Space Weather

RESEARCH ARTICLE
10.1029/2024SW003982

Toward Enhanced Prediction of High-Impact Solar Energetic Particle Events Using Multimodal Time Series Data Fusion Models

Pouya Hosseinzadeh[1], Soukaina Filali Boubrahimi[1], and Shah Muhammad Hamdi[1]

[1]Department of Computer Science, Utah State University, Logan, UT, USA

Key Points:

- Combining proton flux data and solar *X*-ray images for precise Solar Energetic Particle event prediction
- Utilizing six specialized machine learning models to improve SEP event prediction accuracy through time series analysis
- Exploring the impact of varying observation window sizes on classification accuracy in data fusion

Correspondence to:

P. Hosseinzadeh,
pouya.hosseinzadeh@usu.edu

Citation:

Hosseinzadeh, P., Filali Boubrahimi, S., & Hamdi, S. M. (2024). Toward enhanced prediction of high-impact solar energetic particle events using multimodal time series data fusion models. *Space Weather*, *22*, e2024SW003982. https://doi.org/10.1029/2024SW003982

Received 1 MAY 2024
Accepted 23 MAY 2024

Abstract Solar energetic particle (SEP) events, originating from solar flares and Coronal Mass Ejections, present significant hazards to space exploration and technology on Earth. Accurate prediction of these high-energy events is essential for safeguarding astronauts, spacecraft, and electronic systems. In this study, we conduct an in-depth investigation into the application of multimodal data fusion techniques for the prediction of high-energy SEP events, particularly ∼100 MeV events. Our research utilizes six machine learning (ML) models, each finely tuned for time series analysis, including Univariate Time Series (UTS), Image-based model (Image), Univariate Feature Concatenation (UFC), Univariate Deep Concatenation (UDC), Univariate Deep Merge (UDM), and Univariate Score Concatenation (USC). By combining time series proton flux data with solar X-ray images, we exploit complementary insights into the underlying solar phenomena responsible for SEP events. Rigorous evaluation metrics, including accuracy, F1-score, and other established measures, are applied, along with *K*-fold cross-validation, to ensure the robustness and generalization of our models. Additionally, we explore the influence of observation window sizes on classification accuracy.

Plain Language Summary This study is centered on forecasting solar energetic particle (SEP) events, which can pose serious risks to astronauts and technology. We employed advanced machine learning (ML) techniques to make these predictions. Our research involved integrating various data sources, including information about protons and *X*-rays emitted by the Sun, to enhance the accuracy of our forecasts. We evaluated the performance of four distinct multimodal time series data fusion models along with two distinct unimodal time series models to determine the most effective approach for predicting these solar events. Additionally, we examined how the choice of time window for predictions influenced their accuracy. To aid in interpreting the results, we utilized visual representations known as heatmaps, which provide a graphical view of the data. Our findings have significant implications for improving the safety and success of space missions. By achieving precise predictions of SEP events, we can better protect astronauts, spacecraft, and vital electronic systems both in space and on Earth.

1. Introduction

Solar energetic particle (SEP) events, arising from the intense activities of the Sun, specifically solar flares and coronal mass ejections (CMEs), represent a formidable challenge and concern for space exploration (Forbush, 1946; Reames, 1995). These events release torrents of high-energy particles, including protons, electrons, and heavy ions, which ripple through the vast expanse of space, posing substantial risks to astronauts, spacecraft, and terrestrial electronic devices (Wu et al., 2009). Accurate prediction of SEP events is crucial to ensure the safety and success of space missions and interplanetary exploration. In light of the National Aeronautics and Space Administration's (NASA) ambitious plans to explore the Moon and Mars in the coming decades (Smith et al., 2020; Watson-Morgan et al., 2023), the urgency of developing advanced monitoring and predictive systems for SEP events cannot be overstated.

SEP events are uniquely characterized by the rapid expulsion of these energetic particles from the Sun. These particles, when they interact with astronauts, present health risks such as cancer and skin reactions (Fogtman et al., 2023; Simonsen & Slaba, 2021; Walsh et al., 2019). Furthermore, they can wreak havoc on electronic devices and communication systems here on Earth (J.-z. Wang et al., 2022; Chen, Carver, et al., 2021). Additionally, the consequences of SEP events extend into the realm of space missions, potentially jeopardizing their success. This has been notably emphasized in the context of Mars missions, where the unpredictability of SEP

events has led to critical mission planning challenges. Hence, the accurate forecasting and proactive mitigation of risks associated with SEP events call for the development of advanced monitoring and predictive systems.

The scientific community has approached the prediction of SEP events through two principal methodologies: physics-based models and data-driven models (Whitman et al., 2022). Physics-based models delve into the intricate mechanisms underpinning SEP events, seeking to understand the conversion of magnetic energy into thermal energy during solar flares, among other phenomena. While these models, exemplified by the SOLar Particle Engineering Code (SOLPENCO), have proven effective in predicting gradual SEP events and have been widely adopted, they are constrained by their reliance on a priori assumptions and simplified representations of the complex solar processes (Aran et al., 2006). However, factors like computational time and the utilization of non-real-time input observations can hinder physics-based models' capacity to deliver forecasts with the immediacy essential for operational decision-making (Lario, 2005).

In contrast, data-driven models, often leveraging machine learning (ML) techniques, extract knowledge from historical observations to discern patterns and precursors of SEP events for predicting future occurrences. These models circumvent the need for explicit physical models and have exhibited promise in predicting SEP events, particularly in classification tasks. For instance, the work of Boubrahimi et al. (2017) illuminated that multivariate time series of X-ray and proton flux data can serve as precursors for predicting SEP events exceeding 100 MeV (Boubrahimi et al., 2017). Moreover, Aminalragia-Giamini et al. (2021) achieved success by training a deep learning model using solar flare soft X-ray measurements to predict the onset of SEP events (Aminalragia-Giamini et al., 2021). Also, the UMASEP (University of MAlaga Solar particle Event Predictor) model incorporates ML in its prediction process (Núñez & Paul-Pena, 2020). This model has been operational in real-time at the CCMC (Community Coordinated Modeling Center) for many years.

Lavasa et al. (2021) utilized a variety of ML techniques for predicting SEPs, including logistic regression|Linear Regression (LR), fully connected multi-layer perceptron neural networks, random forests (RF), support vector machines (SVM), decision trees (DTs), extremely randomized trees (XT), and extreme gradient boosting (XGB). Additionally, Stumpo et al. (2021) employed SVM and LR methods to achieve SEP predictions similar to the empirical model for solar proton events real time alert concept (Laurenza et al., 2009). Furthermore, multivariate time series data augmentation techniques have been effectively applied to ML models to enhance the prediction of SEP events (Bahri et al., 2023; Hosseinzadeh et al., 2023, 2024).

In a bid to further elevate the precision of SEP event prediction, we have turned to data fusion techniques, which combine information from multiple data sources or modalities (Baltrušaitis et al., 2018). In our study, we concentrate on time series data fusion methods, which combine time series proton flux data and solar X-ray image data. These modalities furnish complementary insights into the solar phenomena underpinning SEP events. Prior research has underscored the efficacy of multimodal approaches in enhancing prediction accuracy (Ismail et al., 2020; Ngiam et al., 2011; W. Wang et al., 2020). By harnessing the insights derived from both proton flux data and solar X-ray images, our objective is to fashion a comprehensive and robust predictive system tailored for SEP events exceeding 100 MeV.

This paper represents a comprehensive exploration of time series data fusion techniques for the prediction of ∼100 MeV SEP events. We employ six ML models, meticulously tailored for time series analysis, comprising univariate time series (UTS), Image-based model (Image), univariate feature concatenation (UFC), univariate deep concatenation (UDC), univariate deep merge (UDM), and univariate score concatenation (USC). These models have gained attention for their effectiveness in a variety of time series prediction tasks (Sleeman IV et al., 2022), and they potentially provide a wide range of approaches for exploring the domain of SEP event classification.

To assess the performance of our models, we apply stringent evaluation metrics, including accuracy, F1-score, and other pertinent measures that are commonplace in similar studies (Alshammari et al., 2024; Aminalragia-Giamini et al., 2021; Bain et al., 2018; Saini et al., 2024). We conduct 5-fold cross-validation and engage in extensive testing to ensure the robustness of our models' performance evaluation and their capacity for generalization. Additionally, we delve into vector analysis to investigate the influence of different observation window sizes on the classification accuracy of our models, in line with the work of Boubrahimi et al. (2017).

Through the development of accurate prediction models for ∼100 MeV SEP events, our aim is to significantly enhance the safety and success of space missions. Our findings not only contribute to the evolving field of SEP

event prediction but also align seamlessly with NASA's strategic objectives in lunar and Martian exploration, as outlined in Whitman et al. (2022). By using data fusion techniques and advanced ML models, our focus lies in the accurate classification of high-energy SEP events. While our approach enhances the ability to differentiate between these event types, it's crucial to clarify that our model doesn't directly address real-time operational forecasting or provide estimations regarding the timing of SEP occurrences. Instead, our emphasis is on effectively categorizing these events, contributing to the broader understanding of space weather dynamics without delving into the real-time prediction of their occurrences or timing.

2. Data and Methods

In this section, we initiate by providing an overview of the data sources and detailing the data collection procedures. Following this introduction to our data sets, we then proceed to elucidate the various methods employed in this study for the fusion of multimodal data.

2.1. Data Set

Consider a scenario involving multiple data sources, denoted as $M1$, $M2$, …, Mn. Each Mi represents a distinct data modality. In this context, we have a sample, $x = (x0, x1, \ldots, xn)$, where each xi corresponds to the features or representation of the modality Mi. These samples are associated with a set of exclusive classes, $C1$, $C2$, …, Cl, which are used to train a multimodal classifier, denoted as f: $M \rightarrow C$, where M represents the feature space with a dimension of T.

The fusion-based multimodal classifiers in this research were trained using two distinct data modalities: time series data and image data. Modality $M1$ represents the time series proton flux data of the Sun, collected from the Geostatioanry Operational Environmental Satellite (GOES) satellite. These data are publicly accessible at: GOES Space Environment Monitor Data: https://www.ncei.noaa.gov/data/goes-space-environment-monitor/access/avg/. The analysis exclusively utilized data from primary spacecraft within the GOES program. Specifically, data from primary spacecraft were employed, while data from secondary spacecraft were not included. Consequently, the potential difference in particle flux between primary and secondary spacecraft was not a factor in this study. We considered proton flux P6 channel (P6_flux) data recorded by the Energetic Proton, Electron and Alpha Detectors. Time series data play a crucial role in understanding the dynamic behavior of SEP events, providing valuable information about the intensity and temporal characteristics of proton flux emanating from the Sun.

The second modality, $M2$, consists of images of the Sun captured by Solar and Heliospheric Observatory (SOHO) from AIA (Atmospheric Imaging Assembly) 304, and is available at Helioviewer: https://helioviewer.org/. These images provide a visual representation of solar activity and help capture the spatial characteristics of the Sun during SEP events. To effectively utilize these images in our analysis, we employed an autoencoder, a neural network architecture consisting of two main components: an encoder and a decoder.

During the training phase, the encoder is responsible for downsizing input images to a latent space, effectively capturing essential features while reducing dimensionality. The decoder then reconstructs these representations to match the original image size. This process aims to minimize differences between the original images and their reconstructions, effectively learning data compression while preserving critical information. After training, we can use the encoder's hidden layer as a compact and meaningful image representation, highlighting salient features in a lower-dimensional space (Thomas et al., 2016). Therefore, we convert the images into vectors.

Figure 1 illustrates the key steps involved in our data-driven approach, encompassing data collection, data fusion, and the subsequent classification process. The process initiates with the collection of data related to both SEP and NSEP (non-SEP) events. SEP events are obtained from the GSEP catalog (https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DZYLHK), while NSEP events are sourced from the HEK website (https://www.lmsal.com/isolsearch). In this work, non-SEP events are defined as X-ray events (solar flares) whose peak intensity is at least C1.3 but did not lead to any SEP event. Therefore, we assume that a ∼100 MeV impulsive event may occur if the parent X-ray event peak is at least C1.3 since SEP events are associated with solar flares as was suggested in Boubrahimi et al. (2017); Núñez (2011). These data sources provide essential information about the occurrences of SEP and NSEP events. Subsequently, from each of these event lists, we

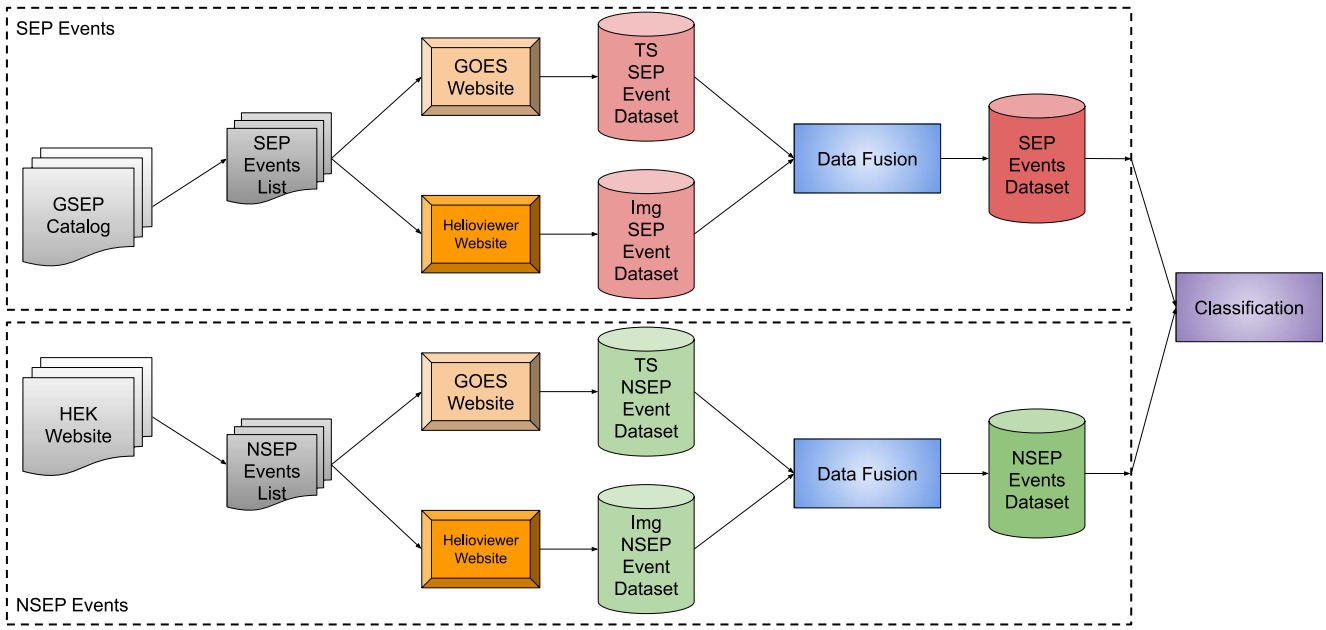

Figure 1. Data collection, fusion, and classification diagram.

collect the respective data sets from appropriate sources. These data sets serve as the foundational information for our subsequent analysis.

Data fusion, a process integral to our study, involves the strategic combination of information from both time series and X-ray images data sets. This merging technique is pivotal as it integrates the diverse yet complementary perspectives provided by the time series and image data modalities. By synthesizing these distinct sources of information, our approach aims to construct a unified, comprehensive portrayal of the solar events. This method enables us to capitalize on the distinctive strengths of each modality, using the precise temporal details from time series data and the spatial context offered by image data. This combined representation facilitates a more profound and nuanced understanding of the underlying solar phenomena associated with these events, enhancing the robustness and depth of our analysis. Finally, the fused data sets are employed in the classification process, where ML models, as detailed in the previous section, are applied to predict SEP events associated with ∼100 MeV. This integrated approach, spanning data collection, fusion, and classification, forms the core of our research methodology and contributes to the accurate prediction of high-impact SEP events.

Figure 2 shows two different processes employed for time series and image-based data. Focusing on proton flux channel P6 in the time series data suggests a targeted analysis approach in ∼100 MeV events prediction. This channel likely holds specific relevance or exhibits distinctive patterns crucial to the study. Further investigation into the behavior of this channel could lead to insights valuable for space weather forecasting and radiation hazard mitigation strategies. On the other hand, employing autoencoders to convert images to vectors for image data processing indicates a transformational approach. Autoencoders compress the image data into lower-dimensional vectors, capturing essential features while reducing noise and irrelevant information. This method can facilitate easier processing and analysis of the images by condensing them into a more manageable format. The contrast in methodologies highlights the tailored strategies for handling different data types.

Figure 3 provides a visual representation of our data using t-SNE, a technique that reduces the dimensionality of data and projects it into two dimensions while preserving data patterns. In this visualization, we combined 5-hr time series observation windows with 60-Dimensional image transformation vectors to create the t-SNE projection. Specifically, we employed t-SNE to visualize the extracted features, including mean and standard deviation, from both time series and vector data.

The *t*-SNE projection effectively separates time series data from image data, underscoring the diverse nature of the data modalities and the complementary information they offer. While it is expected that the 2D projections of completely diverse data samples, such as vector embeddings of single-shot images and time-series observations,

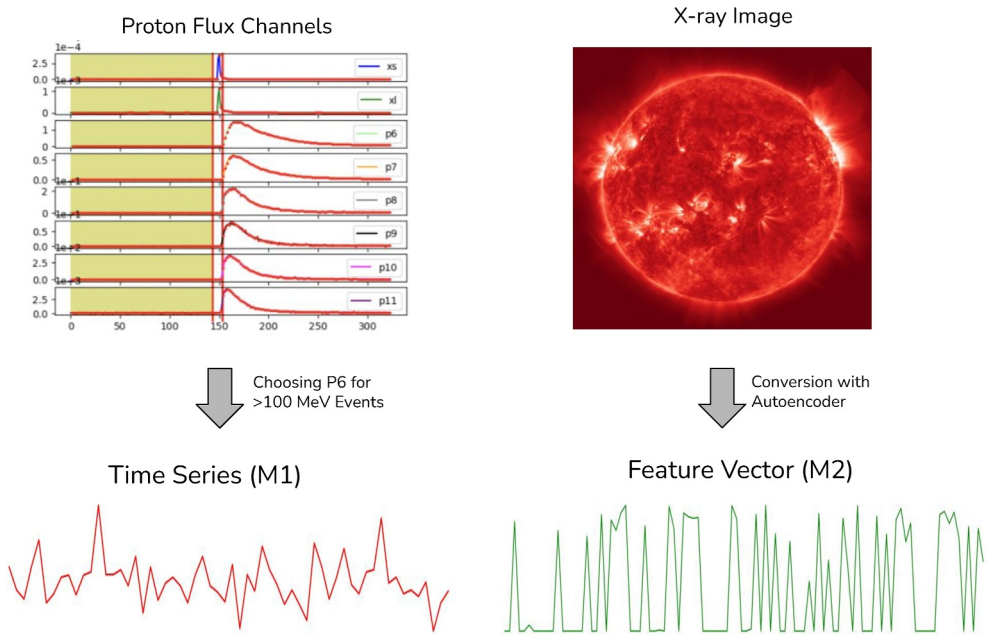

Figure 2. Selecting proton flux channel P6 for time series data alongside image vectorization via Autoencoders for X-ray images using toy data.

may not be similar, the visualization serves as a valuable tool for understanding the inherent structures within each modality. Essentially, the takeaway from Figure 3 is the insight it provides into the underlying data distribution.

Table 1 highlights critical information such as the SEP start time, flare start time, flare peak time, flare class, and NOAA active region for SEP events used in our study. Notably, the flare start time and flare peak time signify the initiation time of the solar flare and the moment of maximum X-ray flux during the flare event, respectively. It is important to note that for this work, we adopt the flare start time as the last observed time within our specified and fixed observation window for time series data. This specific data from the observation window serves as the input to our time series data to be fed to the models, ensuring consistency and relevance.

Likewise, we selected an X-ray image captured before the start time of the corresponding solar flare, for both SEP and NSEP events. However, it is essential to acknowledge the limitations in the data provided by SOHO that may affect our image data collection process. Due to the nature of SOHO's data availability, these single-snapshots were collected approximately a few hours before the solar flare start time. While there is no specific time mentioned for each sample, we ensured that the data were collected very close to the solar flare start time and did not surpass it. We made sure that we collect the image no more than 4 hr away from the solar flare start time. This is because SOHO provides images at specific times each day (but not frequently), and we aimed to select an image captured as close to the solar flare start time as possible. It is important to note that this variability in the timing of the X-ray images can indeed make it slightly harder for our predictive models to accurately predict SEP/NSEP events.

2.2. Methodology

In this study, we employed four distinct multimodal fusion techniques to harness the combined power of time series proton flux data and solar X-ray image data, as shown in Figure 4, and two unimodal techniques, for the accurate prediction of SEP events that associate with ∼100 MeV events. We utilized Time Series Forest (TSF) model as the primary classifier across all scenarios to train and test the data. TSF, implemented using the "sktime" library (Löning et al., 2019), operates as an ensemble technique, employing multiple DTs organized into a forest. Each DT within TSF is constructed via random interval-based sampling from the time series data, ensuring robustness and flexibility in capturing temporal patterns (Rigatti, 2017). This approach allows TSF to effectively model complex temporal relationships present in the data. To optimize the TSF model's performance, we conducted hyper-parameter tuning, focusing on adjusting the number of estimators. Through this process, we found that setting the number of estimators to 150 consistently yielded the most favorable results across all experiments.

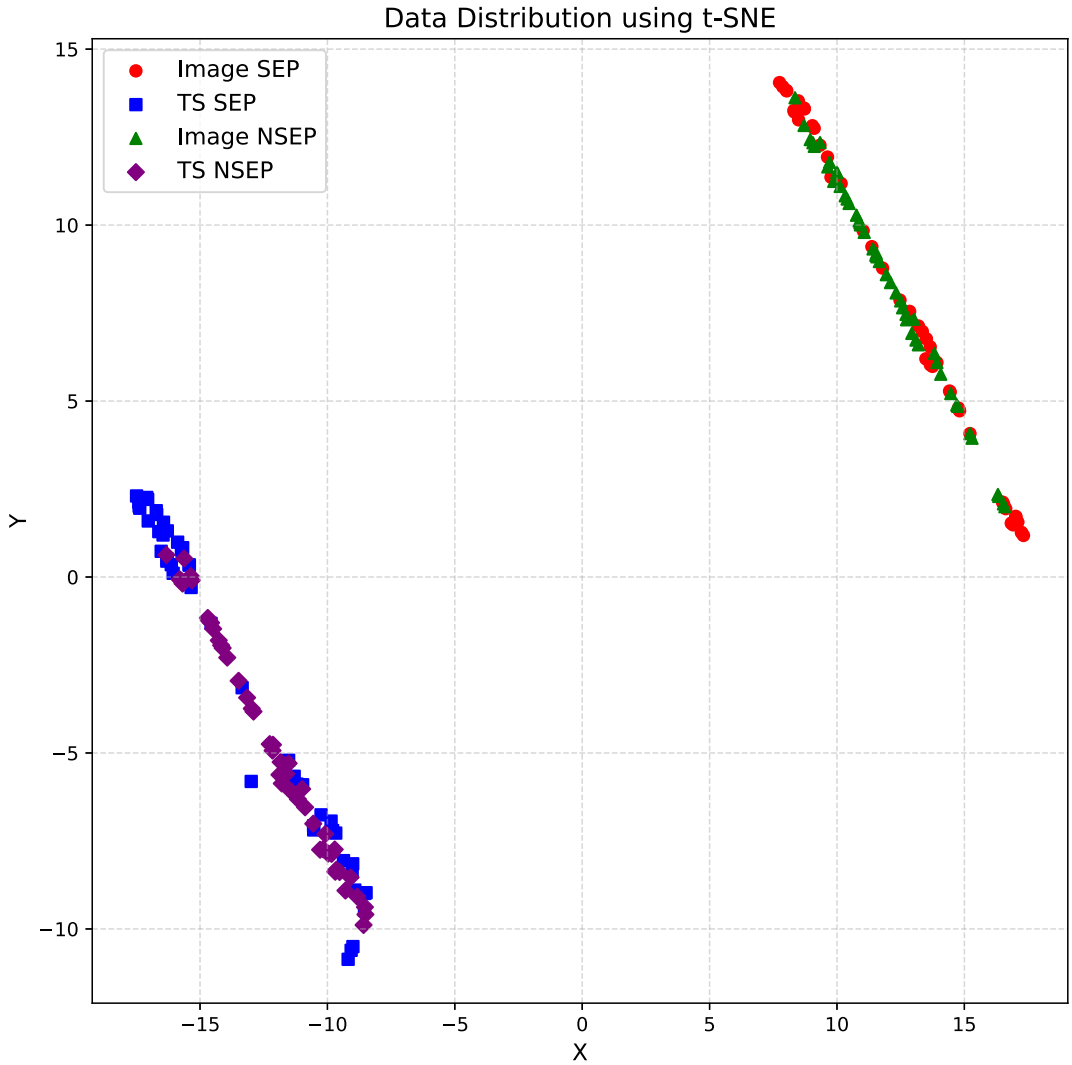

Figure 3. Data visualization of two modalities using 5 h time series observation window and 60-Dimensional image transformation vector data.

The findings presented in this paper stem from the model optimization process applied to the entire data set. However, we illustrate a scenario wherein hyper-parameter tuning has been exclusively conducted on the train/validation sets.

The multimodal fusion techniques employed in this study are derived from a review paper by Sleeman IV et al. (2022), which outlines a comprehensive framework for integrating diverse data sources using multimodal fusion approaches (Sleeman IV et al., 2022). These methods offer different strategies for integrating and exploiting information from multiple modalities to improve predictive accuracy. Let us delve into the details of each of these multimodal fusion techniques, along with two unimodal approaches.

- Univariate Time Series (UTS): The UTS method utilizes only unimodal data, specifically time series data. This approach focuses solely on the temporal characteristics of the proton flux data, offering a unimodal perspective to predict SEP events based on time series patterns.
- Image-based (Image): Similarly, the Image-based approach relies solely on unimodal data, specifically solar X-ray image data. By concentrating exclusively on the spatial characteristics observed in the images, this method provides insights into SEP event prediction through image analysis.
- Univariate feature concatenation (UFC): Feature concatenation involves merging features from multiple modalities into a single unified vector representation. In our case, this means combining the information from

Table 1
List of Solar Energetic Particle Events Associated With ∼100 MeV Energy Band and Their Corresponding Solar Flare Characteristics Based on GSEP Catalog

ID	SEP start time	Flare start time	Flare peak time	Flare class	Active region
1	1997–11–04 06:05:00	1997–11–04 05:52:00	1997–11–04 05:58:00	X2.1	8100
2	1997–11–06 12:20:00	1997–11–06 11:49:00	1997–11–06 11:55:00	X9.4	8100
3	1998–04–20 11:00:00	1998–04–20 09:38:00	1998–04–20 10:21:00	M1.4	8202
4	1998–05–06 08:15:00	1998–05–06 07:58:00	1998–05–06 08:09:00	X2.7	8210
5	1998–11–14 06:10:00	1998–11–14 05:00:00	1998–11–14 05:08:00	C1.3	8375
6	2000–06–10 17:00:00	2000–06–10 16:40:00	2000–06–10 17:02:00	M5.2	9026
7	2000–07–14 10:15:00	2000–07–14 10:03:00	2000–07–14 10:24:00	X5.7	9077
8	2000–07–22 11:50:00	2000–07–22 11:17:00	2000–07–22 11:34:00	M3.7	9085
9	2000–09–12 14:00:00	2000–09–12 11:31:00	2000–09–12 12:13:00	M1.0	9163
10	2000–10–16 08:00:00	2000–10–16 06:40:00	2000–10–16 07:28:00	M2.5	9182
11	2000–11–24 06:00:00	2000–11–24 04:55:00	2000–11–24 05:02:00	X2.0	9236
12	2000–11–24 15:55:00	2000–11–24 14:51:00	2000–11–24 15:13:00	X2.3	9236
13	2000–11–26 00:05:00	2000–11–25 00:59:00	2000–11–25 01:31:00	M8.2	9240
14	2001–01–28 16:45:00	2001–01–28 15:40:00	2001–01–28 16:00:00	M1.5	9313
15	2001–04–09 16:05:00	2001–04–09 15:20:00	2001–04–09 15:34:00	M7.9	9415
16	2001–04–15 13:50:00	2001–04–15 13:19:00	2001–04–15 13:50:00	X14.4	9415
17	2001–04–18 02:35:00	2001–04–18 02:11:00	2001–04–18 02:14:00	C2.2	9415
18	2001–05–20 06:55:00	2001–05–20 06:00:00	2001–05–20 06:03:00	M6.4	9455
19	2001–06–15 15:55:00	2001–06–15 16:15:00	–	C2.2	9494
20	2001–09–24 11:15:00	2001–09–24 09:32:00	2001–09–24 10:38:00	X2.6	9672
21	2001–10–22 16:35:00	2001–10–22 14:27:00	2001–10–22 15:08:00	M6.7	9672
22	2001–11–22 09:15:00	2001–11–22 20:18:00	2001–11–22 20:36:00	M3.7	9704
23	2001–12–26 05:45:00	2001–12–26 04:32:00	2001–12–26 05:40:00	M7.1	9742
24	2002–04–21 01:40:00	2002–04–21 00:43:00	2002–04–21 01:51:00	X1.5	9906
25	2002–08–24 01:15:00	2002–08–24 00:49:00	2002–08–24 01:12:00	X3.1	10,069
26	2003–05–28 05:25:00	2003–05–28 00:17:00	2003–05–28 00:27:00	X3.6	10,365
27	2003–05–31 02:40:00	2003–05–31 02:13:00	2003–05–31 02:24:00	M9.3	10,365
28	2003–10–26 17:40:00	2003–10–26 17:21:00	2003–10–26 18:19:00	X1.2	10,484
29	2003–10–28 11:35:00	2003–10–28 09:51:00	2003–10–28 11:10:00	X17.2	10,486
30	2003–10–29 21:45:00	2003–10–29 20:37:00	2003–10–29 20:49:00	X10.0	10,486
31	2003–1–04 21:40:00	2003–11–04 19:29:00	2003–11–04 19:50:00	X28.0	10,486
32	2011–06–07 06:55:00	2011–06–07 06:16:00	2011–06–07 06:41:00	M2.5	11,226
33	2011–08–09 08:10:00	2011–08–09 07:48:00	2011–08–09 08:05:00	X6.9	11,263
34	2011–09–06 01:45:00	2011–09–06 22:12:00	2011–09–06 22:20:00	X2.1	11,283
35	2012–01–23 04:10:00	2012–01–23 03:38:00	2012–01–23 03:59:00	M8.7	11,402
36	2012–01–27 17:55:00	2012–01–27 17:37:00	2012–01–27 18:37:00	X1.7	11,402
37	2012–03–13 17:35:00	2012–03–13 17:12:00	2012–03–13 17:41:00	M7.9	11,429

Note: Unknown data are denoted as: –.

time series proton flux data and solar X-ray images into a single feature vector. This concatenated vector serves as the input for a classification model. UFC is known for its simplicity and ease of implementation, making it a suitable choice for the initial exploration of multimodal data fusion (Usman & Rajpoot, 2017).

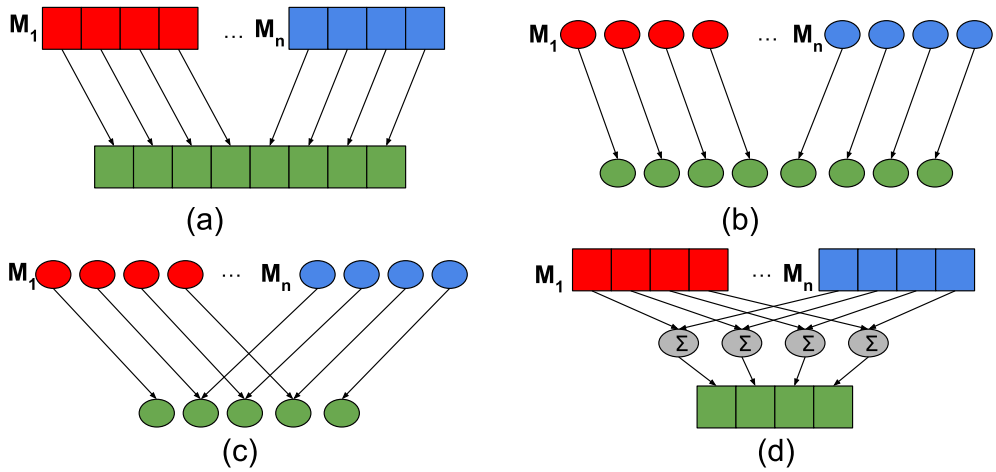

Figure 4. Architecture of multimodal data fusion models. (a) univariate feature concatenation, (b) univariate deep concatenation, (c) univariate deep merge, and (d) univariate score concatenation.

- Univariate deep concatenation (UDC): Deep concatenation, also known as deep feature concatenation, goes a step further by merging features at a deep representation level. Separate deep neural networks are used to extract high-level representations from each modality's data. These learned high-level representations are then concatenated before being used for the final prediction. UDC enables the model to capture complex interactions and dependencies between modalities, making it a powerful approach for multimodal fusion (Kang & Kang, 2017).
- Univariate deep merge (UDM): Deep Merge integrates features from different modalities at a lower level within the neural network. Typically, element-wise operations like summation or multiplication are used to combine these features. UDM aims to effectively capture both shared and unique information present in each modality, allowing for a more nuanced representation of the data. This technique can be particularly valuable when modalities contain complementary information (Yang et al., 2020).
- Univariate score concatenation (USC): The Score Concatenation technique involves training separate models for each modality. These individual models generate prediction scores. These scores are then concatenated and utilized as input to a final classification model. USC leverages the strengths of each modality-specific model and effectively combines their predictions, similar to an ensemble learning approach. This method is beneficial when modalities have distinct predictive power and characteristics (Guggenmos et al., 2020).

The selection of these multimodal fusion techniques and the inclusion of unimodal methods, UTS and Image, allows us to explore various strategies for leveraging the information present in time series proton flux data and solar X-ray images. Each method brings its unique advantages to the table, and our study aims to determine which model yields the most accurate predictions for SEP events ~100 MeV. These techniques are integral to our research methodology, contributing to the development of robust predictive models for high-impact SEP events.

We explain the metrics used in this paper to evaluate the classification of our models. Accuracy, as defined in Equation 1, quantifies the proportion of correct predictions made by the model. In contrast, the F1-score, as expressed in Equation 4, serves as a comprehensive metric that balances precision and recall, effectively identifying positive cases while minimizing the occurrence of both false positives and false negatives. The specific formulas for F1-score, precision (Equation 2), and recall (Equation 3) provide a quantitative representation of these critical performance measures.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1\_score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

True skill statistic (TSS) and updated Heidke Skill Score (HSS2) have been particularly used in the task of flare forecasting problem to assess the models in Chen, Kempton, et al. (2021). The TSS and HSS2 metrics are defined by Equations 5 and 6.

$$TSS = \frac{TP}{TP + FN} - \frac{FP}{FP + TN} \tag{5}$$

$$HSS2 = \frac{2 * ((TP * TN) - (FN * FP))}{(TP + FN) * (FN + TN) + (FP + TN) * (TP + FP)} \tag{6}$$

where, TP, FP, TN, and FN represent True Positive, False Positive, True Negative, and False Negative, respectively.

In our analyses, we employed a five-fold cross-validation approach to rigorously evaluate the performance of our predictive models. In this scheme, the data set is randomly partitioned into five equal-sized subsets, with each subset serving as a validation set once while the remaining data is used for training. This process is repeated five times, ensuring that every data point is included in the validation set exactly once. With five-fold cross-validation, we ensure that each fold contains a reasonable portion of the data for both training and testing in each fold while still providing a reasonable number of values for evaluating model performance metrics. Specifically, we used 37 SEP and 37 NSEP samples in the balanced setting, leaving 14 samples out of 74 for testing purposes. We also used stratified 5-fold cross-validation for imbalanced data to preserve class ratios across folds. In the imbalanced setting, we utilized 37 SEP samples and 104 NSEP samples, resulting in an SEP/NSEP ratio of 0.35.

3. Results

In this section, we present the experimental results of our study in detail. Our analysis involved the utilization of six ML models, each tailored for time series analysis, to predict SEP events with proton flux data and solar X-ray images. The first set of results focuses on the influence of different time series observation window sizes and feature vector lengths on model accuracy, represented through box plots and heatmaps.

We examined the effect of varying observation window sizes and feature vector lengths on the accuracy of two model types: UTS and Image. To assess the performance of UTS in the balanced setting, we created box plots showing the accuracy variations for different observation window sizes: 15 hr, 12 hr, 9 hr, 6 hr, and 3 hr, as shown in Figure 5. The results reveal that the 6-hr observation window size yields the best performance, with an accuracy of 79%, accompanied by minimal variance. Notably, the 9-hr and 12-hr observation window sizes also show strong performance. For the Image-based models, the box plots depict the accuracy variation with different feature vector lengths: 100, 200, 300, 400, and 500. In our study, the implementation of autoencoders was crucial for dynamically resizing images to various dimensions, specifically 100, 200, 300, 400, and 500. Leveraging the capabilities of autoencoders, we devised a sophisticated process to encode the original images into a latent space representation. This method not only ensures preservation of crucial image features but also enables seamless and efficient resizing without compromising the integrity or quality of the visuals. The adaptability of autoencoders in this context has significantly streamlined our workflow, allowing for versatile image manipulation while maintaining fidelity across multiple dimensions. Among these, image sizes of 300 and 500 exhibit the highest accuracy, indicating their superior predictive capabilities.

Moving beyond UTS and Image-based models, we employed heatmaps to visualize the performance of the remaining models: UFC, UDC, UDM, and USC. These heatmaps represent the average accuracy of each model (along with their variance estimates) while considering both observation window size (M1) and feature vector length (M2) as variables. As shown in Figure 5, the heatmap for UFC shows that the model's performance has been influenced by both observation window size and image size. It shows that the 6-hr observation window size and smaller image sizes result in higher accuracy, reaching 77%. UDC exhibits a pattern where smaller image sizes are beneficial when combined with a 6-hr observation window size, achieving an accuracy of 77%. Surprisingly, UDC performs exceptionally well with a larger observation window size of 15 hr, reaching an accuracy

10.1029/2024SW003982

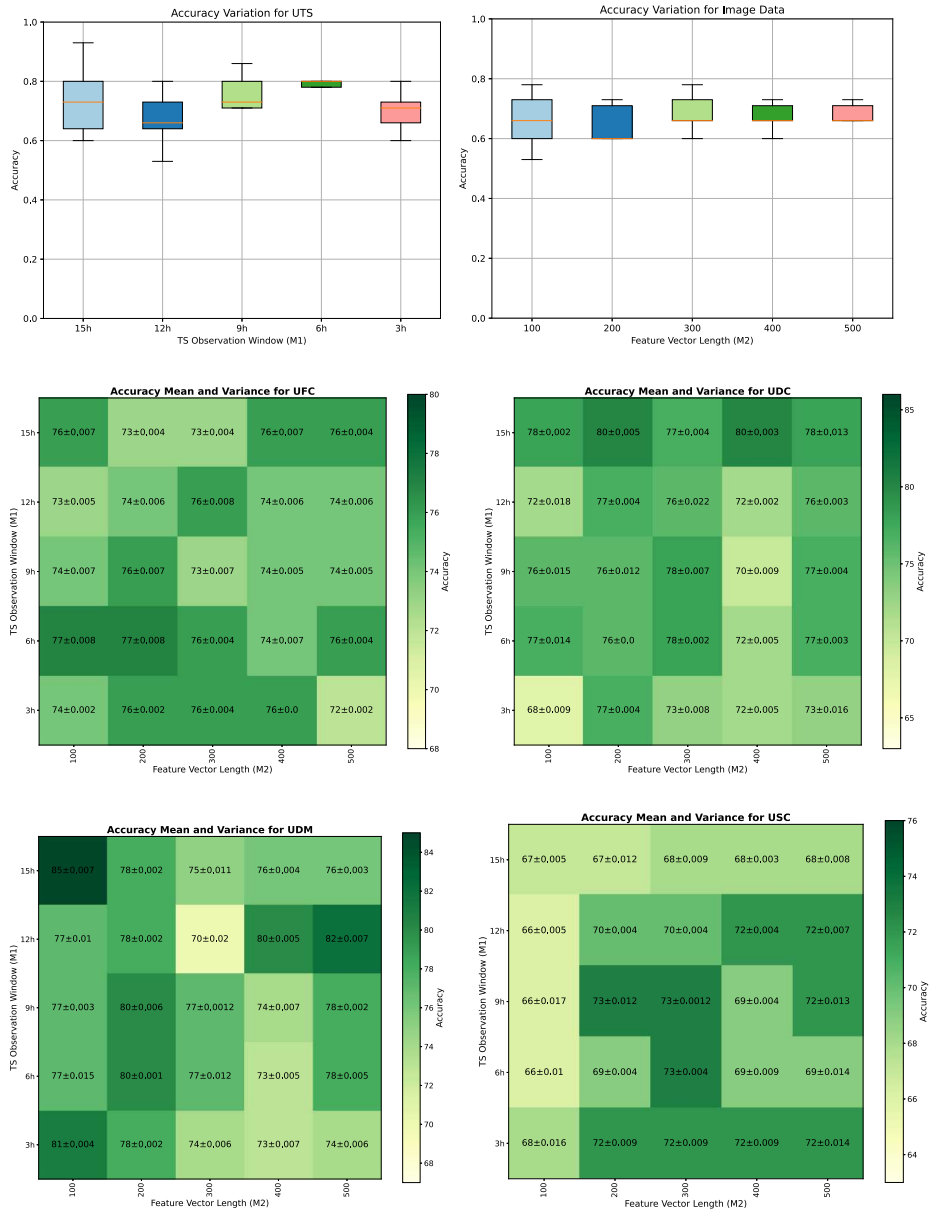

Figure 5. Box plots and average accuracy heatmaps with variance for all models using different time series observation window sizes and different feature vector lengths. The results are from five-fold cross validation performed on balanced data.

of 84%. In the case of UDM, a 15-hr observation window combined with a small image size of 100 leads to the highest accuracy, with an impressive 85%. USC did not demonstrate outstanding performance in our analysis, with accuracy variations depending on the choice of observation window size and image size. Nevertheless, it performed relatively better with a 9-hr observation window and image sizes of 200 or 300.

These results provide valuable insights into the optimal configurations for different ML models to predict high-energy SEP events. The data fusion techniques, as shown by these experiments, highlight the importance of selecting appropriate observation window sizes and image data lengths to enhance prediction accuracy, ultimately contributing to the safety and success of space missions and the protection of astronauts, spacecraft, and terrestrial electronic systems.

It is important to note that the inconsistency in the results and the variations in performance across these diverse dimensions underscore the intricate relationship between the data modalities and the fusion models employed.

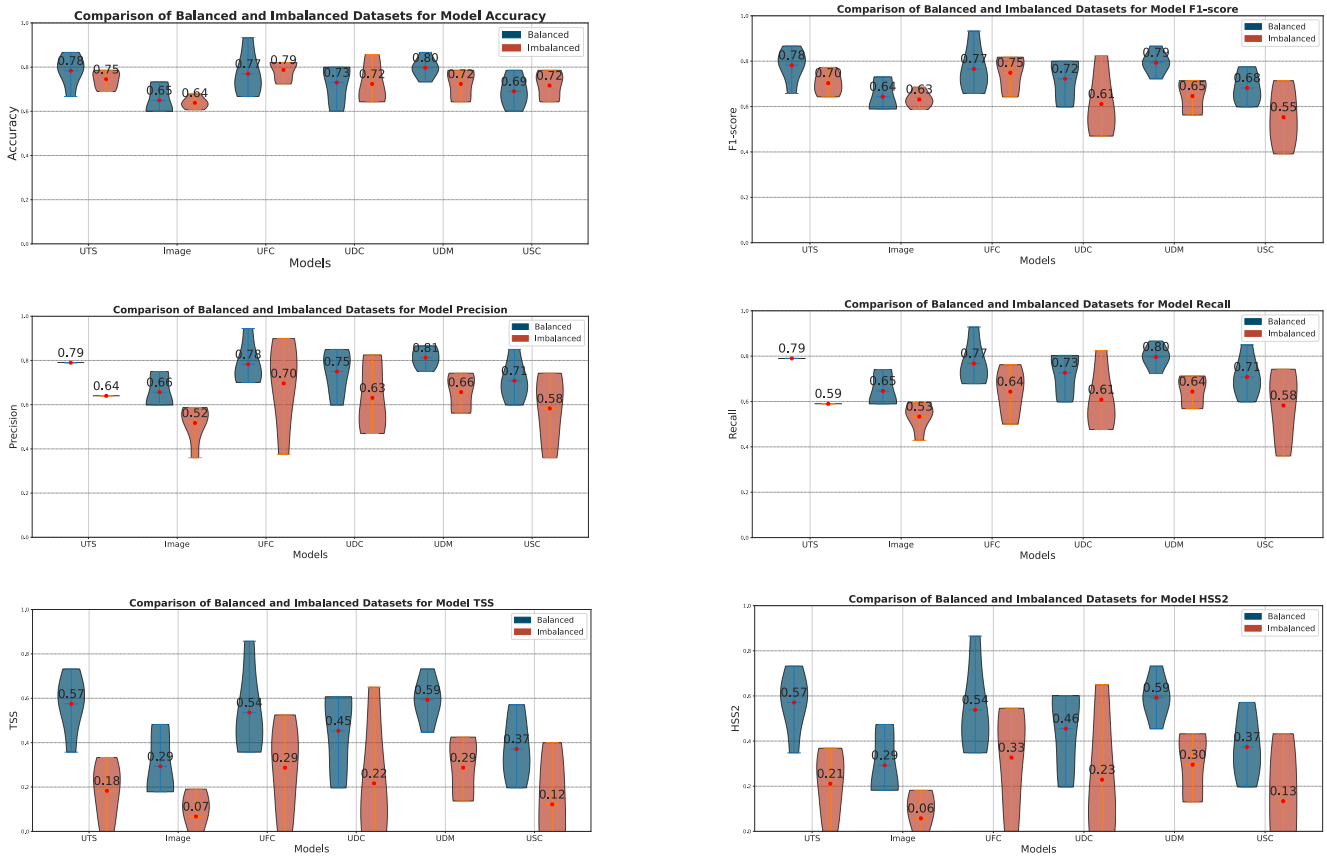

Figure 6. Evaluation of different models using 6 h time series observation window and 200-Dimensional image transformation vector data on balanced and imbalanced settings.

The observed fluctuations can be attributed to the inherent complexities arising from the interaction between the temporal granularity of time series data and the spatial resolution of image data. Different fusion models might exhibit varying degrees of sensitivity or robustness to changes in these dimensions, leading to the observed performance disparities. This inconsistency elucidates the intricate interplay between data granularity and fusion model adaptability, emphasizing the need for a more nuanced understanding of how varying data dimensions influence fusion model performance. Further analysis and interpretation of these results will be incorporated to offer deeper insights into the implications of data dimensionality on the efficacy of fusion models in capturing the complexities of solar phenomena.

In Figure 6, we present the results of our evaluation of various ML models using a specific configuration: a 6-hr time series observation window and 200-Dimensional image transformation vector data. This configuration was chosen based on the insights gained from Figure 5 of box plots and heatmaps, which identified it as one of the most promising settings for prediction accuracy.

The figure consists of six violin plots for balanced and imbalanced data, each representing different performance metrics, namely Accuracy, F1-score, Precision, Recall, TSS, and HSS2. These plots provide a comprehensive view of how each model performed under the specified conditions. Each plot shows the results of the 5-fold cross-validation training and testing process. As it can be seen in Figure 8, all models have higher performance on the balanced setting.

Notably, the UDM model emerges as the top performer across multiple metrics. It achieved an accuracy of 80% on the balanced setting, indicating its proficiency in accurately predicting high-energy SEP events. Furthermore, the F1-score, which balances precision and recall, highlights UDM's ability to strike a harmonious equilibrium, scoring an impressive 79% on the balanced setting.

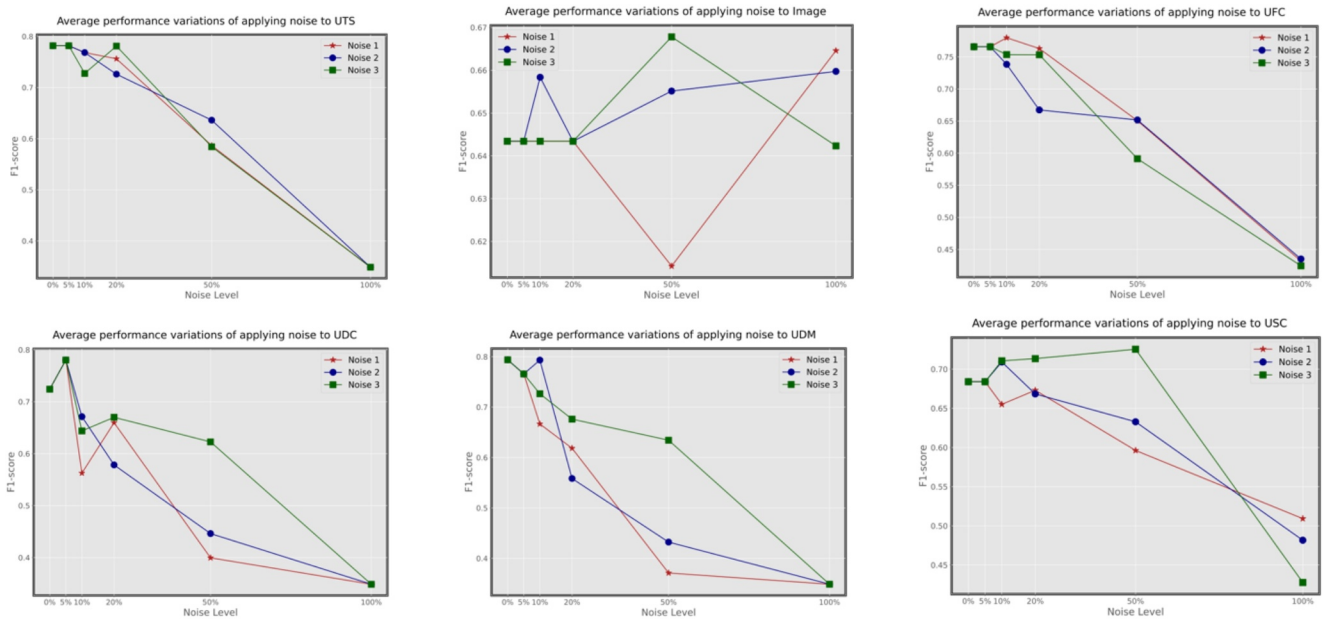

Figure 7. Average F1-score for models after applying different Gaussian noise levels to parts of test data samples. 0% defines the case where there is no noise. 100% signifies that every test data sample is impacted by noise.

Precision is another crucial metric, as it gauges the model's capacity to minimize false positive predictions. Univariate Deep Merge shines in this regard with an 81% precision score, implying that it excels in reducing erroneous predictions, thus enhancing the reliability of its forecasts.

In addition to these key metrics, the figure also provides insights into Recall, TSS, and HSS2 for all the evaluated models. UDM's strength becomes evident in its exceptional recall rate, affirming its capability to correctly identify positive instances, as well as its ability to discriminate between positive and negative cases, as indicated by its high TSS score. Moreover, its high HSS2 score underscores its skill in distinguishing between positive and negative cases while considering random chance. However, UTS shows significant precision and recall with the least variant among other models.

It is important to highlight the comparative performance of the Image-based model in the context of the results presented in Figure 6. The Image-based model exhibits the least favorable performance among the models evaluated. This outcome is largely anticipated due to the model's exclusive reliance on X-ray images, without considering the time series proton flux data. In the context of predicting high-energy SEP events, the absence of crucial time-dependent information is a significant limitation, and the results clearly reflect this constraint.

In summary, the results from Figure 6 reveal that, among the six models, UDM consistently shows superior performance with 80% accuracy, 79% F1-score, and 81% precision. These findings emphasize UDM's effectiveness in predicting high-energy SEP events in the specified configuration, reinforcing its potential to enhance the safety and success of space missions and to safeguard astronauts, spacecraft, and terrestrial electronic systems.

In this part of our experimental results, we investigate the performance of our ML models under varying levels of Gaussian noise. The aim is to assess how these models respond to different types of noise and different noise levels when predicting high-energy SEP events. We introduced three types of Gaussian noise—Noise 1, Noise 2, and Noise 3, each with a mean of 0 and different standard deviations, with Noise three exhibiting the highest standard deviation.

The results are presented in Figure 7, which consists of six subplots, each corresponding to one of our six models: UTS, Image, UFC, UDC, UDM, and USC. The *x*-axis in each subplot represents the levels of added noise, ranging from 0% (no noise) to 100% (maximum noise affecting all test data samples). The *y*-axis displays the F1-score, which provides a balanced measure of precision and recall.

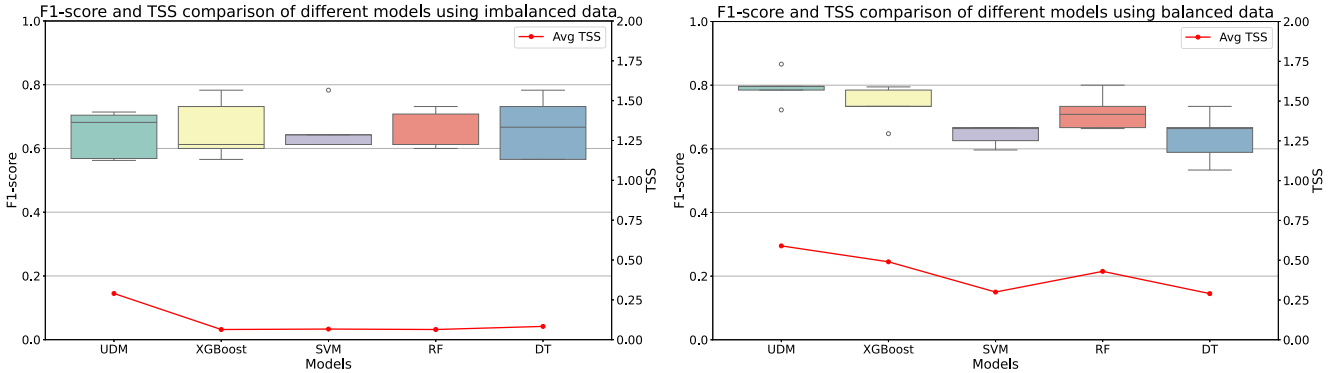

Figure 8. Comparison on model F1-score and average true skill statistic across different models on imbalanced and balanced settings.

The results show a clear trend where, as noise levels increase, the performance of all models consistently deteriorates. This decline is observed in all six models, reaffirming the sensitivity of the models to the presence of noise in the input data.

Furthermore, different types of noise (Noise 1, Noise 2, and Noise 3) influence model predictions differently. Among these, Noise 3 stands out as it has the most significant impact on model performance, especially when applied to all test data samples at the highest noise level. It notably degrades predictions in all models, with the exception of the Image-based model, which displays some fluctuations in performance. This distinction can be attributed to the fact that the Image-based model does not rely on time series data, and as such, the impact of noise on its performance exhibits different characteristics.

These findings are based on the average F1-scores obtained from 5-fold cross-validation, ensuring robust and reliable assessments. The standard deviation values used for Noise 1, Noise 2, and Noise 3 are 0.2, 0.4, and 0.6, respectively, highlighting the varying intensities of noise introduced to the training data.

The primary goal of this analysis is to show how our models behave in the presence of different types of noise and varying noise levels or amounts. These insights are essential for understanding the robustness of the models and their potential to make accurate predictions under real-world conditions, where noise in data is often inevitable. Specifically, we applied noise to distinct types of input data, enabling us to evaluate each model's response to noisy time series and image data separately.

Finally, our investigation concludes with a comparison between our leading model, the UDM, and a variety of state-of-the-art models prevalent in the SEP event prediction field, as delineated in Figure 8. This plot provides a comprehensive overview, illustrating the comparative performance of various models concerning both F1-score and average TSS. In this case, we decided to perform all classifiers on imbalanced data with 100 samples including both SEP and NSEP events, as shown in Figure 8a. Furthermore, we performed hyper-parameter tuning on train/validation sets and received the best hyper-parameters for each model, then applied the classification solely on test data sets. Notably, our evaluation encompasses influential models widely cited in SEP event prediction research, including XGBoost (Lavasa et al., 2021), SVM (Brea et al., 2018), RF classifier (Arnaut et al., 2023), and DT (DT) models (Boubrahimi et al., 2017; Ma et al., 2017). These models, recognized within the SEP and solar flare prediction community for their efficacy, contribute to the depth of our analysis, enriching the contextual understanding of our UDM's performance. univariate deep merge uses two different modalities (i.e., time series proton fluxes and extracted vectors) while other classifiers shown in Figure 8 use only time series proton flux data. We also show the results for balanced data in Figure 8b for the sake of comparison.

As our findings reveal, our UDM remarkably outperforms the other models in terms of both F1-score and average TSS. Specifically, in comparison with XGBoost, SVM, RF, and DT, our UDM consistently show superior performance. XGBoost exhibits a narrower variance in F1-score boxplots compared to other unimodal models. Notably, XGBoost emerges as the frontrunner among the evaluated models, showcasing higher average F1-score and average TSS than SVM. Moreover, RF and DT exhibit relatively lower performance in comparison, with RF displaying greater variance. This comprehensive comparison underscores the unequivocal efficacy of our UDM in predicting SEP events, highlighting its dominance across diverse performance metrics.

4. Discussion

4.1. Model Performance and Fusion Methodology

The performance disparities observed among the various fusion models, as illustrated in Figures 5 and 6, are reflective of the intricate interplay between data fusion methodologies and the inherent characteristics of distinct data modalities. The discrepancies in Figure 5 arise from the unique fusion strategies employed, where models like UDM excel in capturing complex interdependencies within the data. Conversely, the lower performance of the image-based model depicted in Figure 6 is indicative of the challenges encountered in leveraging image data for solar event prediction. The inherent complexities in image representation and the sensitivity of the model to variations in visual features contribute to this observed performance. Moreover, the exclusive reliance of the image-based model on visual cues might hinder its ability to capture nuanced temporal characteristics present in the solar event data, potentially leading to fluctuations in performance under different noise conditions.

4.2. Flare Class Distribution and Solar Cycle Insights

The flare class distribution, as shown in Figure 9, provides insights into the occurrence of solar flares across different classes, namely C, M, and X. It is essential to acknowledge that our focus on showcasing ∼100 MeV SEP events might influence the representation of flare classes in the plot, emphasizing the instances associated with higher-intensity flares. While the plot suggests a relatively higher representation of X-class flares among the observed ∼100 MeV SEP events, this selection bias could potentially amplify the visibility of these events. In reality, C-class flares are more frequently observed in solar activity. However, Figure 9 only demonstrates that SEPs are more correlated with strong flares. Therefore, the visualization, by nature of highlighting ∼100 MeV SEP events, accentuates the occurrences of higher-intensity flares without fully representing the overall distribution of solar flares across different classes.

We coupled this information with solar cycle data to gain a deeper understanding of the relationship between solar activity and SEP events. The barplots show the number of SEP events occurring within each solar cycle for each solar flare class. This analysis is crucial as it provides evidence of the cyclical nature of solar activity and its direct impact on space weather. It is worth noting that the solar cycle counts can be influenced by various factors, including the Sun's magnetic field dynamics, sunspot activity, and solar flare occurrence. We also present the peak intensity of solar X-ray radiation, or X-ray Flux, measured in W/m^2. This data has been collected from the GOES database, accessible at https://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-features/solar-flares/x-rays/goes/xrs/.

4.3. Addressing Class Imbalance and Refining Data Selection

One of the core challenges in predicting SEP events is the extreme class imbalance within the data set, which significantly complicates the training and validation of predictive models. In practical scenarios, SEP events are rare compared to the vast number of non-SEP events, a fact that is mirrored by our selective data extraction from the HEK database. Out of a potential 19,568 flares ≥C1.3 from 1997 to 2012, only 37 flares led to ∼100 MeV SEP events, as cataloged by GSEP. However, our analysis did not utilize all possible non-SEP events. Instead, we selectively included non-SEP events that met specific criteria to ensure relevance and manageability, resulting in a data set that does not reflect the entire spectrum of available data but is tailored for more focused analysis.

Initially, we employed a balanced data set as a proof-of-concept to establish a performance baseline in controlled conditions. This approach allowed us to fine-tune our models effectively without the overwhelming influence of non-SEP events. However, it is crucial to acknowledge that this method does not perfectly simulate the challenges encountered in operational environments where the data is highly imbalanced. Therefore, our results, while promising, serve primarily as a preliminary demonstration of the models' capabilities rather than a definitive assessment of their operational readiness.

In our study, we further refined our approach to handling real-world data distributions by employing careful selection criteria for non-SEP events. We excluded flares that, although not resulting in recognized SEP events, occurred close in time to significant SEP events and might have been associated with lower energy SEP events (e.g., 10, 30, 60 MeV). This careful curation aimed to reduce potential biases in model training and ensure the reliability of our predictions. Moving forward, future research should continue to explore additional techniques

10.1029/2024SW003982

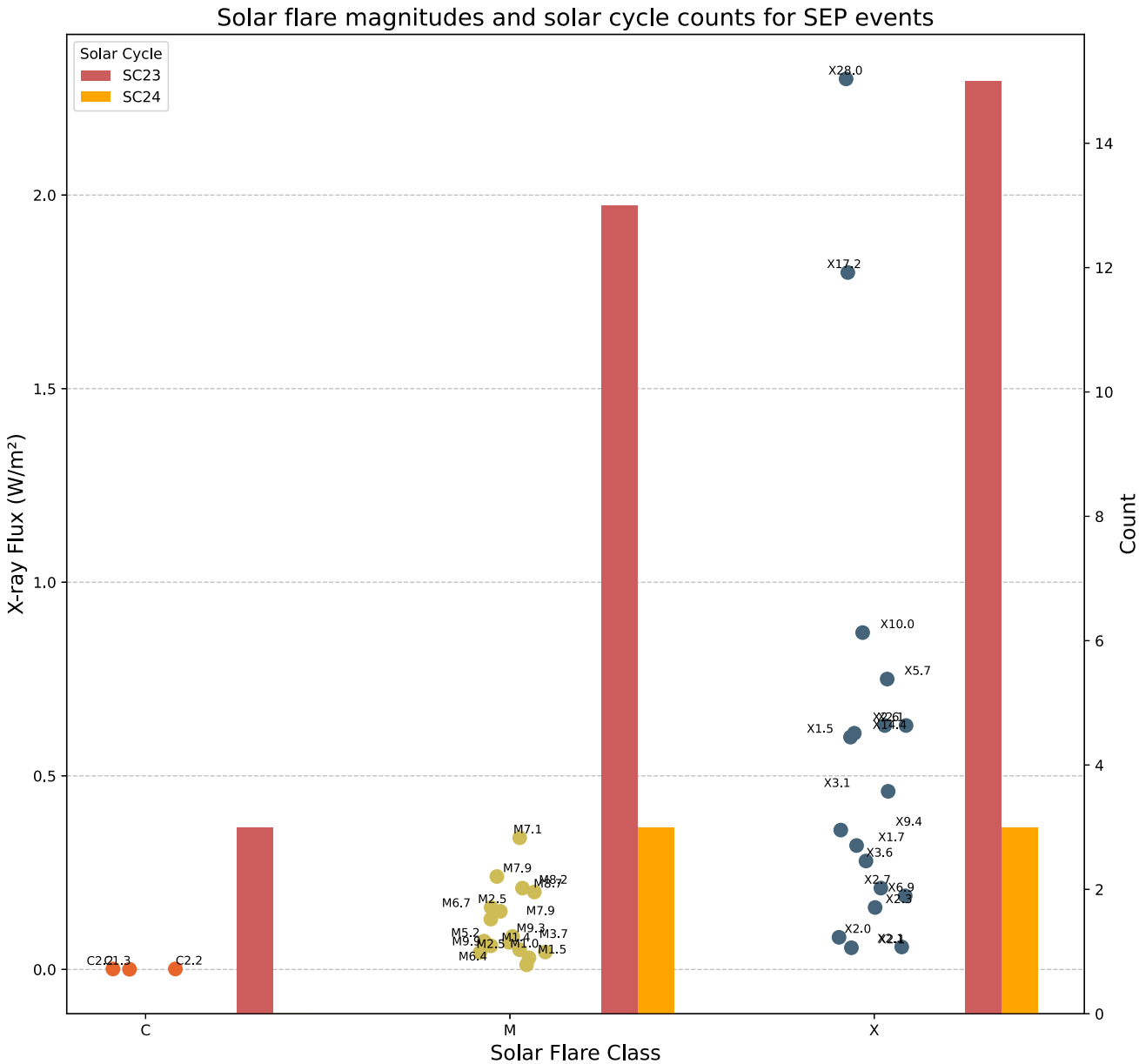

Figure 9. SEP X-ray flux and flare class distribution, and their corresponding solar cycle counts.

such as synthetic minority over-sampling and other data augmentation methods to enhance the models' sensitivity to rare SEP events.

4.4. Data Collection Challenges and Model Validation

A critical limitation in the field arises from the scarcity of historical data, primarily stemming from the protracted duration of solar cycles. The availability of comprehensive records for ∼100 MeV SEP events is notably limited, posing challenges in constructing reliable ML models. Such data scarcity may lead to the issue of overfitting, which can undermine prediction accuracy. To address this challenge, we performed our final evaluations on unseen test data, as shown in Figure 8. Furthermore, we used a higher number of data samples by collecting more NSEP events (reaching to 141 samples including both events), creating an imbalanced setting to evaluate the models on a more realistic application. The quality and accessibility of data constitute a pivotal challenge, as solar monitoring instruments have inherent limitations, potentially resulting in data gaps or inaccuracies. Therefore, addressing data quality concerns is of utmost importance in bolstering the trustworthiness of predictive models.

4.5. Prospects for Real-Time Prediction and Future Research

The demand for real-time prediction in the context of space mission planning amplifies the complexity of the field. In contemplating the potential transition of our methodology to real-time forecasting, several crucial elements come into play. Firstly, the availability of real-time data sources, such as ongoing streams of solar activity observations, is imperative. Secondly, the development of predictive models optimized for real-time applications would be necessary, considering the dynamic nature of space weather phenomena. Furthermore, a thorough analysis and classification on highly imbalanced setting is needed prior to begin with a real-time classification scenario. Lastly, establishing a robust infrastructure capable of continuous monitoring and processing of incoming data would facilitate the timely and effective utilization of classification methodologies for real-time forecasting. While our current study establishes effective classification techniques, the practical implementation of these methods for real-time forecasting warrants dedicated research and collaboration within the space weather monitoring community.

4.6. Addressing Anomalies and Data Quality

The susceptibility of space weather data to noise and anomalies constitutes a challenge that necessitates effective management. The detection and mitigation of data irregularities are pivotal to preserving prediction accuracy and reliability. Lastly, the effectiveness of predictive models is intimately connected to the quality and accessibility of data. Continuous endeavors to refine data collection infrastructure and space weather monitoring systems represent an ongoing challenge within the field.

It is essential to acknowledge that several of these challenges and limitations were actively addressed in our work. Sensitivity analysis was employed, utilizing noise injection, to assess the robustness of our models under varying conditions. Additionally, it is noteworthy that the data used for X-ray observations were comparatively more rare and less abundant than the time series proton flux data, underscoring the importance of addressing these disparities in data availability and quality. Addressing these challenges is important in advancing the accuracy and reliability of ∼100 MeV SEP event predictions, and it remains a cornerstone of ensuring the safety and success of space missions. Future research endeavors should focus on refining ML models and enhancing data collection infrastructure.

5. Conclusions

In this study, we conducted a thorough analysis of predicting high-energy Solar Energetic Particle (SEP) events, particularly those within the approximately 100 MeV energy range, which pose significant hazards to space exploration and terrestrial technology. Our research has leveraged the power of multimodal data fusion by combining time series proton flux data with solar X-ray images, thereby enhancing the prediction accuracy of these potentially catastrophic events. Through the application of six finely-tuned ML models, including UTS, image-based models (image), univariate feature concatenation (UFC), UDC, UDM, and USC, we have strived to provide precise predictions and a deeper understanding of the underlying solar phenomena responsible for SEP events.

The results of our investigation have yielded valuable insights. We have demonstrated that prediction accuracy is influenced by the choice of observation window size and image transformation vector size. Notably, the 6-hr observation window size and smaller image transformation vector sizes have shown promising accuracy levels, thus indicating their potential for enhancing prediction outcomes. However, it is essential to recognize that different models exhibit varying performance based on these parameters. univariate deep merge, for instance, excels with a 6-hr observation window and small image data, achieving 80% accuracy. These findings emphasize the significance of tailoring model configurations to specific input data characteristics for optimal performance. Furthermore, In addition to examining balanced settings, we also tested our models in scenarios with imbalanced data to ensure their applicability to real-world situations. Our study has also explored the impact of noise levels on model performance, revealing the sensitivity of our models to noise in the training data. This sensitivity analysis underlines the importance of data quality and preprocessing in the prediction of SEP events.

In conclusion, our study exemplifies the innovative application of data fusion techniques and advanced ML models in enhancing the prediction of high-energy SEP events. While we have made significant strides in addressing the challenges and limitations, there is still much work to be done to refine and improve the accuracy

of these predictions. Our research aligns with the objectives of safeguarding astronauts, spacecraft, and terrestrial electronic systems, ultimately contributing to the success and safety of space missions in lunar and Martian exploration. As we continue to advance our understanding of space weather, our work paves the way for more robust and reliable SEP event prediction systems.

Future research directions in SEP event prediction encompass data collection enhancements, advancements in ML techniques, feature engineering, and interdisciplinary collaborations. These efforts can lead to more robust and accurate predictions, further contributing to improved space weather forecasting systems.

Data Availability Statement

The research presented in this manuscript relies on various data sources. The space environment monitoring data used can be accessed from the National Centers for Environmental Information (NCEI, 2020). Solar observation data was obtained from the Helioviewer Project, and details can be found at Helioviewer Project (2024). We also utilized solar observation data from Lockheed Martin Solar and Astrophysics Laboratory, available at (Lockheed Martin Solar & Astrophysics Laboratory, 2024). Furthermore, the GSEP data set directly related to this research is archived and accessible on Dataverse Harvard (Rotti et al., 2022). Finally, our source codes and experimental evaluations are publicly available (Hosseinzadeh et al., 2024a, 2024b). These data sources are provided to ensure transparency, accessibility, and reproducibility in support of collaborative research and the advancement of knowledge in the field.

Acknowledgments
This project has been supported by funding from the Division of Atmospheric and Geospace Sciences within the Directorate for Geosciences, under NSF awards #2204363 and #2240022. The authors also acknowledge all those involved with the GOES missions as well as the SOHO mission.

References

Alshammari, K., Hamdi, S. M., & Filali Boubrahimi, S. (2024). Identifying Flare-indicative Photospheric Magnetic Field Parameters from Multivariate Time-series Data of Solar Active Regions. *The Astrophysical Journal Supplement Series*, *271*(2), 39. https://doi.org/10.3847/1538-4365/ad21e4

Aminalragia-Giamini, S., Raptis, S., Anastasiadis, A., Tsigkanos, A., Sandberg, I., Papaioannou, A., et al. (2021). Solar energetic particle event occurrence prediction using solar flare soft x-ray measurements and machine learning. *Journal of Space Weather and Space Climate*, *11*, 59. https://doi.org/10.1051/swsc/2021043

Aran, A., Sanahuja, B., & Lario, D. (2006). Solpenco: A solar particle engineering code. *Advances in Space Research*, *37*(6), 1240–1246. https://doi.org/10.1016/j.asr.2005.09.019

Arnaut, F., Kolarski, A., & Srećković, V. A. (2023). Random forest classification and ionospheric response to solar flares: Analysis and validation. *Universe*, *9*(10), 436. https://doi.org/10.3390/universe9100436

Bahri, O., Li, P., Hosseinzadeh, P., Boubrahimi, S. F., & Hamdi, S. M. (2023). Shapelet-Preserving Bootstrapping For Time Series Data Augmentation. *2023 International Conference on Machine Learning and Applications (ICMLA)*. https://doi.org/10.1109/icmla58977.2023.00069

Bain, H., Brea, P., & Adamson, E. (2018). Using machine learning techniques to forecast solar energetic particles. *Agu fall meeting abstracts*, *2018*, SM31D–3530.

Baltrušaitis, T., Ahuja, C., & Morency, L.-P. (2018). Multimodal machine learning: A survey and taxonomy. *IEEE transactions on pattern analysis and machine intelligence*, *41*(2), 423–443. https://doi.org/10.1109/tpami.2018.2798607

Boubrahimi, S. F., Aydin, B., Martens, P., & Angryk, R. (2017). On the prediction of> 100 mev solar energetic particle events using goes satellite data. In *2017 IEEE international conference on big data (big data)* (pp. 2533–2542).

Brea, P., Bain, H. M., & Adamson, E. T. (2018). *Using machine learning techniques to forecast solar energetic particles*. Agu fall meeting.

Chen, Y., Carver, M. R., Morley, S. K., & Hoover, A. S. (2021). Determining ionizing doses in medium earth orbits using long-term gps particle measurements. In *2021 IEEE aerospace conference* (pp. 1–21). https://doi.org/10.1109/aero50100.2021.9438516

Chen, Y., Kempton, D. J., Ahmadzadeh, A., & Angryk, R. A. (2021). Towards synthetic multivariate time series generation for flare forecasting. *Artificial intelligence and soft computing: 20th international conference, icaisc 2021, virtual event, june 21–23, 2021, proceedings, part i*, *20*, 296–307. https://doi.org/10.1007/978-3-030-87986-0_26

Fogtman, A., Baatout, S., Baselet, B., Berger, T., Hellweg, C. E., Jiggens, P., et al. (2023). Towards sustainable human space exploration—Priorities for radiation research to quantify and mitigate radiation risks. *npj Microgravity*, *9*(1), 8. https://doi.org/10.1038/s41526-023-00262-7

Forbush, S. E. (1946). Three unusual cosmic-ray increases possibly due to charged particles from the sun. *Physical Review*, *70*(9–10), 771–772. https://doi.org/10.1103/physrev.70.771

Guggenmos, M., Schmack, K., Veer, I. M., Lett, T., Sekutowicz, M., Sebold, M., et al. (2020). A multimodal neuroimaging classifier for alcohol dependence. *Scientific reports*, *10*(1), 298. https://doi.org/10.1038/s41598-019-56923-9

Helioviewer Project. (2024). Helioviewer.org. [Dataset]. https://helioviewer.org/

Hosseinzadeh, P., Bahri, O., Li, P., Boubrahimi, S. F., & Hamdi, S. M. (2023). Metforc: Classification with meta-learning and multimodal stratified time series forest. In *2023 international conference on machine learning and applications (icmla)* (pp. 1248–1252). Retrieved from https://ieeexplore.ieee.org/document/10459728

Hosseinzadeh, P., Boubrahimi, S. F., & Hamdi, S. M. (2024). Improving solar energetic particle event prediction through multivariate time series data augmentation. *The Astrophysical Journal - Supplement Series*, *270*(2), 31. https://doi.org/10.3847/1538-4365/ad1de0

Hosseinzadeh, P., Boubrahimi, S. F., & Hamdi, S. M. (2024). Source code. [Software]. https://doi.org/10.5281/zenodo.10966803

Ismail, A. A., Hasan, M., & Ishtiaq, F. (2020). *Improving multimodal accuracy through modality pre-training and attention*. arXiv preprint arXiv:2011.06102.

Kang, H.-W., & Kang, H.-B. (2017). Prediction of crime occurrence from multi-modal data using deep learning. *PLoS One*, *12*(4), e0176244. https://doi.org/10.1371/journal.pone.0176244

Lario, D. (2005). Advances in modeling gradual solar energetic particle events. *Advances in Space Research*, *36*(12), 2279–2288. https://doi.org/10.1016/j.asr.2005.07.081

Laurenza, M., Cliver, E., Hewitt, J., Storini, M., Ling, A., Balch, C., & Kaiser, M. (2009). A technique for short-term warning of solar energetic particle events based on flare location, flare size, and evidence of particle escape. *Space Weather*, *7*(4). https://doi.org/10.1029/2007sw000379

Lavasa, E., Giannopoulos, G., Papaioannou, A., Anastasiadis, A., Daglis, I., Aran, A., et al. (2021). Assessing the predictability of solar energetic particles with the use of machine learning techniques. *Solar Physics*, *296*(7), 107. https://doi.org/10.1007/s11207-021-01837-x

Lockheed Martin Solar and Astrophysics Laboratory. (2024). Isolde data search. [Dataset]. https://www.lmsal.com/isolsearch

Löning, M., Bagnall, A., Ganesh, S., Kazakov, V., Lines, J., & Király, F. J. (2019). *sktime: A unified interface for machine learning with time series*. arXiv preprint arXiv:1909.07872.

Ma, R., Boubrahimi, S. F., Hamdi, S. M., & Angryk, R. A. (2017). Solar flare prediction using multivariate time series decision trees. In *2017 IEEE international conference on big data (big data)* (pp. 2569–2578).

NCEI. (2020). Goes space environment monitor - Access average data. [Dataset]. https://www.ncei.noaa.gov/data/goes-space-environment-monitor/access/avg/

Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., & Ng, A. Y. (2011). Multimodal deep learning. In *Proceedings of the 28th international conference on machine learning (ICML-11)* (pp. 689–696).

Núñez, M. (2011). Predicting solar energetic proton events (e> 10 mev). *Space Weather*, *9*(7). https://doi.org/10.1029/2010sw000640

Núñez, M., & Paul-Pena, D. (2020). Predicting> 10 MEV Sep events from solar flare and radio burst data. *Universe*, *6*(10), 161. https://doi.org/10.3390/universe6100161

Reames, D. V. (1995). Solar energetic particles: A paradigm shift. *Reviews of Geophysics*, *33*(S1), 585–589. https://doi.org/10.1029/95rg00188

Rigatti, S. J. (2017). Random forest. *Journal of Insurance Medicine*, *47*(1), 31–39. https://doi.org/10.17849/insm-47-01-31-39.1

Rotti, S., Aydin, B., Georgoulis, M., & Martens, P. (2022). GSEP dataset. [Dataset]. https://doi.org/10.7910/DVN/DZYLHK

Saini, K., Alshammari, K., Hamdi, S. M., & Filali Boubrahimi, S. (2024). Classification of Major Solar Flares from Extremely Imbalanced Multivariate Time Series Data Using Minimally Random Convolutional Kernel Transform. *Universe*, *10*(6), 234. https://doi.org/10.3390/universe10060234

Simonsen, L. C., & Slaba, T. C. (2021). Improving astronaut cancer risk assessment from space radiation with an ensemble model framework. *Life Sciences and Space Research*, *31*, 14–28. https://doi.org/10.1016/j.lssr.2021.07.002

Sleeman, W. C., IV, Kapoor, R., & Ghosh, P. (2022). Multimodal classification: Current landscape, taxonomy and future directions. *ACM Computing Surveys*, *55*(7), 1–31. https://doi.org/10.1145/3543848

Smith, M., Craig, D., Herrmann, N., Mahoney, E., Krezel, J., McIntyre, N., & Goodliff, K. (2020). The artemis program: An overview of nasa's activities to return humans to the moon. In *2020 IEEE aerospace conference* (pp. 1–10).

Stumpo, M., Benella, S., Laurenza, M., Alberti, T., Consolini, G., & Marcucci, M. F. (2021). Open issues in statistical forecasting of solar proton events: A machine learning perspective. *Space Weather*, *19*(10), e2021SW002794. https://doi.org/10.1029/2021sw002794

Thomas, S. A., Race, A. M., Steven, R. T., Gilmore, I. S., & Bunch, J. (2016). Dimensionality reduction of mass spectrometry imaging data using autoencoders. In *2016 IEEE symposium series on computational intelligence (SSCI)* (pp. 1–7).

Usman, K., & Rajpoot, K. (2017). Brain tumor classification from multi-modality mri using wavelets and machine learning. *Pattern Analysis & Applications*, *20*(3), 871–881. https://doi.org/10.1007/s10044-017-0597-8

Walsh, L., Schneider, U., Fogtman, A., Kausch, C., McKenna-Lawlor, S., Narici, L., et al. (2019). Research plans in europe for radiation health hazard assessment in exploratory space missions. *Life Sciences and Space Research*, *21*, 73–82. https://doi.org/10.1016/j.lssr.2019.04.002

Wang, J.-z., Huo, Z.-x., & Wang, F.-f. (2022). Tid evaluation based on variabilities of space radiation and device failure dose in typical navigation satellite orbits. *Microelectronics Reliability*, *137*, 114747. https://doi.org/10.1016/j.microrel.2022.114747

Wang, W., Tran, D., & Feiszli, M. (2020). What makes training multi-modal classification networks hard? In *Proceedings of the ieee/cvf conference on computer vision and pattern recognition* (pp. 12695–12705).

Watson-Morgan, L., Hawkins, L., Carpenter, L., Gagliano, L., Means, L., Percy, T. K., et al. (2023). Landing humans and human-class cargo on the moon and mars. In *2023 IEEE aerospace conference* (pp. 1–10).

Whitman, K., Egeland, R., Richardson, I. G., Allison, C., Quinn, P., Barzilla, J., et al. (2022). *Review of solar energetic particle models*, Advances in Space Research. https://doi.org/10.1016/j.asr.2022.08.006

Wu, H., Huff, J. L., Casey, R., Kim, M.-H., & Cucinotta, F. A. (2009). Risk of acute radiation syndromes due to solar particle events. In *The human health and performance risks for space explorations* (pp. 171–190). NASA Human Research Program.

Yang, X., Feng, S., Wang, D., & Zhang, Y. (2020). Image-text multimodal emotion classification via multi-view attentional network. *IEEE Transactions on Multimedia*, *23*, 4014–4026. https://doi.org/10.1109/tmm.2020.3035277