Space-Based Earthquake Precursor Monitoring A Framework for Seismo-Ionospheric Anomaly Detection using SWARM Satellite Data

F. Ait Benkhali[1-2 *] and S. Wagner[1-2 *]

[1] ZAH Landessternwarte, Universität Heidelberg, Germany
[2] German Center for Astrophysics (DZA), Görlitz , Germany

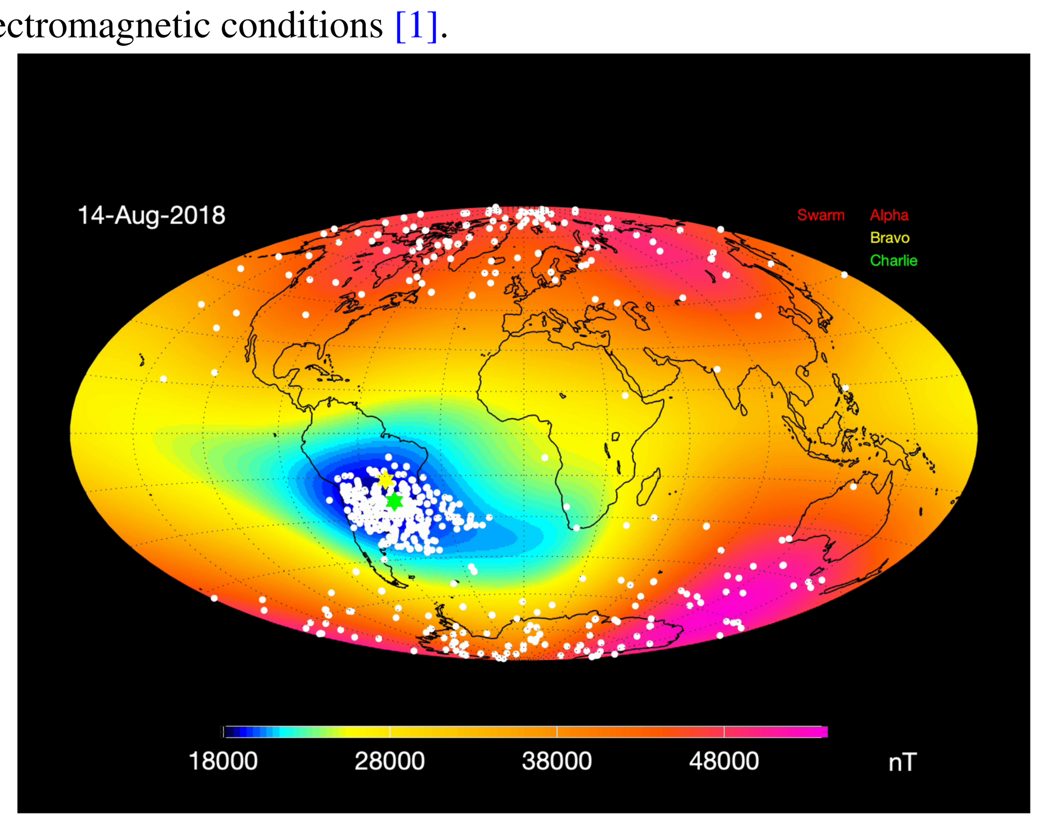
Correspondence to: faical.ait_benkhali@dzastro.de

Abstract

The lithosphere–atmosphere–ionosphere coupling (LAIC) mechanism posits that pre-seismic electromagnetic disturbances may manifest as detectable anomalies in the ionospheric magnetic field. The European Space Agency's SWARM constellation offers a unique platform for probing these phenomena through high-precision, multi-satellite magnetic field observations. We introduce a comprehensive computational framework for identifying potential earthquake precursor signatures in SWARM magnetic field data, leveraging advanced statistical and machine learning methodologies. Multiple anomaly detection algorithms are implemented and systematically validated to detect electromagnetic perturbations that may precede seismic events within the Dobrovolsky preparation zone. The framework adopts a multi-tiered analytical approach that integrates Statistical Process Control (SPC), Isolation Forest machine learning, Principal Component Analysis (PCA), and spectral analysis. It processes magnetic field residuals derived from the CHAOS geomagnetic field model, with stringent quality control guided by geomagnetic activity indices . The pipeline accommodates user-defined temporal windows for pre- and post-seismic evaluation, with spatial constraints determined by the empirical Dobrovolsky scaling $R = 10^{0.43M}$. The system efficiently handles multi-satellite magnetic field datasets from SWARM Alpha, Bravo, and Charlie through parallelized computation, enabling large-scale event analyses. Statistical validation demonstrates the framework's capability to detect anomalous magnetic signatures while accounting for inherent geophysical variability. The Isolation Forest algorithm achieves robust multivariate anomaly detection with tunable contamination parameters, while PCA facilitates dimensionality reduction and interpretation of complex magnetic relationships. This framework constitutes a important advancement in computational infrastructure for earthquake precursor studies, offering versatile tools for investigating lithosphere–atmosphere–ionosphere coupling processes. Its multi-method design permits rigorous evaluation of precursor hypotheses. Future developments will focus on (1) integrating formal statistical significance testing to quantify anomaly reliability and (2) implementing machine learning–based predictive models for near-real-time earthquake probability forecasting. These enhancements aim to transition the framework from a retrospective analysis tool to a predictive platform for operational seismic monitoring and early warning.

Introduction

Electromagnetic anomalies preceding seismic events represent a critical yet unresolved aspect of earthquake physics. Although numerous studies have reported correlations between ionospheric disturbances and seismic activity, the underlying physical mechanisms of LAIC remain poorly constrained. The SWARM mission, launched in 2013, provides a unique platform to investigate these phenomena through high-precision, multi-satellite measurements of the Earth's magnetic and electric fields. The SWARM constellation comprises three identical satellites in near-polar orbits, enabling continuous, high-resolution observations of ionospheric electromagnetic conditions [1].

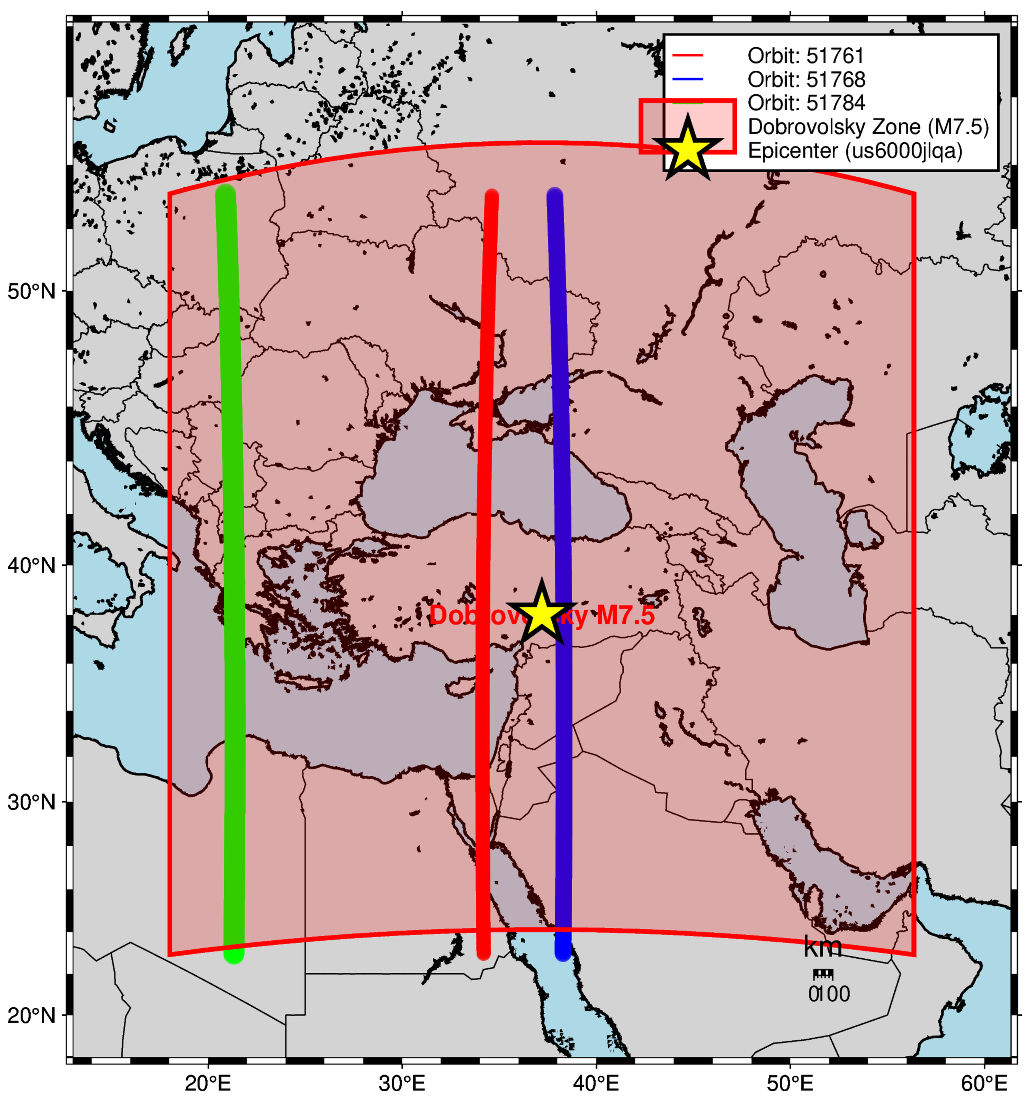


Global distribution of magnetic field strength on 14 August 2018, overlaid with satellite ground tracks. The color map represents geomagnetic field intensity in nanotesla (nT), with warmer colors (red/magenta) indicating stronger fields and cooler colors (blue/cyan) lower field strengths, highlighting the South Atlantic Anomaly (SAA) over South America. The positions of SWARM Alpha (red), Bravo (green), and Charlie (blue) satellites are shown.white dots on the map are the locations where the SWARM Satellites observed the geomagnetism.

SWARM data have become instrumental for investigating potential pre-seismic electromagnetic signatures, complementing the mission's primary objectives of probing core dynamics, crustal magnetization, and space weather phenomena. Despite multiple reports of ultra-low-frequency magnetic variations and ionospheric perturbations preceding earthquakes, reproducibility remains limited due to confounding influences such as geomagnetic storms and atmospheric variability [2]. To address these challenges, we present an integrated computational framework for analyzing SWARM magnetic field data. By combining Statistical Process Control, machine learning–based anomaly detection, Principal Component Analysis [3], and spectral analysis, the framework enhances the detection and interpretation of potential earthquake precursors. This approach aims to provide robust insights into LAIC processes and their relevance for seismic forecasting, offering a methodological advance in pre-seismic electromagnetic research.

Software Architecture and Methodology

The computational framework adopts a modular architecture designed for flexibility, maintainability, and efficient large-scale data analysis. It integrates data acquisition, storage, analytical processing, and visualization within a unified system. SWARM Level 1B magnetic field data are retrieved via the Virtual Internet Resource for Earth and Space Science (`VIRES`) API, employing automated error handling and validation routines to ensure continuous and reliable access. Data are managed in an SQLite relational database, optimized for efficient querying and integrity preservation. Analytical processing is conducted by a unified engine supporting parallel execution of statistical and machine learning algorithms, enabling comparative evaluation across multiple seismic events. The visualization module generates publication-quality plots and analytical reports compatible with collaborative workflows. Data preprocessing incorporates rigorous quality control procedures to enhance reproducibility and data integrity. Contributions from the core, crust, and magnetosphere are removed using the CHAOS geomagnetic field model, isolating residual signals for subsequent analysis. Magnetic field components are consistently transformed across coordinate systems to account for orbital and attitude variations. Periods of elevated geomagnetic activity, identified through indices such as Kp and Dst, are excluded to minimize contamination from space weather effects. Outlier detection methods further mitigate spurious measurements arising from instrumental or environmental factors.

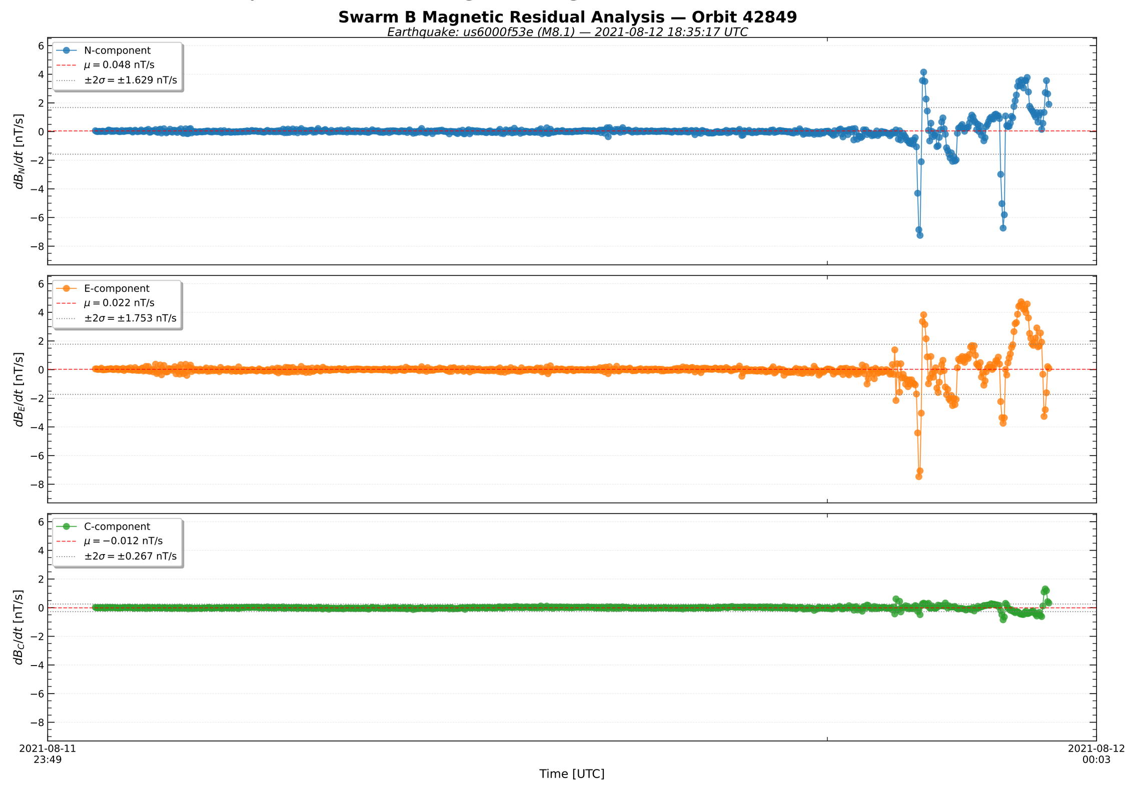


Trajectories over the Dobrovolsky Area. SWARM Satellite **B** trajectories over the Dobrovolsky Area for earthquake ID us6000jlqa. Multiple ground tracks of SWARM Satellite **B** (colored lines, labeled by orbit number in the legend) traverse the defined Dobrovolsky Area (shaded red), representing the potential ionospheric coupling zone for the earthquake. The epicenter is indicated by a yellow star, centrally located within the Dobrovolsky Area [4]. Repeated satellite crossings provide multiple opportunities to observe potential pre-seismic ionospheric anomalies. The x-axis represents Longitude and the y-axis Latitude.

Anomaly detection integrates statistical process control with machine learning methodologies. Exponentially weighted moving average (EWMA) control charts monitor residual field variations relative to baseline statistics, while Isolation Forest algorithms [5] detect multivariate anomalies through random partitioning and path length analysis. Principal Component Analysis (PCA) reduces dimensionality, capturing dominant variance structures and identifying deviations via reconstruction errors. Spectral analysis, employing Welch's method for power spectral density estimation, characterizes frequency-domain features and temporal evolution of dominant frequencies potentially associated with seismic precursors. Collectively, these methods constitute an integrated analytical framework capable of detecting and interpreting subtle geophysical anomalies in satellite magnetic field data. This framework provides a robust computational infrastructure for investigating potential pre-seismic electromagnetic signatures, complementing SWARM's primary objectives of probing core dynamics, crustal magnetization, and space weather processes. By systematically combining statistical, machine learning, and spectral approaches, the framework enhances reproducibility and interpretability of potential earthquake precursors, advancing our understanding of lithosphere–atmosphere–ionosphere coupling and its implications for seismic forecasting.

Data Validation and Analysis

The analysis utilizes SWARM Level 1B data products, providing 1 Hz vector magnetic field measurements with precision better than 0.1 nT, accompanied by calibration metadata and uncertainty estimates. Data are available in multiple coordinate systems (GCI, GSM and NEC), enabling consistent geophysical interpretation. The dataset spans the full SWARM mission from November 2013 onward, supporting both short-term and long-term investigations of magnetic field variability. Earthquake information is integrated from global seismic catalogs, including detailed temporal, spatial, and magnitude parameters, as well as focal mechanisms and tectonic classifications. Configurable selection criteria permit targeted studies by region, depth, or magnitude range, facilitating correlation analyses between geomagnetic anomalies and seismic events.



Residuals magnetic Data: Time series of dB/dt residuals for SWARM Satellite B during orbit 42820, associated with earthquake ID us6000f53e (Magnitude 8.1). The plot displays residuals for the North (dN/dt), East (dE/dt), and Central (dC/dt) components of the magnetic field time derivative in nT/s. Each subplot shows the raw residual data, the mean (red dashed line), and $\pm 2\sigma$ thresholds. Significant transient anomalies, characterized by sharp deviations exceeding $\pm 2\sigma$ limits, are observed primarily in the East and North components around 00:00 UTC on 12 August 2021—approximately 18 hours prior to the earthquake. These deviations depart from expected geomagnetic behavior and may indicate pre-seismic ionospheric perturbations. Legends specify each component, mean value, and corresponding 2σ thresholds.

Analytical validation employs multiple strategies to ensure robustness and reproducibility. Cross-validation assesses generalization performance, estimates confidence intervals, and evaluates statistical significance. These procedures establish the reliability, reproducibility, and scientific rigor of the framework's findings, supporting the credible identification of potential pre-seismic electromagnetic signatures.

Summary and Conclusions

We introduce an integrated computational framework for detecting potential pre-seismic electromagnetic anomalies in SWARM satellite magnetic field data. By combining Statistical Process Control, Isolation Forest machine learning, Principal Component Analysis, and spectral analysis, the system enables robust identification of ionospheric perturbations potentially associated with LAIC processes. Rigorous preprocessing and geomagnetic filtering ensure data integrity, while parallelized computation allows efficient analysis of multiple seismic events. The framework effectively distinguishes statistically significant anomalies within Dobrovolsky preparation zones, accounting for natural geophysical variability. Its multi-method design enhances both sensitivity and interpretability, providing a reproducible foundation for investigating earthquake precursor phenomena. In summary, this pipeline advances computational tools for space-based earthquake precursor research. It establishes a scalable, data-driven methodology that supports scientific exploration of LAIC mechanisms and lays the groundwork for future applications in predictive seismic monitoring and early warning systems.

Future Work

Future work will focus on enhancing the framework's robustness and predictive capabilities. The first objective is to implement rigorous statistical significance testing to validate detected anomalies. Comparative analyses between pre and non-seismic periods, combined with nonparametric tests, Monte Carlo simulations, and false detection rate, will quantify the probability that anomalies arise by chance, ensuring statistical reliability and interpretability. The second objective is to develop a machine learning–based predictive model for near-real-time earthquake forecasting. Statistically validated anomalies will serve as input features, capturing spatial, temporal, and multi-parameter characteristics of pre-seismic activity. Advanced sequence models, including recurrent neural networks and transformer architectures, will be explored to estimate seismic probabilities within defined spatio-temporal windows. Ultimately, the framework aims to evolve into a real-time processing system that integrates continuous satellite data ingestion, adaptive anomaly detection, and probabilistic event scoring. By combining rigorous statistical validation with data-driven prediction, the system will transition from retrospective analysis to operational forecasting, providing a reproducible and quantitative foundation for earthquake hazard assessment.

References

[1] S. Harrigan et al. Detection of electromagnetic seismic precursors from swarm satellite magnetic field data. *Remote Sensing*, 16(16):2985, 2024.

[2] De Santis et al. Precursory worldwide signatures of earthquake occurrences on swarm satellite data. *Scientific Reports*, 9:20287, 2019.

[3] Ian T Jolliffe. Principal component analysis. *Springer Series in Statistics*, 2016. Standard reference for PCA and dimensionality reduction methods.

[4] Dobrovolsky et al. Estimation of the size of earthquake preparation zones. *Pure and Applied Geophysics*, 117(5-6):1025–1044, 1979.

[5] Liu et al. Isolation forest. *2008 Eighth IEEE International Conference on Data Mining*, pages 413–422, 2008.