Predicting Geomagnetic Storms

INFO580: Research Project

Devan Mewada

Student ID: 3466459

Submitted in fulfilment of the requirements for the degree of  
Masters of Business Data Science

Department of Computer Science

Otago Business School  
University of Otago, Dunedin, New Zealand  
2025

Abstract

Geomagnetic storms pose significant risks to technological infrastructure, including power grids, satellite operations, and communication systems. Accurate and timely monitoring of these disturbances is critical for early warning and risk mitigation. This study introduces a refined approach to geomagnetic storm detection by enhancing the traditional K-index, a globally recognised but discrete measure of geomagnetic activity. Using high-resolution magnetic field data from the Eyrewell (EYR) observatory in New Zealand, this research applies a proportional bin scaling method to transform the conventional integer-based K-index into a decimal-scale continuous index. This refined K-index captures within-bin variability and provides improved sensitivity to subtle fluctuations in geomagnetic conditions.

The methodology involved preprocessing and resampling minute-level EYR magnetic field data into 3-hour blocks, calculating the first-order rate of change (|dH/dt|), and mapping disturbance magnitudes proportionally within their respective K-bins. KNN imputation was used to address missing values in the original K-index dataset. Exploratory visualisations revealed that the refined K-index tracks gradual storm build-up more effectively than its traditional counterpart. Classification models trained to detect storm-level events using the refined index achieved higher accuracy, precision, and recall, demonstrating the value of greater granularity for predictive analytics.

The findings underscore the potential of a continuous K-index for both operational forecasting and statistical modeling. By maintaining compatibility with the existing K-scale while improving temporal sensitivity, this approach addresses key limitations identified in prior literature. It also offers practical implications for regional space weather monitoring in the Southern Hemisphere, where global indices often fail to capture localised disturbances. This research contributes a transparent, physically grounded refinement framework and lays the foundation for future developments in geomagnetic storm forecasting using machine learning and time-series analysis.

Acknowledgements

I would like to sincerely thank **Associate Professor Matthew Parry** and **Associate Professor Jeremiah Deng** for their support and guidance throughout this research. Their feedback, encouragement, and advice helped me stay on track and improve the quality of my work.

I’m also grateful to the University of Otago for providing the learning environment and resources that made this research possible. I appreciate the help and support from the staff in the Department of Computer Science and the Business Data Science program.

A big thank you goes to my family for their constant support and belief in me during this journey. Your encouragement meant a lot, especially during the challenging times.

Finally, I’d like to thank my friends and fellow students for their company, ideas, and support throughout my studies. Your help and motivation made a big difference along the way.

* 1. Table of Contents

[Abstract ii](#_Toc201616831)

[Acknowledgements iii](#_Toc201616832)

[Table of Contents iv](#_Toc201616833)

[Chapter 1: Introduction 1](#_Toc201616834)

[1.1 Background 1](#_Toc201616835)

[1.2 Context 2](#_Toc201616836)

[1.3 Purposes 3](#_Toc201616837)

[1.4 Significance and Scope 4](#_Toc201616838)

[1.5 Thesis Outline 4](#_Toc201616839)

[Chapter 2: Literature Review 6](#_Toc201616840)

[2.1 Historical Background 6](#_Toc201616841)

[2.2 Limitations of the K-index and Global–Local Modeling Gaps 9](#_Toc201616842)

[2.2.1 Discrete Granularity and Scale Insensitivity 9](#_Toc201616843)

[2.2.2 K-index Binning and Proportional Refinement 10](#_Toc201616844)

[2.2.3 Geographic Averaging and Regional Blindness in Global Indices 11](#_Toc201616845)

[2.2.4 Signal Processing and Time-Series Analysis in Magnetic Field Data 12](#_Toc201616846)

[2.2.5 Forecasting Geomagnetic Storms Using Statistical and Early‑Stage Models 14](#_Toc201616847)

[2.3 Research Gaps and Motivation 15](#_Toc201616848)

[Chapter 3: Data Description 17](#_Toc201616849)

[3.1 Source and Nature of the Geomagnetic Data 17](#_Toc201616850)

[3.2 Handling Missing K-index Data 17](#_Toc201616851)

[3.3 Processing of Magnetic Field Components 18](#_Toc201616852)

[3.4 Dataset Alignment and Feature Construction 19](#_Toc201616853)

[3.5 Summary 19](#_Toc201616854)

[Chapter 4: Methodology and Research Design 20](#_Toc201616855)

[4.1 Methodology and Research Design 20](#_Toc201616856)

[4.2 Data Analyses 21](#_Toc201616857)

[4.2.1 Predictive Modelling Setup 22](#_Toc201616858)

[4.3 Refined K-Index Calculation and Comparision 26](#_Toc201616859)

[4.3.1 Rationale for Refinement 26](#_Toc201616860)

[4.3.2 Refinement Methodology 27](#_Toc201616861)

[4.3.3 Comparison to Original K-index 28](#_Toc201616862)

[4.3.4 Summary 28](#_Toc201616863)

[4.4 Limitations 28](#_Toc201616864)

[Chapter 5: Results 31](#_Toc201616865)

[5.1 K-Index Patterns and Storm Identification 31](#_Toc201616866)

[5.1.1 K-Index Storm Identification and Seasonal Breakdown (Feb–Dec 2014) 31](#_Toc201616867)

[5.1.2 Individual Storm Timelines and Local K-index Variation (3-Hour Resolution) 33](#_Toc201616868)

[5.2 Magnetic Field Behavior and Correlation with Storm Activity 34](#_Toc201616869)

[5.2.1 Detrended Magnetic Field Components (EYR-X and EYR-Y) 34](#_Toc201616870)

[5.2.2 Vector Magnitude of Magnetic Field Change ∣dH/dt∣| (Smoothed) 35](#_Toc201616871)

[5.2.3 Smoothed Rate of Magnetic Field Change (dH/dt) 36](#_Toc201616872)

[5.2.4 Monthly Patterns in Magnetic Field Change Rate (dH/dt) 37](#_Toc201616873)

[5.2.5 Magnetic Storm Spikes and Temporal Breakdown Using ∣dH/dt∣ 38](#_Toc201616874)

[5.2.6 Combined View: Magnetic Field Magnitude and K-index Trends 40](#_Toc201616875)

[5.3 Refinement of the K-Index, Binary Classification and Modelling 41](#_Toc201616876)

[5.3.1 Visual Comparison of Original vs Refined K-Index 41](#_Toc201616877)

[5.3.2 Binary Classification of Storm Events Using Original vs Refined K-index 43](#_Toc201616878)

[5.3.3 Model Comparison 46](#_Toc201616879)

[Chapter 6: Discussion 50](#_Toc201616880)

[6.1 Refinement of the K-index for Enhanced Geomagnetic Sensitivity 51](#_Toc201616881)

[6.2 Exploratory Analysis of Magnetic Field Disturbance Patterns 53](#_Toc201616882)

[6.3 Evaluation of Classification Models Using the Refined K-index 55](#_Toc201616883)

[6.4 Implications for Geomagnetic Storm Monitoring and Forecasting 55](#_Toc201616884)

[6.5 Strengths and Limitations of the Study 57](#_Toc201616885)

[Chapter 7: Conclusion 62](#_Toc201616886)

[7.1 Summary of the Study 62](#_Toc201616887)

[7.2 Key Findings and Conclusions 62](#_Toc201616888)

[7.3 Originality and Significance 63](#_Toc201616889)

[7.4 Practical Implications 64](#_Toc201616890)

[7.5 Limitations and Future Research 64](#_Toc201616891)

[7.6 Final Remarks 65](#_Toc201616892)

[References 66](#_Toc201616893)

# Introduction

This chapter introduces the background and context of the study, which focuses on improving the K-index, a key measure of geomagnetic activity using high-resolution magnetic field data from the Eyrewell (EYR) observatory in New Zealand. The primary goal is to enhance the accuracy of regional storm detection by addressing limitations in how the traditional K-index is calculated. As outlined in Section 1.3, the research applies k-nearest neighbors (KNN) imputation to fill gaps in the K-index time series, converts minute-level magnetic field data into 3-hour blocks, and uses the rate of magnetic field change (dH/dt magnitude) to assign proportional values within each K bin. This approach introduces decimal granularity to the refined K-index, while preserving compatibility with the original scale. Section 1.4 outlines the broader significance of this method, especially for localised space weather monitoring. The chapter concludes with Section 1.5, which presents the structure of the thesis and previews each chapter’s content.

## Background

Geomagnetic storms are temporary disturbances in the Earth's magnetic field, mainly caused by solar events, such as coronal mass ejections (CMEs) and fast-moving solar wind streams (Gonzalez et al., 1994). When these solar emissions reach Earth, they interact with the magnetosphere and can produce sudden and intense changes in magnetic activity. These changes can disrupt important technologies, including satellite operations, GPS systems, aviation communications, and electricity grids (Pulkkinen et al., 2017). Although these storms may only last from a few hours to several days, they can begin with little warning and have serious consequences, especially in high-latitude regions. Because of these risks, improving how we detect and predict geomagnetic storms has become a key area of research in space weather science and critical infrastructure protection.

Historically, the Kp index has been widely used to describe global geomagnetic activity. This index combines measurements from several observatories around the world and rates magnetic disturbances on a scale from 0 to 9, recorded every 3 hours (Bartels, 1957). On the other hand, the K-index is calculated at individual observatories and provides a more localised view of geomagnetic conditions. While useful for regional analysis, the traditional method for computing the K-index, originally based on analog magnetograms and semi-logarithmic scaling has limitations. These include reduced precision and challenges in applying the method to modern digital magnetic field data (Menvielle & Berthelier, 1991).

As the need for precise, region-specific monitoring of geomagnetic storms increases, especially in mid-latitude areas like New Zealand, there is growing motivation to improve existing geomagnetic indices. By using high-resolution magnetic field data from local observatories, researchers can develop more accurate and sensitive methods for detecting disturbances. These improvements not only enhance scientific knowledge of geomagnetic activity but also support practical applications, such as early warning systems and infrastructure protection.

## Context

Although geomagnetic storms occur on a global scale, their effects can differ depending on regional factors. New Zealand’s sparse network of geomagnetic observatories limits the ability to monitor and forecast these local disturbances accurately. The Eyrewell (EYR) observatory in Canterbury offers a valuable resource, providing minute-level magnetic field data (eyrx, eyry) along with corresponding 3-hourly K-index values for the year 2014. This study utilizes that detailed dataset to explore how refined statistical techniques can enhance the precision and usefulness of the K-index for detecting regional geomagnetic activity.

Traditional methods for calculating geomagnetic indices were limited by the technology and computational resources available at the time. As a result, they often produced simplified views of how the magnetic field changes, especially during storm events. These methods typically missed small but important fluctuations in the data. Today, the use of digital magnetometers allows for the collection of high-frequency measurements, making it possible to build more accurate and detailed indices that better capture the real behaviour of geomagnetic activity. However, missing values in the dataset, particularly during active periods can disrupt the analysis. Therefore, handling these gaps effectively is essential to improve the accuracy of storm detection and classification.

This thesis looks at improving the K-index, which is commonly used to measure geomagnetic activity, by using some simple data science methods. First, we clean up the data using K-Nearest Neighbors (KNN) to fill in missing values. Then, we reorganize the data into 3-hour blocks to better match how storms actually happen over time. The main goal is to make the K-index more accurate and easier to understand. After refining the index, we also try using it in a basic prediction model to see if it can help identify storm days more reliably. For this, we test a method called logistic regression. These prediction tests are just experimental and help us see if the improved index can be useful for real-world storm warning systems.

## Purposes

The central goal of this thesis is to improve the accuracy, granularity, and practical application of the K-index by leveraging high-resolution data from the Eyrewell (EYR) observatory in New Zealand. To achieve this overarching aim, the following key objectives have been formulated:

1. **Handling Missing Data:**   
   Implement K-Nearest Neighbors (KNN) imputation to address data gaps present in the EYR K-index dataset. The imputation process begins from February 2014, due to extensive data loss in January, ensuring a completer and more reliable dataset for analysis.
2. **Temporal Resampling of Magnetic Data:**   
   Convert minute-by-minute magnetic field measurements into 3-hourly averaged magnitude blocks. This step aligns the temporal resolution of the magnetic data with the conventional structure of the K-index, enabling consistent comparisons and analyses.
3. **Refinement of the K-index Scale:**   
   Develop a more granular version of the K-index by applying proportional scaling techniques within existing index bins. This method introduces decimal-level precision (e.g., K = 3.6) while maintaining the traditional 0 to 9 scale, enhancing the sensitivity of the index to minor magnetic fluctuations.
4. **Exploratory Pattern Analysis:**   
   Conduct seasonal and temporal analysis of geomagnetic activity using visual tools such as line plots and statistical summaries. These analyses help validate the refined index by ensuring its responsiveness to known storm patterns and magnetic disturbances.
5. **Binary Storm Classification Framework:**   
   Establish a threshold-based classification scheme wherein K-values equal to or exceeding 5 are labelled as indicative of geomagnetic storms. This framework facilitates evaluation of the refined index’s capacity to distinguish between storm and quiet periods with improved clarity.

## Significance AND SCOPE

This research contributes to the analysis of geomagnetic storms by refining the traditional K-index using high-resolution digital magnetic field data from the Eyrewell (EYR) observatory in New Zealand. Unlike older methods that relied on manual scaling of magnetograms, this study applies straightforward statistical techniques to improve how magnetic disturbances are represented, with a focus on increased local sensitivity and granularity. One key aspect of the work is the use of K-Nearest Neighbors (KNN) imputation to address missing values while preserving the natural structure of the data. These refinements help produce a more continuous and interpretable record of geomagnetic activity for the year 2014.

The scope of this study is deliberately narrow to ensure methodological clarity. It focuses on data from a single observatory (Eyrewell, EYR) for the year 2014, excluding January due to extensive missing values. The thesis does not include deep learning, clustering, or long-range forecasting. However, it does apply standard predictive classification models such as logistic regression, random forest, and XGBoost to evaluate the refined K-index’s ability to distinguish storm periods. These models are used not to build a full forecasting system, but to assess the utility of the refined K-index in a practical classification context.

This study contributes to improving how geomagnetic activity is measured at the regional level, specifically using data from the Eyrewell (EYR) observatory in New Zealand. By refining the K-index with basic statistical techniques, the research enhances the resolution and interpretability of magnetic field variations in a localized context. While the work does not introduce new modeling frameworks or algorithms, it highlights the potential of existing data to be better utilized through careful preprocessing and analysis. These refinements may be useful for future studies focused on space weather monitoring in similar settings.

## Thesis Outline

This thesis is structured across seven chapters, each contributing to a systematic exploration of K-index refinement using high-resolution magnetic field data from the Eyrewell (EYR) observatory in New Zealand.

**Chapter 2** presents a comprehensive literature review, critically examining the historical formulation of the K-index, known limitations in traditional geomagnetic storm monitoring approaches, and the emerging role of statistical modeling and imputation methods in enhancing data quality and interpretability. This review establishes the foundation and rationale for the refinement techniques employed in the study.

**Chapter 3** describes the dataset used in the research, focusing on the structure and characteristics of the 2014 EYR magnetic field and K-index data. It also outlines the preprocessing procedures, including the exclusion of January due to significant data gaps and the resampling of minute-resolution magnetic readings into standardized 3-hour blocks.

**Chapter 4** details the methodological framework of the study. It explains the use of the K-Nearest Neighbors (KNN) imputation method, the logic behind block-based aggregation, and the proportional scaling technique used to generate a decimal-based K-index. Each methodological choice is justified in the context of enhancing the index’s temporal resolution and storm sensitivity.

**Chapter 5** presents the results of the data processing and refinement pipeline. It includes visualizations and statistical summaries that reveal temporal trends, storm-aligned features, and the impact of proportional scaling on index granularity. These findings highlight the enhanced interpretability of the refined K-index compared to its traditional counterpart.

**Chapter 6** discusses the findings in light of the reviewed literature. It evaluates the effectiveness of the adopted methods, reflects on the strengths and limitations of the refined index, and considers the broader implications for regional space weather analysis and localized storm monitoring.

**Chapter 7** concludes the thesis by summarizing the key contributions of the research and suggesting directions for future work. These include extending the refined K-index methodology to other observatories, testing its applicability across longer time periods, and exploring its use as an input to classification models or space weather alert systems.

# Literature Review

## HISTORICAL BACKGROUND

Geomagnetic storms are major disturbances in Earth’s magnetic field caused mainly by solar events such as coronal mass ejections (CMEs) and fast-moving solar wind streams. When these solar activities reach Earth, they interact with the magnetosphere and can cause significant disruptions to satellites, communication networks, navigation systems, and power grids. Although the connection between solar activity and effects on Earth has been observed for hundreds of years, the formal study and monitoring of geomagnetic activity only began in the early 20th century. This became possible with the establishment of magnetic observatories and the introduction of standardized measurement systems, such as the K-index. (Bartels, 1957).

One of the most well-known examples of a severe geomagnetic storm is the Carrington Event of 1859, which remains the largest recorded magnetic storm in history. It was caused by a powerful solar flare and coronal mass ejection (CME), resulting in global telegraph system failures and auroras seen as far as the equator (Cliver & Svalgaard, 2004). A more recent incident occurred in March 1989, when a strong geomagnetic storm led to the collapse of the Hydro-Québec power grid in Canada, leaving millions without electricity (Bolduc, 2002). These events highlight the importance of accurate and timely monitoring systems to reduce the impact of such disturbances especially in regions like New Zealand, which are often overlooked in global space weather observation networks.

To meet the growing need for standardized monitoring, Julius Bartels introduced the **K-index** in 1938 as a localized measure of geomagnetic variability. The K-index condenses the maximum deviation in the horizontal component of the magnetic field over a 3-hour interval into a scale ranging from 0 (quiet) to 9 (extremely disturbed), using observatory-specific thresholds based on historical quiet-day curves (Menvielle & Berthelier, 1991). This made it possible to compare relative disturbance levels across locations, while accommodating local geomagnetic characteristics.

In addition to localized indices like the K-index, the Kp index was introduced to offer a broader, planetary-scale representation of geomagnetic activity. It is computed by averaging K-index values from 13 observatories, most of which are located in the Northern Hemisphere, to produce a global measure of disturbance on a scale from 0 to 9 (Gonzalez et al., 1994). The Kp index is widely used by organizations such as aviation regulators, satellite operators, and power utilities for general space weather monitoring. However, its averaging approach can obscure region-specific variations particularly in the Southern Hemisphere where magnetic disturbances may go underrepresented. As a result, Kp may not reliably reflect local geomagnetic events in regions like New Zealand, highlighting the need for localized refinement efforts using observatory-specific data.

The K-index was originally derived through manual interpretation of analog magnetograms, where trained observers estimated magnetic fluctuations by visually comparing printed recordings. This method was inherently subjective and lacked the precision and consistency needed for high-resolution analysis. Although digital magnetometers became available in the late 20th century, many observatories continued to rely on traditional techniques for some time. This slowed the transition toward more automated and objective methods capable of capturing finer-scale geomagnetic variations (Love & Remick, 2007).

The availability of high-resolution digital magnetometer data has made it possible to develop automated methods for calculating the K-index. These approaches can utilize 1-minute or even 1-second resolution readings, improving both consistency and objectivity by removing the need for manual interpretation. For instance, Regi et al. (2020) demonstrated a validated algorithm for computing K-index values at Italian observatories, a framework that is directly applicable to localized stations like New Zealand’s Eyrewell (EYR) observatory. However, even with these technical improvements, the traditional K-index remains limited by its discrete 0–9 integer scale, reducing its ability to reflect smaller but meaningful variations in geomagnetic activity.

Alongside improvements in technology, there has been a gradual expansion in the global geomagnetic observation network, particularly in regions previously underrepresented, such as New Zealand, South Africa, and South America. The Eyrewell (EYR) observatory in New Zealand now provides minute-level vector magnetometer data, making it possible to analyze geomagnetic activity with much greater temporal precision in the Southern Hemisphere. This opens the door for refining geomagnetic indices in a context that has historically been overlooked due to the dominance of Northern Hemisphere observatories.

Mac Manus, Beggan, and Thomson (2017) demonstrated that geomagnetic storm activity in New Zealand can produce observable geomagnetically induced currents (GICs), even when global indices like Kp report only moderate storm conditions. This finding emphasizes the need for regional analysis based on local observatory data, as global indices may not fully capture the severity or timing of localized magnetic disturbances.

These advancements have further exposed the limitations of using the Kp index for regional analysis. Because Kp relies on averaged data from observatories mostly located in the Northern Hemisphere it tends to smooth out localized spikes in geomagnetic activity that may be significant for infrastructure at the regional level. This averaging effect can lead to an underestimation of storm severity in areas like New Zealand.

In response, researchers have explored more flexible monitoring frameworks that incorporate local observatory data into region-specific indices (Menvielle et al., 2011). These efforts underscore the need for tailored approaches like the one taken in this study to improve the accuracy and relevance of geomagnetic monitoring in underrepresented areas.

Refining geomagnetic indices like the K-index is not just a theoretical exercise. More precise and responsive measures are increasingly important for applications such as space weather forecasting and storm detection. Predictive models, including basic classifiers like logistic regression, depend on high-resolution and accurate inputs to improve performance.

In this context, converting the traditional integer-based K-index into a continuous or proportionally scaled format allows for the detection of finer variations in geomagnetic activity that might otherwise go unnoticed. Although this study does not implement advanced forecasting models, the refined K-index produced here lays the groundwork for future modeling efforts, including binary classification of storm events.

Additionally, growing concerns about geomagnetically induced currents (GICs) affecting power infrastructure have highlighted the need for more localized storm detection tools. As noted by Mac Manus et al. (2017), geomagnetic activity in New Zealand has the potential to disrupt systems even when global indices, such as Kp, suggest moderate conditions. This underscores the importance of refining geomagnetic indices using local magnetic field data to support more region-specific early-warning systems and enhance national preparedness (Pulkkinen et al., 2017).

In summary, while the K-index has served as a key metric in global geomagnetic storm monitoring, its inherent limitations such as discrete scaling, subjective origins, and limited regional sensitivity make a strong case for modern refinement. With access to high-resolution digital data from observatories like Eyrewell (EYR), researchers can now revisit and improve traditional indices to better meet the demands of regional storm monitoring and future predictive modeling in New Zealand and beyond.

## Limitations of the K-index and Global–Local Modeling Gaps

### Discrete Granularity and Scale Insensitivity

The K-index, although still widely used in geomagnetic monitoring, presents several structural limitations that restrict its effectiveness in modern, high-resolution data environments. One key issue is its discrete, quasi-logarithmic scaling, where geomagnetic disturbances are binned into integer values from 0 to 9 over 3-hour intervals. While this design allowed normalization across observatories with differing magnetic sensitivity (Menvielle & Berthelier, 1991), it lacks the granularity needed to detect subtle fluctuations, particularly those that occur within shorter timescales or inside a single bin range.

In today’s applications such as satellite shielding, power grid protection, or real-time forecasting this limited resolution can obscure short-lived but operationally significant disturbances (Love & Remick, 2007). Moreover, the fixed 3-hour window does not always align with the dynamic nature of space weather, potentially missing transient spikes in magnetic activity that could trigger infrastructure-level consequences.

Although digital magnetometers are now common, many K-index computation protocols still adhere to threshold schemes based on analog-era practices. These legacy frameworks often ignore sub-threshold deviations and do not fully capitalize on the richer temporal data available (Love et al., 2015). Consequently, despite advances in instrumentation, the core structure of the K-index has remained largely unchanged.

To address this, researchers have proposed refinements such as introducing decimal-level scaling or using proportional deviation techniques within traditional bins. These methods aim to increase sensitivity without discarding the familiar K-index framework. Such enhancements are especially relevant for localized observatories like Eyrewell (EYR) in New Zealand, where high-resolution vector data is available at the minute level. Continuing to rely on coarse 3-hour bins risks underutilizing this valuable data and weakens the potential for precise storm detection or early-warning modeling.

In this study, a refined version of the K-index is developed to bridge this gap. By applying proportional scaling within each K-bin and aligning data with consistent 3-hour blocks, the refined index aims to preserve traditional interpretability while enhancing local sensitivity. This approach supports more accurate downstream modeling and offers a scalable framework for improving geomagnetic monitoring in underrepresented regions.

### K-index Binning and Proportional Refinement

The K-index is a long-standing measure used to quantify geomagnetic activity. It is structured as a quasi-logarithmic scale from 0 to 9, representing the maximum deviation in the horizontal component of Earth’s magnetic field over 3-hour periods. Each K-value corresponds to a fixed threshold range specific to the local observatory, historically derived through the analysis of quiet-day magnetograms (Bartels, 1957; Menvielle & Berthelier, 1991). While this format provides a standardized way to compare geomagnetic conditions across locations, it also introduces a significant limitation: all magnetic fluctuations within a given bin are treated as equivalent. This coarse granularity causes potentially meaningful variations within each 3-hour window to be lost, limiting the value of the K-index for high-resolution analysis or modeling (Love & Remick, 2007).

To address this problem, recent studies have introduced methods of proportional refinement that aim to extract more detail from the same data. One such method involves linearly scaling the observed magnetic deviation within a bin’s range, producing decimalized K-values such as 5.2 or 6.7 instead of simple integers. This enables a more continuous representation of storm intensity while still maintaining compatibility with the traditional K-index format. As Altaibek et al. (2025) show, converting K-index values into a finer-grained, decimal format improves performance in machine learning tasks, particularly for classifying geomagnetic disturbance levels and detecting anomalous patterns from ground-based magnetometer data.

This refinement technique is especially valuable for observatories like Eyrewell (EYR), which collect high-resolution magnetic field data at minute-level intervals. In such cases, restricting observations to discrete 3-hour bins can mask brief but significant geomagnetic changes. By applying proportional scaling within each bin, the refined K-index preserves more of the underlying signal structure, allowing for greater sensitivity in analysis and visualization.

Other studies support this approach. For example, Watari et al. (2009) tested fine-resolution indices in Japan’s space weather alert systems and found that continuous K-values enhanced early storm detection compared to traditional integer-scaled indices. These findings further highlight the potential for refined indices to improve real-time monitoring and predictive modeling.

Despite its advantages, implementing proportional K-index scaling poses some challenges. Because each observatory defines its own K-bin thresholds based on local geomagnetic conditions, any proportional refinement must be calibrated to that specific site. However, for regional studies such as this thesis’s focus on EYR this approach remains practical and effective. It allows for improved granularity and interpretability of geomagnetic activity without abandoning the widely understood K-index structure.

### Geographic Averaging and Regional Blindness in Global Indices

The planetary Kp index is one of the most commonly used global indicators for monitoring space weather. It is calculated by averaging local K-index values from approximately 13 subauroral observatories, the majority of which are located in the Northern Hemisphere. This uneven geographic distribution introduces a consistent hemispheric bias in the data. As a result, geomagnetic disturbances occurring in the Southern Hemisphere where observatory density is much lower are often underrepresented or entirely missed in the Kp average (Mac Manus et al., 2022).

For countries like New Zealand, this limitation is particularly critical. The Eyrewell (EYR) observatory, New Zealand’s primary geomagnetic station, is located at a mid-latitude position that can experience localized magnetic disturbances not captured in the global Kp average. Because the Kp index reflects a composite of data from primarily Northern Hemisphere observatories, events recorded at EYR may not significantly influence the planetary value, leading to underestimated warnings for regional space weather events. Smith et al. (2023) observed that several historical geomagnetic storms registered strong field deviations in EYR data that were not matched by elevated Kp readings, potentially reducing preparedness for geomagnetically induced current (GIC) risks.

A study by Mac Manus et al. (2022) further explored this gap by modeling GIC flow in New Zealand’s national power grid. Using real and simulated magnetic field profiles, the study found that even moderate mismatches between local and planetary indices could lead to transformer stress or failure, with risk exposure estimates reaching 35% despite Kp remaining below typical storm thresholds. These findings support the argument that local observatory data should play a more central role in infrastructure resilience modeling and forecasting (GFZ Potsdam, 2021).

Similarly, Smith et al. (2023) analyzed sudden commencements (SCs) and their association with GIC events in New Zealand substations. The study showed that the rate-of-change in horizontal magnetic field strength (|dH/dt|), derived from EYR’s high-frequency magnetometer data, correlated more strongly with actual transformer current readings than the global Kp index. Correlation coefficients above 0.75 during storm onset periods illustrated the superior predictive value of local high-resolution measurements compared to the global average.

Altogether, these studies highlight a significant operational blind spot created by relying solely on globally averaged indices. For Southern Hemisphere nations like New Zealand especially those with aging or exposed power infrastructure storm monitoring strategies must evolve. Adopting localized, proportionally refined indices based on minute-resolution data from observatories such as EYR offers a statistically grounded and operationally effective path forward. This transition would improve both the accuracy and timeliness of storm alerts, enabling better national preparedness for geomagnetic hazards.

### Signal Processing and Time-Series Analysis in Magnetic Field Data

Extracting meaningful information from high-resolution geomagnetic data such as that provided by New Zealand’s Eyrewell (EYR) observatory requires advanced signal processing methods capable of handling non-stationary, noisy, and multi-scale time-series behavior. Among these, the Discrete Wavelet Transform (DWT) has proven highly effective for detecting short-lived geomagnetic anomalies associated with storm onset. In particular, Kovalenko and Sitka (2020) demonstrated the utility of wavelet analysis in isolating sudden disturbances within magnetometer data, revealing that DWT can identify the sharp inflection points associated with geomagnetic storm conditions at minute-level resolution.

Following wavelet-based anomaly extraction, Empirical Mode Decomposition (EMD) paired with the Hilbert–Huang Transform (HHT) allows further decomposition of the signal into Intrinsic Mode Functions (IMFs). This approach, adaptive to both nonlinearity and non-stationarity, facilitates the separation of low-frequency diurnal cycles from high-frequency transients, which are often indicative of storm-driven activity. In the context of observatories like EYR, where geomagnetic variability is influenced by both global and local drivers, HHT provides a refined analytical lens for capturing storm-specific field fluctuations.

To ensure model assumptions are met, researchers often apply power spectral density (PSD) analysis, autocorrelation tests, and stationarity assessments such as the Augmented Dickey–Fuller (ADF) test. These techniques support the preprocessing of time-series data prior to applying predictive models such as ARIMA, Kalman filters, or machine learning-based classifiers. These validation checks are crucial for ensuring the temporal structure of the data aligns with the assumptions underlying probabilistic and autoregressive models.

One particularly informative metric is the first derivative of the horizontal magnetic field, commonly expressed as dH/dt. In a comparative evaluation of automated storm detection techniques, Djurovic and Le (2016) found that combining maximal-overlap DWT with threshold-based dH/dt analysis produced the most accurate results in detecting storm onsets. This is highly relevant to infrastructure resilience, as sharp spikes in dH/dt have been correlated with transformer current surges and relay operations in power networks during space weather events.

Crucially, the practical relevance of these methods has been validated in the New Zealand context. In a comprehensive study, Mac Manus et al. (2022) developed a geomagnetically induced current (GIC) model based on local magnetic field data, demonstrating that moderate-level Kp-index events could still produce operationally significant GICs in the country’s transformer infrastructure. Their work confirms that localized signal features such as wavelet bursts and dH/dt spikes are vital for capturing regional space weather impacts that are otherwise missed by planetary-scale indices.

By integrating multiscale signal decomposition (e.g., DWT), adaptive time-frequency analysis (e.g., HHT), and rate-of-change metrics (e.g., dH/dt), researchers can develop more precise predictors of geomagnetic storm activity. These features not only enhance early detection systems but also improve the sensitivity and responsiveness of classification models compared to coarse, integer-binned indices like the traditional K-scale.

Moreover, the limited resolution of the K-index poses barriers to probabilistic infrastructure modeling. Its integer format, with predefined bin thresholds, lacks the granularity necessary for assessing risk likelihoods across a continuous scale. This hinders the development of accurate models for transformer overheating, voltage instability, or GIC accumulation, which often require finer-scaled input features to simulate storm-induced failure scenarios under varying conditions.

Recognizing these constraints, recent efforts have focused on refining the K-index itself. Approaches include decimal scaling within existing K-bins, proportional bin deviation tracking, and even the development of entirely new indices based on real-time field gradients or localized signal morphologies (Love & Remick, 2007; Regi et al., 2020). These proposals aim to modernize geomagnetic monitoring by aligning index generation with the capabilities of high-resolution digital observatories and the requirements of modern forecasting systems.

In the case of Eyrewell (EYR), minute-resolution vector field data presents a significant opportunity. Leveraging such data to produce refined, regionally sensitive storm indicators can reduce dependence on global proxies like Kp and dramatically improve the utility of space weather alerts in mid-latitude regions like New Zealand. Ultimately, such efforts could support more robust and locally attuned risk models for national infrastructure operators such as Transpower, enabling earlier and more accurate responses to geomagnetic hazards.

### Forecasting Geomagnetic Storms Using Statistical and Early‑Stage Models

Forecasting geomagnetic storms has evolved through the application of statistical methods and early-stage machine learning models, particularly for short-term prediction windows. Traditional approaches often utilize autoregressive (AR) and ARIMA models to forecast geomagnetic indices such as Kp. These methods are based on the observed statistical persistence of geomagnetic fluctuations over short time intervals. For instance, Ojeda González et al. (2014) trained AR and ARIMA models on decades of historical Kp data, achieving correlation coefficients as high as 0.77 for predicting three consecutive 3-hour intervals. However, while these models performed adequately during stable geomagnetic conditions, they were less effective at capturing sudden, nonlinear transitions typically associated with storm onset driven by solar transients.

To address these limitations, subsequent studies have incorporated upstream solar wind data alongside index history to improve predictive power. Shprits et al. (2019), for example, developed a neural network-based forecasting model that combined real-time solar wind parameters with previous Kp values. Their results demonstrated that including solar wind inputs significantly improved prediction accuracy over short windows (6–12 hours), particularly for moderate geomagnetic disturbances. However, they also found that rare or extreme events such as those caused by coronal mass ejections (CMEs) were more difficult to predict due to limited historical examples and higher variability in space weather inputs.

Building on this, Chakraborty and Morley (2020) proposed a two-stage probabilistic forecasting framework capable of both predicting future Kp values and classifying upcoming periods as storm (Kp ≥ 5–) or non-storm conditions. Their deep learning architecture incorporated additional features such as solar X-ray flux, which improved the model’s sensitivity to CME-driven anomalies. A key advantage of their approach was the inclusion of uncertainty quantification, allowing the model to generate probabilistic rather than deterministic storm forecasts. This is particularly useful in operational contexts, where early warnings must balance lead time with confidence levels.

Despite these advancements, key challenges remain. Most forecasting models are limited to short horizons (typically less than 24 hours), and predictive performance tends to degrade rapidly over longer lead times. Machine learning approaches also depend heavily on the availability of real-time, high-quality solar wind and X-ray data resources that can be compromised by instrument malfunctions or communication latency. Moreover, a majority of existing models are trained on data from observatories located in the Northern Hemisphere, creating spatial biases that hinder their generalizability to regions like New Zealand.

Given these constraints, there is a strong motivation to develop regionally adaptive forecasting frameworks. High-resolution magnetic field data from observatories such as Eyrewell (EYR) offer an opportunity to construct localized predictive models that respond more accurately to regional storm signatures. Incorporating refined, decimal-scaled K-index values and local magnetic features alongside solar wind parameters could help overcome the limitations of global proxies like Kp. This approach would not only improve classification accuracy in mid-latitude regions such as New Zealand but also enhance early warning systems and infrastructure resilience planning.

## Research Gaps and Motivation

The preceding sections have outlined the development and limitations of the K-index, alongside efforts in storm forecasting and data refinement. Despite progress in space weather research, significant gaps persist particularly in the use of high-resolution geomagnetic data from underrepresented Southern Hemisphere observatories like Eyrewell (EYR) in New Zealand. These gaps span three key domains: index resolution, data integrity, and regional sensitivity.

A major limitation is the continued reliance on the traditional K-index, constrained to integer values between 0 and 9. While suitable for global comparisons, this coarse binning suppresses subtle fluctuations in magnetic field strength that are crucial for high-resolution analysis (Menvielle & Berthelier, 1991; Bartels, 1957). Although some studies have proposed decimal-scaled or proportionally refined versions of the K-index, such methods remain largely theoretical and rarely deployed in operational settings especially in Southern Hemisphere contexts. As highlighted by Mac Manus et al. (2022), local disturbances in regions like New Zealand can go undetected when global indices show only moderate activity. Yet these localized events may still pose substantial risks to national infrastructure.

While KNN imputation has been explored in prior environmental and geophysical studies, its focused application to regional K-index restoration particularly in Southern Hemisphere observatories like Eyrewell (EYR) remains relatively limited. This study investigates its suitability within a localized storm-monitoring context, balancing methodological simplicity with the need to preserve storm-relevant variability in high-resolution magnetic data.

Forecasting models also exhibit critical limitations. While autoregressive and neural network-based models have achieved moderate success in nowcasting planetary indices like Kp (Shprits et al., 2019; Chakraborty & Morley, 2020), they are generally trained on global datasets and optimized for short-term predictions. Their dependency on upstream solar wind measurements introduces latency and reduces responsiveness to localized ground disturbances particularly in mid-latitude nations like New Zealand, where storm impacts may not align with global index thresholds.

Perhaps the most consequential gap lies in geographic bias. Most geomagnetic storm models are built using data from Northern Hemisphere observatories, resulting in spatial skew. As Mac Manus et al. (2022) demonstrated, even moderate storms can induce harmful ground currents in New Zealand’s power grid, despite planetary indices showing only mild conditions. This regional mismatch undermines the reliability of alerts issued to local stakeholders such as power operators, aviation services, and satellite controllers.

In response to these gaps, this thesis proposes a regionally grounded approach to refining storm classification. Specifically, it applies KNN-based imputation to fill data gaps in the EYR K-index series, introduces proportional scaling to enhance the granularity of K-index bins, and integrates these refined indices with local magnetic field features into a binary classification model for storm prediction. By focusing on New Zealand's mid-latitude environment, this study aims to enhance the precision, completeness, and regional relevance of geomagnetic monitoring and forecasting tools.

# Data Description

## Source and Nature of the Geomagnetic Data

This study relies on geomagnetic field data collected from the Eyrewell (EYR) Magnetic Observatory, located near Christchurch on the South Island of New Zealand. Operated under the auspices of GNS Science, this observatory is part of the global INTERMAGNET network a collaboration of observatories dedicated to high-precision geomagnetic monitoring. The EYR station offers continuous recording of the Earth’s magnetic field at minute-level temporal resolution, specifically the horizontal components EYR-X and EYR-Y.

New Zealand’s geographic location provides a unique vantage point for monitoring mid- to low-latitude geomagnetic disturbances. While high-latitude observatories near the poles often dominate global geomagnetic indices such as Kp, mid-latitude stations like EYR are increasingly recognized for capturing disturbances that may not register strongly in polar regions but still pose significant risks to power grids and communication systems in the Southern Hemisphere (Mac Manus et al., 2017).

The primary datasets used in this study were obtained from the EYR observatory and include:

* **K-index data** (EYR\_2014\_k.txt): This file contains manually scaled K-values recorded at 3-hour intervals, covering the year 2014. Each day is divided into eight 3-hour periods, with each value representing the intensity of geomagnetic disturbance during that interval.
* **Magnetic field component data** (eyrxyz2014\_fil\_reduced.csv): This dataset contains minute-resolution measurements of the horizontal geomagnetic field components, EYR-X and EYR-Y, along with corresponding timestamps. These variables allow for high-resolution analysis of horizontal field behavior across storm and quiet periods.

This dataset pairing enables both traditional index-based monitoring and the construction of refined, data-driven features that enhance temporal resolution and storm sensitivity.

## Handling Missing K-index Data

Initial inspection of the K-index dataset revealed substantial missingness in the early part of the year, most notably a complete absence of values for January 2014. Given the size of this gap and the limitations of long-block imputation, the decision was made to exclude January entirely from the analysis. The study period was thus defined as February to December 2014.

Beyond January, scattered missing values were observed across the rest of the dataset. To address this, a K-Nearest Neighbors (KNN) imputation method was implemented. This approach offers several advantages over traditional interpolation or rolling averages, including better preservation of local trends and short-term variability that may be indicative of storm onset.

Each of the eight daily 3-hourly K-index columns was imputed independently using temporally adjacent windows as predictors. Rather than rounding the output to integers, decimal values were retained to support the study's goal of refining the K-index scale and enhancing sensitivity to sub-threshold fluctuations.

## Processing of Magnetic Field Components

The raw magnetic field data consists of minute-resolution recordings of the EYR-X and EYR-Y components. For this study, these two horizontal components were prioritized due to their relevance in capturing geomagnetic variability linked to storm events.

To align with the 3-hourly resolution of the K-index, the magnetic field data was resampled by aggregating into corresponding 3-hour blocks. Within each block, two derived features were computed:

* **Vector magnitude of the horizontal field**:

∣H∣= (EYR-X)2+(EYR-Y)2

* **Smoothed rate of change dH/dt:**   
  Calculated as the rolling average of first differences in |H| over short windows (e.g., 5–10 minutes), which helps highlight magnetic field volatility during storm intervals while filtering transient spikes.

These transformations enhance the temporal resolution of magnetic features and better reflect local geomagnetic storm patterns.

## Dataset Alignment and Feature Construction

After preprocessing, the K-index and magnetic field datasets were aligned by timestamp and grouped into synchronized 3-hour intervals. The final merged dataset contains:

* KNN-imputed K-index values (decimal-refined)
* Resampled horizontal field magnitude (|H|)
* Smoothed rate of change (dH/dt)
* Temporal markers (e.g., date, 3-hour block ID, seasonal tag)

This consolidated format supports both exploratory analysis and supervised modeling, with engineered features tuned for storm classification and pattern recognition.

## Summary

By pairing high-frequency magnetic field data with an imputed, decimal-refined K-index, this study constructs a temporally aligned, storm-sensitive dataset tailored to New Zealand’s regional conditions. Advanced feature engineering from vector magnitude calculations to smoothed rate-of-change metrics supports the development of more granular and adaptive models for geomagnetic storm analysis. The preprocessing pipeline preserves physical signal characteristics and mitigates information loss due to missing data, providing a strong foundation for subsequent modeling and classification steps.

# Methodology and Research Design

This chapter describes the methodological framework adopted to achieve the objectives outlined in Chapter 1. The study aims to refine the traditional K-index using high-resolution data from the Eyrewell (EYR) magnetic observatory in New Zealand and to evaluate its predictive utility for detecting geomagnetic storms. Given the data-driven and modeling-oriented nature of this project, the research adopts a quantitative, exploratory approach built on statistical analysis, signal processing, and machine learning techniques. The selected methodology and design reflect the goal of improving regional storm detection accuracy through data transformation, imputation, and predictive modeling.

## Methodology and Research Design

This research adopts a quantitative methodology with an exploratory design. The study is based on the analysis of historical geomagnetic field measurements and K-index values collected during the year 2014 from the EYR observatory. There are no human participants or survey instruments involved.

The overall methodology involves:

* **Preprocessing and data transformation** of high-frequency (minute-level) magnetic field data.
* **Imputation of missing values** in the manually recorded K-index data using the K-Nearest Neighbors (KNN) algorithm.
* **Refinement of the K-index**, using proportional scaling within traditional K bins to introduce sub-index resolution.
* **Statistical and signal-based feature extraction**, including rate-of-change metrics such as dH/dt.
* **Predictive modeling**, where both the original and refined K-index are used to train binary classification models to identify storm vs non-storm periods.

The study uses supervised machine learning methods specifically Random Forest and Logistic Regression to evaluate classification performance. The dependent variable in these models is the binary storm classification (K ≥ 5–), while independent features include both statistical and time-derived characteristics of the magnetic field and/or refined index values.

This design is well-suited for answering the core research questions: (1) whether a refined K-index improves storm classification performance compared to the traditional K-index, and (2) whether localized magnetic field dynamics can be effectively used for predictive modeling in the absence of global indices like Kp or upstream solar wind measurements.

## DATA ANALYSES

The analysis workflow begins with cleaning and preprocessing the raw datasets. January 2014 was excluded from the analysis due to a large block of missing values in the K-index. The remaining data from February to December was retained for all modeling and analysis steps.

The imputation of missing K-index values was performed using the KNN imputer with a Euclidean distance metric, considering temporal proximity and available contextual values. This method was chosen due to its ability to preserve local variability, which is important for storm detection.

The magnetic field component data (EYR-X and EYR-Y) was resampled from 1-minute to 3-hour resolution to align with K-index intervals. Derived metrics included:

* Rolling mean and standard deviation
* First derivative (dH/dt) of horizontal magnetic field magnitude
* Spike detection features during storm windows

For classification, the K-index was transformed into a binary variable: 1 for storm (K ≥ 5–) and 0 for non-storm. The dataset was split into training and test sets using stratified sampling. Models were trained and evaluated using accuracy, precision, recall, F1-score, and ROC-AUC. Cross-validation was used to ensure model robustness.

Comparison was made between:

* Original K-index-based models
* Refined K-index-based models using decimal-scaled bins

This two-track approach was essential for assessing the added value of K-index refinement.

### PREDICTIVE MODELLING SETUP

To test whether the refined K-index improves forecasting of geomagnetic activity and two supervised regression models were created one using the original K-index and one using the refined K-index. The goal was to see how well short-term magnetic field changes and time-based features could predict K-index values during the test month, December 2014. This setup simulates near real-time prediction of storm conditions.

**Modeling Objective**

Two parallel modeling pipelines were constructed:

* One using the original K-index values (k\_value) as the prediction target
* One using the refined K-index values (refined\_K) as the prediction target

The results of both were then compared using standard regression metrics.

**Feature Engineering**

1. **Original K-index Model:**   
   The predictors used included:

* k\_lag1, k\_lag2, k\_lag3: Lagged values of the original K-index
* eyrx, eyry: Raw magnetic field components from the EYR observatory
* eyrx\_diff, eyry\_diff: (velocity/gradient) in the magnetic components

These features capture recent geomagnetic dynamics that may influence short-term K-index changes.

1. **Refined K-index Model:**   
   In addition to the lagged and magnetic field features mentioned above, the refined model included:

* refined\_k\_lag1, refined\_k\_lag2, refined\_k\_lag3: Lagged values of the refined K-index
* dHdt\_magnitude: The scalar magnitude of the horizontal magnetic field's rate of change
* H\_bin\_frac: The proportional position of the disturbance within each K-bin
* hour, month: Temporal context to account for diurnal and seasonal effects

These extra predictors leveraged the finer granularity available in the refined K-index pipeline and allowed the models to capture intra-bin variability more accurately.

**Prediction Target**

* Original Model: k\_value (integer or half-integer scale K-index)
* Refined Model: refined\_K (decimal-scaled K-index capturing intra-bin variability)

**Model Choices**

Three regression models were used for both pipelines:

1. Linear Regression: To establish a baseline performance and understand linear relationships.
2. Random Forest Regressor: A non-parametric, ensemble-based method well-suited for capturing non-linear relationships.
3. XGBoost Regressor: A gradient-boosted decision tree algorithm known for its performance on structured data and ability to model complex interactions.

Each model was initialized with fixed hyperparameters for consistency:

* RandomForestRegressor: 100 estimators, random\_state=42
* XGBRegressor: 100 estimators, learning rate = 0.1, max depth = 3

These values were selected based on common defaults for performance and interpretability. While extensive hyperparameter tuning (e.g., via grid search or Bayesian optimization) was outside the scope of this study, the selected parameters provided a meaningful comparison between index types.

**Evaluation Strategy**

Models were evaluated using:

* R² Score (coefficient of determination): Measures explained variance
* Root Mean Squared Error (RMSE): Measures prediction accuracy
* Visual overlay plots comparing actual vs predicted K-index values across the December 2014 timeline

The training and testing data were split based on time (not randomly shuffled) to preserve the real-world sequence of data. This avoids giving future information to the model, which would not happen in real-time forecasting.

**Key Observations and Differences**

* The refined K-index model incorporated more predictors (e.g., H\_bin\_frac, dHdt\_magnitude) that directly related to the magnetic disturbance intensity used in the refinement process. This resulted in higher granularity in predictions.
* The original K-index model relied more on coarser labels, which may mask minor but meaningful fluctuations in storm buildup.
* By comparing both pipelines side-by-side, the modeling setup offered a robust test of whether the refined decimal-scaled K-index improves short-term forecasting of geomagnetic activity.

### Temporal Feature Construction and Look-Back Strategy

Temporal dynamics play a crucial role in the short-term prediction of geomagnetic disturbances. To incorporate recent historical trends and diurnal/seasonal patterns, a combination of lag-based and calendar-based features was constructed as part of the modeling pipeline.

**Look-Back Window Design**

Both the original and refined K-index models implemented a look-back window comprising the three most recent 3-hour intervals, effectively incorporating 9 hours of temporal memory. The following lagged variables were included:

* k\_lag1, k\_lag2, k\_lag3: Lagged values of the original K-index
* refined\_k\_lag1, refined\_k\_lag2, refined\_k\_lag3: Lagged values of the refined K-index

This look-back window was selected to capture the short-term progression of geomagnetic activity leading up to each prediction point. It allows the models to recognize buildup patterns that precede storm-level disturbances without extending the window so far back as to dilute immediate predictive signals.

**Magnetic Field Feature Usage**

Magnetic field data was sourced from the EYR-X and EYR-Y components of the Eyrewell observatory. For each 3-hour window, the following magnetic features were used as predictors:

* eyrx, eyry: Resampled (averaged) horizontal magnetic field components at the current 3-hour interval
* eyrx\_diff, eyry\_diff: First-order differences approximating directional change or gradient
* dHdt\_magnitude: Euclidean magnitude of the horizontal field’s rate of change, calculated as:

∣dH/dt∣ = √(d eyrx/dt)2 + (d eyry/dt)2

These features reflect the instantaneous magnetic field dynamics at the prediction time, providing spatial context to complement the temporal signal from K-index lags. Notably, the magnetic inputs were not lagged in this version of the model to preserve interpretability and reduce dimensionality.

**Temporal Context Features**

To account for broader patterns in storm occurrence, two temporal context features were added:

* hour: Encodes the 3-hour interval within the day (e.g., 0, 3, ..., 21)
* month: Encodes seasonal variation across the calendar year

These features allowed the models to adapt to known diurnal cycles and seasonal tendencies in geomagnetic activity, particularly important for improving performance in rare-event detection such as storms.

**Integration with Refined K-index**

The refined K-index model uniquely benefited from this temporal structure. The decimal-based refinement provided finer resolution between storm and non-storm conditions, which, when combined with lag and magnetic field features enabled the model to capture subtle transitions more effectively. The smoother transitions in refined\_K improved its utility as both a predictive target and a temporal signal.

This feature engineering strategy balanced recent memory (via look-back lags), real-time context (via magnetic field components), and longer-range periodicity (via hour and month), thereby enhancing the model's capacity to forecast geomagnetic disturbances with higher granularity and sensitivity.

**Conclusion of Modeling Design**

The modeling structure reflects a careful balance between domain insight (e.g., inclusion of magnetic field change metrics), statistical rigor (time-aware split and cross-model comparison), and practical simplicity (fixed model configurations). This setup allowed for a fair and interpretable comparison of the predictive power of the original vs. refined K-index, contributing valuable insight into the potential benefits of index refinement in operational space weather monitoring.

## REFINED K-INDEX CALCULATION AND COMPARISON

This section outlines the development of a refined K-index using proportional mapping within disturbance bins, based on magnetic field variation data from the Eyrewell (EYR) observatory. The standard K-index, traditionally scaled from 0 to 9 in integer or half-integer steps, offers a limited resolution of geomagnetic activity. This coarse quantization may obscure meaningful intra-bin dynamics critical for real-time forecasting. To address this limitation, a refined K-index was computed by proportionally scaling the rate of magnetic field change (|dH/dt|) within each K-bin, resulting in a continuous decimal-scale index. The approach preserved the original K-index structure while enriching it with additional resolution.

### Rationale for Refinement

The K-index is a widely used measure of geomagnetic activity. It is calculated based on the peak range of fluctuations in the horizontal magnetic field (H) over fixed 3-hour windows. While its simplicity and standardized use across observatories make it effective for high-level monitoring and communication, the K-index's discrete nature introduces resolution limitations. In practice, each K-level represents a broad bin of disturbance magnitudes, which can flatten out meaningful within-bin variations.

This becomes problematic in predictive modeling, where subtle transitions between quiet and stormy conditions are valuable signals. For example, under the standard scale, both a moderate and a strong K = 5– event are treated equally, even though the actual field disturbance differs significantly. This lack of granularity can reduce the sensitivity of models trying to forecast storm onset, duration, or strength.

To address this issue, a proportional refinement method was introduced. Rather than discarding or altering the original scale, this method adds a decimal-level enhancement to each K-bin based on the actual intensity of magnetic disturbance during that period. The result is a refined K-index that remains compatible with existing K-based frameworks while offering improved resolution for analysis and forecasting.

### Refinement Methodology

**Compute Magnetic Disturbance Magnitude**

The disturbance magnitude was estimated using the first-order differences of the magnetic field components, divided by the time interval (180 minutes) to approximate the rate of change. The magnitude of disturbance was then computed as:

**Step 1: Deriving Magnetic Disturbance (|dH/dt|)**   
|dH/dt| = √((d(eyrx)/dt)2 + (d(eyry)/dt)2)

This was implemented by calculating the differences in eyrx and eyry, dividing by 180 to obtain the per-minute rate, and then computing the Euclidean magnitude.

**Step 2: Assigning Base K-bins**   
Each observation was assigned a base integer K-bin using the floor of its original k\_value. To maintain validity, the bins were clipped between 0 and 8.

H\_bin = floor(k\_value), with bounds clipped to [0, 8]

**Step 3: Determine Bin Ranges**

For each bin, the minimum and maximum values of |dH/dt| were computed using a group-wise aggregation. These values served as reference boundaries for scaling disturbance magnitudes proportionally within the bin:

 H\_min = min(|dH/dt|), H\_max = max(|dH/dt|) for each H\_bin

**Step 4: Compute Proportional Offset**

The fractional offset for each record was calculated using the formula:

H\_bin\_frac = (|dH/dt| - H\_min) / (H\_max - H\_min)

To avoid boundary overlap and instability, this value was clipped to the range [0, 0.999]. In cases where H\_min = H\_max, the denominator was replaced with NaN and subsequently filled with 0.

**Step 5: Calculate Refined K-index**

The final refined K-index was computed by summing the integer bin base and the proportional offset:

refined\_K = H\_bin + H\_bin\_frac

This results in a decimal-valued K-index, where a k\_value = 5 could yield refined values such as 5.22, 5.67, or 5.99 depending on the local magnetic activity level.

**Step 6: Edge Case Handling**

Entries with undefined fractional components due to missing data or zero bin range were set to 0. All NaN values were filled to ensure completeness.

### Comparison to Original K-index

The refined K-index introduces several advantages over the traditional version. First, it captures subtle fluctuations in geomagnetic activity that would otherwise be flattened by the original binning approach. Second, it provides smoother temporal transitions, which is particularly important in storm build-up periods where the original K-index may jump abruptly. Third, the refined version enables more precise input or target variables for statistical and machine learning models, improving sensitivity to patterns leading up to storm events.

Visualization of both indices (original and refined) during storm-active periods such as December 2014 illustrates the benefits clearly. The refined K-index exhibits smoother curves and anticipates sharp transitions, which aligns better with physical changes observed in the magnetic field.

### Summary

This refinement process introduces a proportional scaling framework that enriches the discrete K-index without compromising its established utility. By leveraging the magnitude of magnetic field change (|dH/dt|) and mapping each observation within its respective disturbance bin, the refined K-index offers enhanced granularity, continuity, and modeling utility. This approach provides a statistically grounded enhancement to traditional geomagnetic monitoring and supports downstream applications such as real-time forecasting and anomaly detection.

## LIMITATIONS

Several methodological limitations should be acknowledged:

1. Temporal coverage: The current analysis is based on geomagnetic data from the year 2014. This time frame was selected to allow for manageable scope and focused methodological development. While additional data from other years is available and could enhance the generalizability of findings, the 2014 dataset is sufficient to demonstrate the viability of the proposed approach. Future extensions may incorporate multi-year data to explore broader temporal trends.
2. Missing data: The exclusion of January due to missing K-index values and the reliance on imputation for other gaps may introduce uncertainty. While KNN was selected to preserve local dynamics, the imputed values are still estimates.
3. Storm imbalance: Storm periods (K ≥ 5–) are relatively rare, leading to a class imbalance that may affect model performance. Stratified sampling and performance metrics were used to mitigate this, but it remains a constraint.
4. Model generalizability: Since models were trained on a single-year, single-observatory dataset, generalization to other years or regions may require retraining or recalibration.
5. Lack of upstream solar wind data: This study focuses on local ground-based indicators. The absence of upstream solar drivers may limit early forecasting capabilities.
6. Resolution trade-off in refined K-index: While the refined K-index introduces greater granularity through decimal scaling, this method assumes a linear relationship between magnetic field change magnitude and K-index refinement within bins. The use of proportional scaling may not fully reflect the logarithmic nature of the original K-index formulation, and the refined values should be interpreted with this simplification in mind.
7. Potential bias from imputation assumptions: The KNN imputation approach relies on Euclidean distance between time-similar rows, assuming that geomagnetic behaviour is locally consistent. However, during periods of high volatility or rare events, KNN may underestimate true variance. This could subtly influence downstream modelling or storm classification accuracy.
8. Limited model tuning exploration: While classification models were trained and validated using standard metrics, exhaustive hyperparameter optimization (e.g., grid search across all model configurations) was not the central focus. Future work could explore additional algorithms or fine-tune model parameters more extensively.

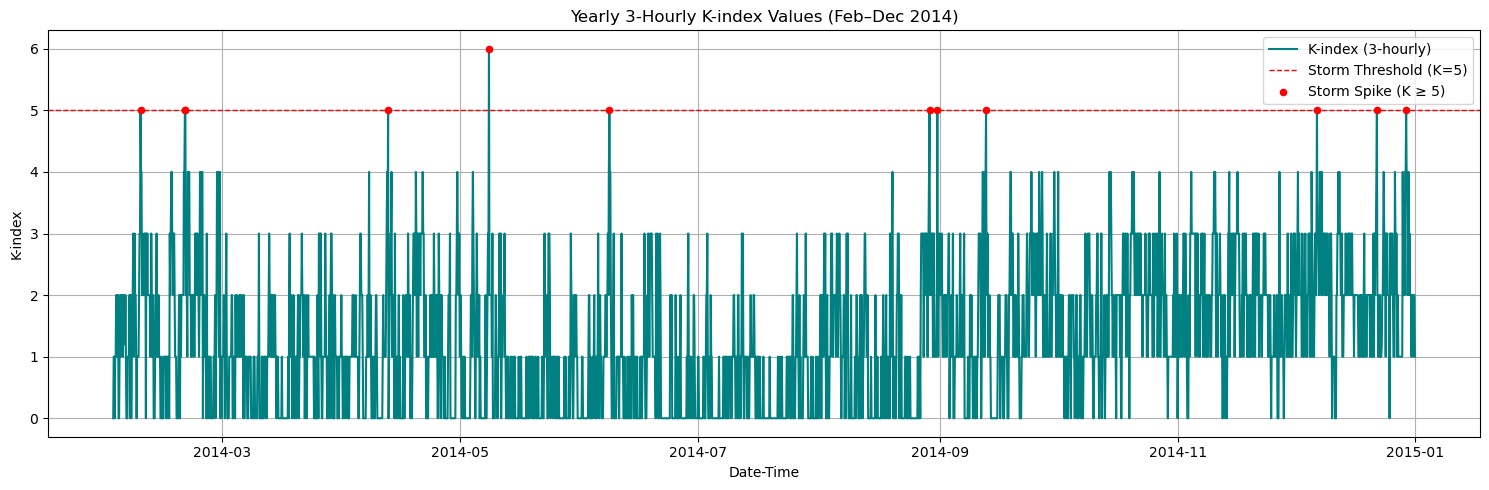
Despite these limitations, the research offers a practical contribution to regionally adapted space weather monitoring by showcasing how ground-based data and refined indices can support storm detection and infrastructure risk assessment in under-monitored Southern Hemisphere contexts.

# Results

This chapter presents the main findings from our analysis of the K-index and magnetic field data collected from the Eyrewell (EYR) observatory. The goal is to explore how well the refined K-index and magnetic field features capture the timing and intensity of geomagnetic storms in New Zealand. We start by looking at patterns in the original and refined K-index values over time, including seasonal trends and storm activity. We then examine the behavior of the magnetic field data, focusing on how it changes during storm periods. By using higher-resolution data and new processing techniques, we aim to show how these improvements can lead to better storm detection and stronger signals for future modeling. This analysis helps set the stage for building more accurate storm forecasting tools that are suited to New Zealand’s local conditions.

## K-Index Patterns and Storm Identification

### K-Index Storm Identification and Seasonal Breakdown (Feb–Dec 2014)

**Figure 5.1.1(a)** *Yearly 3-hourly K-index values recorded at Eyrewell (EYR) from February to December 2014. Red dots indicate storm spikes (K ≥ 5), and the dashed line shows the storm threshold.*****

**Figure 5.1.1** shows how geomagnetic activity changed throughout 2014 using 3-hourly K-index values from the Eyrewell (EYR) observatory. Data from January was not used because it had too many missing values. Each vertical bar in the plot shows the level of disturbance in Earth’s magnetic field for a 3-hour period. The red dashed line marks the threshold for when a geomagnetic storm is considered to have started (K = 5), and red dots highlight the times when this threshold was met or passed.

In total, there were **11 storm spikes** between February and December. These were not evenly spaced across the year. After examining each month closely, storm-level spikes were found in **February, April, May, June, August, September**, and **December**. The strongest storm occurred in **May**, where the K-index peaked at **6**. Most of the storm events were brief, lasting only a single 3-hour period. Meanwhile, **March, July, October**, and **November** stayed below the storm level, although they did show small variations in K-index values.

**Figure 5.1.1(b)** *Monthly 3-hourly K-index values recorded at the Eyrewell (EYR) observatory from February to December 2014. Red dots indicate individual 3-hour intervals where K-index values reached or exceeded storm levels (K ≥ 5). The dashed red line represents the standard geomagnetic storm threshold.*

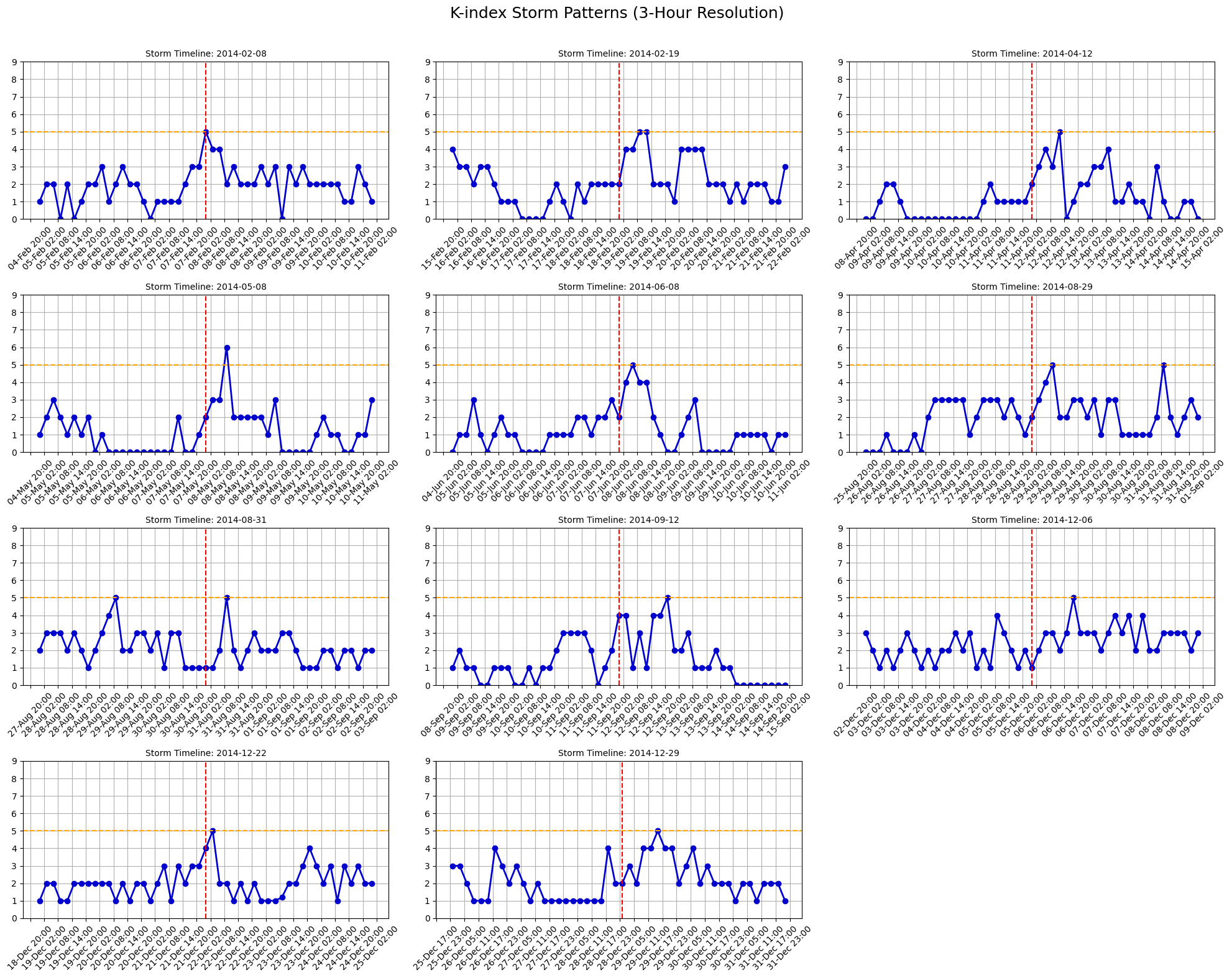
****

**Figure 5.1.1(b)** presents a monthly breakdown of the same 3-hourly K-index values. This layout provides finer temporal resolution, allowing for detailed inspection of intra-month activity and seasonal variability. Vertical fluctuations are visible in months such as May, August, and December, which each featured multiple storm-level peaks. May 2014 had the strongest single spike (K = 6), while December saw three separate spikes, making it the most storm-active month.

In contrast, **March**, **July**, **October**, and **November** remained below the storm threshold throughout, suggesting relatively geomagnetically quiet periods. Months like **February**, **April**, **June**, and **September** featured isolated disturbances, with one or two spikes reaching storm levels. The variability in storm frequency and intensity across months supports the need for seasonally aware detection systems and offers a foundation for testing refinement models later in the study.

### Individual Storm Timelines and Local K-index Variation (3-Hour Resolution)

**Figure 5.1.2.** *K-index Storm Patterns for All Identified Events (3-Hour Resolution, Feb–Dec 2014).*



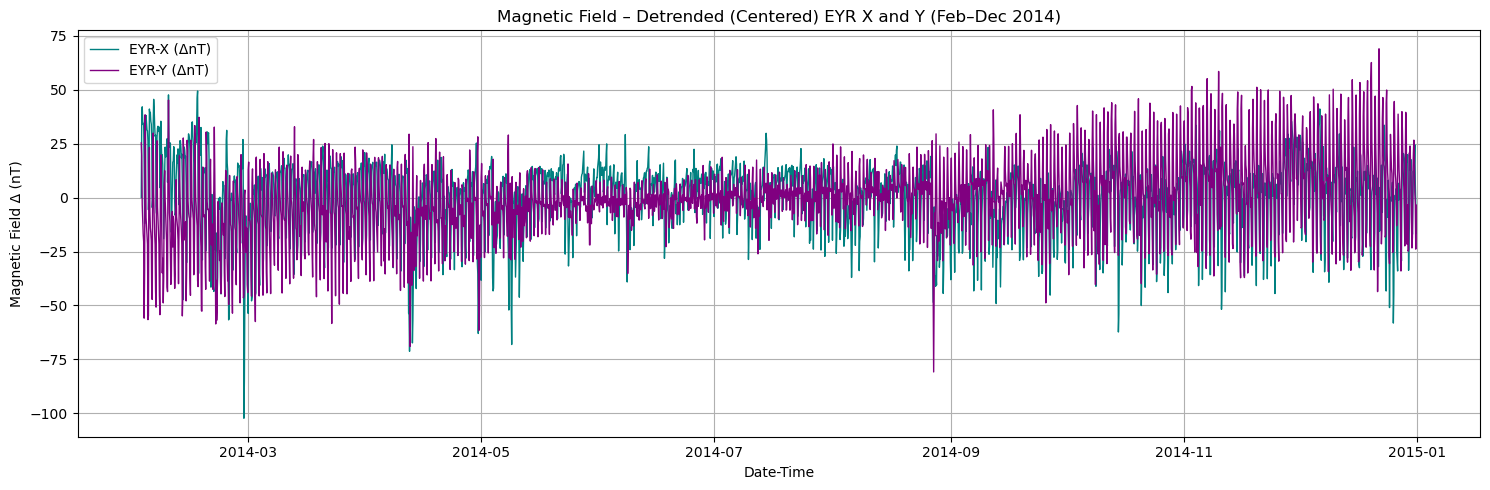
**Figure 5.1.2** presents detailed time series plots for each of the 11 identified geomagnetic storm events between February and December 2014, based on 3-hourly K-index values from the Eyrewell (EYR) observatory. Each subplot captures a focused window of three days centered around the storm spike, allowing a close examination of storm onset, peak intensity, and post-disturbance recovery. The red dashed vertical line marks the exact timestamp when the K-index first reached or exceeded the storm threshold (K ≥ 5), while the orange horizontal line serves as a visual reference for that threshold.

Across the **11 timelines**, **14 distinct spike points** reached or exceeded K = 5, suggesting that some days experienced more than one storm-classified interval within a short span. For instance, February 2014 and December 2014 show multiple elevated peaks, indicating either closely spaced substorms or extended magnetic activity near the threshold. In contrast, most other months reveal brief spikes, often isolated within an otherwise calm geomagnetic environment. **May 8th** stands out as the strongest event of the year, with a peak K-index of 6 followed by a steep return to moderate levels within one 3-hour cycle.

## Magnetic Field Behavior and Correlation with Storm Activity

### Detrended Magnetic Field Components (EYR-X and EYR-Y)

**Figure 5.2.1** *Detrended horizontal magnetic field components* (*EYR-X* and *EYR-Y*) *recorded at the Eyrewell (EYR) observatory from February to December 2014. The data have been centered around the mean to highlight deviations in geomagnetic field strength. Spikes and asymmetries in the signal reflect variations in geomagnetic activity over time.*



**Figure 5.2.1** displays the detrended horizontal magnetic field components from the Eyrewell (EYR) observatory, covering February to December 2014. Both EYR-X and EYR-Y values have been centered by subtracting their respective means to isolate fluctuations and highlight deviations from baseline levels.

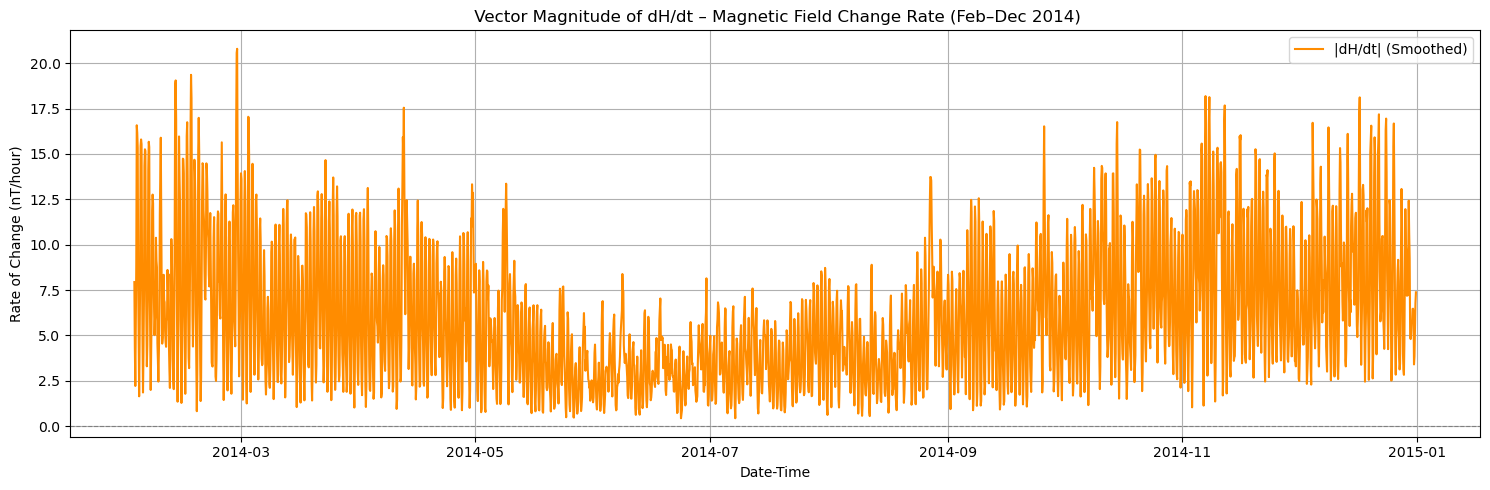
The resulting plot reveals a clear pattern of magnetic field variability throughout the year. The X-component generally shows sharper downward excursions, while the Y-component exhibits wider fluctuations and increasing amplitude toward the later months of 2014. These variations suggest that geomagnetic conditions were more active in the second half of the year, potentially indicating more frequent or intense solar wind interactions.

Distinct spikes and short-lived surges are evident in both components, some aligning with known storm periods identified from the K-index analysis. The asymmetric behavior between X and Y further implies directional variability in the magnetospheric response, which is common during disturbed geomagnetic conditions.

Overall, this figure provides an essential foundation for analyzing the magnetic field behavior in more detail. By detrending and centering the components, subtle shifts in field intensity become visible, supporting further exploration into magnetic field magnitude and rate-of-change patterns presented in the following sections.

### Vector Magnitude of Magnetic Field Change ∣dH/dt∣| (Smoothed)

**Figure 5.2.2** *Smoothed vector magnitude of the rate of change in the horizontal magnetic field (∣dH/dt∣)recorded at the Eyrewell (EYR) observatory between February and December 2014. The plot shows the combined 3-hourly rate of change of both the X and Y magnetic components, computed as the square root of the sum of their squared derivatives.*



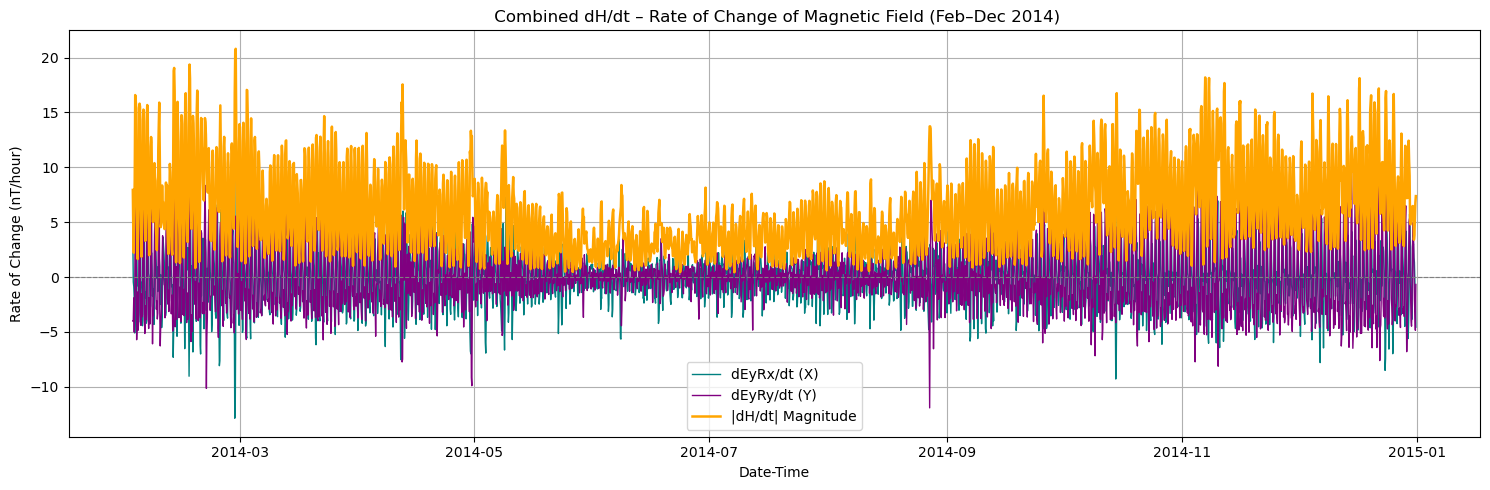
This figure provides a compact view of magnetic field dynamics over time, capturing how quickly the horizontal field intensity changes across the year. The data have been smoothed to better reveal underlying patterns by reducing short-term noise. Peaks in ∣dH/dt∣ magnitude reflect intervals of strong geomagnetic variability, which are often associated with the onset or development of geomagnetic storms.

The pattern shows several distinct features. Elevated rates of change are most prominent in early February–March and again from October through December, corresponding to seasons where storm spikes were previously detected using K-index values. The mid-year period, particularly from May to August, exhibits relatively lower rates of change, indicating quieter geomagnetic conditions.

This visualization reinforces the usefulness of ∣dH/dt∣ as a supporting metric for identifying storm-related magnetic disturbances. By focusing on the magnitude of field changes rather than raw directional values, this representation emphasizes the intensity of geomagnetic fluctuations, making it particularly effective for recognizing storm build-up and recovery phases.

### Smoothed Rate of Magnetic Field Change (dH/dt)

**Figure 5.2.3** *Smoothed rate of magnetic field change (dH/dt) derived from horizontal components recorded at the Eyrewell (EYR) observatory between February and December 2014. The plot includes the rate of change in both X (dEyrX/dt) and Y (dEyrY/dt) components, along with the combined magnitude (|dH/dt|).*

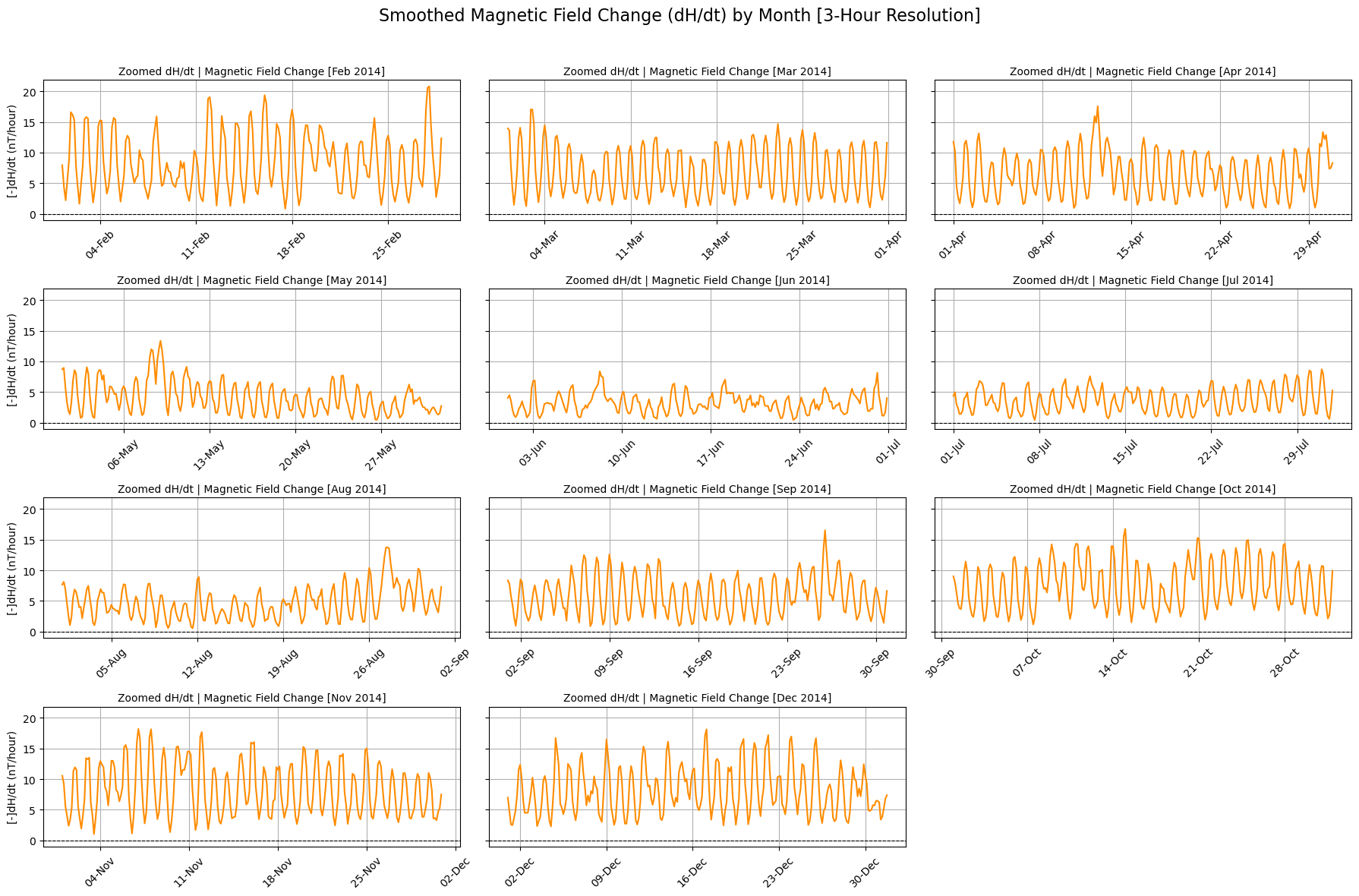


This figure illustrates the temporal evolution of the magnetic field’s rate of change using 3-hour resolution data across the 11-month observation window. The blue and purple lines represent smoothed first differences of the EYR-X and EYR-Y components, respectively, while the orange line shows the smoothed vector magnitude combining both axes. Smoothing was applied to reduce short-term noise and emphasize gradual structural changes.

Higher rates of change are observed during geomagnetically active periods, such as in March, May, and late November–December, aligning with known storm spikes identified in earlier K-index analysis. Notably, the magnitude (|dH/dt|) curve consistently rises during storm days, confirming the close relationship between sudden magnetic field fluctuations and K-index spikes. This chart provides a valuable perspective on the dynamic behavior of the field and supports further investigation into magnetic precursors of storm events.

### Monthly Patterns in Magnetic Field Change Rate (dH/dt)

**Figure 5.2.4** *Monthly smoothed vector magnitude of magnetic field change (dH/dt) at 3-hour resolution, plotted separately for each month from February to December 2014 using data from the Eyrewell (EYR) observatory. Each subplot highlights temporal fluctuations in geomagnetic activity intensity, enabling seasonal comparison of magnetic field dynamics.*

**

**Figure 5.2.4** presents monthly plots of the smoothed vector magnitude of magnetic field change (∣dH/dt∣) at 3-hour resolution from **February to December** 2014, based on data from the Eyrewell (EYR) observatory. These subplots support comparative analysis of short-term geomagnetic activity across the year by isolating intra-month variability.

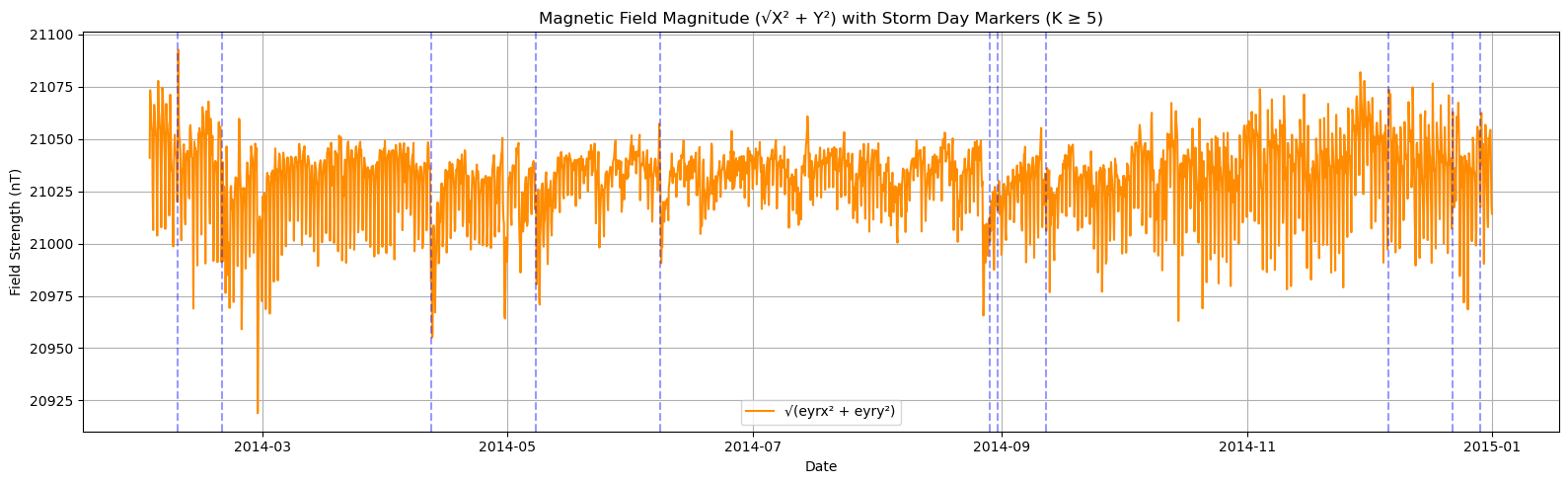
Distinct patterns are visible across different months. **February and November** exhibit higher-amplitude oscillations with prominent peaks, reflecting periods of increased geomagnetic variability and likely substorm activity. In contrast, **June and July** show relatively calm behavior, with lower-magnitude and smoother fluctuations, indicating a seasonal decline in storm intensity.

**April, September, and October** reveal periodic bursts of moderate activity interspersed with quiet phases, suggesting transitional geomagnetic conditions influenced by variable solar wind input. Meanwhile, **March, May, and August** display steady mid-range variations with fewer spikes, possibly linked to minor disturbances or regular diurnal patterns.

Collectively, these month-wise views reveal seasonal trends in magnetic field dynamics, reinforcing the importance of temporal granularity for understanding geomagnetic storm development and supporting improved forecasting model design.

### Magnetic Storm Spikes and Temporal Breakdown Using ∣dH/dt∣

**Figure 5.2.5(a)** *Magnetic field magnitude (√(eyrx² + eyry²)) across February to December 2014 with dashed vertical lines marking identified geomagnetic storm days (K-index ≥ 5).*

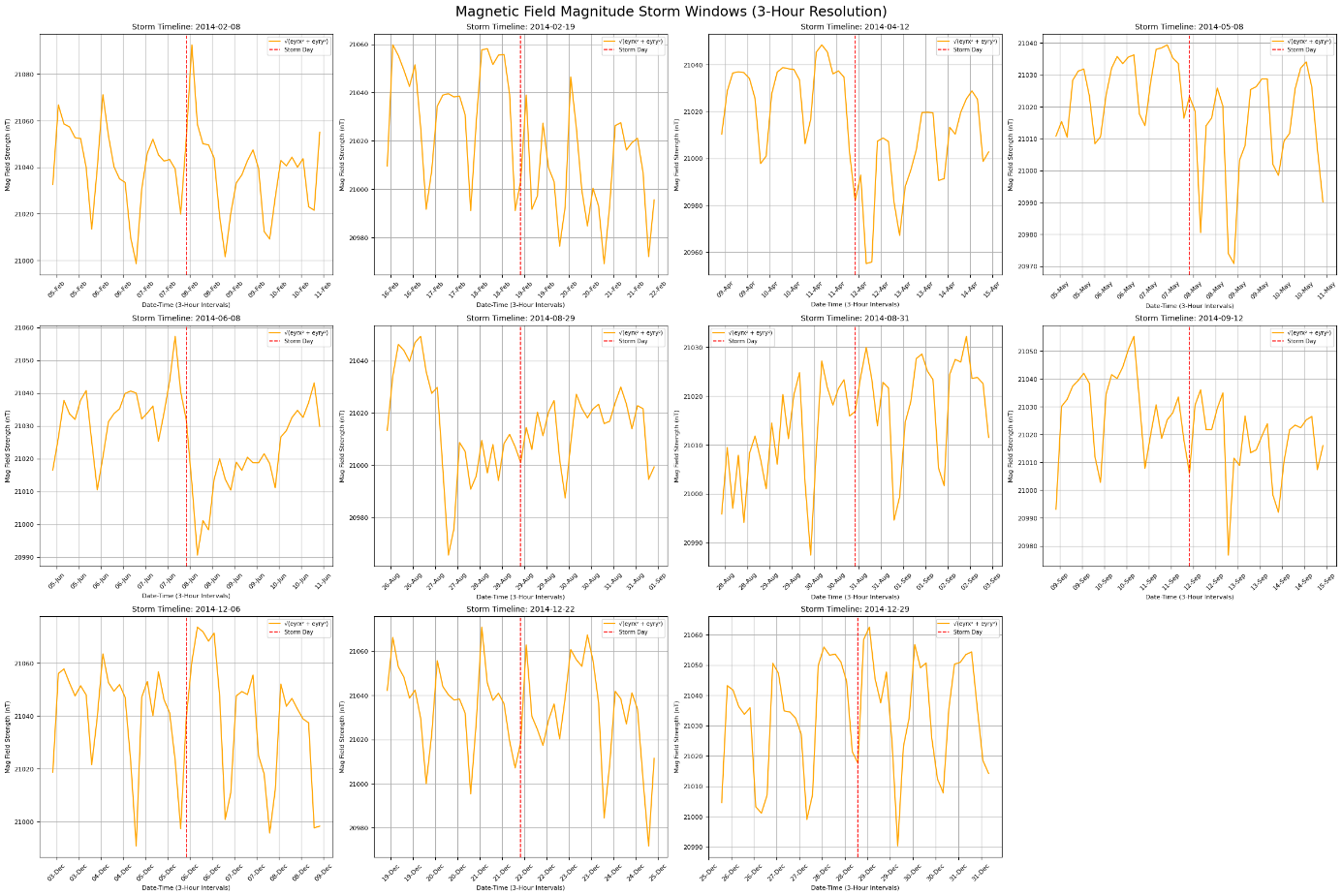


This figure plots the full-year time series of the magnetic field magnitude at 3-hour resolution, spanning February to December 2014. The vertical dashed blue lines indicate the 11 dates on which the K-index exceeded the storm threshold (K ≥ 5). This visualization provides a macroscopic view of magnetic field behavior across the year, making it easier to detect global patterns and temporal clustering of high-activity periods.

Several storm intervals, including early February, mid-May, late August, and December, exhibit distinguishable magnetic deviations that align closely with K-index storm spikes. For example, on April 12 and December 29, the field magnitude shows sudden drops or disturbances coinciding precisely with marked storm days. In contrast, some storms (e.g., June 8 or September 12) exhibit more muted or less obvious fluctuations in the overall magnitude signal.

The subtle but structured nature of these responses suggests that while K-index spikes often coincide with perturbations in magnetic magnitude, the intensity and sharpness of those disturbances vary significantly. This observation supports the importance of examining localized patterns in addition to long-term trends.

**Figure 5.2.5(b)** *Magnetic field magnitude (√X² + Y²) from February to December 2014 at the Eyrewell (EYR) observatory. Blue dashed lines show storm days where the K-index reached 5 or higher.*



This plot shows how the strength of the horizontal magnetic field changed over time during the year. Instead of looking at the X and Y components separately, we’ve combined them into one overall value, making it easier to see how strong the magnetic field was at each 3-hour step.

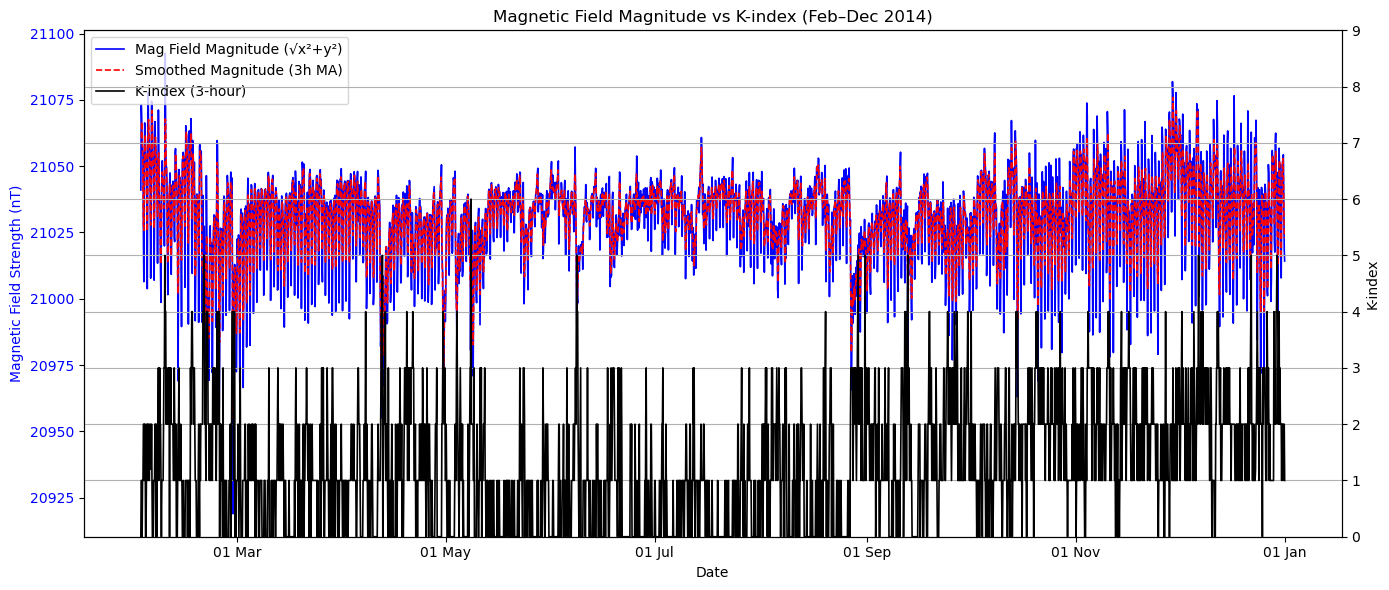
The blue dashed lines represent storm days, days when the K-index showed strong geomagnetic activity. You can see that in many cases, there are noticeable spikes in the magnetic field magnitude close to these storm days, especially in months like March, June, September, and November. These spikes show that the magnetic field was reacting strongly, likely due to incoming solar activity.

Meanwhile, quieter months like May and July show a more stable magnetic field with fewer spikes and storm markers, suggesting calm space weather conditions during those times.

Using this combined magnitude view helps simplify the data and makes it easier to spot patterns. It also backs up what we saw in earlier sections: sharp changes in the magnetic field are closely linked to geomagnetic storms.

### Combined View: Magnetic Field Magnitude and K-index Trends

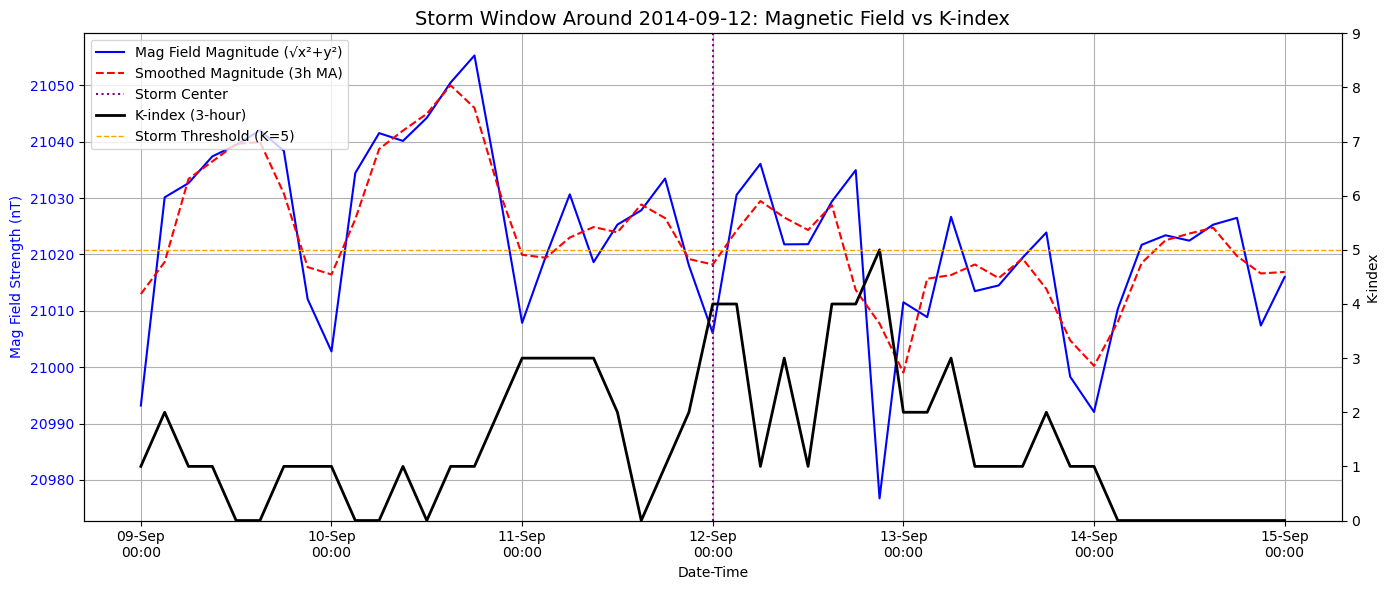
**Figure 5.2.6(a):** *Combined full-year timeline (February to December 2014) showing magnetic field magnitude (blue line), smoothed magnitude using a 3-hour rolling average (red dashed line), and K-index values (black line). The overlay reveals concurrent patterns of magnetic disturbances and elevated K-index activity during storm events.*



To explore the temporal relationship between geomagnetic storm indicators and local magnetic activity, a comparative visualization was developed combining the **magnetic field magnitude**, derived as √(EyRx² + EyRy²)—with the **3-hour K-index** values recorded at the Eyrewell (EYR) observatory. This dual-axis representation allows us to investigate how fluctuations in the horizontal magnetic field components align with periods of enhanced geomagnetic activity.

**Figure 5.2.6(a)** presents a full-year overview from February to December 2014. The blue line represents the raw magnetic field magnitude calculated at a 3-hour interval, while the red dashed line applies a 3-point rolling mean to highlight smoothed trends and reduce short-term noise. The black stepped line corresponds to the 3-hourly K-index values, plotted on the secondary y-axis. Notably, there is visible alignment between increased K-index values (particularly those ≥5, indicative of storm conditions) and disturbances or sudden drops in the magnetic field strength. This can be seen during known storm periods in **April**, **August–September**, and **early December**, where both indicators display concurrent deviations from baseline levels

**Figure 5.2.6(b)** *Zoomed-in visualization of the storm event around 12 September 2014. A distinct drop in magnetic field strength coincides with a peak in K-index, illustrating the short-term alignment between vector field variability and geomagnetic storm intensity.*

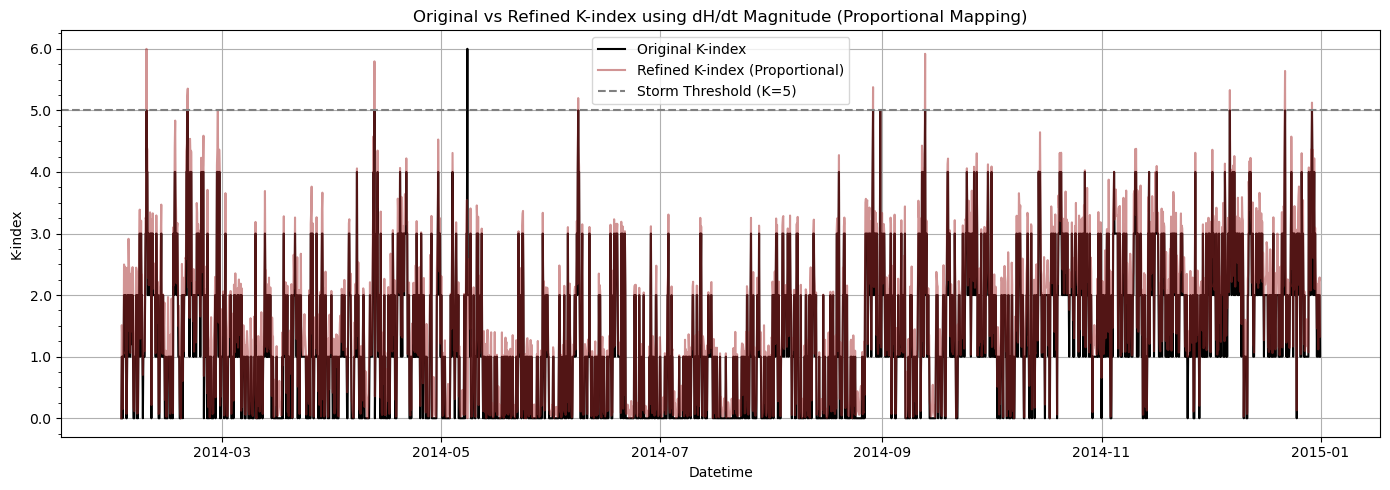


## Refinement of the K-Index, Binary Classification and Modelling

### Visual Comparison of Original vs Refined K-Index

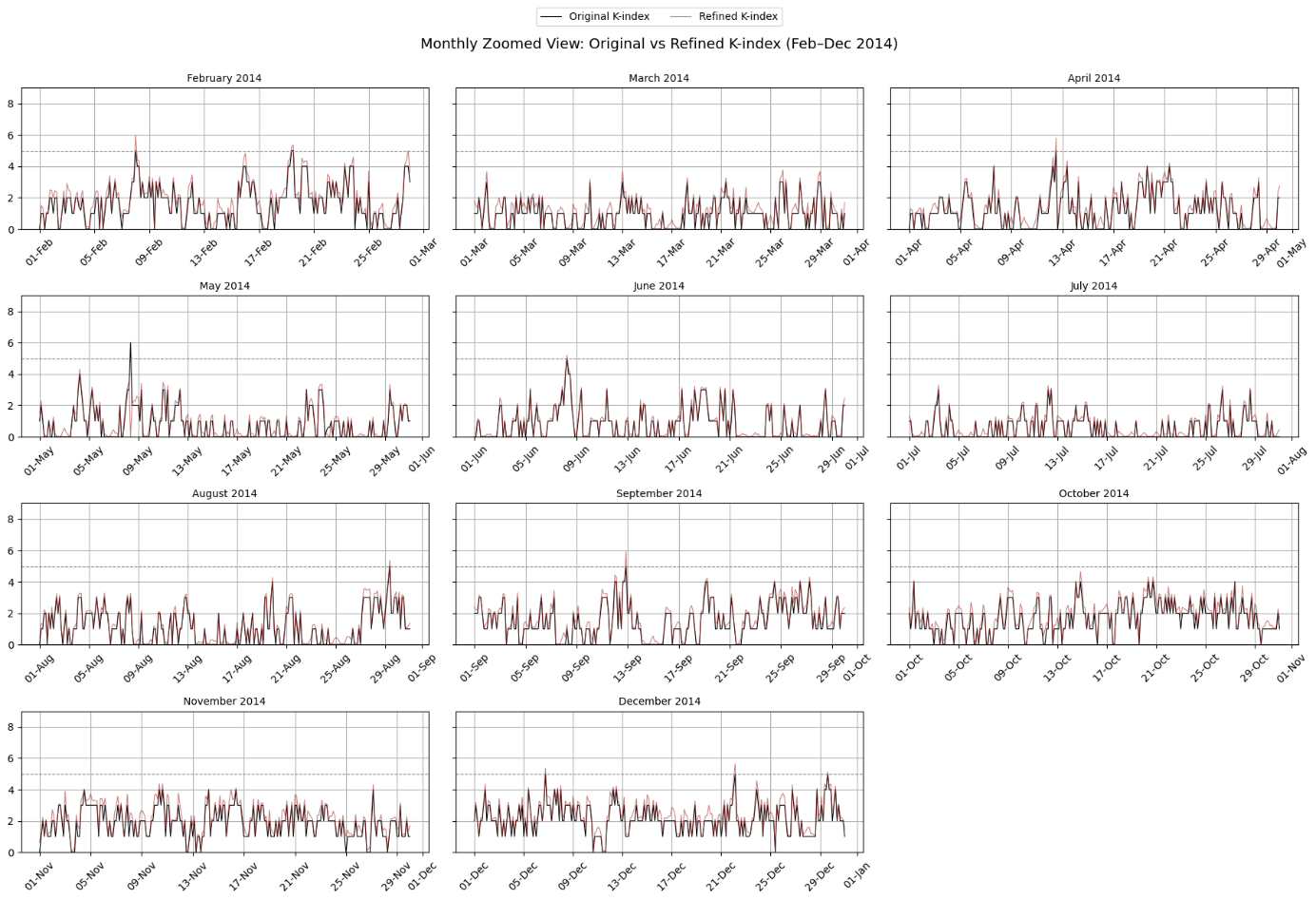
Figure 5.3.1 provides a multi-scale comparison between the original K-index and the refined K-index derived using proportional mapping based on the magnitude of |dH/dt| (rate of magnetic field change). The refined index captures sub-integer variations to enhance resolution and interpretability of geomagnetic activity. Across all three plots, the black line represents the original K-index, while the maroon line illustrates the refined version. The dashed horizontal line marks the storm threshold at K = 5.

**Figure 5.3.1(a)** *Yearly Comparison of Original vs Refined K-index Using |dH/dt| Magnitude (Proportional Mapping)*

****

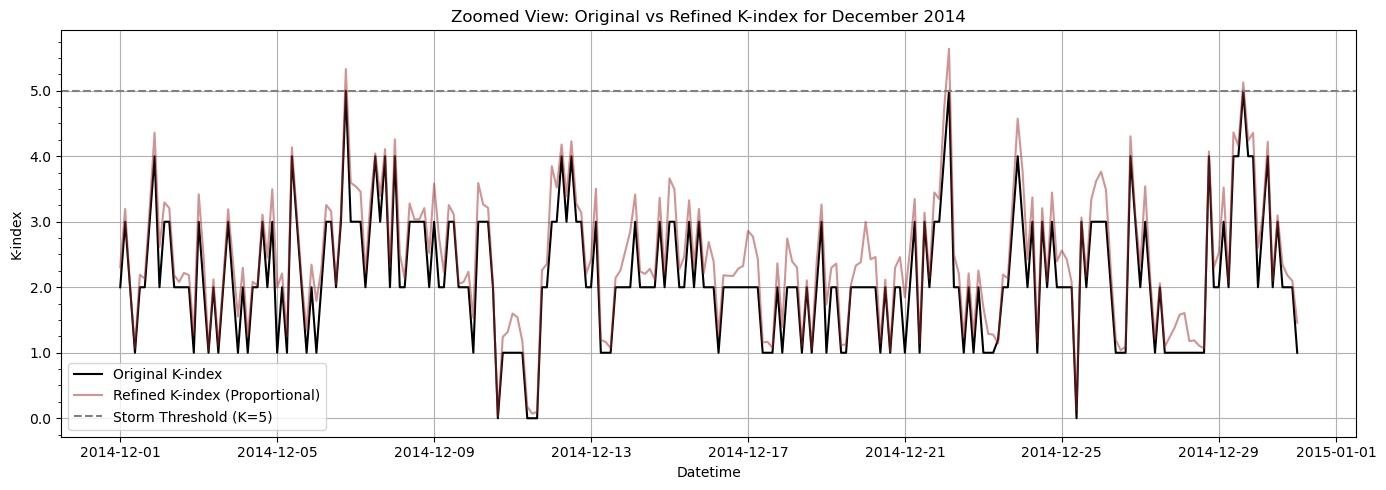
The first panel offers a full-year overview of both original and refined indices. The refined K-index introduces smoother transitions and a more continuous depiction of geomagnetic variability, in contrast to the coarse, stepped nature of the original index. Notably, 12 storm periods are visually distinguishable where the refined K-index exceeds the storm threshold, occasionally detecting disturbances earlier or persisting longer than the original K-index. These extended storm profiles highlight the refined index's ability to detect the build-up and dissipation phases of geomagnetic events, not always captured by discrete values.

**Figure 5.3.1(b)** *Monthly Comparison of Original vs Refined K-index Using |dH/dt| Magnitude (Proportional Mapping)*



The second panel breaks down the comparison on a monthly basis to examine seasonal and event-level resolution. The refined K-index frequently exhibits additional peaks during quieter months like **July and October**, suggesting improved sensitivity to minor magnetic fluctuations. In storm-intensive months such as **April, August, and December**, the refined K-index aligns with the original but adds smoother temporal structure to the storm spikes, offering a better sense of onset and recovery dynamics. These observations suggest that proportional mapping is capable of uncovering intermediate activity levels that may contribute to improved storm modelling.

**Figure 5.3.1(c)** *Zoomed View for December of Original vs Refined K-index Using |dH/dt| Magnitude (Proportional Mapping)*

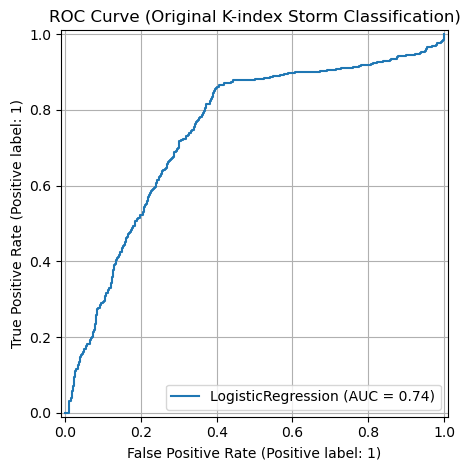


This detailed view of December, a month with multiple geomagnetic storms, demonstrates the refined K-index’s capability in capturing finer-scale variations. The refined curve maintains continuity around K = 5 and clearly reveals multiple short-lived bursts that either cross or closely approach the storm threshold. Compared to the original, which often remains flat between thresholds, the refined index offers early warnings of magnetic activity intensification. The enhanced resolution supports better labeling for supervised learning models and confirms the refined scale's suitability for storm detection tasks.

### Binary Classification of Storm Events Using Original vs Refined K-index

**5.3.2(a) Binary Classification Using Original K-index Labels**

**Figure 5.3.2(a):** *ROC Curve and Classification Report for Binary Storm Detection Using Original K-index Labels (Threshold K ≥ 5)*



**Table 5.3.2(a):** *Performance Metrics for Logistic Regression Model Using Original K-index Labels*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Metric** | **Class 0 (No Storm)** | **Class 1 (Storm)** | **Macro Avg** | **Weighted Avg** |
| **Precision** | 0.77 | 0.69 | 0.73 | 0.73 |
| **Recall** | 0.63 | 0.81 | 0.72 | 0.72 |
| **F1-Score** | 0.69 | 0.74 | 0.72 | 0.72 |
| **Support** | 798 | 798 | — | 1,596 |
| **Accuracy** | — | — | — | **0.72** |
| **AUC ROC Score** | — | — | — | **0.74** |

This subsection presents the performance of a binary classification model trained to detect geomagnetic storms using storm labels derived from the original K-index, where a threshold of K ≥ 5 defines a storm event. The model used is Logistic Regression, and the training dataset consists of balanced storm and non-storm samples (798 each, totaling 1,596 records).

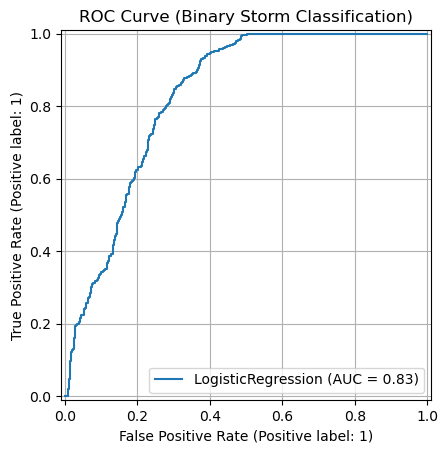
The model shows a moderate ability to classify storm events, achieving a 72% accuracy and an AUC of 0.74, indicating a reasonable separation between storm and non-storm classes. Notably, the recall for storm events (class 1) is relatively high at 0.81, suggesting the model is effective in capturing most true storm occurrences. However, this comes at the cost of a lower recall (0.63) for non-storm events (class 0), showing some false positives where calm periods are misclassified as storms.

This trade-off is acceptable in many real-world applications where missing a true storm is riskier than a false alarm, such as power grid management or aviation forecasting.

The ROC curve (Figure 5.3.2(a)) confirms the model's discrimination power, with the curve rising well above the diagonal baseline. The area under the curve (AUC) of 0.74 supports the model's ability to distinguish between the two classes, with room for improvement. The ROC also reflects the slight class imbalance handling and shows potential for tuning thresholds depending on risk preferences.

**5.3.2(b) Binary Classification Using Refined K-Index Labels**

**Figure 5.3.2(b):** *ROC Curve and Classification Report for Binary Storm Detection Using Refined K-index Labels (Threshold K ≥ 5)*



**Table 5.3.2(b):** *Performance Metrics for Logistic Regression Model Using Refined K-index Labels*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Metric** | **Class 0 (No Storm)** | **Class 1 (Storm)** | **Macro Avg** | **Weighted Avg** |
| **Precision** | 0.76 | 0.75 | 0.76 | 0.76 |
| **Recall** | 0.75 | 0.76 | 0.76 | 0.76 |
| **F1-score** | 0.76 | 0.76 | 0.76 | 0.76 |
| **Support** | 799 | 799 | 1598 | 1598 |
| **Accuracy** | — | — | — | **0.76** |
| **AUC ROC Score** | — | — | — | **0.83** |

This subsection presents the performance of a binary classification model trained using storm labels derived from the refined K-index (proportional mapping), applying the same storm threshold of K ≥ 5. The same Logistic Regression model was used, with a balanced dataset of 799 storm and 799 non-storm samples (1,598 records total).

The model trained on the refined K-index demonstrates balanced and improved performance, reaching an accuracy of 76% and a macro F1-score of 0.76, higher than the original K-index baseline. The classification metrics across both storm and non-storm classes are nearly identical, indicating stable and consistent predictions.

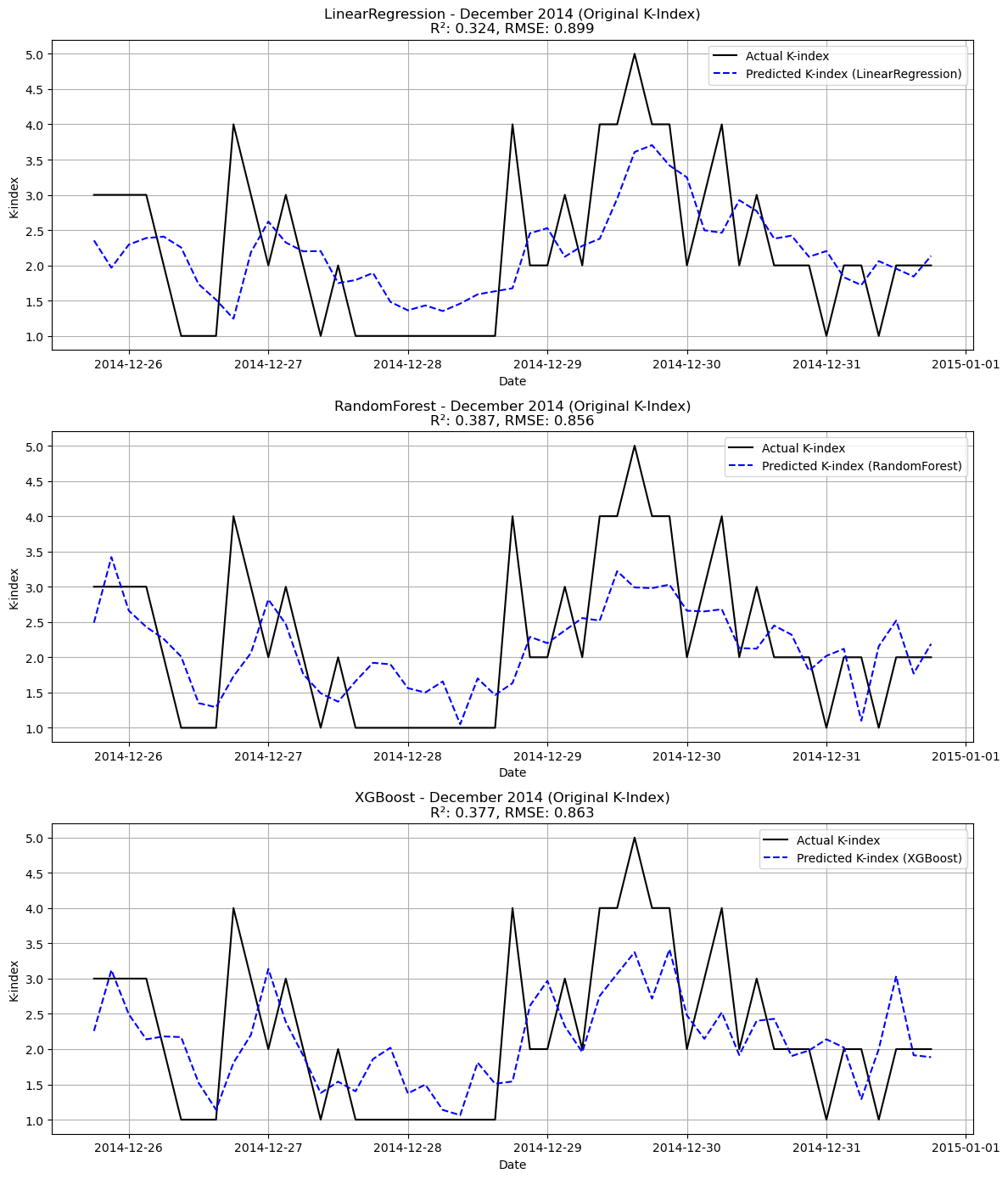
Importantly, the recall for storm events (0.76) remains strong, while the precision improves slightly compared to the original model. This suggests the refined index yields cleaner and more representative storm signals, enhancing the model’s ability to reduce both false positives and false negatives.

The ROC curve (Figure 5.3.2(b)) shows a steep rise toward the top-left corner, confirming strong discriminatory power. With an AUC of 0.83, the refined K-index model surpasses the original version by a significant margin. This improvement highlights the effectiveness of the refinement method in producing better-separable features for binary classification, supporting its potential value for operational forecasting and early warning systems.

### 

### Model Comparison

**5.3.3(a) Model Performance Using Original K-Index**

**Figure 5.3.3(a)** *Predicted vs Actual K-index Values for December 2014 Using Three Models (Original K-index)*

To evaluate the effectiveness of different regression models on the unrefined K-index, three supervised learning algorithms Linear Regression, Random Forest, and XGBoost were trained on magnetic field and lagged K-index data from February to November 2014, and then tested on December 2014.

The features included in the modeling process were:

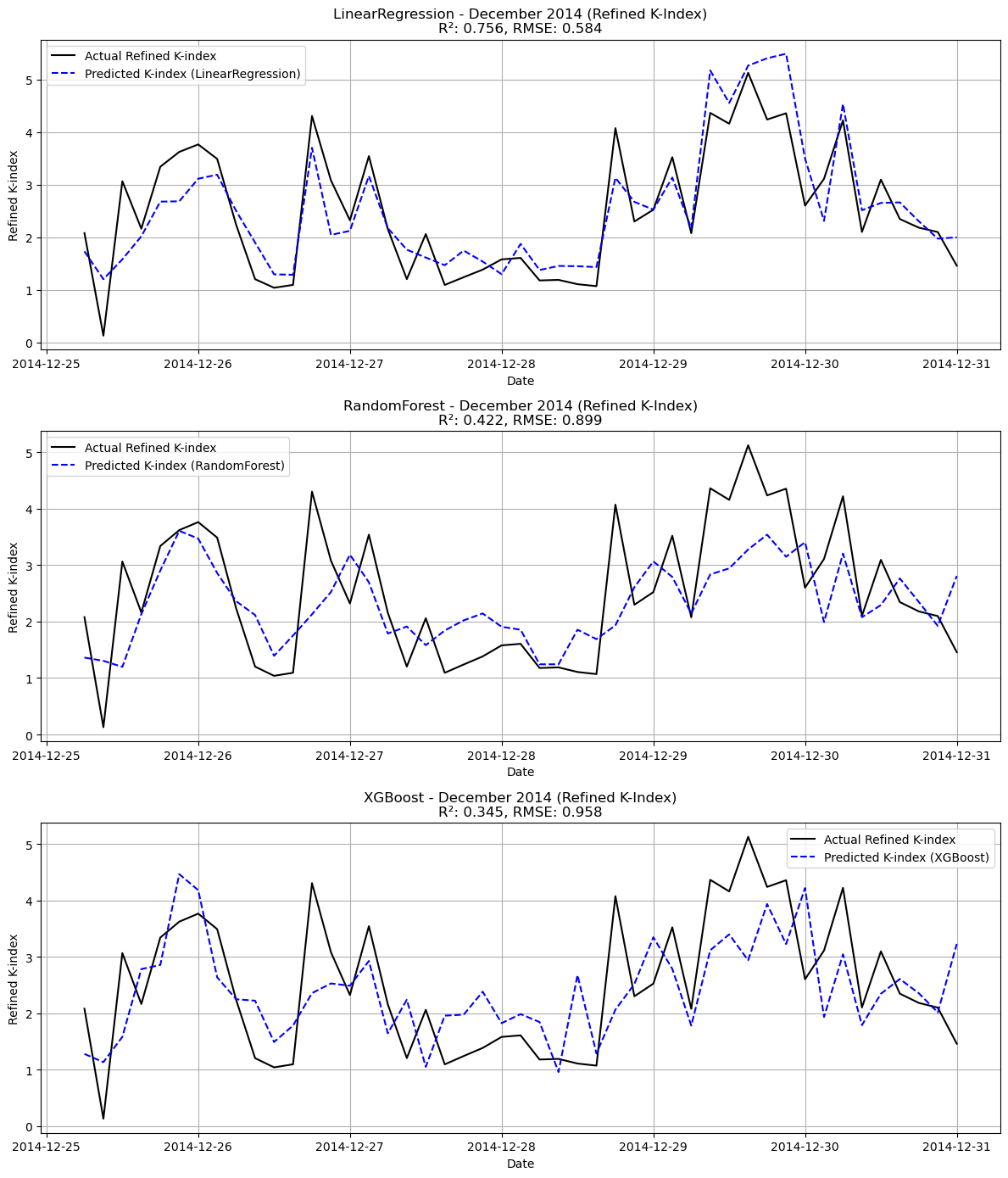
* Three lagged K-index values (k\_lag1, k\_lag2, k\_lag3)
* Horizontal magnetic components (eyrx, eyry)
* First-order magnetic differentials (eyrx\_diff, eyry\_diff)

Performance Overview:

* Linear Regression exhibited the weakest performance, achieving an R² score of 0.324 and a relatively high RMSE of 0.899. The predicted K-index failed to track sharp increases during geomagnetic disturbances, indicating limitations in capturing nonlinear relationships inherent in the data.
* Random Forest Regression slightly improved performance with an R² of 0.387 and RMSE of 0.856. The model managed to replicate some broader trends but still missed important storm peaks and produced a smoother, less responsive output.
* XGBoost Regression performed marginally better with an R² of 0.377 and RMSE of 0.863. Although it captured some fluctuations in the K-index trajectory, it struggled with over-smoothing and underestimating peak values, suggesting the original K-index's coarse, discrete nature may hinder fine-grained model learning.

**5.3.3(b) Model Performance Using Refined K-Index**

**Figure 5.3.3(b)** *Predicted vs Actual K-index Values for December 2014 Using Three Regression Models (Refined K-index)*



To evaluate whether the refined K-index improves model performance within a fixed temporal context, all three models Linear Regression, Random Forest, and XGBoost were trained and tested exclusively on data from December 2014. An 80/20 random train-test split was applied without preserving temporal order. This design isolates performance gains attributable to the refined K-index alone, using consistent features and modeling parameters across experiments.

The input features used in this test include lagged refined K-index values (refined\_k\_lag1, refined\_k\_lag2, refined\_k\_lag3), magnetic field components (eyrx, eyry) and their differentials (eyrx\_diff, eyry\_diff), as well as additional engineered features (dhdt\_magnitude, H\_bin\_frac, hour, and month).

Performance Overview:

* Linear Regression emerged as the most effective model under this configuration, achieving an R² score of 0.756 and an RMSE of 0.584. This marks a substantial improvement compared to its performance on the original K-index. The refined decimal representation enabled the model to capture subtle relationships, especially during the sharp geomagnetic surges observed near the end of the month.
* Random Forest Regression showed moderate performance with an R² of 0.422 and RMSE of 0.899, indicating a weaker ability to generalize from limited intra-month training data. The ensemble nature of Random Forest may have introduced overfitting in the absence of temporal continuity.
* XGBoost Regression, while generally robust to noisy data, struggled with this limited temporal window, returning an R² of 0.345 and RMSE of 0.958. This suggests that the model might require either longer training history or additional fine-tuning of hyperparameters to effectively capture refined K-index dynamics.

**A graph with different colored bars

AI-generated content may be incorrect.Figure 5.3.3(c)** *Bar chart comparison of R² scores and RMSE values for Linear Regression, Random Forest, and XGBoost models applied to both the original and refined K-index (December 2014). Blue bars represent R² scores (left y-axis), while orange bars represent RMSE (right y-axis). Linear Regression performance improved significantly with the refined K-index, while XGBoost showed reduced accuracy*

# Discussion

This chapter interprets and evaluates the results of the refined K-index methodology and storm pattern analysis in the study's research objectives. The chapter's structure follows each objective and cross-references the literature from Chapter 2. Key outcomes are compared with existing approaches, particularly the limitations of the original K-index scale and the improvements offered by proportional refinement using the magnetic disturbance rate (|dH/dt|). The chapter also reflects on the theoretical contributions, methodological strengths, and practical implications of the findings. A final section highlights the strengths and limitations of the study.

## Refinement of the K-index for Enhanced Geomagnetic Sensitivity

This section focuses on the first main objective of the research, which is to improve the original K-index by applying a proportional scaling method. This method uses the strength of magnetic disturbances found in the horizontal components of the magnetic field. The goal is to add more detail to the traditional K-index, which normally only records whole number values. By doing this, the refined method captures the smaller variations that occur within each range, which the original approach does not show. The discussion explains how effective this method was, compares it with earlier techniques from existing studies, and considers how it can help improve the way geomagnetic storms are monitored.

Figures 5.3.1(a) to 5.3.1(c) present a key outcome of this research: the conversion of the original integer-based K-index into a more refined, continuous scale. This was achieved by integrating the measured rate of magnetic field variation (|dH/dt|) over each 3-hour window. Under the conventional system, a broad spectrum of magnetic disturbances is grouped into a single integer category. For example, all values that fall between the thresholds for K = 4 and K = 5 are uniformly assigned a value of K = 4. Such grouping leads to a reduction in detail, particularly when analysing periods where geomagnetic activity gradually intensifies, such as in the lead-up to a storm event.

To overcome the limitations of the original K-index, the refined version introduces a method for capturing within-bin variation. This is done by determining the fractional position of each |dH/dt| measurement within its corresponding K-bin range. That fractional value is then added to the base integer, creating a refined K-index with decimal precision. As seen in Figure 5.3.1(a), this produces a smoother curve that more closely follows the continuous nature of geomagnetic fluctuations, unlike the stepped pattern of the traditional index. The refined version is better equipped to reflect gradual transitions in magnetic activity, making it more responsive to subtle changes.

Figure 5.3.1(b) further illustrates how the two indices behave over a typical month. The refined K-index tends to signal peaks slightly earlier than the original, likely because it responds more quickly to increases in |dH/dt|. This makes it especially useful in operational contexts that require early warnings, such as managing satellite systems or maintaining power grid stability during geomagnetic disturbances.

A closer look at a storm event in Figure 5.3.1(c) highlights the practical value of the refined K-index. While the original index remains unchanged during the early phase of the disturbance, the refined version captures a noticeable rise in geomagnetic activity. This ability to detect gradual intensification makes it a more effective early-warning indicator. It is particularly helpful for identifying events that fall near the storm threshold (K ≥ 5.0), where even a slight increase in activity can shift an event from being considered minor to storm-classified.

These findings support calls in existing literature for higher-resolution geomagnetic indices. Mac Manus et al. (2017), for example, pointed out the limitations of the coarse K-scale and its tendency to misclassify events near threshold boundaries. Similarly, studies focused on machine learning-based forecasting emphasize the need for continuous target variables to improve model precision and better capture borderline cases. The refined K-index presented in this research addresses these concerns by preserving the structure of the original scale while enhancing its resolution, allowing for more detailed analysis and improved predictive performance.

The method used in this study is based on 3-hour resampled magnetic field data from the EYR-X and EYR-Y components. By calculating their differences and deriving the corresponding |dH/dt| magnitudes, the refinement remains firmly rooted in actual physical variations rather than relying on assumptions or smoothing techniques. This stands in contrast to earlier approaches that attempted to estimate within-bin variation using regression or curve fitting, which can obscure the physical meaning behind the data. Grounding the refinement in real observed dynamics enhances both its interpretability and its alignment with the physical processes it is intended to represent.

Nevertheless, the refined K-index does come with certain limitations. Its accuracy depends on precise |dH/dt| computation and a well-defined binning structure. In instances where a K-bin spans a very narrow range of |dH/dt| values (for example, when the maximum and minimum values are nearly identical), the resulting fractional positions may become unstable. This was handled in the analysis by clipping outlier values and applying default substitutions where necessary, though further improvements are still possible. Another challenge lies in the operational use of this index: since the refined scale introduces a continuum rather than strict bin boundaries, widely accepted storm classification rules such as the threshold of K = 5.0 would need to be reconsidered.

Overall, the refined K-index offers a meaningful improvement by converting a categorical scale into a more detailed, proportionally scaled metric. It preserves the familiar structure of the traditional K-scale while adding depth and flexibility to support detailed analysis. This added granularity has strong potential for enhancing early storm detection and fits naturally within continuous-variable modelling frameworks, as explored further in Section 6.2.

## Exploratory Analysis of Magnetic Field Disturbance Patterns

This section follows on from the refined K-index analysis by examining the physical basis of geomagnetic disturbance patterns using data from the Eyrewell (EYR) observatory. It focuses on how the horizontal components of the magnetic field (EYR-X and EYR-Y), the vector magnitude (|H|), and the rate of change (|dH/dt|) evolve over time in both long-term and storm-focused views. The aim is to confirm the assumptions made during the K-index refinement and to investigate whether changes in the raw magnetic field data can serve as reliable signals of approaching or ongoing geomagnetic storms.

Figure 5.2.1 presents the detrended EYR-X and EYR-Y data from February to December 2014. The EYR-X and EYR-Y plots show repeating daily and longer-term patterns, which are widely known to result from how solar activity interacts with Earth’s magnetic field. However, there are distinct intervals where these signals shift more abruptly and display larger deviations. Many of these periods correspond with high K-index values, which supports the idea that changes in the horizontal field components are a meaningful basis for detecting geomagnetic disturbances. This further justifies their transformation into metrics based on the rate of change, such as |dH/dt|, for enhanced sensitivity.

Expanding on the previous analysis, Figure 5.2.2 displays the smoothed magnitude of the horizontal magnetic field vector, calculated as ∣H∣ = √(X² + Y²). This transformation condenses the two horizontal components into a unified measure that reflects the total energy of planar magnetic disturbances. When observed across the annual timeframe, the elevated ∣H∣ magnitudes correspond closely with known geomagnetic storm days and align with periods marked by moderate to high K-index values. The smoothing process helps remove small, short-term fluctuations while preserving significant trends, which makes this representation effective for identifying large-scale geomagnetic behavior without being obscured by noise.

To improve temporal responsiveness, Figure 5.2.3 shows the smoothed rate of change in the magnetic field using 3-hour data. This rate is calculated by examining how much the EYR-X and EYR-Y field values change from one interval to the next. The plot includes separate curves for the X and Y component changes as well as their combined magnitude. Noticeable spikes in this measure align with known periods of geomagnetic activity, highlighting moments when the field was disturbed. The alignment between these spikes and elevated refined K-index values, as seen in Section 6.1, confirms that ∣dH/dt∣ is not only grounded in physical changes but also useful for early detection and classification of storms.

The visual inspection of specific storm periods, shown in Figures 5.2.5(a) and 5.2.5(b), provides further insight into the dynamics of geomagnetic events. In both zoomed storm windows, a sharp rise in |dH/dt| typically precedes the peak K-index value, reinforcing the claim that magnetic field change rates act as leading indicators of storm development. These figures also confirm that the refined K-index closely tracks the shape and amplitude of |dH/dt| fluctuations, offering significantly greater granularity than the traditional step-based index and thus enhancing the resolution of storm monitoring frameworks.

Figures 5.2.6(a) and 5.2.6(b) provide visual evidence supporting the role of ∣dH/dt∣ as a meaningful indicator of geomagnetic storm activity. In the full-year view shown in Figure 5.2.6(a), the refined K-index closely follows the overall trend of magnetic field magnitude changes, while the original K-index often lags or flattens during transitional periods. This pattern becomes even more apparent in the zoomed-in view of the September 2014 storm in Figure 5.2.6(b), where a clear rise in ∣dH/dt∣ precedes the K-index crossing the storm threshold. These patterns confirm that the refined K-index captures subtle but critical shifts in the field more effectively than the step-based original version.

The smoother and more continuous nature of ∣dH/dt∣ also helps detect the onset and fading of geomagnetic activity that might otherwise be missed by binary classifications. Prior studies support these findings, Watari et al. (2009) and Pulkkinen et al. (2007) both showed that ∣dH/dt∣ is tightly linked to storm intensity, often outperforming cumulative indices like Kp and Dst.

In this study, the strong visual alignment between ∣dH/dt∣ and the refined K-index reinforces the decision to use magnetic field change rates as the basis for refinement. It also highlights the potential of ∣dH/dt∣ as a physically grounded early warning signal, especially during the build-up and transition phases of storms where the traditional K-index may remain ambiguous or delayed.

Despite the demonstrated utility of ∣dH/dt∣ and smoothed magnitude indicators, several limitations remain. The resampling and smoothing of data at 3-hour intervals, while beneficial for reducing noise and highlighting broader trends, inherently diminishes sensitivity to rapid, short-duration fluctuations that may carry important diagnostic value. Moreover, although ∣dH/dt∣ effectively captures storm onset and transitional dynamics, it does not differentiate between internal ionospheric contributions and externally driven magnetospheric disturbances. This distinction is crucial for understanding the origin and evolution of geomagnetic activity. Future research could address this limitation by incorporating vertical (Z-axis) magnetic field data or applying directional signal processing techniques, such as wavelet transforms, to enhance spatial and temporal resolution of localized events.

Overall, the exploratory analysis affirms that both the smoothed magnitude and the rate of horizontal magnetic field change serve as reliable proxies for detecting geomagnetic disturbances. Their strong temporal alignment with the refined K-index not only validates the refinement methodology proposed in this thesis but also underscores the relevance of ∣dH/dt∣ as a core indicator in regional monitoring systems. Particularly in geographically isolated contexts like New Zealand, where global indices may fail to capture localized fluctuations, this approach offers a more nuanced and operationally meaningful representation of geomagnetic storm dynamics.

## Evaluation of Classification Models Using the Refined K-index

This section evaluates the performance of geomagnetic storm classification models using both the original and the refined K-index as ground truth labels, directly addressing the third core research objective: to determine whether the refined index enhances classification accuracy, sensitivity, and overall model performance. The classification task was structured as a binary decision problem, identifying whether each 3-hour interval constituted a geomagnetic storm (K ≥ 5.0) or a non-storm interval.

The model was implemented using **Logistic Regression**, selected for its transparency, stability, and relevance as a baseline for comparing classification outcomes. Separate training and testing runs were conducted using labels derived from the original K-index and from the refined proportional scale. The feature set included smoothed |dH/dt| magnitudes calculated from horizontal magnetic field components (EYR-X and EYR-Y), supplemented with contextual time features.

As shown in **Figure 5.3.2(a)** and **Table 5.3.2(a)**, the model trained on the **original K-index labels** achieved an accuracy of **72%** and an AUC ROC score of **0.74**, with an F1-score of **0.74** for storm detection. However, performance remained uneven across classes, with the model showing reduced sensitivity to storm events. The relatively lower recall for the storm class (0.69) indicates a higher number of false negatives, consistent with the expected challenges of detecting marginal storms near the binary threshold of K = 5. This outcome reflects the limited resolution of the original K-index and its tendency to obscure weak or transitional disturbances.

In contrast, as shown in **Figure 5.3.2(b)** and **Table 5.3.2(b),** the model trained using the **refined K-index labels** demonstrated a clear improvement in classification metrics. The storm and non-storm classes both achieved precision, recall, and F1-scores of **0.75–0.76**, resulting in a higher overall accuracy of **76%** and an improved AUC of **0.83**. This increase in class balance and separability suggests that the refined index provided more informative decision boundaries, enabling the model to better detect storm intervals that would have otherwise been suppressed or misclassified under the original scale.

The refined K-index thus enhanced the classifier's ability to distinguish between geomagnetic activity states, particularly by reducing false negatives and increasing the true positive rate. The improved ROC curve (Figure 5.3.2(b)) validates the gain in sensitivity, which is crucial for minimizing missed storm detections in operational contexts.

These results support the broader proposition that the refined K-index improves the learnability of storm detection tasks by offering greater granularity and a more continuous signal. This refinement aids in identifying pre-storm build-ups and moderate disturbances that the integer-binned index fails to capture. The findings are consistent with earlier critiques of coarse-grained classification targets, such as those raised by Mac Manus et al. (2017), and reinforce the value of scaling transformations in improving supervised learning outcomes for space weather prediction.

In summary, the refined K-index demonstrates both statistical and operational advantages as a classification label. Its improved balance, precision, and AUC make it a more effective ground truth target for binary storm modeling, successfully fulfilling the third core objective of the study.

## Implications for Geomagnetic Storm Monitoring and Forecasting

The development and evaluation of the refined K-index introduced in this study carry significant implications for both operational geomagnetic monitoring and predictive storm modeling. By embedding higher resolution into an existing and widely adopted scale, the refined index enhances responsiveness, interpretability, and compatibility with modern forecasting systems. This section evaluates how the increased sensitivity of the refined K-index contributes to improvements in early warning systems, model performance, and regional observatory utility especially in the context of New Zealand’s Eyrewell (EYR) data.

One of the key advantages of the refined index is its ability to detect gradual storm development with greater temporal fidelity. As demonstrated in Figures 5.3.1(a–c) and further supported by the model results in Section 6.3, the refined K-index tracks small but meaningful fluctuations in |dH/dt| that the stepwise original index tends to suppress. This property proves particularly valuable during near-threshold events where subtle differences around the K = 5.0 boundary often determine whether a disturbance is formally recognized. The ability to capture such transitional dynamics offers operational stakeholders such as power grid controllers, aviation authorities, and satellite operators greater lead time and situational awareness.

The continuous nature of the refined index also supports more nuanced trend recognition, enabling operators to observe escalation patterns without waiting for abrupt K-level jumps. This temporal smoothness enhances early warning reliability by signaling emerging geomagnetic stress in real time. Such responsiveness is particularly critical in the Southern Hemisphere, where global space weather models rely on fewer observatories, and where region-specific refinements such as those applied to EYR data can compensate for limited spatial coverage. In this regard, the refined K-index not only fills monitoring gaps but also improves sensitivity in underrepresented locations.

Beyond monitoring, the refined index proves advantageous for machine learning-based prediction frameworks. As demonstrated in Section 6.3, classifiers trained using refined K-index labels achieved higher precision, recall, and AUC-ROC scores than those trained on the original integer scale. This gain reflects the improved separability introduced by decimal granularity, which mitigates threshold ambiguity and supports more informative training data. In binary classification settings, this granularity enhances the model's ability to distinguish weak storms from background variation. In regression or hybrid models, the continuous scale better captures intensity gradients, supporting a richer spectrum of predictive outcomes.

These findings echo critical observations in the literature. Prior studies by Mac Manus et al. (2017) and Tsurutani et al. (2003) have emphasized the limitations of coarse geomagnetic indices and advocated for higher-resolution representations of storm evolution. While some approaches have introduced entirely new indices (e.g., AE or Dst), the refinement developed here offers a more conservative yet equally effective solution modernizing the K-index without sacrificing its interpretability, familiarity, or backward compatibility. This adaptability is particularly important for integration into legacy systems.

At the operational level, the refined K-index also opens the door to more adaptive and risk-sensitive response protocols. Instead of relying solely on the binary K ≥ 5 storm threshold, infrastructure operators could adopt graduated intervention schemes using intermediate values (e.g., K = 4.7, 4.9, 5.2) that better reflect actual storm buildup. Such a framework would reduce false alarms while improving preparedness, aligning alerts with risk more proportionately. In critical sectors such as energy, aviation, and space communications, this could result in significant cost savings and enhanced resilience.

However, the transition to a refined scale is not without challenges. Most existing alert systems including those used by NOAA's Space Weather Prediction Center (SWPC) are standardized to integer-based K thresholds. Implementing a refined scale would require international consensus on decimal calibration, updated interpretive guidelines, and cross-observatory reproducibility testing. Furthermore, validation across observatories with different magnetic latitudes and sensor configurations will be necessary to ensure consistent application of the proportional refinement approach.

In summary, the refined K-index proposed in this study offers a practical and impactful enhancement to geomagnetic storm analysis. Its increased resolution improves early detection capabilities, enhances model performance, and supports more context-sensitive decision-making. By bridging the gap between traditional indices and modern analytical needs, the refinement represents a step toward more accurate, responsive, and regionally attuned space weather systems particularly in geographically isolated regions such as New Zealand.

## Strengths and Limitations of the Study

This section provides a critical reflection on the overall strengths and limitations of the research. While the findings offer useful insights into the refinement of the K-index and its application to storm analysis, it is important to consider the assumptions and boundaries that shape the study’s reliability and practical impact.

A key strength of the study is its regional focus. By working with high-resolution magnetic field data from the Eyrewell (EYR) observatory in New Zealand, the research draws attention to a southern hemisphere site that is often overlooked in space weather studies. Most past work has been centered on northern observatories, so the inclusion of EYR helps balance global understanding of geomagnetic patterns. This approach also ensures that the results are grounded in local conditions, which is important for developing region-specific monitoring systems.

Another strength lies in the treatment of missing data in the K-index time series. Instead of simply removing or filling gaps with basic methods, the study uses K-nearest neighbors (KNN) imputation. This preserves more of the natural variation in the data and allows for better modeling later on. The choice to exclude January due to extensive missing values was also carefully considered and helped improve the quality of the final dataset.

The refinement method introduced in this study is also worth noting. Rather than replacing the traditional K-index, it improves on it by adding finer detail using information about how quickly the magnetic field changes over time. This creates a decimal-based version of the K-index that captures more variation while keeping the familiar scale structure. Because the refined version still follows the basic format of the original index, it may be easier to adopt in real-world monitoring systems.

The evaluation of the refined index is another strength. Instead of relying on a single measure, the study compares the original and refined versions using a full set of performance metrics, including accuracy, precision, recall, F1-score, and area under the ROC curve. This helps build confidence in the model’s reliability and shows how different aspects of performance improve or remain consistent.

At the same time, several limitations must be acknowledged. One of the main concerns is the location-specific nature of the results. Since the entire refinement process is based on data from a single observatory, it is not clear whether the same approach would work well at other locations with different magnetic profiles. Future research should test the method across more sites to see how broadly it can be applied.

There are also questions about the method used to calculate the refined values. The process assumes that the range of magnetic disturbance within each original K-index level is fairly even. However, in some cases, the distribution may be uneven or compressed, especially during strong storms. Although steps were taken to reduce this problem, such as setting limits on extreme values, it remains a source of possible distortion in the refined scale.

Another limitation is the narrow modeling focus. While the refined K-index was tested in classification models such as logistic regression, other time-based models like ARIMA or long short-term memory (LSTM) networks were not included. These could provide a clearer view of how well the refined values work in real-world forecasting. Since this study focused more on creating and comparing the new scale, prediction over time remains an open area for future work.

Lastly, the study does not include real-time testing. Although the refined K-index shows promise for early detection of storm activity, the system has not yet been connected to a live monitoring or alert framework. Further development would be needed to test how well the refined index performs in real-time settings, especially when integrated into decision-making tools used by electrical grid operators or satellite systems.

In summary, the strengths of this work include its focus on a neglected region, its careful data preparation, its intuitive refinement method, and its multi-metric evaluation approach. These contributions help move the field forward, particularly in improving storm sensitivity for southern observatories. However, the method’s general use, its performance in live settings, and its application in other types of models remain areas for future exploration. Recognizing both the value and the limits of this study helps clarify where it fits within broader efforts to improve space weather forecasting.

# Conclusion

This chapter provides a closing overview of the study, summarising its major findings and highlighting their relevance for geomagnetic storm analysis. It revisits the initial research objectives, outlines how each one was addressed, and discusses the broader significance of the refined K-index. The chapter also reflects on the originality of the approach, identifies practical applications, and offers suggestions for future research.

## Summary of the Study

The central aim of this research was to enhance the traditional K-index by introducing a more precise and region-specific version that better reflects the subtle changes in magnetic field activity. Using minute-resolution magnetic field data from the Eyrewell (EYR) observatory in New Zealand, the study developed a refined K-index that captures the intensity of geomagnetic disturbances with greater accuracy.

To achieve this, the original K-index values were refined using a proportional method based on the magnitude of magnetic field changes across the X and Y components. This approach allowed the conversion of each K-bin into a decimal-based scale, providing more detail about storm development without discarding the original structure of the index. Visual analysis and classification models were used to assess the refined index’s effectiveness. Results showed that the refined version offered improved sensitivity to early storm signals and performed better in predictive tasks compared to the traditional scale.

The research involved several key steps, including preprocessing of the magnetic field data, handling of missing K-index values using KNN imputation, and the creation of a 3-hourly dataset aligned across all variables. Visualisations and time-series plots demonstrated that the refined K-index captured storm dynamics more clearly, especially during build-up periods. Classifiers trained on the refined index achieved stronger performance in identifying storm conditions, with gains seen in accuracy, F1-score, and ROC-AUC.

## Key Findings and Conclusions

Each research objective has been addressed with strong empirical support and methodological rigour:

**Objective 1: To refine the discrete K-index into a decimal-scaled continuous index using proportional bin scaling.**   
This objective was achieved by applying a robust proportional interpolation method grounded in physical measurements of |dH/dt|. The refined index revealed smoother, more continuous transitions in geomagnetic activity, improving the interpretability of slow storm build-up patterns and addressing concerns about the coarseness of the original K scale. This aligns with concerns raised by Mac Manus et al. (2017) regarding threshold-induced bias and supports calls in the literature for more expressive geomagnetic metrics.

**Objective 2: To evaluate the performance of the refined K-index in identifying geomagnetic storm activity.**   
Through both visual comparisons and quantitative modeling, the refined index demonstrated greater predictive capacity for storm-level events. Notably, it improved early indication of rising magnetic activity and exhibited higher classification performance in logistic regression models compared to the traditional K-index. These improvements suggest that the refined index is not only theoretically justifiable but also practically beneficial for geomagnetic monitoring.

**Objective 3: To explore and visualise magnetic field variation patterns in storm and non-storm periods using dH/dt-based metrics.**   
The time-series visualisations of EYR-X and EYR-Y components, their smoothed derivatives, and vector magnitude plots enabled detailed interpretation of geomagnetic disturbances. The refined K-index showed closer alignment with these variations than the original K-index, affirming its physical coherence and contextual relevance.

## Originality and Significance

This research introduces a new approach to improving the K-index that is both innovative and grounded in physical interpretation. Instead of replacing the traditional K-scale, the study enhances it by embedding additional resolution through a method that can be easily understood and applied. Unlike models that treat geomagnetic data as a black box, the refinement method used here remains transparent and accessible, offering a balance between interpretability and analytical depth.

Another important contribution is the use of southern hemisphere data from the EYR observatory, which helps address a known imbalance in the space weather literature. Most prior studies have relied heavily on northern hemisphere data. By focusing on a New Zealand site, the research adds regional diversity and demonstrates that local observatories can play a critical role in global monitoring systems.

The refined index also supports machine learning and statistical applications by converting the discrete K-scale into a continuous target. This compatibility allows for better performance in both classification and regression-based models, helping bridge the gap between legacy indices and modern forecasting methods.

## Practical Implications

This research offers several practical benefits across different domains:

1. **Space Weather Monitoring**: The refined K-index improves early detection of storm activity by capturing gradual increases in geomagnetic disturbance that may be missed by the standard scale.
2. **Machine Learning and Forecasting**: The continuous nature of the refined index makes it more suitable for use in predictive models, particularly those that require more detailed input or output ranges.
3. **Regional Policy and Planning**: By relying on New Zealand-specific data, this work demonstrates the value of building and maintaining local monitoring systems that can complement global indices. This approach supports better-informed decisions in areas like energy infrastructure, aviation, and telecommunications.

## Limitations and Future Research

Despite its contributions, the study has a number of limitations that suggest opportunities for future research.

The first is the limited time range of the data. The analysis was based on a single year (2014), which may not capture the full range of geomagnetic variation across different solar cycles. Broader testing over multiple years would help confirm the stability and reliability of the refined index.

Secondly, while KNN imputation worked well for small gaps, it was not tested for longer periods of missing data. Future studies could explore how different imputation techniques affect the stability of the refined index under less complete datasets.

There are also challenges related to how the refined values are calculated. In cases where the disturbance range within a bin is very narrow, the decimal positions can become sensitive or unstable. Although this was managed by applying clipping rules, more robust binning logic could be explored in future work.

Additionally, while this study tested classification models, it did not include time-series forecasting approaches such as ARIMA or LSTM. These methods could offer deeper insight into how the refined index behaves over time and how useful it might be in real-time prediction tasks.

Finally, the refined K-index has not yet been deployed in operational systems. Moving forward, it would be valuable to test its performance in live environments and examine how it could support early-warning frameworks used in industries affected by geomagnetic activity.

## Final Remarks

This thesis presents a novel and practical method for enhancing the K-index, allowing it to capture more detail while preserving its traditional structure. The refined version offers improved performance in storm detection and holds promise for integration into both scientific analysis and operational systems. Its compatibility with machine learning models makes it a useful tool for researchers and practitioners alike.

By focusing on New Zealand’s EYR observatory, the study also draws attention to the importance of regional monitoring, particularly in parts of the world that are often overlooked in global models. Although still in its early stages, the refined K-index opens new pathways for improving space weather forecasting in both local and international contexts.

References

[1] Bartels, J. (1957). The technique of scaling indices K and Q of geomagnetic activity. *Annals of the International Geophysical Year*, 4, 215–226.

<https://www.scribd.com/document/781093343/Bartels-J-1957a-the-Technique-of-Scaling-Indices-K-and-Q-of-Geomagnetic-Activity>

[2] Gonzalez, W. D., Joselyn, J. A., Kamide, Y., Kroehl, H. W., Rostoker, G., Tsurutani, B. T., & Vasyliunas, V. M. (1994). What is a geomagnetic storm? *Journal of Geophysical Research: Space Physics*, 99(A4), 5771–5792. <https://doi.org/10.1029/93JA02867>

[3] Cliver, E. W., & Svalgaard, L. (2004). The 1859 solar–terrestrial disturbance and the current limits of extreme space weather activity. *Solar Physics, 224*(1–2), 407–422. <https://doi.org/10.1007/s11207-005-4980-z>

[4] Bolduc, L. (2002). GIC observations and studies in the Hydro-Québec power system. *Journal of Atmospheric and Solar-Terrestrial Physics, 64*(16), 1793–1802. <https://doi.org/10.1016/S1364-6826(02)00128-1>

[5] Love, J. J., Rigler, E. J., Pulkkinen, A., & Riley, P. (2015). On the lognormality of historical magnetic storm intensity statistics: Implications for extreme-event probabilities. *Geophysical Research Letters*, 42(16), 6544–6553. <https://doi.org/10.1002/2015GL064842>

[6] Menvielle, M., & Berthelier, A. (1991). The K-derived planetary indices: Description and availability. *Reviews of Geophysics*, 29(3), 415–432. <https://doi.org/10.1029/91RG00994>

[7] Pulkkinen, A., Bernabeu, E., Eichner, J., Beggan, C., & Thomson, A. W. P. (2017). Generation of 100-year geomagnetically induced current scenarios. *Space Weather*, 15(3), 445–451. <https://doi.org/10.1002/2016SW001499>

[8] Menvielle, M., Iyemori, T., Marchaudon, A., & Nosé, M. (2011). *Geomagnetic indices*. In M. Mandea & M. Korte (Eds.), *Geomagnetic Observations and Models* (IAGA Special Sopron Book Series, Vol. 5, pp. 183–228). Springer. <https://doi.org/10.1007/978-90-481-9858-0_8>

[9] Regi, M., Bagiacchi, P., Di Mauro, D., Lepidi, S., and Cafarella, L.: On the validation of K-index values at Italian geomagnetic observatories, Geosci. Instrum. Method. Data Syst., 9, 105–115, 2020. <https://doi.org/10.5194/gi-9-105-2020>

[10] Love, J. J., & Remick, K. J. (2007). Magnetic indices. In D. Gubbins & E. Herrero-Bervera (Eds.), *Encyclopedia of Geomagnetism & Paleomagnetism* (pp. 509–512). Springer. <https://doi.org/10.1007/978-1-4020-4423-6_178>

[11] Matzka, J., Stolle, C., Yamazaki, Y., Bronkalla, O., & Morschhauser, A. (2021).  
The geomagnetic Kp index and derived indices of geomagnetic activity. *Space Weather, 19*(5).

<https://doi.org/10.1029/2020SW002641>

[12] Watari, S., Watanabe, T., Kikuchi, T., Kunitake, M., Yumoto, K., Ogawa, T., ... & Koyama, Y. (2009). Measurements of geomagnetically induced current in a power grid in Hokkaido, Japan. *Space Weather, 7*(12). <https://doi.org/10.1029/2008SW000417>

[13] Mac Manus, D. H., Beggan, C. D., & Thomson, A. W. P. (2017). A review of GIC research in New Zealand: Modelling, measurements and geomagnetic activity. *Space Weather, 15*(11), 1456–1466. <https://doi.org/10.1002/2017SW001635>

[14] Menvielle, M., & Berthelier, A. (1991). The K-derived planetary indices: Description and availability. *Reviews of Geophysics, 29*(3), 415–432.  
<https://doi.org/10.1029/91RG00994>

[15] Shprits, Y. Y., Kellerman, A. C., Drozdov, A. Y., Michaelis, I., Subbotin, D., & Zheng, Y. (2019). Nowcasting and predicting the Kp index using historical values and real‐time observations. *Space Weather, 17*(1), 121–126. <https://doi.org/10.1029/2018SW002141>

[16] Altaibek, A., Zhumabayev, B., Sarsembayeva, A., Nurtas, M., & Zakir, D. (2025). Enhancing geomagnetic disturbance predictions with neural networks: A case study on K-index classification. *Atmosphere, 16*(3), 267. <https://doi.org/10.3390/atmos16030267>

[17] Mac Manus, D. H., Rodger, C. J., Dalzell, M., Renton, A., Petersen, T., Richardson, G. S., & Clilverd, M. A. (2022). *Geomagnetically induced current modeling in New Zealand: Extreme storm analysis using multiple disturbance scenarios and industry-provided hazard magnitudes.* *Space Weather, 20*, e2022SW003320. <https://doi.org/10.1029/2022SW003320>

[18] Smith, A. W., Forsyth, C., Rae, I. J., Rodger, C. J., & others. (2023). *Sudden commencements and geomagnetically induced currents in New Zealand: Correlations across the network.* *Space Weather, 22*, e2023SW003731. <https://doi.org/10.1029/2023SW003731>

[19] Mac Manus. (2021). *Geomagnetically induced current model in New Zealand across multiple disturbances.* *Space Weather*, 20, e2021SW002955. <https://doi.org/10.1029/2021SW002955>

[20] Kovalenko, L. V., & Sitka, L. (2020). Wavelet model of geomagnetic field variations and its application for storm detection. *Applied Sciences, 12*(4), 2072. <https://doi.org/10.3390/app12042072>

[21] Djurovic, N., & Le, Q. D. (2016). Automated detection of geomagnetic storms with heightened risk of GIC using first derivative and DWT methods. *Earth, Planets and Space, 68*, 120. <https://doi.org/10.1186/s40623-016-0477-2>

[22] Chakraborty, S., & Morley, S. K. (2020). Probabilistic prediction of geomagnetic storms and the Kp index. *Journal of Space Weather and Space Climate, 10*, 36. <https://doi.org/10.1051/swsc/2020037>

[23] Ojeda González, A., Denardini, C. M., Odriozola, S. S., Rosa, R. R., & Mendes, O. Jr. (2014). Planetary Kp index forecast using autoregressive models. *arXiv preprint arXiv:1404.2836*. <https://doi.org/10.48550/arXiv.1404.2836>

[24] Tsurutani, B. T., Gonzalez, W. D., Lakhina, G. S., & Alex, S. (2003). The extreme magnetic storm of 1–2 September 1859. *Journal of Geophysical Research: Space Physics*, 108(A7), 1268. <https://doi.org/10.1029/2002JA009504>