Space Weather AI into Warfighter Kill Chains Toward Operational Dominance

Cordula A Robinson, Ph.D., S. Mark Leidner,
Collin Trail, Ph.D., Kevin Lind, Ph.D., Kishan Shetty
JANUS Research Group, AER
131 Hartwell Avenue, Lexington, MA 02421-3126
cordula.robinson@janusresearch.us;
mark.leidner@janusresearch.us;

Benjamin D Prince, Ph.D.
Air Force Research Laboratory
Space Vehicles Directorate
afrl.rvborgmailbox@us.af.mil

ABSTRACT

As global investment in space accelerates, the need for resilient, intelligent space systems has become critical. With over 100,000 satellites projected to orbit Earth by 2030, diagnosing, predicting, and responding to space environment threats in real time is essential for sustaining operations in a contested, congested domain.

To address this challenge, the authors, in collaboration with the Department of the Air Force, have developed an AI/ML-powered tool for real-time space hazard detection and satellite anomaly diagnostics. This system integrates multi-data streams using a middle fusion neural network architecture that merges latent representations from temporally synchronized but spatially offset sources. This significantly reduces computational cost and enables real-time inference.

Demonstrated using >800 keV flux data from the Van Allen Probes (RBSPA and RBSPB), the pipeline includes semi-automated hazard labeling, cross-sensor alignment, and adaptive normalization strategies. Key performance metrics such as Average Precision (AP) with varying IoU are presently used to evaluate model effectiveness. The best results correspond with 304 days and 1 class using high-quality training data, which produces metrics for RBSPA/RBPSB of mAP@50 = 0.94/0.95, and F1 = 0.76/0.76. Performance is influenced by input channel selection, normalization methods, choice of architecture, and training data curation.

Future development includes: simulated data augmentation; unsupervised learning techniques; integration of additional datasets (e.g., GOES, GPS); and deployment in DevSecOps environments. Establishing a Space Weather for Training and Simulation (SWFTS) framework for operational use and potentially Kafka-based event-driven architecture are two scenarios that can facilitate practical for real-time deployment.

By embedding AI inferences into the F2T2EA kill chain, this tool offers a unique discriminator to advance space domain awareness and survivability by identifying and contextualizing hazards. It provides a scalable, real-time approach to protecting critical space infrastructure in an era of growing operational complexity and threat.

ABOUT THE AUTHORS

Cordula A. Robinson, Ph.D., is Vice President of Space Weather with decadal experience advancing artificial intelligence and deep learning applications. In this role, she leads research and development initiatives that integrate AI/ML, advanced data processing, and analytical modeling to accelerate scientific discovery and operational readiness. She is also dedicated to building high-performing teams by fostering collaboration, optimizing workflows, and aligning resources to deliver mission-focused results. In this capacity, she has supported organizations including DAF, IARPA, NGA, DEVCOM, NASA, NSF, and other international agencies.

Benjamin D. Prince,.Ph.D., is a Principal Research Chemist with 14 years of experience at the Air Force Research Laboratory. In this capacity, he leads research and development efforts exploring next-generation space domain awareness capabilities from basic science to prototype fielding and demonstration. Relevant activities in this area

include energetic charged particle sensing of the near-Earth space environment and the development of models and decision aids to support safety-of-flight operations for space vehicles.

S. Mark Leidner is a Senior Engineering Manager for JANUS at Kirtland Air Force Base in support of Research-to-Operations efforts for the U.S. Space Force. His career spans more than 30 years of scientific research, including numerical modeling of oceans, lakes, the Martian atmosphere, severe convective storms, hurricanes, and global ocean winds from satellites. He previously served as a Certified Consulting Meteorologist with the American Meteorological Society (2012–2022) and began his career as a research metallurgist with the Department of Energy in support of the U.S. nuclear weapons program.

Radhakishan Shetty is a Director with JANUS Research Group's Strategic Solutions Team, bringing more than 20 years of experience delivering innovative, cost-effective solutions for government and defense customers. He has supported organizations including the U.S. Army, U.S. Air Force, NATO Allied Command Transformation, NIST's Public Safety Communications Research Division, the National Institutes of Health, and the Federal Bureau of Investigation. His work spans the design and development of Interactive Multimedia Instruction, serious games, simulations, and extended reality applications. He also serves on the Serious Game Showcase and Challenge IPT and holds a B.S. in Computer Science from the Georgia Institute of Technology.

Kevin Lind, Ph.D., is a Senior Staff Scientist at JANUS with expertise in space weather research and high-performance computing. He collaborates with AFRL scientists on projects spanning solar-based space hazard detection, magnetospheric modeling for nowcasting and anomaly resolution, and calibration of particle flux measurements. His background includes extensive experience in computational science applied to a variety of disciplines, with a focus on leveraging high-performance computing to advance research outcomes. Trained as an astrophysicist, he has worked across both observational and theoretical domains, bringing a broad scientific perspective to space environment challenges.

Collin M. Trail, Ph.D., is a Staff Scientist with JANUS specializing in space weather research and modeling. His work spans the application and development of key models such as the International Reference Ionosphere, Versatile Electron Radiation Belt, Magnetospheric Specification and Forecasting Model, and Far-field Antenna Radiation Model, as well as analytical tools including the AFRL Power Raytracer. He collaborates closely with AFRL scientists on projects that advance validation and verification, automate and optimize model output, and improve scientific simulations. His research includes modeling and analysis of events such as VLF wave detection from the DSX satellite, supporting both scientific advancement and operational objectives in space environment forecasting.

INFORMATION

The views expressed are those of the authors and do not necessarily reflect the official policy or position of the Department of the Air Force, the Department of Defense or the U.S. government.

BACKGROUND

Strategic/Tactical Need

The growing complexity of satellite operations, particularly in contested or congested environments, has amplified the need for resilient, scalable, and intelligent systems capable of diagnosing and mitigating hazard anomalies in real time. In response, this paper describes a first-gen prototype software implementation for assessing energetic charged particle (ECP) hazards to spacecraft and their payloads. This tool, known as the Space Anomaly AI Nextgen Tool (SAAINT), represents a first-of-its-kind application of artificial intelligence (AI) in real-time space hazard diagnostics for both Department of Defense (DoD) and commercial use cases. A single satellite version is currently operational at Technology Readiness Level (TRL) 8/9 for an existing DoD mission.

Proposed Solution

SAAINT provides an AI-driven framework for the diagnosis of on-orbit hazard anomalies applicable for large satellite constellations by leveraging ECP data (Figure 1). Disruptions caused by space weather phenomena would be expressed as temporally varying hazards within the magnetosphere (O'Brien, 2009). The purpose of the present research is to detect and assess those risks to satellite constellation health across LEO, MEO, HEO, and GEO. Impacts may manifest dramatically like a safe mode trip, but impacts are often innocuous, like a register getting scrambled. SAAINT is designed to attribute anomalies to these impacts. Anomaly attribution is helpful if an operator transmits a command to correct an abnormality or if flight software is updated to automatically reset a subsystem.

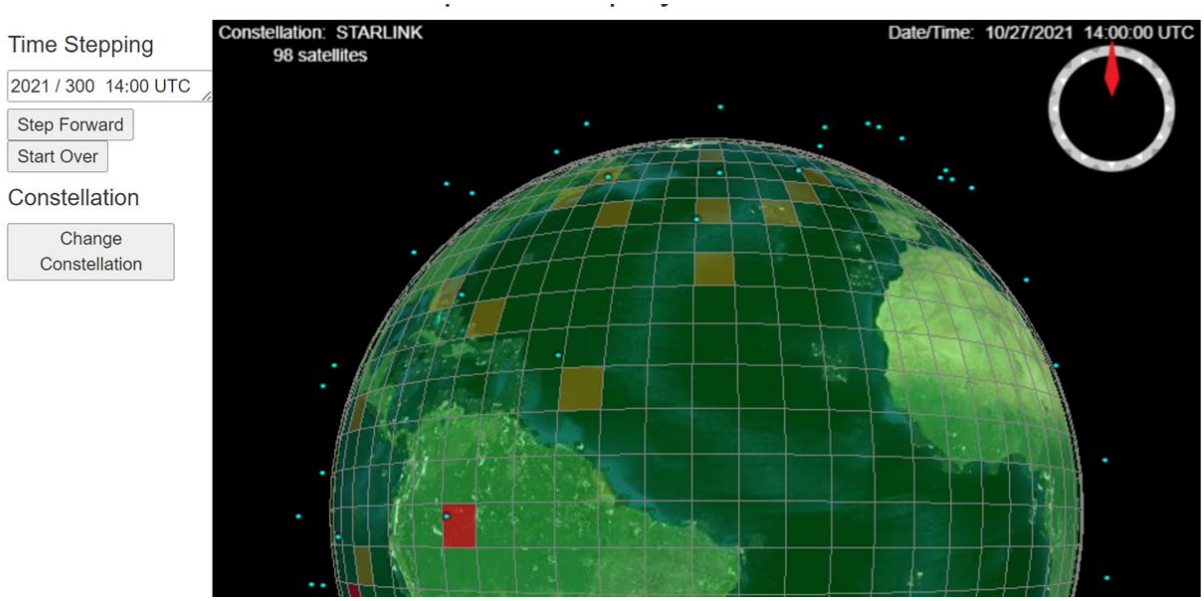

Figure 1. POC 3D interface visualization for SAAINT, showing satellite constellations with environment hazards highlighted

This capability is critical, not only for individual satellite health, but also for ensuring the integrity of broader satellite networks supporting strategic and tactical operations. Consequently, understanding and mitigating these space environmental effects is essential for operational continuity and mission assurance.

Requirements

To support AI workflows that underpin SAAINT, a standardized, scalable, and reliable data processing framework is required, as well as a Common Operating Picture (COP), a single repository or display of relevant information. This is a key element of the existing AI backbone. Satellite anomaly analysis typically requires ingestion and harmonization of diverse datasets. This necessitates common data formats, consistent units, and interoperable

protocols to allow for accurate interpolation and cross-calibration. Recent advances in computational tooling support these requirements. Libraries such as PyCDF and Xarray enable effective metadata management (using HDF5 file formats) and long-term maintainability, while modular programming approaches help separate environment configurations, constants, and data operations for greater flexibility and maintainability. These methods enhance scalability for multi-satellite applications.

A key technical challenge in the AI approach is spatial and temporal asynchrony inherent in multi-data sources (Feng et. Al. 2014). Fusion is the integration of diverse data at different cadences into a single stream for inference. Neural network (NN) fusion offers a promising approach to mitigate this issue by aligning temporally co-observed data streams across shared fields of view. Fusion approaches typically fall into three categories:

- Early fusion, where raw inputs are combined prior to model ingestion, enabling access to all modalities throughout the model's decision process. While effective when spatial alignment is present, this method is highly sensitive to misalignment (Wagner et al., 2016; Wu et al., 2006).
- Late fusion, which combines outputs at or near the decision stage, often through ensemble methods such as majority voting or decision averaging (Morvant et al., 2014; Shutova et al., 2016). While generally more robust, it lacks shared learning during inference and is cumbersome.
- Middle (hybrid) fusion, the approach employed in this study, provides shared representation learning during runtime and is known to facilitate synergistic gains across data streams (Lan et al., 2014).

To implement middle fusion, a custom auxiliary module, the Multi-Modal Region Proposal Network (MMRPN), is utilized, which incorporates a projection function to enhance interpretability and manage spatial alignment (Robinson et al., 2021). This projection function explicitly fuses latent feature representations, referred to here as hazard or risk proposals, by converting spatial coordinates from one modality into the reference space of another.

This approach reduces latency and computational bottlenecks by avoiding the need to process raw inputs in their entirety. The function resolves spatial offsets caused by temporal variations in magnetic fields across L* shells and magnetic latitudes (MLAT), ensuring accurate alignment of corresponding proposal regions across modalities. The aligned outputs, or proposal sets, contain m co-registered data elements (where m is the number of input streams), which are then propagated through the remainder of the neural network for robust final prediction.

Summary

This fusion architecture enables more accurate, interpretable, and mission-ready assessments of satellite system health in real time and supports future integration into training and operational planning tools. For example, the space environment potentially exerts significant influence over command, control, and communications (C3) systems, particularly in high-latitude theaters where radiation hazards may pose risks to personnel safety. Thus, the relevance of SAAINT for the Army Synthetic Training Environment (STE) is also considered to model space effects and its impact on sensors, weapons, systems, and soldiers. This research supports the development of a Space Weather for Training and Simulation (SWFTS) prototype that can be integrated into digital training exercises to improve decision-making under physically plausible space conditions.

METHODS

Data Sources and Preprocessing

This study employs data from robust satellite sources, namely the Van Allen Probes (RBSPA and RBSPB), to capture energetic charged particle (ECP) fluxes and associated ephemerides. These in situ radiation measurements are used to calculate adiabatic invariants and define the magnetic coordinates (L* and magnetic latitude, MLAT) necessary for modeling particle behavior within the magnetosphere.

To prepare these data for machine learning (ML), a suite of computational tools that streamline data ingestion, transformation, and harmonization were developed. Emphasis is placed on reproducibility, scalability, and modularity. A common challenge is addressing mismatched cadences, incomplete ephemeris records, and inconsistent energy bin definitions (e.g., across MagEIS and REPT instruments). Both measure differential electron fluxes at various energies: MagEIS (Magnetic Electron Ion Spectrometer) is configurable in energy channels and uses magnetic focusing and pulse height analysis to provide the cleanest possible energetic electron measurements

over the critical energy range of 30 keV to 4 MeV for electrons and 20 keV to 1 MeV for ions. REPT (Relativistic Electron Proton Telescope) has fixed energy channels and covers the challenging electron range of 4-10 MeV and proton energy range of 20-75 MeV to capture most intense events (JHUAPL, 2025). These cover the entire range of energies required for a full representation of the electron energy spectrum which are relevant to both satellite charging and single-particle event risk anomalies.

These are resolved through expert-driven preprocessing routines that ensure standardized formats, units, and fill values, with metadata preserved in derived datasets. Effective management of both incoming and outgoing API connections supports integration with existing workflows.

To align with the current ML algorithm, datasets are converted into a 2D grid structure: particle fluxes across energy channels are mapped as pixel intensity, with L* and MLAT serving as the x- and y-axes, respectively. We assume azimuthal symmetry for the present study. Equatorial pitch angles θ at each L* are used to calculate omnidirectional fluxes f

using functional fits of the form

$$j = f_o \sin^n(\theta). \quad (1)$$

The off-equator omnidirectional fluxes are extrapolated from the equatorial value, with results binned in 0.1 L* intervals and 2° MLAT increments. This structure enables interpolation across differential energy channels and accurate representation of high-energy electron behavior. These truth magnetosphere datasets are generated daily for both RBSPA and RBSPB individually (for use in the training dataset), and jointly (for hazard identification).

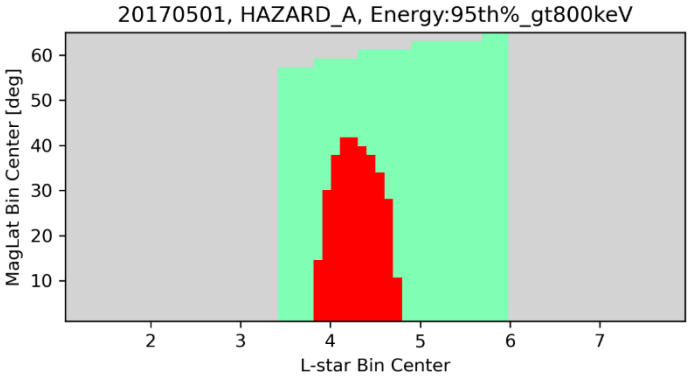

Figure 2a. Example hazard classification using RBSPA >800keV flux from 05/01/2017. Bins with omnifluxes above 95th percentile depicted red, values below 95th percentile depicted green, missing values (due to insufficient observations or loss cone) shown in grey. L* bin width .1, magnetic latitude bin width 2 degrees.

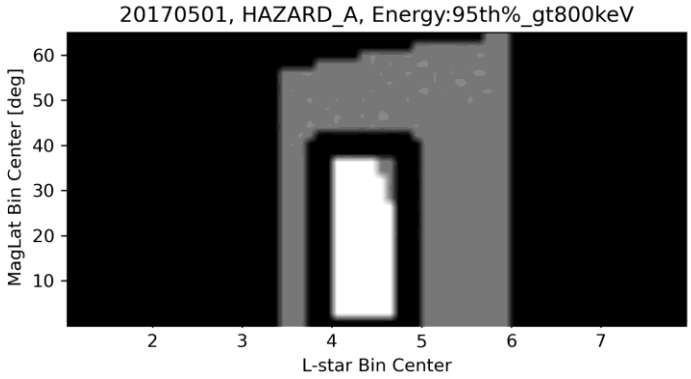

Figure 2b. Example hazard bounding box using RBPSA >800keV flux from 05/01/2017. Bins with omnifluxes above 95th percentile depicted as white hazard region, black rectangles around hazard region depicts bounding box used in NN training. Values below 95th percentile shown as grey region, unavailable data shown as black outer region.

In preparation for training the joint truth magnetospheres are integrated across select energy channels to compute >800 keV omniflux. This consolidates the differential energy spectrum into a single number which provides a reasonable risk quantity for learning. The L* range is constrained between 3.5 and 5.5 to match the region of interest for radiation belt dynamics. Magnetic latitude is truncated at 50° to avoid complications related to loss cone modeling (Borovsky et al., 2022), which is a complication related to the loss of particles at large pitch angles during significant disturbances of the magnetic field. While this is significant in terms of risk evolution, it is beyond the scope of the feasibility study presented here and will be deferred for further development.

A thresholding procedure identified “hazard” bins using the 90th percentile of omniflux values, considered as hazardous because satellites are designed to function during lower flux values but must be secured against charging and single particle events for unusually high levels of electron flux. These are grouped into bounding boxes for model input. Dates with no detected hazards or with inadequate pitch angle coverage are excluded (Figures 2a, b).

The resulting curated dataset includes 2,166 dates. After filtering on L* and hazard presence, 304-685 dates are retained for the >800 keV channel (electron charging risk anomalies, damaging electronic components). For this study, MagEIS channels at 749 keV, 904 keV, 1064 keV, 1072keV, 1575 keV, 1728 keV, 2254 keV, and 2590 keV, and REPT channels at 2.1 MeV through 20 MeV are used since they share a Common Operating Picture (COP; as per Buzzi et al., 2019), thus potentially provide\ synergistic hazard-related diagnostics. GOES satellites were considered as an additional source but their lack of overlap in L* (as shown in Figure 3) created challenges for putting them in a Common Operating Picture and will be revisited as an additional data source to extend the coverage of the model in later work.

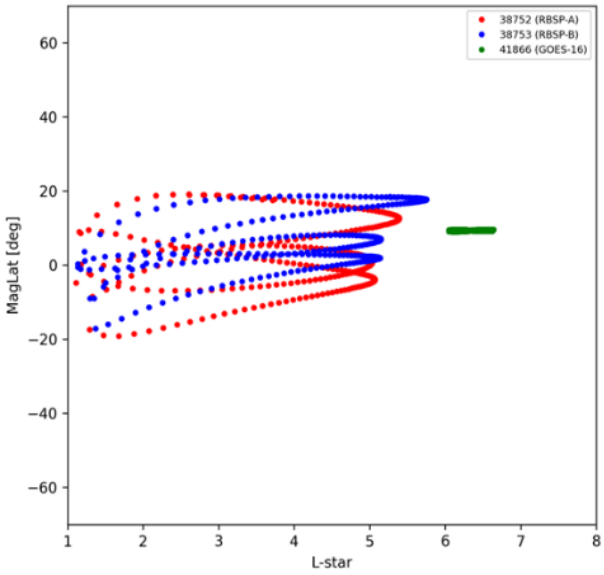

Figure 3. Tracks of satellites in L* vs Magnetic Latitude space

Model Architecture and Training

Currently, the ML architecture is based on a modified object detection network with a Feature Pyramid Network (FPN)-enabled ResNet101 backbone. The FPN module – called a Feature Pyramid Network (FPN)- is added to the backbone component to further improve hazard extraction enabling the accurate detection of small features (Wu et al., 2019). Future versions may consider transformer backbones. The network was originally designed to process three modalities for 2D tensors for AFRL (Robinson et al., 2021) and is reconfigured for this study to accept RBSPA and RBSPB as separate modalities and their differential energy channels as the 2D tensors. Each satellite modality produces a 19-channel image, with each pixel representing the flux at a given L*-MLAT coordinate. Missing values were replaced with a baseline flux of 0. This number is sufficiently low to indicate a non-hazard, which is admittedly biasing missing data as non-hazardous, but for the purposes of this validation study, was considered a suitable choice so that only demonstrable hazards would be flagged.

To enable training on using this middle fusion NN, the network's first convolutional layer was extended to accept 19-channel inputs. Other adjustments include disabling the original modality alignment projection (as RBSP data share a common coordinate frame) and modifying transformation parameters to accommodate small image sizes (20×25 pixels).

Hazard detection is presently treated as a supervised object recognition task. Bounding boxes are labeled based on hazard thresholds using OpenCV. A subset of outputs were visually inspected to confirm stability of regions across days. Our bounding box algorithm groups adjacent hazardous bins into minimal rectangular regions. These labels and input tensors are used to train the neural network. Early performance issues, such as negative or undefined metrics, are resolved by setting MLAT cutoffs and ensuring valid fill values.

Data Integrity and Formatting

Several issues were encountered related to data formatting and consistency:

• Empty bounding boxes caused training to fail - the current HDF5 format does not support frames with no hazards.
• Modality mismatch in hazard detection required matching the number of hazards between RBSPA and RBSPB inputs. An initial approach using separate RBPSA and RBPSB datasets for hazard identification resulted in inconsistent counts. This was resolved by switching to a joint RBSPA+RBSPB dataset for hazard identification.
• Input dimensionality posed challenges; the 20×25 resolution is considerably smaller than expected by the base model, causing sparsity after transformation. Scaling techniques are under evaluation.
• Normalization procedures require ongoing evaluation for physics-based hazard detection training. It appears that traditional techniques, rescaling input based on means and standard deviation, perform poorly.

Despite these issues, the model successfully maps: • Satellite → modality; • Energy channel → image channel; • 2D truth magnetosphere → 2D image; Hazard box → object box; and • Hazard class → object class. Future iterations will incorporate physics-based adjustments (e.g., loss cone modeling) to refine hazard designation and real-time inference on unseen data, including GOES inputs, through a production-ready API.

Multi-Modal Fusion and System Deployment

This AI hazard detection system implements a middle fusion strategy, allowing latent feature alignment across modalities to reduce computational latency and improve robustness in data-sparse conditions.. Temporally synchronous data enables spatial registration, facilitating 2D tensor alignment for hazard classification even when one modality has weak or missing signals.

This architecture positions the model as a leap-ahead capability for space domain awareness, enabling early warning, anomaly attribution, and support for simulation environments like the Army Synthetic Training Environment (STE). The training and inference systems can be containerized, e.g., via Miniconda, to provide software environment control, dependency management, and computational reproducibility.

Initial training datasets are constructed from publicly available RBSP datasets hosted by the New Mexico Consortium (Boyd et al., 2019; 2021). To ensure consistent energy channel definitions, the dataset is restricted to the operational period from August 3, 2013, to July 16, 2019. Days with inconsistent channels are excluded, yielding a stable dataset for object detection training.

RESULTS AND OBSERVATIONS

Our machine learning (ML) pipeline demonstrates the critical role of high-quality, multi-source training data in enabling effective hazard extraction and accurate model inference. To ensure robustness and generalizability, the training dataset is curated to scan over the diverse geophysical conditions available, spanning the wide range of possibilities for particle environment differences, to learn and identify hazard indicators

Preprocessing and Training Pipeline

Data preprocessing utilized semi-automated hazard labeling via OpenCV, with ground-truth annotations stored in HDF5 format. An Index Generator matched temporally aligned data across channels, supporting strict synchronization required for neural network (NN) fusion. Data are split into training, validation and testing sets using a 70:10:20 ratio. HDF5 chunk sizes are matched to model batch sizes for optimal performance.

Evaluation Metrics

Model performance is evaluated using the following metrics:

IoU (intersection over Union) is a measure of how well the bounding boxes define the hazard, where 1 means that the actual hazard is precisely bounded by the inferred or predicted boundary.

Average precision refers to the fraction of positive detections that are true positives (true positives divided by true positives plus false positive), whereas recall is the fraction of true positives over actual positives (true positives plus false negatives). These metrics and the resulting F1 score defined below are a measure of the ability of the learning model to train from the data and predict anomalies in a reliable way.

• Average Precision (mAP50-95): Calculated over IoU thresholds from 0.50 to 0.95

• mAP50: AP at IoU threshold of 0.50
• F1 Score: Harmonic mean of precision and recall

Model Performance and Observations

The best results correspond with a 304-day run using high-quality training data and linear flux values producing metrics for RBSPA/RBPSB of mAP@50 = 0.95/0.95, and F1 = 0.76/0.76. These results reflect performance is influenced by input channel selection, normalization methods, and training data curation (Table 1).

Table 1. Results of Training

Classes	Stopping Condition	Number of Days	Backbone Architecture	Trainable Layers	Energy Channels	Norm	mAP50 (A/B)	mAP50-95 (A/B)	F1 (A/B)
>90th / <10th	Fixed 2k	685	Resnet101	3/5	Full	Log	.81/.80	.57/.58	.60/.60
>90th / <10th	Patience 200	685	Resnet101	5/5	Full	Log	.81/.82	.56/.56	.59/.59
>90th / <10th	Fixed 2k	685	Resnet50	3/5	Full	Log	.82/.83	.54/.56	.58/.60
>90th / <10th	Patience 200	685	Resnet34	3/5	Full	Log	.82/.83	.58/.57	.60/.60
>90th / N/A	Patience 200	304	Resnet101	3/5	Full	Log	.94/.95	.71/.70	.72/.71
>90th / <10th	Patience 200	685	Resnet101	3/5	Drop 4 Highest	Log	.82/.84	.56/.58	.59/.61
>90th / <10th	Fixed 2k	685	Resnet101	3/5	Drop 1064 keV and 1072 keV	Log	.83/.85	.57/.57	.60/.61
>90th / <75th	Fixed 2k	480	Resnet101	3/5	Drop 1064 keV and 1072 keV	Log	.88/.89	.72/.72	.73/.74
>90th / <10th	Fixed 2k	685	Resnet101	3/5	Drop 1064 keV and 1072 keV	Log	.83/.85	.56/.58	.59/.61
>90th / N/A	Patience 200	304	Resnet101	3/5	Full	Unit	.94/.95	.72/.76	.73/.77
>90th / N/A	Fixed 2k	304	Resnet101	3/5	Full	Unit	.95/.95	.75/.74	.76/.76

Where:

Classes: Percentile computed from distribution of >800 keV fluxes within modeled L* and MLAT range used as upper (lower) bound for hazard class, second hazard class where applicable.

Stopping Condition: Runs terminated after either 200 epochs without improvement of the F1 value for RBSPA during validation stage, or for a fixed 2000 epochs. The best result for F1 in validation is tracked and used for testing.

Normalization: Normalization used means and standard deviations of logarithm of flux, or unit values.

Input Resolutions and Computational Improvements

Initial models maintained internal representations near 1000×1000 pixels, far exceeding the 20×25-pixel input size. This resulted in computational inefficiencies due to sparsity. By modifying the architecture to maintain smaller internal representations, we achieve substantial speed improvements without compromising model accuracy.

Hazard Class Variants

We explored a second class defined by the region with fluxes below the 10^{th} percentile, in order to help the NN learn the high flux class better through contrast. The average performance of the two classes was lower than the single class case. One possible explanation is that setting fill values to 0 in hazard identification and training data hurt our ability to identify low flux cases- modifying the hazard identification to exclude these values during classification did not improve performance. By contrast, adding a second class defined by fluxes above the 75^{th} percentile resulted in similar average performance to the single class case.

Energy Channel Variants

Finally, we explore excluding the highest energy differential channels (9.9 MeV–20 MeV), and/or the 1064 keV and 1072 keV channels, which had limited data availability. The differences between runs including and excluding these energy channels were minor and could easily be due to differences in initialization – these channels don't appear to contain critical information needed for good performance but including them doesn't seem to harm performance either.

In sum, experiments across ResNet variants (101, 50, 34) showed minor sensitivity to backbone depth but confirmed performance degradation under shorter training durations, reduced data quality, and limited normalization. Introducing a second class created lower-performing models with mAP@50 values lowering to 0.81/0.80 and F1 scores down to 0.59/0.59. Thus, expanding the classification framework to include other feature types (e.g., low-flux regions) does not improve detection accuracy.

CONCLUSIONS

This study demonstrates the feasibility of using neural network (NN)-based models for identifying high-energy particle hazards from satellite observations, leveraging curated, multi-source datasets and custom preprocessing pipelines.

A key limitation of the current work was limited data quantity and diversity for both ground truth labeling and model input. With access to additional satellite datasets, or simulated data, a more rigorous separation can be implemented: using one dataset to define the ground truth while using others to test the model's ability to infer hazards independently. This would more accurately assess generalization across satellite platforms.

The temporal and spatial resolution of the input data is presently constrained by low observation density, one frame per day and coarse L* and magnetic latitude binning. Expanding the volume and diversity of observational data would allow for finer-grained inputs and, likely, improved, more stringent predictive performance. Similarly, introducing unsupervised learning methods for transfer learning or using transformer/generative NN.

Key performance drivers include the number of training days, normalization method (unit outperforming log), and careful channel selection. Top models used full energy channels with no channel drops and benefited from unit-normalized flux, highlighting the importance of comprehensive input data and normalization strategy.

FUTURE WORK

Several avenues can extend this work toward a more operational space hazard inference. Refining energy channel selection and normalization remains a priority, as current methods struggle with flux values spanning several orders of magnitude. Unit normalization shows promise, yet refined alternatives also merit exploration.

Improving treatment of missing data is essential. Current fill strategies may bias the model, especially when data absence is unrelated to low flux. Interpolation or neutral fill values like channel means could reduce false negatives. Adjusting L* and latitude bounds warrant evaluation as it relates to spatial generalization while maintaining data coverage.

Cross-dataset inference to train using one dataset (e.g., RBSP) and test on another (e.g., GOES or GPS) can help assess model transferability and generalization. Simulated data to expand class diversity and support edge-case learning could enrich data for learning. A Space Weather for Training and Simulation (SWFTS) framework is currently being evaluated to support space domain awareness and multi-domain operations. This includes building an API-based Space Weather Scenario Manager, capable of publishing simulated weather impacts via the Distributed Interactive Simulation (DIS) protocol for use in training environments.

Together, these steps provide a foundation for transitioning from a promising prototype to a reliable, real-time space hazard monitoring system to detect and assess risks to satellite constellation health.

ACKNOWLEDGMENTS

Original AI Data-driven Defeat System for small objects (DS2) knowledge prototype was developed for AFRL, Rome (2019-2022), and is available for use within the DoD. Current patent application no.: 17/709,279.

REFERENCES

Borovsky, J. E., Yakymenko, K. N., & Delzanno, G. L. (2022). Modification of the Loss Cone for Energetic Particles in the Earth's Inner Magnetosphere. Journal of Geophysical Research: Space Physics, 127, e2021JA030106. https://doi.org/10.1029/2021JA030106

Boyd, A. J., Reeves, G.D., Spence, H.E., Funsten, H.O., Larsen, B.A., Skoug, R.M., Blake, J.B., Fennell, J. F., Claudepierre, S. G., Baker, D. N., Kanekal, S. G., and Jaynes, A. N. (2019), RBSP-ECT Combined Spin-Averaged Electron Flux Data Product, Journal of Geophysical Research: Space Physics. https://doi.org/10.1029/2019ja026733

Boyd, A. J., Spence, H.E., Funsten, H.O., Skoug, R.M., Larsen, B.A., Blake, J.B., Fennell, J. F., Claudepierre, S. G., Baker, D. N., Kanekal, S. G., and Jaynes, A. N. (2021), RBSP-ECT Combined Pitch Angle Resolved Electron Flux Data Product, Journal of Geophysical Research: Space Physics. https://doi.org/10.1029/2020JA028637.

Buzzi, P.G., Selva, D., Hitomi, N., Blackwell, W.J. (2019) Assessment of Constellation Designs for Earth Observation: Application to the TROPICS Mission, Acta Astronautica, Volume 161, 166-182, ISSN 0094-5765. https://doi.org/10.1016/j.actaastro.2019.05.007.

Feng, D., Haase-Schütz, C., Rosenbaum, L., Hertlein, H., Glaeser, C., Timm, F., Dietmayer, K. (2020). Deep Multi-modal Object Detection and Semantic Segmentation for Autonomous Driving: Datasets, Methods, and Challenges. Retrieved from https://arxiv.org/abs/1902.07830

JHUAPL (2025), Van Allen Probes, Spacecraft Instruments, Van Allen Probes: Spacecraft Accessed May 16th, 2025.

Lan, Z.Z. L. Bao, S. I. Yu, W. Liu, and A. G. Hauptmann, "Multimedia Classification and Event Detection using Double Fusion," Multimedia Tools and Applications, Vol. 71, 2014, pp. 333–347. https://link.springer.com/article/10.1007/s11042-013-1391-2

Morvant, E., Habrad, A., and S. Avache, "Majority Vote of Diverse Classifiers for Late Fusion," Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR), 2014. DOI:10.1007/978-3-662-44415-3_16

O'Brien, T. P. (2009). SEAES-GEO: A Spacecraft Environmental Anomalies Expert System for Geosynchronous Orbit, Space Weather, Vol. 7, S09003. https://doi.org/10.1029/2009SW000473.

Robinson, C.A., A. Tapley, M. Franco, and S. Ioannidis (2021), "Data-driven Defeat System for Small Objects: DS2", Accepted Military Communications Conference 29 November - 2 December 2021, San Diego, USA.

Shutova, E., D. Kelia, and J. Maillard, "Black Holes and White Rabbits: Metaphor Identification with Visual Features," in Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2016, pp. 160–170. https://aclanthology/N16-1020

Wagner, J., Fischer, V., Herman, M., & Behnke, S. (2016). Multispectral Pedestrian Detection using Deep Fusion Convolutional Neural Networks. Retrieved from http://ais.uni-bonn.de/papers/ESANN_2016_Wagner.pdf

Wu, Z., Cai, L., Meng, H., "Multi-level Fusion of Audio and Visual Features for Speaker Identification". In: International Conference on Advances in Biometrics, 2006, pp. 493–499. https://www1.se.cuhk.edu.hk/~hccl/publications/pub/200601_Multi-levelFusion.pdf

Wu, Y., Kirillov, A., Massa, F., Lo, W., & Girshick, R. (2019). Detectron2 (Version 0.2.1) [Computer software]. Github/detectron2