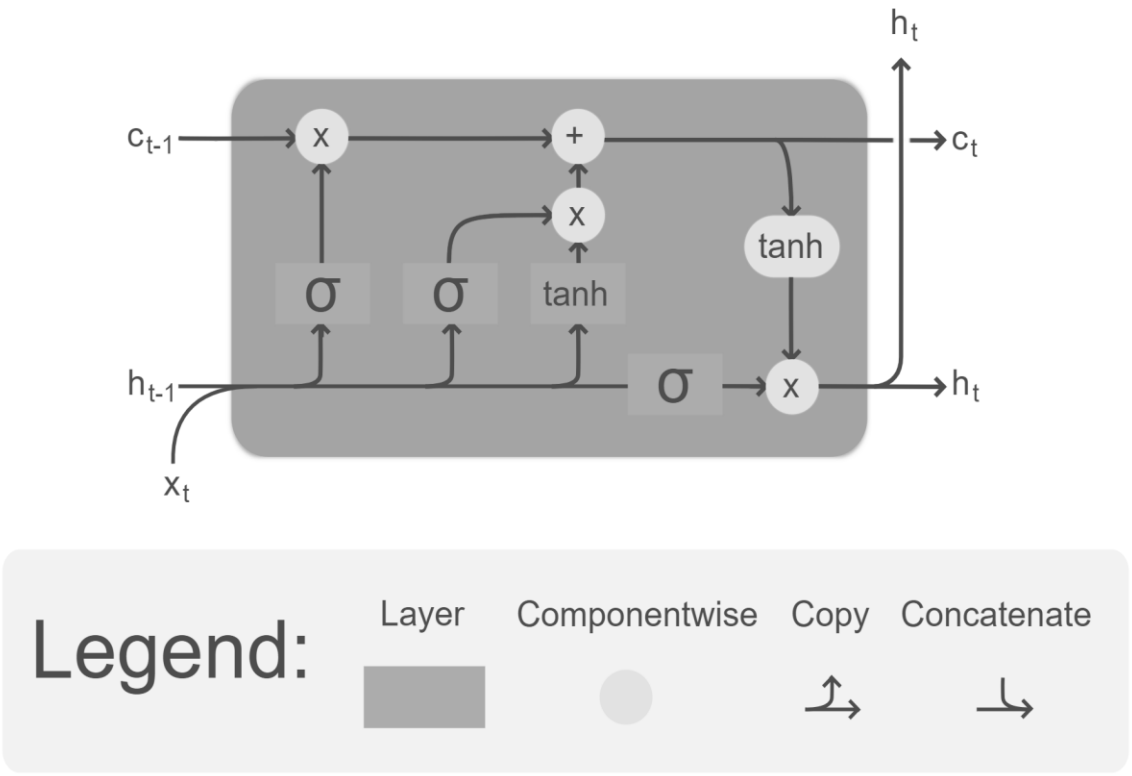

Figure 4.4. The Long Short-Term Memory (LSTM) cell, depicted here, can sequentially process data, and retain its hidden state across multiple time steps.

The Long Short-Term Memory (LSTM) network, a type of recurrent neural network (RNN), addresses the vanishing gradient problem inherent in traditional RNNs, offering a solution that is relatively insensitive to gap length. This resilience renders it advantageous over other RNN variants, hidden Markov models, and sequence learning methods. The LSTM architecture aims to equip RNNs with a short-term memory capable of persisting over extended time intervals, hence the moniker "long short-term memory." Widely applicable across various domains, including handwriting, speech recognition, machine translation, speech activity detection, robot control, video games, and

healthcare, LSTM networks excel in classifying, processing, and predicting data based on time series. Comprising a cell (Figure 4.4) along with input, output, and forget gates, the LSTM unit orchestrates the flow of information within the network. The cell retains values across arbitrary time spans, while the gates regulate the inflow and outflow of information.

Forget gates determine the relevance of past information relative to current inputs, deciding what to retain or discard. Similarly, input gates select which new information to integrate into the current state. Output gates control the release of information from the current state, considering both previous and current states. This selective processing enables the LSTM network to preserve vital long-term dependencies, facilitating accurate predictions across current and future time steps.

A flowchart illustrating the processing pipeline for the LSTM Network is given in Figure 4.5. The pipeline commences with the sequence input layer, which assumes responsibility for feeding all the input parameters, consisting of 16 layers simultaneously. Subsequently, these input parameters traverse through the LSTM layer, which facilitates the modeling of temporal dependencies and sequential patterns within the data.

The architecture incorporates additional layers aimed at enhancing the network's performance. Batch normalization is employed to standardize and accelerate the training process by normalizing the input activations, thereby reducing internal covariate shift, and improving convergence. Following batch normalization, a fully connected layer is utilized to integrate information from the LSTM layer and prepare it for the final classification step. This layer serves as a bridge between the LSTM output and the subsequent classification layer. Before reaching the final classification stage, where the

network determines whether the output corresponds to a substorm or not, the non-linear activation function SoftMax is applied. SoftMax normalizes the output scores across different classes, converting them into probabilities. This normalization process streamlines the output for classification purposes, facilitating the decision-making process.

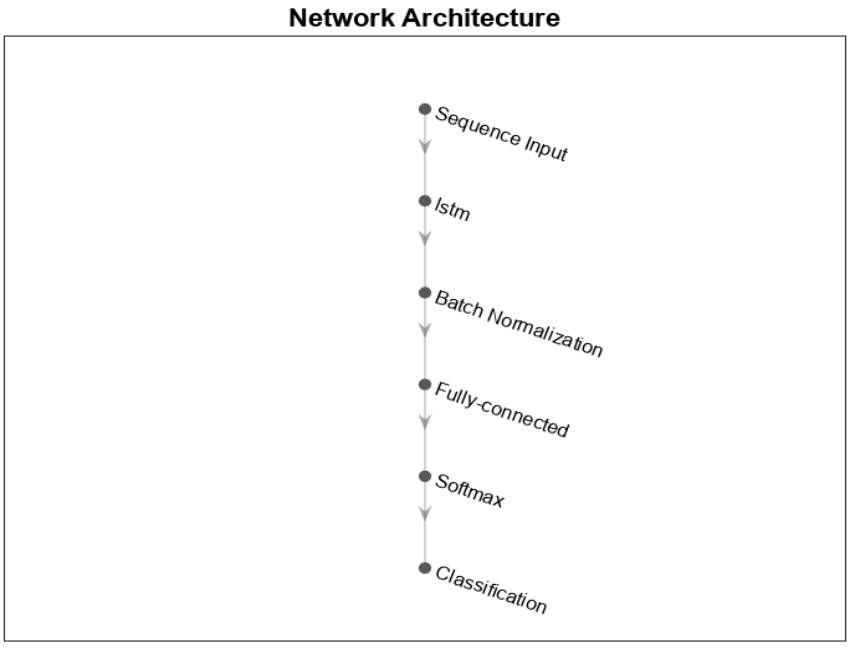

Figure 4.5. A flowchart of processing pipeline for the LSTM Network, sequence input is the layer responsible for feeding all the input parameters (16 layers at once) and then passes it through the LSTM layer. The architecture also has the following layers, batch normalization, fully connected layer. Before final classification where it classifies whether the output is a substorm or not a substorm, the non-linear activation function SoftMax is adopted to streamline the output.

In summary, Figure 4.5 provides an overview of the processing pipeline for the LSTM Network, delineating the sequence of operations involved in temporal modeling, feature integration and classification within the network architecture.

4.3 CNN Architecture

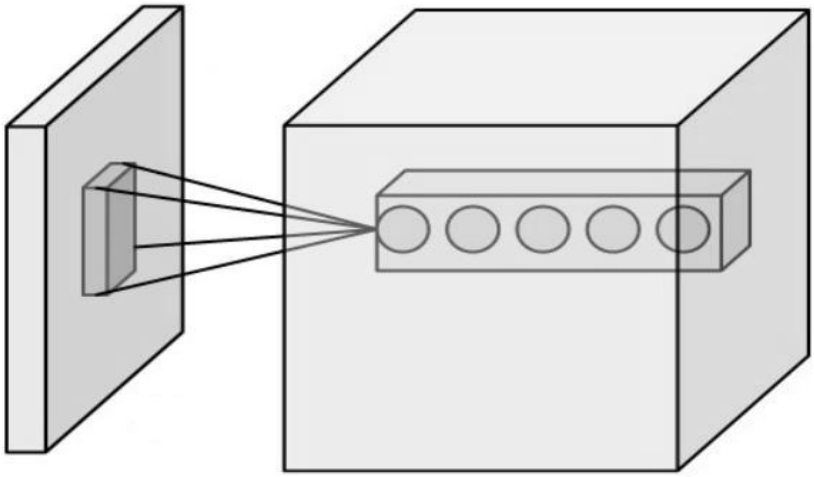

Figure 4.6. Neurons depicted in blue belong to a convolutional layer, each connected to their respective receptive field highlighted in red.

A Convolutional Neural Network (CNN) is a specialized form of feed-forward neural network designed for effective feature extraction from input data, particularly in tasks involving images and spatial data. CNNs automatically learn feature representations through the optimization of filters or kernels applied to input data, mitigating issues like vanishing, and exploding gradients seen in earlier neural network architectures. Unlike traditional fully connected networks, where each neuron requires a vast number of weights to process high-dimensional inputs like images, CNNs leverage cascaded convolutional kernels to dramatically reduce the number of connections needed per neuron, enabling efficient processing of input tiles Figure 4.6. This architecture comprises several layers, with the convolutional layer being the fundamental building block. In the convolutional layer, learnable filters with small receptive fields are applied

across the input volume to generate activation maps, capturing spatial features. Stacking these activation maps forms the output volume, where each entry corresponds to the output of a neuron analyzing a specific region in the input. Through this process, CNNs can effectively learn hierarchical representations of input data, extracting higher-level features from wider context windows as the network progresses through deeper layers. Overall, CNNs excel in tasks such as image classification, object detection, and image segmentation, owing to their ability to learn complex feature hierarchies directly from raw data.

A flowchart illustrating the processing pipeline for the Convolution 2D Neural Network is presented in Figure 4.7. The pipeline begins with the sequence input layer, responsible for feeding all the input parameters, consisting of 16 layers simultaneously. These input parameters are then passed through the Convolution 2D layer, initiating the feature extraction process.

The architecture further incorporates additional layers to enhance the network's performance. These layers include batch normalization, which aids in stabilizing and accelerating the training process by normalizing the input activations. The linear activation function Rectified Linear Unit (RELU) is applied to introduce non-linearity and increase the model's capacity to learn complex patterns.

Following the RELU activation, the Flatten layer is employed to transform the multi-dimensional feature maps into a one-dimensional vector, facilitating the transition from convolutional layers to fully connected layers. Subsequently, a Long Short-Term Memory (LSTM) layer is incorporated, enabling the network to capture temporal dependencies and sequential patterns within the data. A fully connected layer follows the

LSTM layer, serving as a bridge between the LSTM output and the final classification layer. This layer integrates information from the previous layers and prepares it for the final classification step.

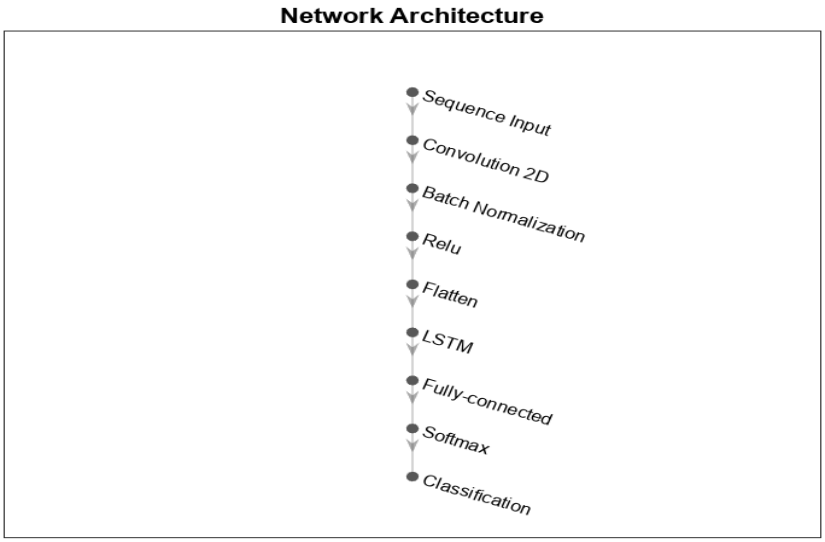

Figure 4.7. A flowchart of processing pipeline for the Convolution 2D Neural Network. The sequence input is the layer responsible for feeding all the input parameters (16 layers at once) and then passes it through the Convolution 2D layer. The architecture also has the following layers, batch normalization, linear activation function RELU, Flatten layer, a LSTM layer, fully connected layer. Before final classification where it classifies whether the output is a substorm or not a substorm, the non-linear activation function SoftMax is adopted to streamline the output. The input to the 2D-Convolution is represented as a sequence. The data can be further processed as a 1D Convolution or temporal convolution network.

Before reaching the final classification stage, where the network determines whether the output corresponds to a substorm or not, the non-linear activation function SoftMax is applied. SoftMax normalizes the output scores across different classes, converting them into probabilities, thereby streamlining the output for classification purposes. In summary, Figure 4.7 provides an overview of the processing pipeline for the

Convolution 2D Neural Network, detailing the sequence of operations involved in feature extraction, temporal modeling, and classification within the network architecture.

4.4 Data processing and Model Training

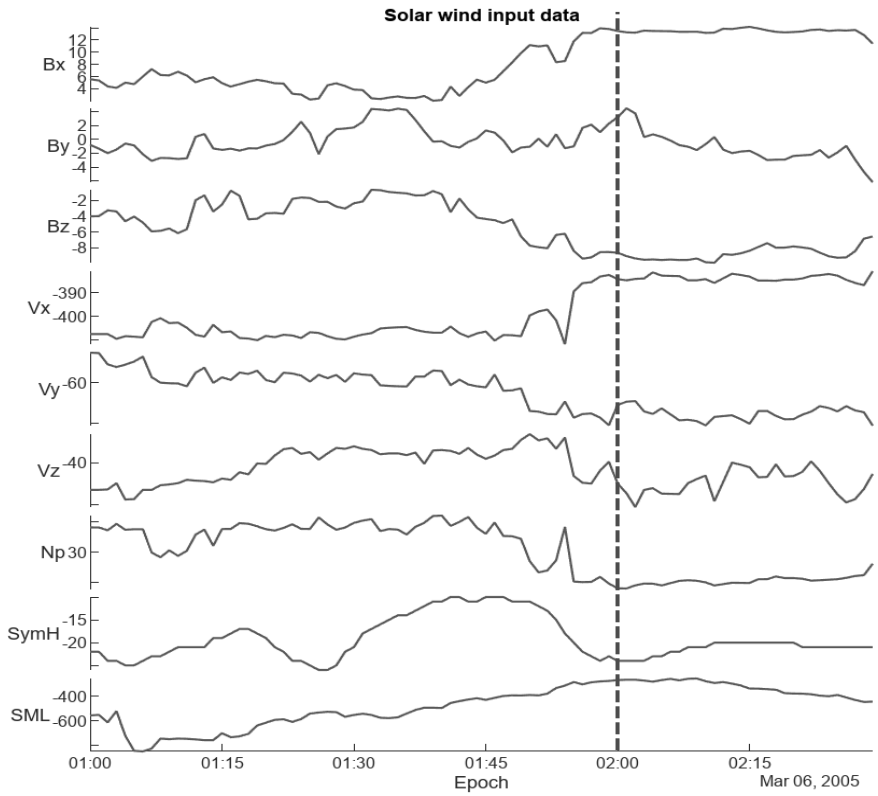

Figure 4.8. An instance of the Solar Wind input data showing 9 parameters (case 13) segmented into 60 minutes training and 30 minutes prediction window.

Figure 4.8 illustrates an instance of data when the parameters B_x, B_y, B_z, V_x, V_y, V_z, N_p, SML, and SYMH (case 13) were provided as input to the neural network model. The plot depicts the duration for 1.5 hours, starting at 01:00 and ending at 2:30 on March 06, 2005. These parameters represent solar wind and ground parameters. The data is segmented into two parts: the first 60 minutes, from 01:00 to 02:00, constitute the input

batch sent to the model, while the subsequent 30 minutes, from 02:00 to 02:30, represent the prediction period. Notably, this data instance captures a substorm onset occurring within the prediction period, highlighting the model's ability to predict such events based on the provided input parameters.

In each case, the dataset is partitioned into segments, with each segment spanning (number of parameters) x 60 minutes. The number of rows in each segment varies based on the input provided. Within each segment, the columns represent individual time steps. However, segments containing five or more missing values for each parameter, or a total of 5 x (number of features) missing values across all parameters, are rejected from the dataset. This rigorous filtering process ensures data quality and integrity. Additionally, due to the inherently imbalanced nature of the dataset, where there are more instances of no substorm onsets compared to substorm onsets, some segments labeled as "no substorms" are also excluded. This curation methodology aims to maintain a high standard of data quality and relevance, contributing to the robustness and reliability of subsequent analyses or model training processes. Given an instance of the data, comprising 1.5 hours (60 minutes of pre-substorm window for model training and 30 minutes labels. The sampling rate used is 4 min resolution.

The number of rows/features varies based on, alongside under sampling is based on criteria. The number of hidden units used is 120. The number of samples used is ~18,000 based on the number of substorms that occurred over the eight years' timeframe between 2000-2007 and a 70/30 split is used for validation. The pre-processing and cleaning of the data were performed in Excel as well as MATLAB. The data went through data partitioning as well as class balancing (an instance of equal number of

“substorm” and “no substorms” samples) before it was provided to the model. The criteria for “no substorm” selection by the model was randomized. Missing values were addressed with scaling and normalization. If there were more than 5 missing data points per input parameter sequence sent to the model, then that entire batch was discarded.

The provided code outlines the process of training different neural network architectures for a given task. It begins by defining several network architectures using various types of layers, such as LSTM, and convolutional layers, each aimed at capturing different aspects of the input data to potentially enhance model performance. Parameters like the number of hidden units, classes, and features are initialized based on the input data characteristics. The configurations of these layers are saved for future reference. The desired network architecture and training options are then selected, and a folder is created to store the training results. Training options, including the optimization algorithm (Adam), maximum epochs, initial learning rate, and verbosity, are specified, along with options for plotting training progress. Subsequently, the selected network architecture is trained using the specified options, with training data used to iteratively update the network parameters over multiple epochs, while validation data are employed to evaluate model performance and prevent overfitting. Overall, the provided code demonstrates a systematic approach to configuring, training, and evaluating neural network models, enabling experimentation with different architectures and training parameters to optimize performance.

Figure 4.9 presents a mind-map illustrating the 16 combinations cases adopted and sent into the two neural networks as sequence inputs of the time series data. Each case comprises either a singular or a multiple combination of parameters, including

ground parameters such as SML and/or SYMH, satellite parameters such as magnetic field components (V_x, V_y, V_z, B_x, B_y, B_z), and physics parameters (I_1 and Θ).

The mind-map visually represents the different combinations of parameters used as input for the neural networks, showcasing the complexity and diversity of input configurations explored in the study. It provides a comprehensive overview of the variations in input parameters investigated, offering insights into the impact of different parameter combinations on the models' performance and predictive capabilities.

Overall, Figure 4.9 serves as a valuable reference for understanding the range of input configurations examined in the study and their relevance to substorm onset prediction.

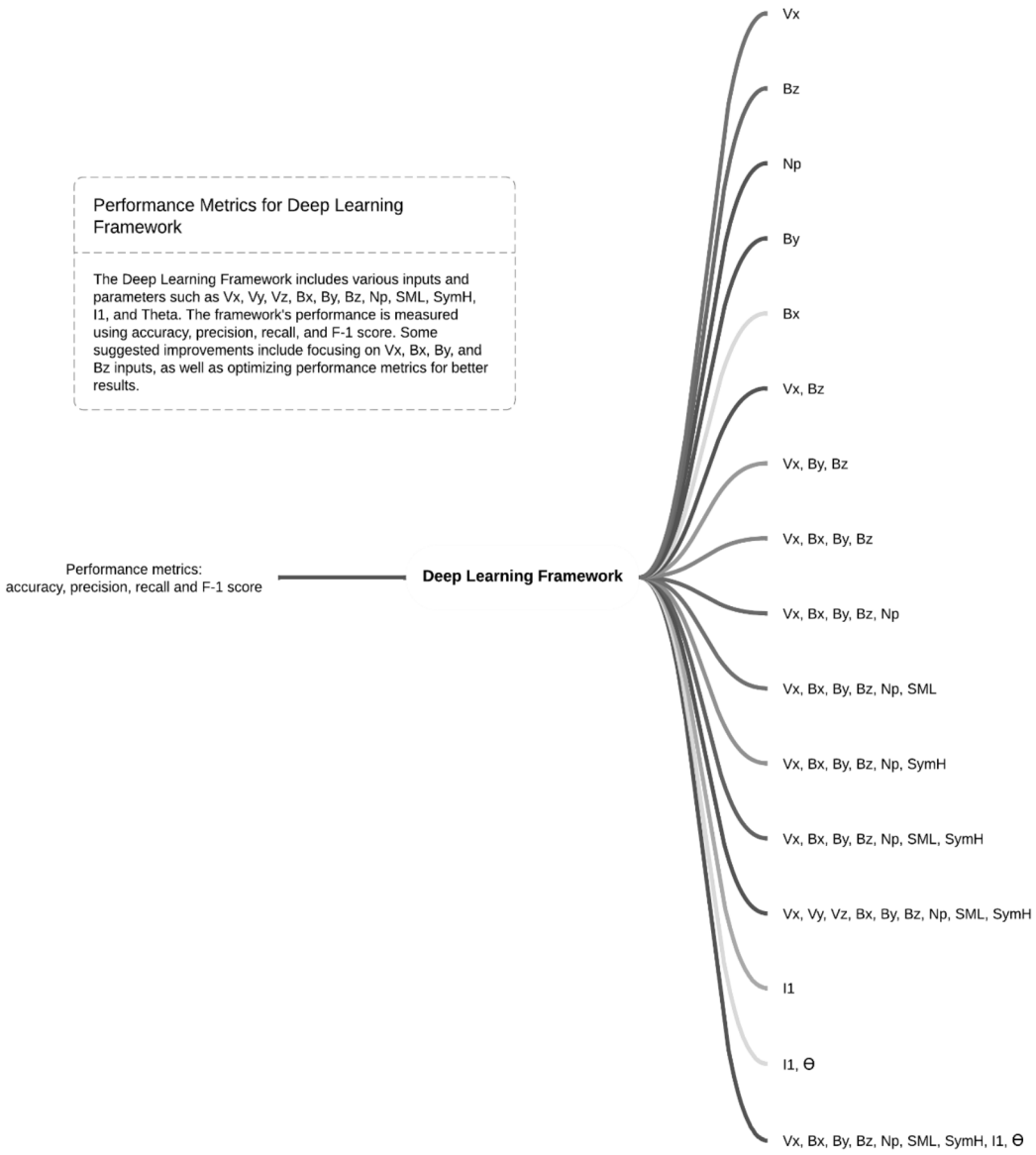

Figure 4.9. A mind-map illustrating the 16 combinations cases adopted and sent into the two neural networks as sequence inputs of the time series data. Each case consists of a singular or a multiple combination of the parameters and they are ground parameters such as SML and/or SYMH, satellite parameters such as magnetic field components V_x, V_y, V_z, B_x, B_y, B_z, N_p and physics parameters I_1 and Θ.

To gauge the performance of the LSTM and CNN neural networks, the following metrics were employed: Accuracy, Precision, Recall, and the F1 Score. These metrics serve as fundamental measures in evaluating the efficacy of classification models.

Each metric provides unique insights into the models' predictive capabilities, offering valuable information for assessing their strengths and weaknesses.

During testing, the trained neural network model classifies the test data to produce predicted labels. These labels are then compared with the true labels to evaluate the model's performance using various accuracy metrics. Accuracy is a measure of the overall correctness of the model. It is the ratio of correctly predicted instances to the total instances. While accuracy is a simple and intuitive metric, it might not be sufficient when dealing with imbalanced datasets, where one class significantly outnumbers the other. The overall accuracy indicates the proportion of correctly classified instances.

Precision measures the ratio of correctly classified positive instances to the total instances predicted as positive, while recall calculates the ratio of correctly classified positive instances to all true positive instances. Precision is important when the cost of false positives is high. In medical diagnoses, you want to be certain that a positive prediction is accurate, meaning we want our False Positives to be as low as possible, giving us a low value for precision.

Recall, also known as Sensitivity or True Positive Rate, measures the ability of the model to capture all the relevant instances. It is the ratio of correctly predicted positive observations to all observations in actual class. Recall is crucial when the cost of false negatives is high. In spam email detection, for e.g. we want to avoid missing any

actual spam, we want False Negatives to have a desired high value, which renders a high value for recall.

The F1 Score combines precision and recall providing a balanced assessment of the model's performance in terms of both false positives and false negatives, also known as the harmonic mean of precision and recall. F1 Score is especially useful when there is an uneven class distribution or when both false positives and false negatives are equally important. These metrics collectively offer insights into the model's ability to accurately classify substorm onset events without referencing specific code implementations.

Here are the formulae for the metrics:

$$accuracy\ (\%) = \frac{correct\ predictions}{all\ predictions} \times 100\% \quad (9)$$

$$precision = \frac{TP}{TP+FP} \quad (10)$$

$$recall = \frac{TP}{TP+FN} \quad (11)$$

$$F_1\ score = \frac{2 \times precision \times recall}{precision+recall} \quad (12)$$

1. **Accuracy (%):** Accuracy (Equation 9) serves as a measure of the overall correctness of predictions made by a model. It is computed by dividing the number of correct predictions by the total number of predictions made and then multiplying the result by 100%.

2. **Precision:** Precision (Equation 10) evaluates the model's capacity to accurately identify positive instances among all instances predicted as positive. It is determined by dividing the number of true positives (TP) by the sum of true positives and false positives (FP).

3. **Recall:** Recall (Equation 11), also referred to as sensitivity or true positive rate, gauges the proportion of actual positive instances correctly identified by the model. It is

calculated by dividing the number of true positives (TP) by the sum of true positives and false negatives (FN).

4. **F1 Score:** The F1 score (Equation 12) represents the harmonic mean of precision and recall, providing a single metric that balances both aspects. It proves particularly useful when dealing with imbalanced datasets.

These metrics collectively offer a comprehensive evaluation of the classification model's performance, facilitating informed decision-making regarding model selection and optimization within the thesis framework.

In our study, we integrated data from the SuperMAG substorm lists compiled by Frey, Newell, Forsyth, and Ohtani to generate a prioritized substorm list, providing comprehensive insights into substorm occurrences. This prioritized list forms the basis for our analysis of various parameter cases, encompassing different combinations of input parameters for a specific study conducted over a time range from the year 2000 to the end of 2007. The parameter cases include:

1. **Single Driver Parameters:** These cases involve individual driver parameters, including velocity components (V_x), magnetic field components (B_x, B_y, B_z), and plasma number density (N_p).

2. **Driver Parameters (Combinations):** Combinations of different driver parameters are explored to understand their combined effects, such as V_x with B_z, or V_x with B_y and B_z.

3. **Driver Parameters + Ground Parameters (Single):** Single ground parameters, namely SML (Substorm Onset Electrojet) and SYMH, are added to the driver parameters to investigate their influence on the study.

4. **Driver Parameters + Ground Parameters (Combination):** Like the previous set, but with a combination of both SML and SYMH added to the driver parameters to explore their joint effects.

5. **V_y or V_z Exploration:** A specific case is dedicated to exploring the potential effects of V_y and V_z (velocity components in the y and z directions, respectively) in addition to the other parameters considered.

6. **WINDMI Parameters:** Cases involving parameters specific to WINDMI, such as I_1 (the index I_1) and Θ.

7. **Mixed Parameters:** Combinations of driver parameters, ground parameters, and WINDMI parameters are considered in these cases, providing a comprehensive exploration of their combined effects on the study.

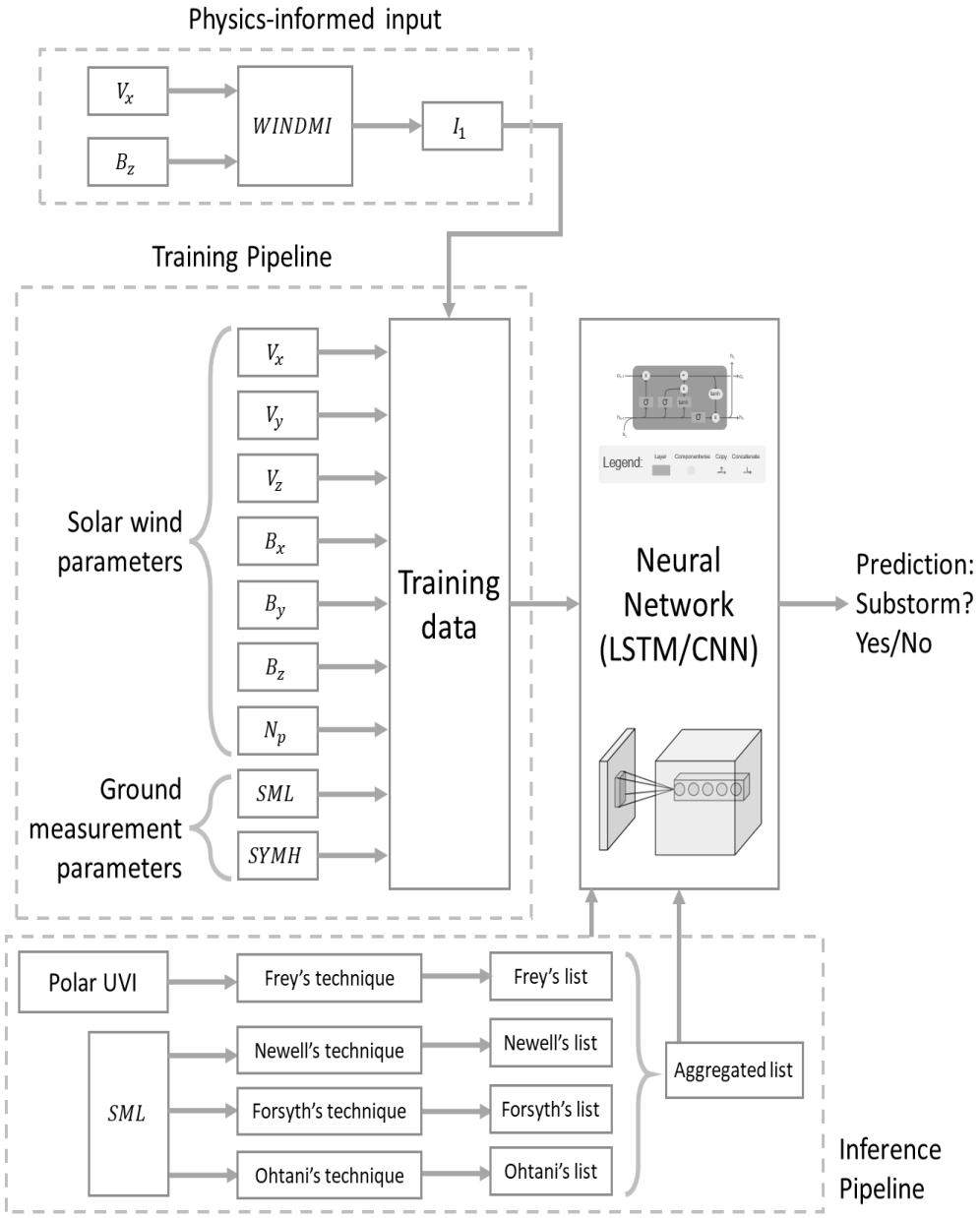

Figure 4.10. Description of Parameters and Physics-Incorporated Neural Network Model observational data and theoretical outputs from the WINDMI model, providing a comprehensive framework for substorm prediction and analysis. The inference pipeline encapsulated SML parameters which is our ground truth. The substorm lists are published volumes of work which serve as the basis for reference.

Figure 4.10 illustrates the Description of Parameters and Physics-Incorporated Neural Network Model. It encompasses observational data and theoretical outputs from the WINDMI model, thereby offering a holistic framework for substorm prediction and analysis. This integrated approach leverages both observational data and theoretical insights to enhance the predictive capabilities of the neural network model. The main instigators are solar wind parameters and forming a novel neural network architecture including these parameters for substorm onset prediction is the foundation of this thesis (Johnson et al., 2022). By incorporating physics-based principles into the model architecture, it facilitates a more comprehensive understanding of substorm dynamics and improves prediction accuracy. Overall, Figure 4.10 serves as a key component in advancing substorm prediction methodologies by integrating empirical observations with theoretical foundations.

The data undergoes preprocessing, cleaning, and time-shifting using MATLAB to ensure consistency and relevance for neural network training. This model integrates both the physics-incorporated neural network model and leverages the capabilities of deep learning, employing one optimizer, such as Adam, to characterize training loss functions and optimize model performance. By integrating observational data with outputs from the WINDMI model, our model offers a holistic approach to substorm prediction, combining empirical observations with theoretical insights into magnetospheric dynamics. Through rigorous training and optimization, we aim to develop a robust predictive model capable of accurately forecasting substorm onset, thereby advancing our understanding of space weather phenomena and improving forecasting capabilities for mitigating potential impacts on technological infrastructure and human activities in space.

In our study, we employed confusion matrices and relevant evaluation alongside data analysis methods to establish the relationship between two proposed deep learning architectures, namely Long Short-Term Memory (LSTM) and 2D Convolutional Neural Network (CNN), and the WINDMI model. The primary goal of our research was to train these deep learning neural networks to serve as benchmarks for improving the WINDMI model, which defines the underlying physics principles to describe geomagnetic activity.

4.5 Performance Analysis

The performance analysis of substorm onset prediction involved evaluating the performance of the LSTM and CNN neural networks through various tests and metrics. Different combinations of ground parameters and drivers from the WINDMI model were incorporated into the neural networks, totaling sixteen combinations. These combinations encompassed a wide range of parameters such as SML, SYMH, magnetic field components, plasma density, ion velocity, and WINDMI model parameters I_1, and Θ. Each combination was tested to determine the model's performance in predicting substorm onset.

Performance metrics including accuracy, precision, recall, and F1-score were generated to assess the validity and reliability of the trained deep learning models. These metrics provided insights into how well the models performed in terms of correctly identifying substorm onsets and distinguishing them from non-substorm periods. Accuracy measured the overall correctness of the predictions, precision indicated the proportion of correctly predicted substorm onsets among all predicted onsets, recall measured the proportion of actual substorm onsets that were correctly predicted, and F1-

score was the harmonic mean of precision and recall, providing a balanced assessment of the model's performance.

Furthermore, the Deep Learning Framework incorporated various inputs and parameters such as velocity components (V_x, V_y, V_z), magnetic field components (B_x, B_y, B_z), plasma density (N_p), and ground parameters like SML, SYMH, I_1, and Θ. These parameters were utilized to train and test the neural networks, and their performance was evaluated using the performance evaluation metrics.

Additionally, a mind-map illustrated the 16 combinations of cases adopted and sent into the two neural networks as sequence inputs of the time series data. This visualization helped in understanding the different parameter combinations tested and their impact on the model's performance.

Moreover, the confusion matrix served as a crucial tool in the data analysis methodology, allowing for the extraction of patterns within the data, particularly the aggregated substorm list. By analyzing the confusion matrix, researchers gained insights into the model's ability to correctly predict substorm onsets and classify them accurately. Overall, the performance analysis provided valuable insights into the effectiveness of the LSTM and CNN neural networks in predicting substorm onset and contributed to further improvements in the WINDMI model.

CHAPTER V

MODEL PREDICTIONS AND RESULTS

This section presents the predictions and results obtained from various models applied to our dataset. We begin by outlining the training progress of a model for a specific case (one out of the 16 combination cases used for NN), providing insights into its learning dynamics and performance metrics. Subsequently, we examine the outcomes of two Long Short-Term Memory (LSTM) cases, followed by an analysis of two Convolutional Neural Network (CNN) cases.

In some instances, our models encountered challenges in learning, resulting in a class bias phenomenon. This bias was particularly evident when the models struggled to discern patterns associated with specific classes, leading to imbalanced predictions.

After discussing individual cases, we present aggregated statistics summarizing the performance across all 80 evaluated cases. These statistics offer an overview of the models' efficacy in capturing underlying patterns within the dataset.

Moreover, to visually represent the model outputs, we present side-by-side heatmaps comparing the predictions of the LSTM and CNN models. Additionally, heatmaps illustrating predictions based on five distinct types of lists are presented in parallel, enabling a comparative analysis of their performance. Accompanying these

visualizations are accuracy percentages represented as fractions, providing insights into the predictive capabilities of each model variant and list type.

Overall, this section provides a detailed exploration of model predictions and results, shedding light on the performance of different models and their ability to capture complex patterns within the dataset.

5.1 Model Training and Convergence

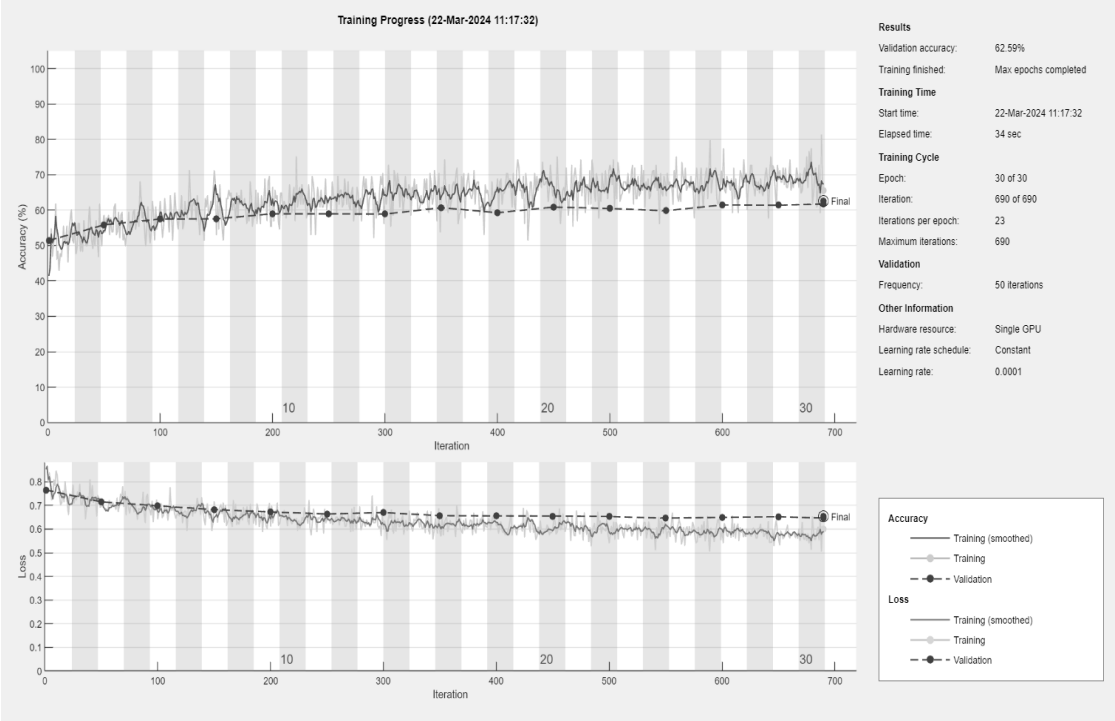

Figure 5.1. Training Progress for one case using the LSTM Model. The blue line represents the accuracy (%) in the top panel, while the black dashed line illustrates the validation accuracy (%) during training. In the bottom panel, depicting the loss per iteration, the red line indicates the training loss, and the black dashed line shows the validation loss as fractions. The plot demonstrates a convergence trend observed after approximately 20 epochs.

Figure 5.1 depicts the training progress for a machine learning model using MATLAB. It provides a detailed overview of the model's performance metrics across epochs, facilitating an understanding of its learning dynamics.

In the top panel of the figure, the blue line represents the accuracy of the model over epochs, expressed as a percentage. This metric indicates the proportion of correctly predicted outcomes relative to the total number of instances in the training set. As training progresses, the blue line reflects fluctuations in accuracy, providing insights into the model's ability to learn from the data.

The black dashed line in the top panel illustrates the validation accuracy, also represented as a percentage. This metric assesses the model's performance on a separate validation dataset, ensuring that the model generalizes well to unseen data. Monitoring the validation accuracy helps prevent overfitting, where the model memorizes the training data but fails to generalize to new instances.

In the bottom panel of the figure, the red line depicts the training loss per iteration. The training loss quantifies the error between the model's predictions and the actual target values during training. A decreasing trend in the training loss indicates that the model is converging towards an optimal solution.

Additionally, the black dashed line in the bottom panel represents the validation loss, expressed as fractions. Like the training loss, the validation loss measures the error on the validation dataset. Monitoring both training and validation losses helps assess the model's performance and identify potential issues such as underfitting or overfitting.

Overall, the figure provides a comprehensive visualization of the model's training progress, offering insights into its accuracy, validation performance, and convergence

behavior across epochs. This information is crucial for optimizing model parameters, improving performance, and ensuring robustness in real-world applications.

5.2 LSTM Cases

Following the processes outlined in the methodology section, two cases of resulting matrices are represented in the subsequent section after running the LSTM model for the aggregated list.

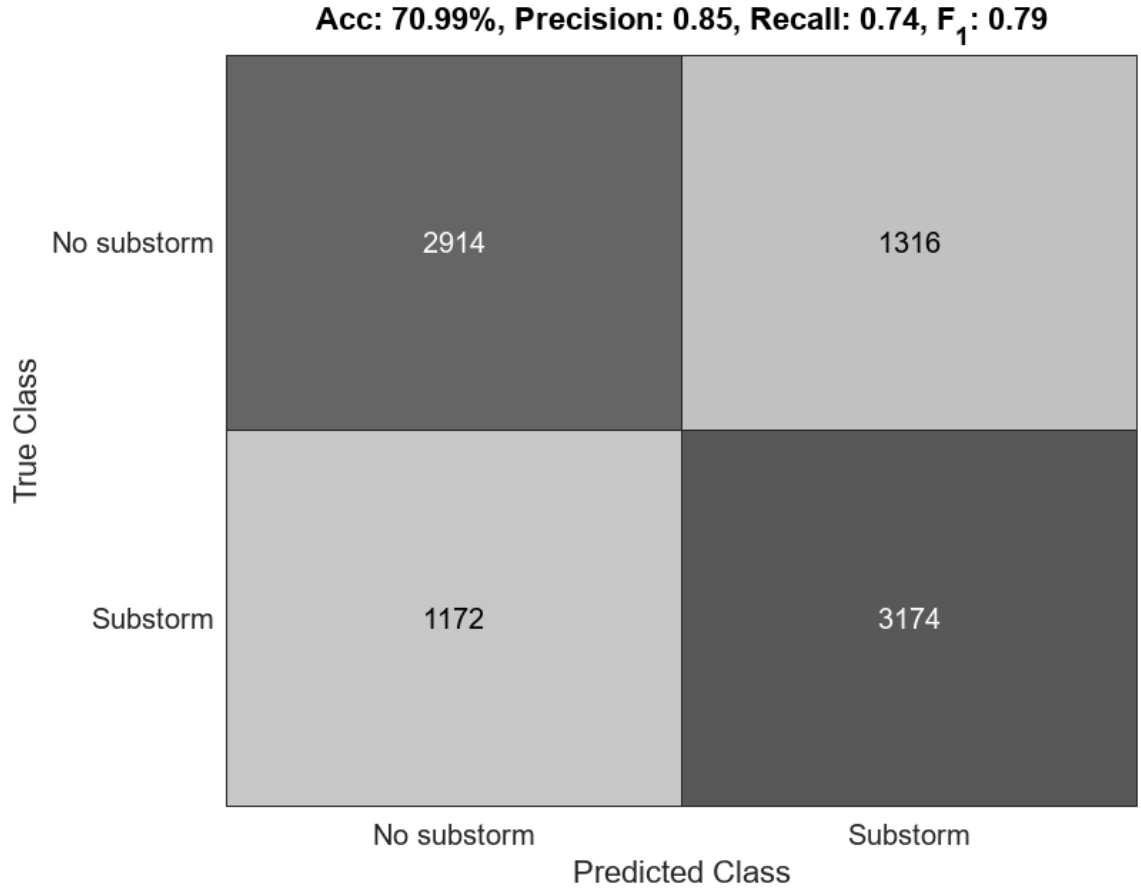

Figure 5.2. A confusion matrix resulting from the evaluation of the LSTM model applied to the aggregated dataset for training case. This specific confusion matrix corresponds to Case 13, where the input features used for training and testing the model are V_x, V_y, V_z, B_x, B_y, B_z, N_p, SML, and SYMH. The matrix showcases the performance of the model in predicting the occurrence of substorms (training set).

Figure 5.2 presents a confusion matrix indicating a 71% accuracy in testing, with a precision of 0.85, a recall of 0.74, and an F-1 score of 0.79. The rows of the confusion matrix represent the true class, signifying the true label of the data, while the columns show the predicted class, indicating the predictions obtained on the test data after training the LSTM model.

The diagonals in blue denote the correctly predicted data, while the red off-diagonal components represent the falsely predicted data. In the top left corner, where there were no substorms, the model correctly predicted all instances as no substorms, resulting in 2914 instances of true negatives. Conversely, the bottom right corner signifies the truly predicted substorms, where the model correctly identified instances as substorms, resulting in 3174 instances of true positives. These two cases correspond to true negatives and true positives, respectively.

The top right segment represents false positives, indicating instances where there were no substorms, but the model predicted them as substorms. Similarly, the bottom left corner signifies false negatives, where there were substorms, but the model predicted them as no substorms. Higher numbers in the top right and bottom left corners indicate the precision and recall values for these test cases. A diagonal-heavy confusion matrix indicates better model performance.

Figure 5.3 shows a confusion matrix that corresponds to the case where the LSTM model was applied solely on feature I_1. Here, a noticeable class bias is observed in predicting no substorms, indicating that the model struggled to learn the correlation between I_1 and predicting substorm onsets within the next 30 minutes.

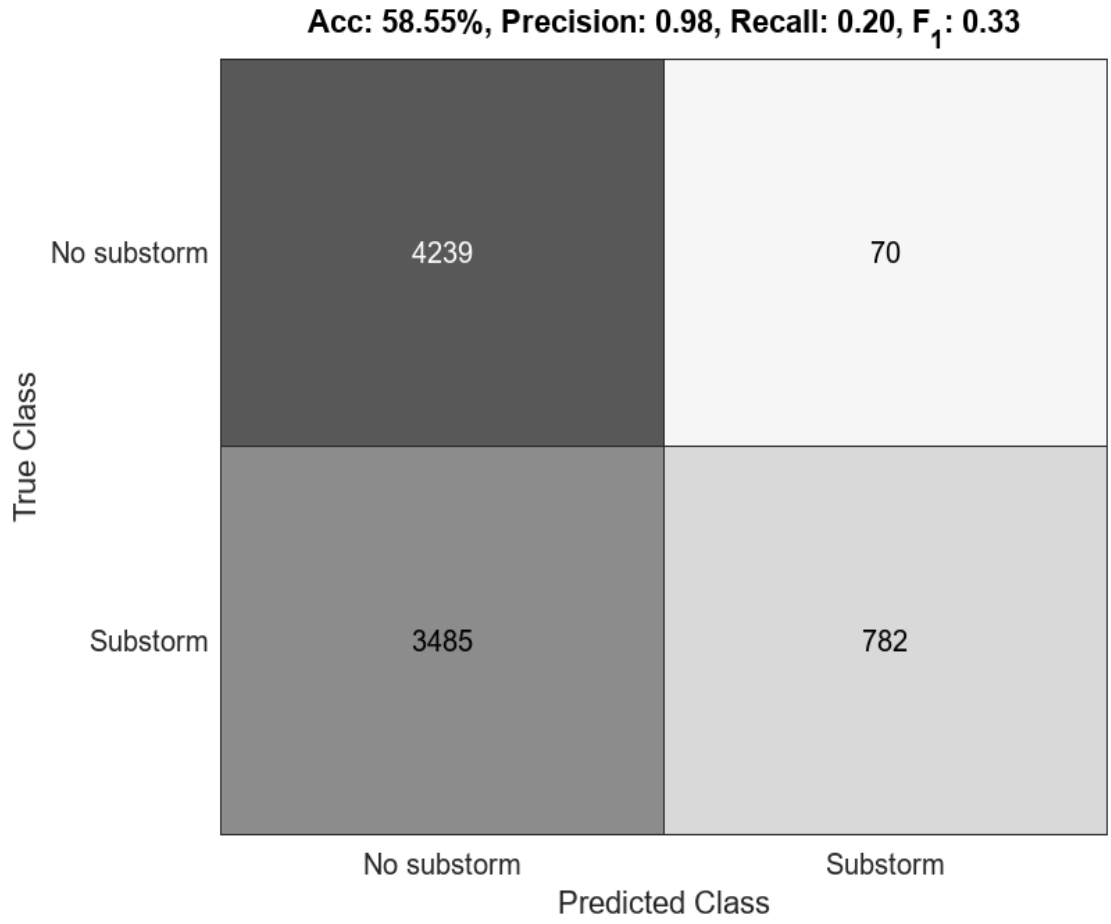

Figure 5.3. A confusion matrix resulting from the evaluation of the LSTM model applied to the aggregated dataset. This specific confusion matrix corresponds to Case 14, where the input features used for training and testing the model is I1. The matrix showcases the performance of the model in predicting the occurrence of substorms.

The high number (4239) of instances correctly predicted as no substorms contributes to a reliable precision score. However, this bias towards predicting no substorms results in a significant discrepancy in the bottom left corner, where substorms were incorrectly predicted as no substorms (3485 instances). Consequently, the recall score is considerably low (0.20) compared to the baseline of 0.50.

Given that the F-1 score is a harmonic mean of precision and recall, the F-1 score for this case being significantly low indicates the imbalance between precision and recall,

as well as reflects the model's inability to effectively capture substorm occurrences based solely on the I_1 feature.

5.3 CNN Cases

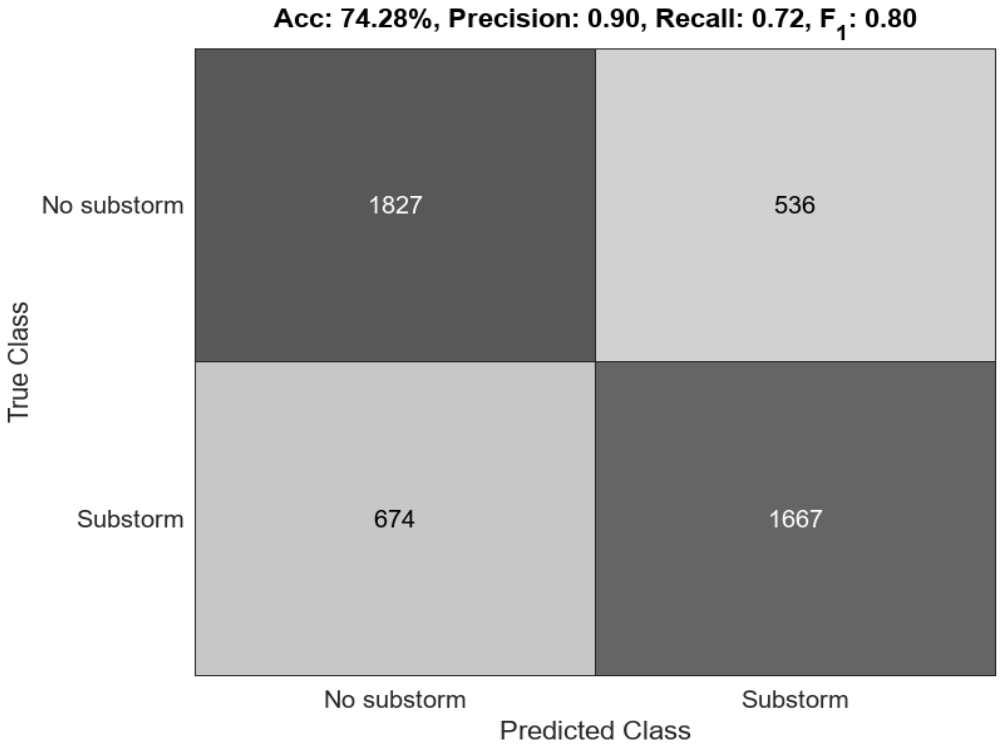

Figure 5.4. A confusion matrix resulting from the evaluation of the 2D-CNN model applied to the Newell dataset. This specific confusion matrix corresponds to Case 13, where the input features used for training and testing the model are V_x, V_y, V_z, B_x, B_y, B_z, N_p, SML, and SYMH. The matrix showcases the performance of the model in predicting the occurrence of substorms.

The presented confusion matrix in Figure 5.4 resulted from the evaluation of the 2D Convolutional Neural Network (CNN) model applied to the Newell dataset. Specifically, this matrix corresponded to Case 13, where the input features utilized for both training and testing the model included V_x, V_y, V_z, B_x, B_y, B_z, N_p, SML, and SYMH.

The matrix served to illustrate the performance of the CNN model in predicting the occurrence of substorms. The values within the matrix indicated the model's predictions compared to the ground truth labels.

In the top left corner of the matrix, there was a count of 1827, representing instances where the model correctly predicted the absence of substorms (true negatives). Conversely, in the top right corner, the count of 536 signified instances where the model incorrectly predicted the presence of substorms when there were none (false positives).

There was a count of 674 at the bottom left corner of the confusion matrix, denoting instances where the model failed to predict substorms when they occurred (false negatives). Finally, in the bottom right corner, the count of 1667 represented instances where the model accurately predicted the presence of substorms (true positives).

The confusion matrix depicted in Figure 5.5 resulted from the evaluation of the CNN model applied to the Newell dataset. Specifically, this matrix corresponds to Case 16, where the input features utilized for training and testing the model included V_x, V_y, V_z, B_x, B_y, B_z, N_p, SML, and SYMH, I_1, and Θ.

In analyzing this confusion matrix, a noticeable class bias is observed, particularly in predicting no substorms. The top left corner of the matrix reveals a count of 2374, representing instances where the model correctly predicted the absence of substorms (true negatives). Conversely, in the top right corner, the count of 4 signifies instances where the model incorrectly predicted the presence of substorms when there were none (false positives).

However, a significant discrepancy arises in the bottom left corner, where the count of 1899 denotes instances where the model failed to predict substorms when they occurred (false negatives). This discrepancy contributes to a considerably low recall score, reflecting the model's struggle to effectively capture substorm occurrences based on the given features.

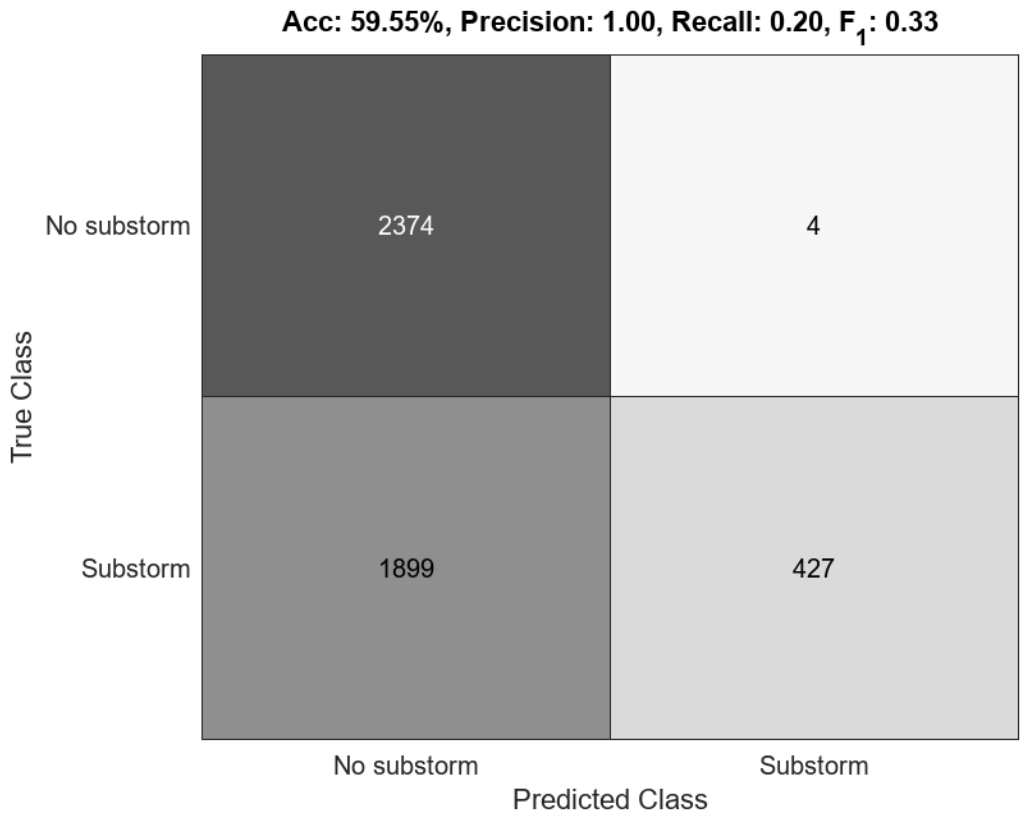

Figure 5.5. A confusion matrix resulting from the evaluation of the 2D-CNN model applied to the Newell dataset. This specific confusion matrix corresponds to Case 16, where the input features used for training and testing the model are V_x, B_x, B_y, B_z, N_p, SML, SYMH, I_1, and Θ. The matrix showcases the performance of the model in predicting the occurrence of substorms.

Furthermore, the bottom right corner of the matrix displays a count of 427, representing instances where the model accurately predicted the presence of substorms

(true positives). Despite this, the imbalance between precision and recall leads to a notably diminished F-1 score for this case.

5.4 Model Performance

Figure 5.6 presents a heatmap illustrating the accuracies obtained from both the LSTM and CNN models across 16 different parameter cases. These cases include variations in driver parameters, combinations of driver parameters, driver parameters combined with ground parameters, as well as cases involving additional parameters such as I_1, Θ, and combinations thereof. The substorm list used for these cases was the aggregated list.

The heatmap showcases an overall better performance of the LSTM model compared to CNN. Notably, both models exhibit significant improvements in accuracy when parameters B_x and B_z are utilized. Conversely, parameters N_p, B_y, and B_x do not contribute to enhanced model performance for either model.

Furthermore, the addition of I_1 and/or Θ as parameters does not improve model performance; rather, it leads to a reduction in accuracy for both LSTM and CNN models. This suggests a lack of correlation between these parameters and the occurrence of substorm onsets.

The top-performing cases for the LSTM model include combinations of parameters such as V_x, B_x, B_y, B_z, N_p, SML, and SYMH, with accuracies peaking in configurations including these parameters. Specifically, cases involving V_x, B_x, B_y, B_z, N_p, SML, and SYMH, as well as those including B_x, B_y, B_z, N_p, SML, and SYMH, yield the highest accuracies.

Similarly, the CNN model achieves high accuracy, particularly in cases involving parameters B_x, B_y, B_z, N_p, SML, and SYMH. The configuration with these parameters consistently performs best for CNN.

Additionally, single parameters such as V_x and B_z demonstrate relatively high accuracy for both LSTM and CNN models when predicting substorm onsets for the aggregated list.

In summary, the heatmap analysis highlights the importance of specific parameters, such as B_x and B_z, in predicting substorm onsets accurately. It also underscores the limited impact of certain parameters like N_p, B_y, and single parameters such as I_1, and Θ. The findings suggest that careful selection and combination of parameters are crucial for improving the predictive performance of both LSTM and CNN models for substorm onset prediction.

Upon closer examination of the heatmap, it is evident that the LSTM model generally outperforms the CNN model in most cases. In particular, the LSTM model consistently achieves higher accuracy scores compared to CNN, as indicated by the prevalence of warmer colors in the LSTM column relative to the CNN column.

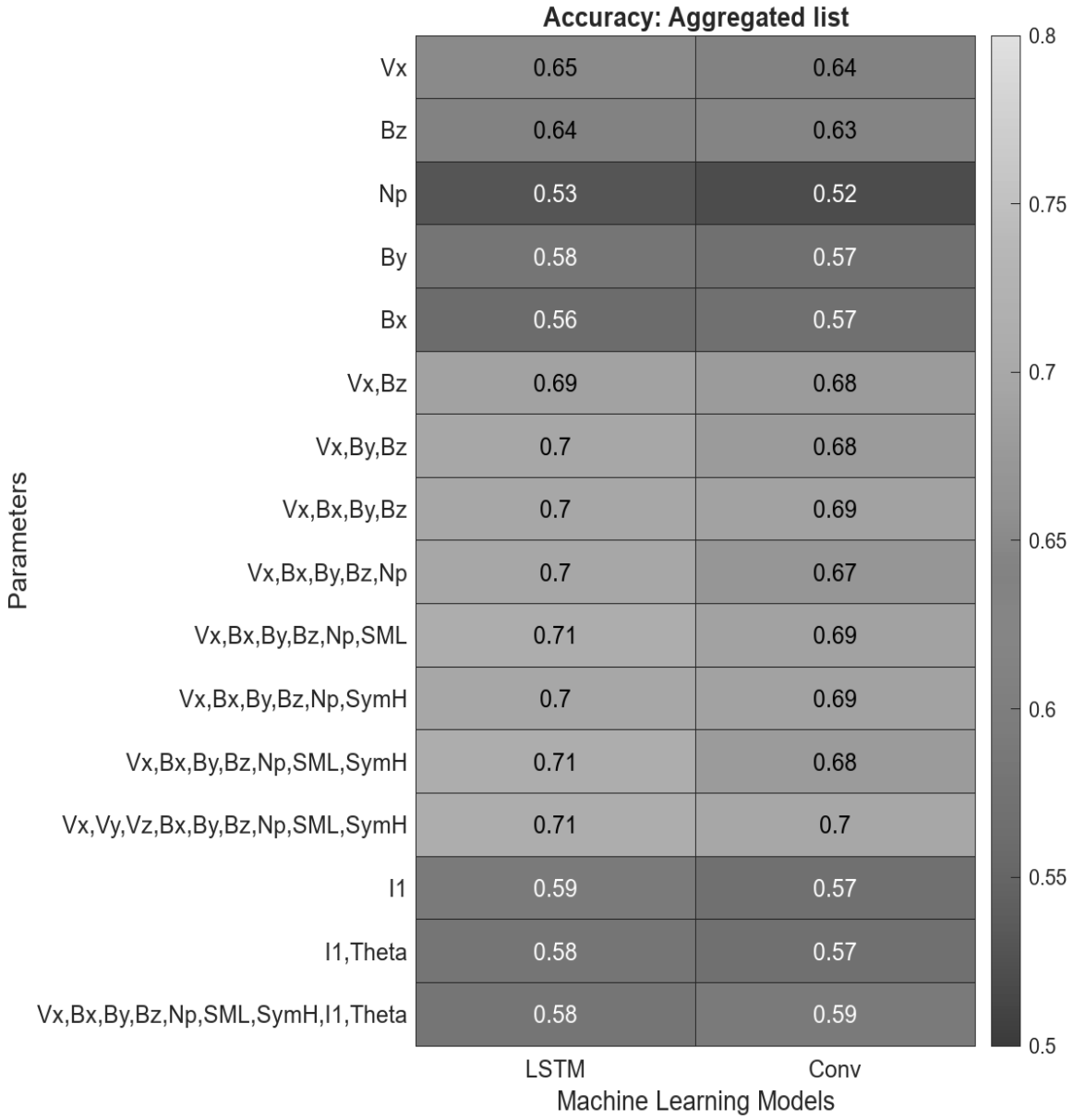

Figure 5.6. Heatmap depicting the comparison of LSTM and CNN model accuracies for the aggregated substorm list across 16 cases. The heatmap illustrates the accuracy scores obtained for each model, with the rows representing the individual cases evaluated and the columns corresponding to LSTM and CNN, respectively. The color gradient within the heatmap provides a visual representation of the accuracy scores, with warmer colors indicating higher accuracy and cooler colors indicating lower accuracy. The comparison highlights the nuanced differences in performance between the LSTM and CNN models, with LSTM demonstrating marginally higher accuracy across most cases, underscoring its relatively stronger predictive capability for substorm onset prediction with the aggregated list.

Across the various cases, the LSTM model consistently maintains accuracy scores above 60%, with some cases reaching accuracy levels of 70% or higher. In contrast, the CNN model demonstrates slightly lower accuracy scores overall, with many cases falling below the 60% accuracy threshold.

Furthermore, there are noticeable fluctuations in accuracy scores across different cases for both models. While some cases yield relatively high accuracy for both LSTM and CNN, others result in lower accuracy, indicating the sensitivity of model performance to specific input parameters or configurations.

5.5 Model Performance for different lists

Figure 5.7 illustrates a heatmap depicting the performance of the LSTM model across five types of substorm lists for the 16 cases that were run. In this heatmap, warmer colors represent higher accuracy values, while cooler colors indicate lower accuracy values.

The analysis reveals that the LSTM model achieves the highest accuracies when provided with the Newell list as the substorm list. This indicates that the Newell list contains features or characteristics that enable the model to make more accurate predictions of substorm onset.

Conversely, when Frey's list is used as the substorm list, the accuracies are observed to be the lowest. This suggests that the features or patterns present in Frey's list may pose challenges for the LSTM model in accurately predicting substorm onsets.

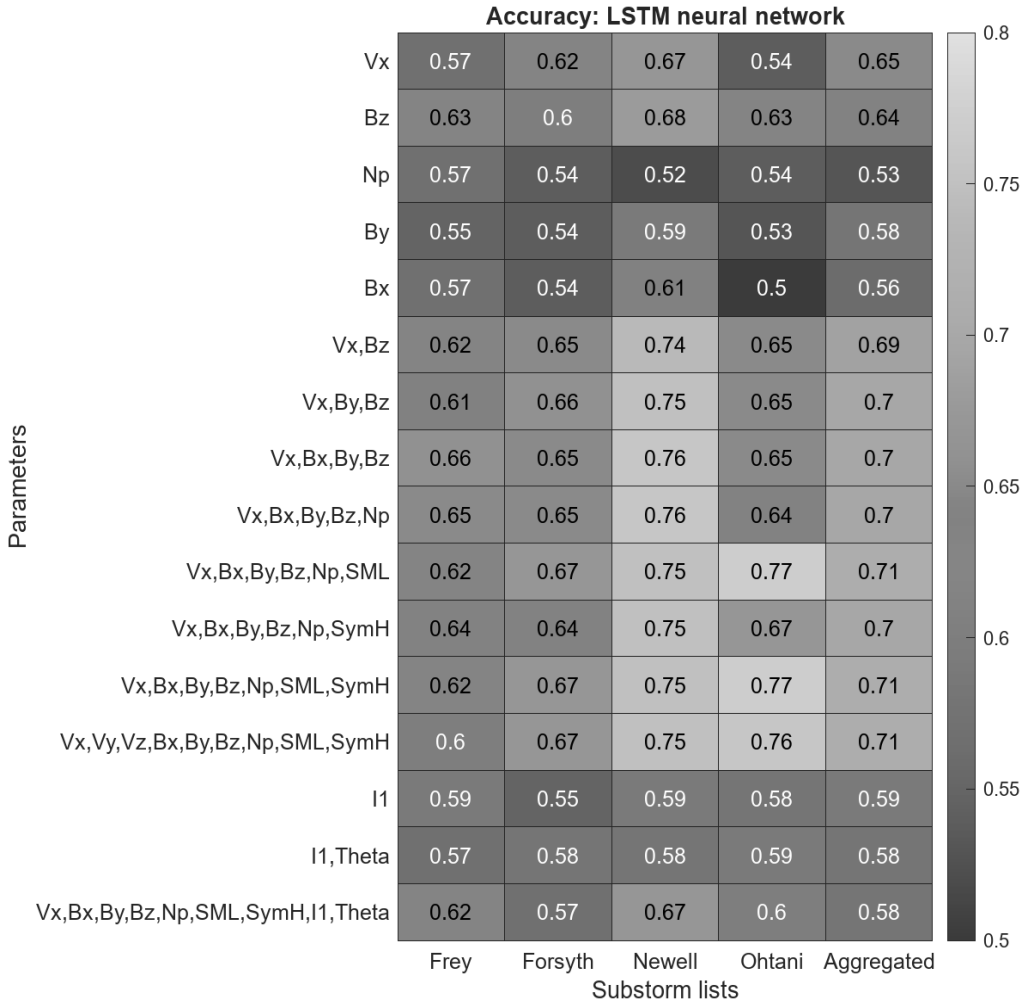

Figure 5.7. Performance heatmap illustrating the results of the LSTM network across 16 different cases for predicting substorm onset with Frey, Forsyth, Newell, Ohtani, and the aggregated substorm lists. The numbers in fraction show accuracy corresponding to each case and list. Cooler colors represent low and warmer colors represent higher accuracy.

The aggregated list also demonstrates significantly high performance (77%) which is higher than Maimati's performance which we selected to be our reference performance indicator, indicating that combining data from multiple sources enhances our model's predictive capabilities. Interestingly, for Ohtani's list, the accuracies vary across different cases. Some cases exhibit high accuracies, while others show lower

accuracies. This variability suggests that Ohtani's list may contain certain features that are conducive to accurate predictions in some cases but not in others.

Notably, despite the challenges observed with other lists, the LSTM model performs well when provided with Newell's list for the last case. This suggests that Newell's list may contain unique features or patterns that enable the model to learn accurately, even when other lists present difficulties.

Overall, the heatmap analysis provides valuable insights into the performance of the LSTM model across different types of substorm lists. It highlights the importance of list selection and underscores the potential impact of list-specific features on model performance.

The results obtained from the LSTM network analysis provide detailed insights into the predictive performance of various parameter combinations. Initial examinations reveal that cases focusing solely on velocity parameters, such as V_x and V_y, exhibit accuracies ranging from 0.57 to 0.63. This suggests a moderate predictive capability when considering only velocity-related data. However, the addition of magnetic field components, including B_x, B_y, and B_z, alongside velocity parameters notably improves predictive performance. Cases incorporating these magnetic field components display accuracies spanning from 0.54 to 0.68, indicating a significant enhancement in predictive power. Interestingly, the inclusion of environmental parameters such as N_p, SML, and SYMH demonstrates varied impacts on prediction accuracy, with accuracies fluctuating between 0.52 and 0.64. This highlights the nuanced influence of environmental factors on predictive outcomes. Moreover, the incorporation of interaction parameters like I_1 and Θ yields mixed results, with accuracies ranging from 0.54 to 0.59, suggesting a complex

relationship between these parameters and the predicted outcomes. Overall, these detailed findings underscore the importance of meticulous parameter selection and model optimization in maximizing predictive accuracy within the LSTM framework, providing valuable insights for future model refinement and application in forecasting scenarios.

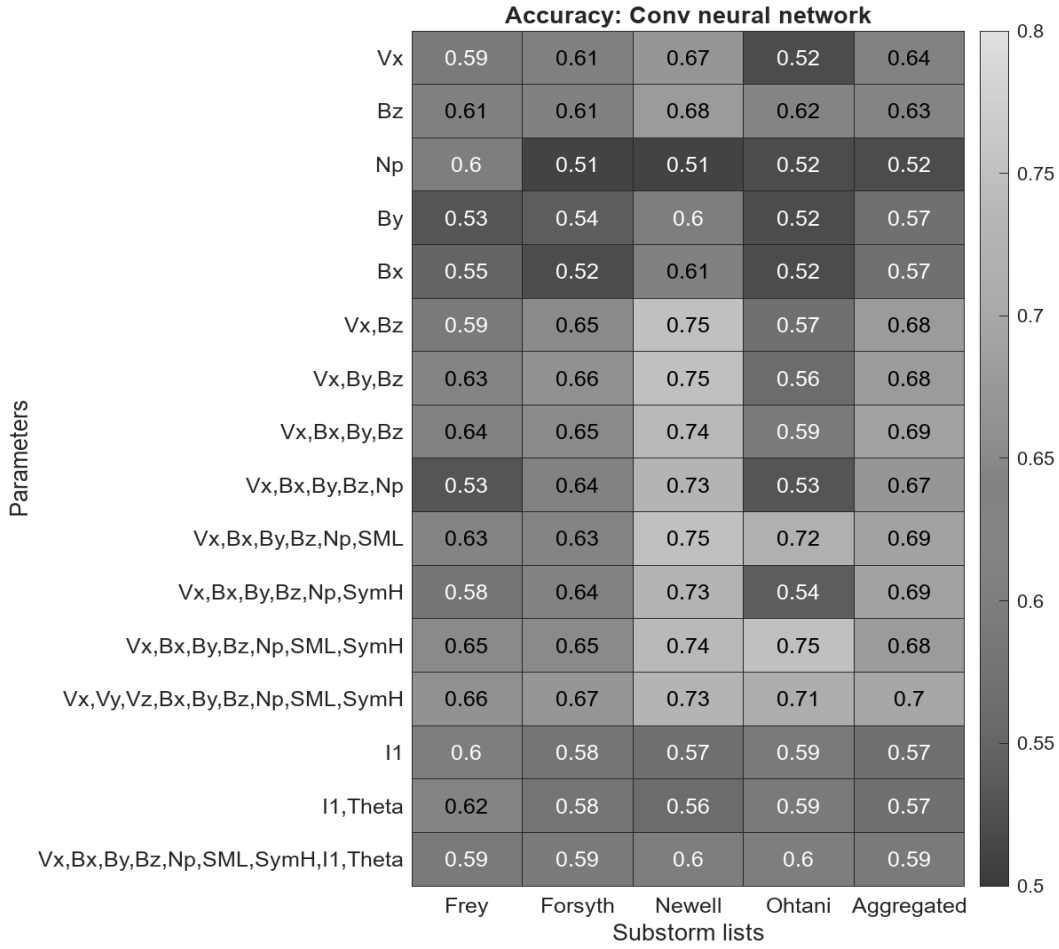

Figure 5.8. Performance heatmap illustrating the results of the CNN network across 16 different cases for predicting substorm onset with Frey, Forsyth, Newell, Ohtani, and the aggregated substorm lists. The numbers in fraction show accuracy corresponding to each case and list. Cooler colors represent low and warmer colors represent higher accuracy.

Figure 5.8 displays a heatmap representing the performance of the Convolutional Neural Network (CNN) across five types of substorm lists for the 16 cases that were

executed. In this heatmap, warmer colors indicate higher accuracy values, while cooler colors signify lower accuracy values.

The analysis reveals that the CNN model achieves its highest accuracies when provided with the Newell list as the substorm list. This suggests that the Newell list contains features or characteristics that enable the CNN model to make more accurate predictions of substorm onset compared to other lists.

Conversely, when Frey's list is utilized as the substorm list, the accuracies are observed to be the lowest. This implies that the features or patterns present in Frey's list may pose challenges for the CNN model in accurately predicting substorm onsets. Like the LSTM model, the aggregated list demonstrates notably high performance for the CNN model as well. This indicates that aggregating data from multiple sources enhances the CNN model's predictive capabilities.

For Ohtani's list, the accuracies vary across different cases, with some cases exhibiting high accuracies and others showing lower accuracies. This variability suggests that Ohtani's list may contain certain features that are conducive to accurate predictions in some cases but not in others.

The analysis of the CNN model's performance provides a detailed understanding of how different parameter combinations influence predictive accuracy. When considering individual parameters, accuracies range from 0.52 for B_y to 0.64 for B_x, indicating variations in predictive power based on the specific parameter under consideration. Combining velocity and magnetic field parameters in different configurations yields accuracies ranging from 0.61 to 0.63, showcasing the potential synergy between these features for enhanced predictions.

However, the introduction of environmental parameters (N_p, SML, SYMH) exhibits mixed effects on prediction accuracy, with accuracies spanning from 0.51 to 0.67 across various cases. Interestingly, the inclusion of interaction parameters (I_1, Θ) showcases a moderate impact, with accuracies ranging from 0.53 to 0.54. These nuanced findings emphasize the intricate interplay between different parameters and highlight the need for meticulous parameter selection to optimize predictive performance within the CNN framework. The insights gained from these results serve as valuable guidance for refining model architecture and enhancing predictive capabilities in future applications of CNN-based predictive analytics.

CHAPTER VI

SUMMARY AND CONCLUSION

This research evaluates the substorm onset criteria based on ground-based magnetic indices SML, field-aligned currents I_1, auroral observations, electrojet indices criteria to establish a relationship and lay the foundation for neural network training and modeling. In the field of Space Physics, this research introduces a solution for substorm onset detection by combining neural network training for classification with hyperparameter tuning—a novel approach that has not been previously explored across multiple substorm datasets. All the onset detection techniques used the SML index to indicate the onset of substorms. This is also true for the WINDMI model.

Derived from assumptions about magnetospheric configuration, the hybrid model demonstrates remarkable accuracy, achieving a highest 77% predictive performance (with parameters V_x, B_x, B_y, B_z, N_p, SML, and SYMH) with the LSTM architecture significantly surpassing the 40% accuracy attained by the WINDMI model (See Tables A1 to A5 and B1 to B5). The aggregated list performs consistently well across both architectures (average value of LSTM 70% and CNN 69%). This is the first time this methodology has been attempted which attests to the novelty of the framework. Substorm lists are integral indicators since they found a sound correlation between solar wind, ground measurements and the Newell as well as aggregated lists.

The goal of this research was to delve into an in-depth exploration of the predictive capacities of Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) models in forecasting substorm onsets. Leveraging a diverse range of input parameters derived from satellite observations and the WINDMI model, the study meticulously analyzed the performance of each model across various scenarios. LSTM emerged as the more proficient model, excelling in capturing intricate temporal dependencies within the dataset. The significance of key parameters such as solar wind speed (V_x) and interplanetary magnetic field strength (B_z) was underscored, with their inclusion leading to notable enhancements in predictive accuracy. However, the introduction of the physics parameter I_1 yielded only marginal improvements, suggesting its limited influence on the models' predictive capabilities. The performance of the models varied across different substorm onset lists, with Forsyth's list consistently yielding reasonable accuracies while Frey and Ohtani's lists, characterized by fewer substorms, posed challenges for effective model learning. Newell's list emerged as the most reliable predictor, showcasing its efficacy in forecasting substorm onsets with the highest average accuracies. Furthermore, the CNN model demonstrated optimal performance when incorporating a comprehensive set of nine solar wind parameters, emphasizing the importance of meticulous parameter selection for accurate predictions. CNN Model also performed well with the aggregated list. The Newell substorm list and model combinations performed well as demonstrated by performance evaluation metrics such as confusion matrix and heatmaps.

In our study, we compared the performance of two neural network architectures: LSTM and CNN. Surprisingly, LSTM outperformed CNN overall. The highest accuracy

achieved was an impressive 77%, which occurred when we used specific parameters—V_x, B_x, B_y, and B_z—alongside Newell's substorm onset list. However, the addition of the physics parameter I_1 didn't enhance accuracy, suggesting a lack of correlation with substorm occurrences. Interestingly, when we looked at the aggregated list, CNN performed well than LSTM. Convolutional layers seemed effective in learning patterns when all lists were combined. Key parameters contributing to improved performance were solar wind speed (V_x) and interplanetary magnetic field (B_z). On the other hand, parameters like B_x, B_y, and N_p didn't significantly enhance accuracy.

We also explored different substorm lists. Forsyth's list yielded relatively good performance, averaging around 65% accuracy. However, sparse data posed challenges for Frey and Ohtani's lists, resulting in lower accuracies. Individual impact analysis revealed that V_x and B_z were crucial, with accuracies around 63% when input individually. Newell's list consistently achieved the highest accuracies (ranging from 74% to 75%) for both models. For CNN, the case with nine solar wind parameters performed best across all lists, including the aggregated one. Interestingly, adding SYMH didn't significantly drop model accuracy. In summary, our findings highlight the importance of specific parameters and the effectiveness of different neural network architectures in predicting substorm occurrences.

REFERENCES

Adhya, P., Spencer, E.A., & Kayode-Adeoye, M. (2024). Substorm Identification With The WINDMI Magnetosphere-Ionosphere Nonlinear Physics Model. Space Weather. [Manuscript submitted for publication].

Akasofu, S. I. (1964). The development of the auroral substorm. Planetary and Space Science, **12**(4), 273-282.

Akasofu, S. I. (1981). Energy coupling between the solar wind and the magnetosphere. Space Science Reviews, **28**, 121-190.

Akasofu, S. I. (2004). Several ‘Controversial' Issues on Substorms. Space Science Reviews, **113**(1), 1-40.

Akasofu, S. I. (2013). Where is the magnetic energy for the expansion phase of auroral substorms accumulated?. Journal of Geophysical Research: Space Physics, **118**(11), 7219-7225.

Baker, D. E., Fritz, T. A., McPherron, R. L., Fairfield, D. H., Kamide, Y., & Baumjohann, W. (1985). Magnetotail energy storage and release during the CDAW 6 substorm analysis intervals. Journal of Geophysical Research: Space Physics, **90**(A2), 1205-1216.

Bittencourt, J. A. (2013). Fundamentals of plasma physics. Springer Science & Business Media.

Forsyth, C., Rae, I. J., Coxon, J. C., Freeman, M. P., Jackman, C. M., Gjerloev, J., & Fazakerley, A. N. (2015). A new technique for determining Substorm Onsets and Phases from Indices of the Electrojet (SOPHIE). Journal of Geophysical Research: Space Physics, **120**(12), 10-592.

Frey, H. U., Mende, S. B., Angelopoulos, V., & Donovan, E. F. (2004). Substorm onset observations by IMAGE-FUV. Journal of Geophysical Research: Space Physics, **109**(A10).

Horton, W., & Doxas, I. (1996). A low-dimensional energy-conserving state space model for substorm dynamics. Journal of Geophysical Research: Space Physics, **101**(A12), 27223-27237.

Johnson, R., Boubrahimi, S. F., Bahri, O., & Hamdi, S. M. (2022, August). Physics-Informed Neural Networks for Solar Wind Prediction. In International Conference on Pattern Recognition (pp. 273-286). Cham: Springer Nature Switzerland.

Kallio, E. I., Pulkkinen, T. I., Koskinen, H. E. J., Viljanen, A., Slavin, J. A., & Ogilvie, K. (2000). Loading-unloading processes in the nightside ionosphere. Geophysical Research Letters, **27**(11), 1627-1630.

Kistler, L. M., Frey, H. U., Möbius, E., Mouikis, C., Quinn, J. M., Klecker, B., ... & Fazakerley, A. N. (2002). Motion of auroral ion outflow structures observed with CLUSTER and IMAGE FUV. Journal of Geophysical

Research: Space Physics, **107**(A8), SMP-17.

Kivelson, M. G., & Russell, C. T. (Eds.). (1995). Introduction to space physics. Cambridge University Press.

Maimaiti, M., Kunduri, B., Ruohoniemi, J. M., Baker, J. B. H., & House, L. L. (2019). A deep learning-based approach to forecast the onset of magnetic substorms. Space weather, **17**(11), 1534-1552.

McPherron, R. L. (2020). Early studies in solar wind coupling and substorms. Journal of Geophysical Research: Space Physics, **125**(5), e2019JA027615.

Mende, S. B., Harris, S. E., Frey, H. U., Angelopoulos, V., Russell, C. T., Donovan, E., ... & Peticolas, L. M. (2009). The THEMIS array of ground-based observatories for the study of auroral substorms. The themis mission, 357-387.

Newell, P. T., & Gjerloev, J. W. (2011). Evaluation of SuperMAG auroral electrojet indices as indicators of substorms and auroral power. Journal of Geophysical Research: Space Physics, **116**(A12).

Newell, P. T., Lee, A. R., Liou, K., Ohtani, S. I., Sotirelis, T., & Wing, S. (2010). Substorm cycle dependence of various types of aurora. Journal of Geophysical Research: Space Physics, **115**(A9).

Ohtani, S., & Gjerloev, J. W. (2020). Is the substorm current wedge an ensemble of wedgelets?: Revisit to midlatitude positive bays. Journal of Geophysical Research: Space Physics, **125**(9), e2020JA027902.

Pu, Z. Y., Chu, X. N., Cao, X., Mishin, V., Angelopoulos, V., Wang, J., ... & Lucek, E. (2010). THEMIS observations of substorms on 26 February 2008 initiated by

magnetotail reconnection. Journal of Geophysical Research: Space Physics, **115**(A2).

Pulkkinen, T. (2007). Space weather: Terrestrial perspective. Living Reviews in Solar Physics, **4**(1), 1.

Spencer, E., Horton, W., Mays, M. L., Doxas, I., & Kozyra, J. (2007). Analysis of the 3–7 October 2000 and 15–24 April 2002 geomagnetic storms with an optimized nonlinear dynamical model. Journal of Geophysical Research: Space Physics, **112**(A4).

Spencer, E., Rao, A., Horton, W., & Mays, M. L. (2009). Evaluation of solar wind-magnetosphere coupling functions during geomagnetic storms with the WINDMI model. Journal of Geophysical Research: Space Physics, **114**(A2).

Spencer, E., Srinivas, P., & Vadepu, S. K. (2019). Global energy dynamics during substorms on 9 March 2008 and 26 February 2008 using satellite observations and the WINDMI model. Journal of Geophysical Research: Space Physics, **124**(3), 1698-1710.

Spencer, E., Vadepu, S. K., Srinivas, P., Patra, S., & Horton, W. (2018). The dynamics of geomagnetic substorms with the WINDMI model. Earth, Planets and Space, **70**, 1-9.

Tsurutani, B. T., Zank, G. P., Sterken, V. J., Shibata, K., Nagai, T., Mannucci, A. J., ... & Akasofu, S. I. (2022). Space plasma physics: A review. IEEE Transactions on Plasma Science, **51**(7), 1595-1655.

Yoon, P. H., & Lui, A. T. (2004). Lower-hybrid-drift and modified-two-stream instabilities in current sheet equilibrium. Journal of Geophysical Research: Space Physics, **109**(A2).

APPENDICES

Appendix A: LSTM Results

Table A1. Performance Evaluation Metrics for Aggregated Substorm List using LSTM

Case	Parameters	Accuracy	Precision	Recall	F1 Score
1	Vx	0.65	0.83	0.59	0.69
2	Bz	0.63	0.80	0.62	0.69
3	Np	0.52	0.72	0.60	0.66
4	By	0.57	0.76	0.52	0.62
5	Bx	0.55	0.78	0.62	0.69
6	Vx, Bz	0.70	0.85	0.46	0.61
7	Vx, By, Bz	0.63	0.84	0.73	0.79
8	Vx, Bx, By, Bz	0.64	0.85	0.61	0.72
9	Vx, Bx, By, Bz, Np	0.53	0.85	0.14	0.24
10	Vx, Bx, By, Bz, Np, SML	0.73	0.81	0.70	0.77
11	Vx, Bx, By, Bz, Np, SYMH	0.69	0.80	0.80	0.82
12	Vx, Bx, By, Bz, Np, SML, SYMH	0.71	0.88	0.51	0.67
13	Vx, Vy, Vz, Bx, By, Bz, Np, SML, SYMH	0.69	0.86	0.61	0.74
14	I_1	0.60	0.99	0.29	0.34
15	I_1, Θ	0.62	0.99	0.24	0.39
16	Vx, Vy, Vz, Bx, By, Bz, Np, SML, I_1, Θ	0.59	0.99	0.82	0.38

Table A2. Performance Evaluation Metrics for Forsyth Substorm List using LSTM

Case	Parameters	Accuracy	Precision	Recall	F1 Score
1	Vx	0.61	0.80	0.59	0.68
2	Bz	0.61	0.78	0.73	0.75
3	Np	0.51	0.72	0.64	0.68
4	By	0.54	0.74	0.59	0.66
5	Bx	0.52	0.75	0.36	0.49
6	Vx, Bz	0.65	0.82	0.67	0.74
7	Vx, By, Bz	0.66	0.82	0.69	0.75
8	Vx, Bx, By, Bz	0.65	0.82	0.68	0.75
9	Vx, Bx, By, Bz, Np	0.64	0.81	0.70	0.75
10	Vx, Bx, By, Bz, Np, SML	0.63	0.85	0.46	0.60
11	Vx, Bx, By, Bz, Np, SYMH	0.64	0.80	0.71	0.76
12	Vx, Bx, By, Bz, Np, SML, SYMH	0.65	0.85	0.55	0.67
13	Vx, Vy, Vz, Bx, By, Bz, Np, SML, SYMH	0.67	0.84	0.70	0.76
14	I_1	0.58	0.99	0.18	0.31
15	I_1, Θ	0.58	0.99	0.18	0.31
16	Vx, Vy, Vz, Bx, By, Bz, Np, SML, I_1, Θ	0.59	0.99	0.19	0.32

Table A3. Performance Evaluation Metrics for Newell Substorm List using LSTM

Case	Parameters	Accuracy	Precision	Recall	F1 Score
1	Vx	0.67	0.85	0.68	0.75
2	Bz	0.68	0.84	0.71	0.77
3	Np	0.51	0.73	0.38	0.50
4	By	0.59	0.77	0.58	0.66
5	Bx	0.61	0.81	0.53	0.64
6	Vx, Bz	0.75	0.89	0.74	0.81
7	Vx, By, Bz	0.75	0.91	0.72	0.81
8	Vx, Bx, By, Bz	0.74	0.87	0.80	0.83
9	Vx, Bx, By, Bz, Np	0.73	0.86	0.82	0.84
10	Vx, Bx, By, Bz, Np, SML	0.75	0.908	0.69	0.78
11	Vx, Bx, By, Bz, Np, SYMH	0.73	0.87	0.77	0.81
12	Vx, Bx, By, Bz, Np, SML, SYMH	0.74	0.90	0.72	0.80
13	Vx, Vy, Vz, Bx, By, Bz, Np, SML, SYMH	0.73	0.85	0.86	0.86
14	I_1	0.56	0.99	0.15	0.25
15	I_1, Θ	0.56	0.99	0.14	0.24
16	Vx, Vy, Vz, Bx, By, Bz, Np, SML, I_1, Θ	0.59	0.99	0.19	0.33

Table A4. Performance Evaluation Metrics for Ohtani Substorm List using LSTM

Case	Parameters	Accuracy	Precision	Recall	F1 Score
1	Vx	0.52	0.76	0.40	0.53
2	Bz	0.62	0.82	0.71	0.76
3	Np	0.52	0.76	0.62	0.69
4	By	0.52	0.75	0.48	0.59
5	Bx	0.52	0.75	0.40	0.52
6	Vx, Bz	0.57	0.76	0.92	0.83
7	Vx, By, Bz	0.56	0.78	0.39	0.52
8	Vx, Bx, By, Bz	0.59	0.78	0.91	0.84
9	Vx, Bx, By, Bz, Np	0.53	0.75	0.57	0.64
10	Vx, Bx, By, Bz, Np, SML	0.72	0.83	0.93	0.88
11	Vx, Bx, By, Bz, Np, SYMH	0.54	0.76	0.65	0.70
12	Vx, Bx, By, Bz, Np, SML, SYMH	0.75	0.85	0.92	0.88
13	Vx, Vy, Vz, Bx, By, Bz, Np, SML, SYMH	0.71	0.83	0.97	0.89
14	I_1	0.59	0.99	0.19	0.32
15	I_1, Θ	0.59	0.99	0.19	0.32
16	Vx, Vy, Vz, Bx, By, Bz, Np, SML, I_1, Θ	0.60	0.99	0.22	0.36

Table A5. Performance Evaluation Metrics for Frey Substorm List using LSTM

Case	Parameters	Accuracy	Precision	Recall	F1 Score
1	Vx	0.58	0.83	0.53	0.65
2	Bz	0.61	0.84	0.62	0.71
3	Np	0.60	0.83	0.47	0.60
4	By	0.52	0.78	0.48	0.59
5	Bx	0.55	0.78	0.63	0.70
6	Vx, Bz	0.59	0.84	0.48	0.61
7	Vx, By, Bz	0.63	0.84	0.75	0.79
8	Vx, Bx, By, Bz	0.64	0.86	0.63	0.72
9	Vx, Bx, By, Bz, Np	0.53	0.86	0.15	0.24
10	Vx, Bx, By, Bz, Np, SML	0.63	0.82	0.73	0.77
11	Vx, Bx, By, Bz, Np, SYMH	0.58	0.80	0.84	0.82
12	Vx, Bx, By, Bz, Np, SML, SYMH	0.65	0.87	0.55	0.67
13	Vx, Vy, Vz, Bx, By, Bz, Np, SML, SYMH	0.66	0.86	0.65	0.74
14	I_1	0.59	0.99	0.21	0.35
15	I_1, Θ	0.62	0.99	0.24	0.39
16	Vx, Vy, Vz, Bx, By, Bz, Np, SML, I_1, Θ	0.59	0.99	0.23	0.38

Appendix B: CNN Results

Table B1. Performance Evaluation Metrics for Aggregated Substorm List using CNN

Case	Parameters	Accuracy	Precision	Recall	F1 Score
1	Vx	0.65	0.85	0.57	0.69
2	Bz	0.64	0.80	0.71	0.75
3	Np	0.53	0.71	0.55	0.62
4	By	0.57	0.78	0.50	0.61
5	Bx	0.56	0.75	0.46	0.57
6	Vx, Bz	0.69	0.85	0.68	0.76
7	Vx, By, Bz	0.69	0.85	0.70	0.77
8	Vx, Bx, By, Bz	0.70	0.86	0.70	0.77
9	Vx, Bx, By, Bz, Np	0.70	0.86	0.70	0.78
10	Vx, Bx, By, Bz, Np, SML	0.71	0.86	0.73	0.79
11	Vx, Bx, By, Bz, Np, SYMH	0.70	0.85	0.71	0.77
12	Vx, Bx, By, Bz, Np, SML, SYMH	0.71	0.85	0.71	0.78
13	Vx, Vy, Vz, Bx, By, Bz, Np, SML, SYMH	0.71	0.85	0.74	0.79
14	I_1	0.59	0.98	0.19	0.33
15	I_1, Θ	0.58	0.98	0.20	0.33
16	Vx, Vy, Vz, Bx, By, Bz, Np, SML, I_1, Θ	0.58	0.78	0.34	0.48

Table B2. Performance Evaluation Metrics for Forsyth Substorm List using CNN

Case	Parameters	Accuracy	Precision	Recall	F1 Score
1	Vx	0.62	0.83	0.58	0.68
2	Bz	0.60	0.77	0.72	0.75
3	Np	0.54	0.74	0.60	0.66
4	By	0.54	0.74	0.51	0.61
5	Bx	0.54	0.74	0.50	0.61
6	Vx, Bz	0.65	0.82	0.68	0.75
7	Vx, By, Bz	0.66	0.83	0.70	0.76
8	Vx, Bx, By, Bz	0.65	0.83	0.67	0.74
9	Vx, Bx, By, Bz, Np	0.65	0.82	0.70	0.75
10	Vx, Bx, By, Bz, Np, SML	0.67	0.84	0.72	0.78
11	Vx, Bx, By, Bz, Np, SYMH	0.64	0.82	0.67	0.74
12	Vx, Bx, By, Bz, Np, SML, SYMH	0.67	0.83	0.72	0.77
13	Vx, Vy, Vz, Bx, By, Bz, Np, SML, SYMH	0.67	0.84	0.69	0.76
14	I_1	0.55	0.72	0.99	0.84
15	I_1, Θ	0.58	0.96	0.19	0.32
16	Vx, Vy, Vz, Bx, By, Bz, Np, SML, I_1, Θ	0.57	0.96	0.17	0.30

Table B3. Performance Evaluation Metrics for Newell Substorm List using CNN

Case	Parameters	Accuracy	Precision	Recall	F1 Score
1	Vx	0.67	0.86	0.66	0.75
2	Bz	0.68	0.82	0.79	0.81
3	Np	0.52	0.76	0.50	0.60
4	By	0.59	0.80	0.53	0.63
5	Bx	0.61	0.83	0.51	0.63
6	Vx, Bz	0.75	0.89	0.76	0.82
7	Vx, By, Bz	0.75	0.90	0.76	0.83
8	Vx, Bx, By, Bz	0.76	0.89	0.79	0.84
9	Vx, Bx, By, Bz, Np	0.76	0.89	0.78	0.83
10	Vx, Bx, By, Bz, Np, SML	0.75	0.89	0.76	0.82
11	Vx, Bx, By, Bz, Np, SYMH	0.75	0.89	0.77	0.83
12	Vx, Bx, By, Bz, Np, SML, SYMH	0.75	0.90	0.76	0.82
13	Vx, Vy, Vz, Bx, By, Bz, Np, SML, SYMH	0.75	0.88	0.77	0.82
14	I_1	0.59	0.96	0.21	0.34
15	I_1, Θ	0.58	0.94	0.22	0.36
16	Vx, Vy, Vz, Bx, By, Bz, Np, SML, I_1, Θ	0.67	0.91	0.45	0.60

Table B4. Performance Evaluation Metrics for Ohtani Substorm List using CNN

Case	Parameters	Accuracy	Precision	Recall	F1 Score
1	Vx	0.54	0.80	0.41	0.54
2	Bz	0.63	0.83	0.65	0.73
3	Np	0.54	0.77	0.68	0.72
4	By	0.53	0.76	0.60	0.67
5	Bx	0.50	0.74	0.46	0.57
6	Vx, Bz	0.65	0.83	0.79	0.81
7	Vx, By, Bz	0.65	0.82	0.75	0.78
8	Vx, Bx, By, Bz	0.65	0.82	0.75	0.79
9	Vx, Bx, By, Bz, Np	0.64	0.82	0.77	0.79
10	Vx, Bx, By, Bz, Np, SML	0.77	0.89	0.85	0.87
11	Vx, Bx, By, Bz, Np, SYMH	0.67	0.84	0.78	0.81
12	Vx, Bx, By, Bz, Np, SML, SYMH	0.77	0.89	0.85	0.87
13	Vx, Vy, Vz, Bx, By, Bz, Np, SML, SYMH	0.76	0.87	0.84	0.86
14	I_1	0.58	0.76	0.99	0.87
15	I_1, Θ	0.59	0.77	0.99	0.87
16	Vx, Vy, Vz, Bx, By, Bz, Np, SML, I_1, Θ	0.60	0.78	0.97	0.86

Table B5. Performance Evaluation Metrics for Frey Substorm List using CNN

Case	Parameters	Accuracy	Precision	Recall	F1 Score
1	Vx	0.57	0.80	0.72	0.75
2	Bz	0.63	0.86	0.65	0.74
3	Np	0.57	0.80	0.56	0.67
4	By	0.54	0.81	0.41	0.54
5	Bx	0.57	0.82	0.52	0.64
6	Vx, Bz	0.62	0.85	0.59	0.70
7	Vx, By, Bz	0.61	0.83	0.67	0.74
8	Vx, Bx, By, Bz	0.66	0.87	0.70	0.78
9	Vx, Bx, By, Bz, Np	0.65	0.86	0.67	0.75
10	Vx, Bx, By, Bz, Np, SML	0.62	0.83	0.68	0.75
11	Vx, Bx, By, Bz, Np, SYMH	0.64	0.85	0.65	0.73
12	Vx, Bx, By, Bz, Np, SML, SYMH	0.62	0.82	0.68	0.75
13	Vx, Vy, Vz, Bx, By, Bz, Np, SML, SYMH	0.60	0.83	0.60	0.69
14	I_1	0.59	0.78	0.99	0.87
15	I_1, Θ	0.57	0.96	0.23	0.37
16	Vx, Vy, Vz, Bx, By, Bz, Np, SML, I_1, Θ	0.62	0.82	0.89	0.85

BIOGRAPHICAL SKETCH

Name of Author: Ruthba Yasmin

Graduate and Undergraduate Schools Attended:

The University of South Alabama, Mobile, Alabama, USA
United International University, Dhaka, Bangladesh.
Mount Holyoke College, South Hadley, Massachusetts, USA.

Degrees Awarded:

Bachelor of Arts in Physics and Mathematics, 2014.
Mount Holyoke College, South Hadley, Massachusetts, USA.

Master of Science in Computer Science and Engineering, 2024.
United International University, Dhaka, Bangladesh.

Master of Science in Electrical and Computer Engineering, 2024.
University of South Alabama, Mobile, Alabama, USA.

Awards and Honors:

University of South Alabama, Mobile, Alabama, USA.
Graduate Research Assistant (2022-2024)
IEEE – HKN Honor Society (Induction Spring 2022)
Acted as Recording Secretary (Fall 2022-Spring 2023)
United International University, Dhaka, Bangladesh.
ICT Innovation Fund for Digital Inventions, Government of Bangladesh
Mount Holyoke College, South Hadley, Massachusetts, USA
Mount Holyoke Global Grant