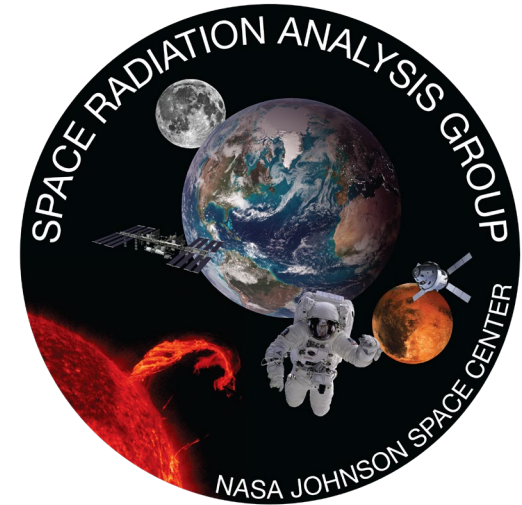


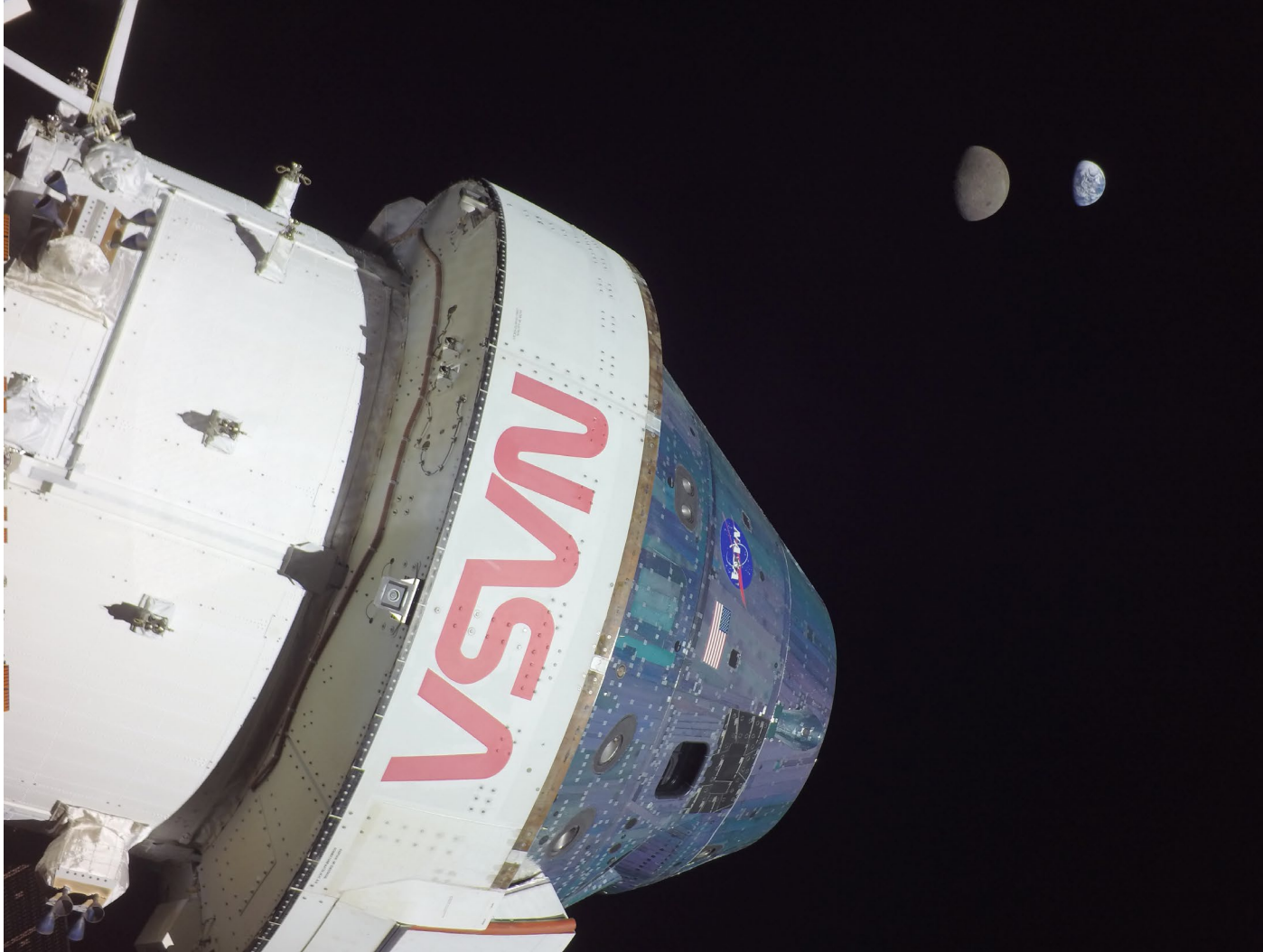
NASA's Ongoing SEP Model Validation Effort Driving an Effective R2O2R Process

Kathryn Whitman on behalf of SRAG and the ISEP Project
PSW.1, COSPAR 2024
Busan, South Korea
July 12 – 20, 2024



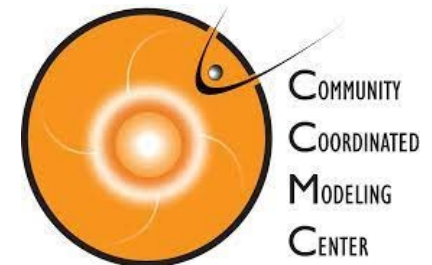
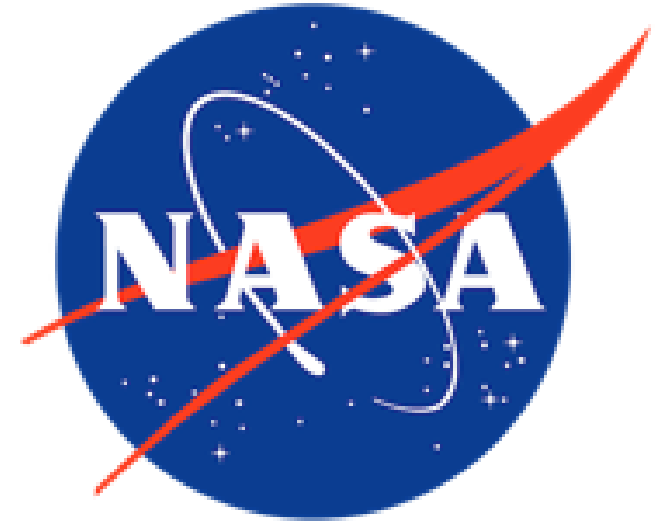
Artemis is Driving New SEP Forecasting Efforts at SRAG

ISEP: Integrated Solar Energetic Proton Event Alert/Warning System



NASA's ISEP Project

- The Integrated Solar Energetic Proton Event Alert/Warning System (ISEP) project is a collaborative effort to enhance space weather alerting capabilities for SRAG's space radiation operations.
- Collaborators:
 - NASA JSC Space Radiation Analysis Group (SRAG)
 - NASA GSFC Community Coordinated Modeling Center (CCMC)
 - NASA GSFC Moon to Mars Space Weather Analysis Office (M2M)
- ISEP works directly with researchers to add capabilities to SEP models, e.g., new features or ability to run in real time
- **This collaboration is an excellent and effective example of a focused R2O2R effort.** The combined expertise and effort from all parties has produced the SEP Scoreboards and has begun to set standards for SEP model prediction and validation.



ISEP Approach

Identified SEP models that already were or could be readily implemented to run in real time

- SEPSTER
- SEPMOD
- UMASEP-10, -100, HESPERIA/UMASEP-500
- HESPERIA/RELeASE
- MAG4
- SPRINTS

Supported larger efforts to improve and adapt promising models

- SAWS-ASPECS
- iPATH
- Mauna Loa early CME alert

Supported development of new models

- SEPSTER2D
- Georgia State University
- UMASEP-30, -50
- HESPERIA/RELeASE+
- MagPy

Tested models in real time during Artemis I

- SRAG console operators and M2M analysts monitored the Scoreboards 24/7

Build on the relationships and products developed through ISEP

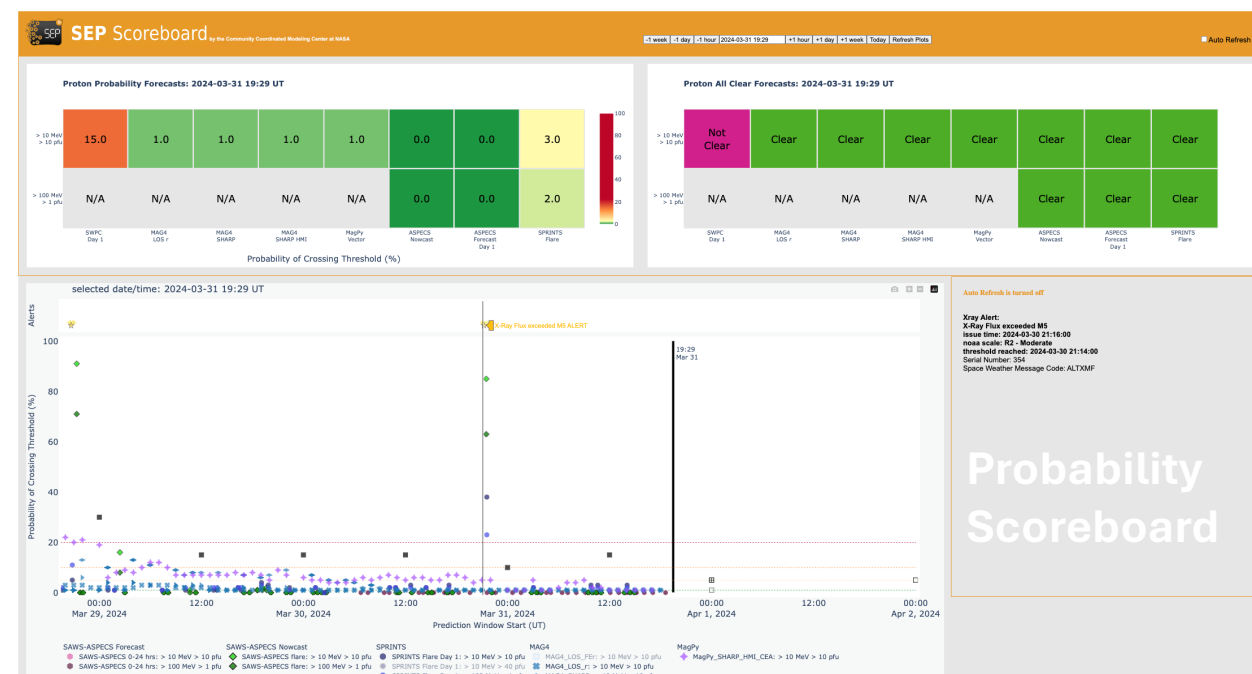
- Update, develop, and tailor models to meet SRAG's needs
- Incorporated UMASEP-100 into dose projection

ISEP Effectiveness

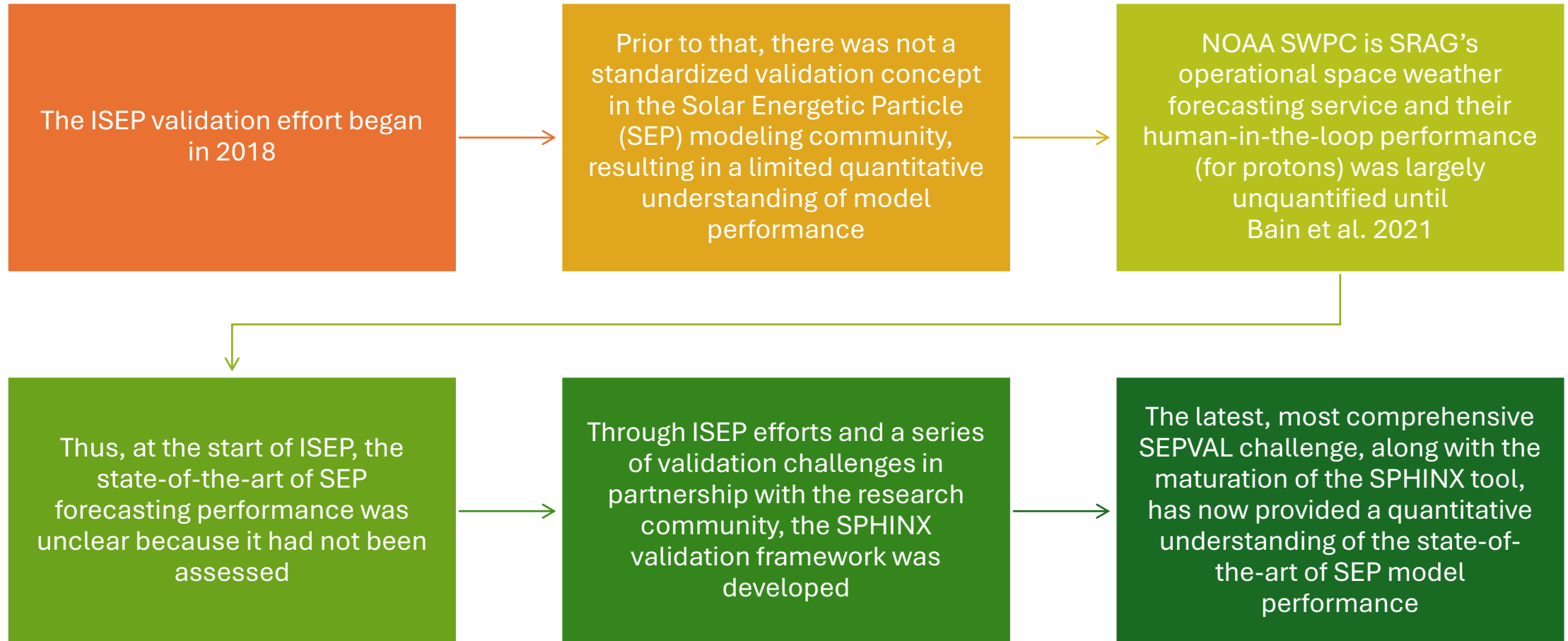
- Work one-on-one with model developers at research institutions
- Point of Contact from SRAG, CCMC, and M2M assigned to each model and developed expertise on each model
- SRAG, CCMC, M2M end-users have a high level of scientific knowledge and understanding regarding model approach
- End-users/operators directly discussing needs with developers, which allows implementation of features that directly address those need (e.g. SEPMOD high energy predictions, HESPERIA/REleASE+ reduction of false alarms)
- End-user interaction – developers don't have to guess what kind of features to add, communicated directly
- Many eyes on the models watching real time performance on the SEP Scoreboards and identifying strange outputs/problems
- Validation of model performance carried out through ISEP – development of tools and leading efforts within the scientific community

SEP Scoreboards – Intensity, Probability, All Clear

- SRAG console operators use the SEP Scoreboards to obtain situational awareness
- SRAG, CCMC, and M2M work with modelers on model development
- CCMC develops the Scoreboards and onboards models
- M2M provides human-in-the-loop space weather analyses to run models in real time
- SRAG, CCMC, and M2M work on the validation of model performance

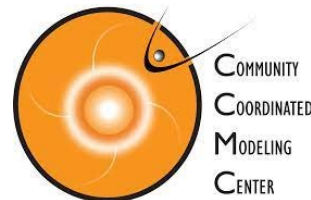


Overview of the ISEP Validation Effort



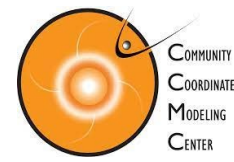
Multi-Pronged Effort towards the Validation of SEP Models

- **Primary Goals:**
 - **Develop a generalized, automated tool that can validate any kind of forecasted quantity from any type of solar energetic particle (SEP) prediction model.**
 - Provide feedback to model developers in an R2O2R cycle
- **Community validation effort through SHINE, ISWAT, ESWW, and SEPVAL**
 - Inspire quantitative validation
 - **Communicate forecasted quantities that would be useful in operations**
 - Encourage broad participation to enable development of a generalized tool
 - Establish benchmark lists
 - Evaluate the current state-of-the-art of model performance across the field
- **The SEP Scoreboards developed by CCMC through the ISEP project**
 - **Run models in real time**
 - Evaluate model performance for real time space radiation operations



And the 20+ participating model developers' institutions

SPHINX Validation Code



And the 17+
participating model
developers' institutions

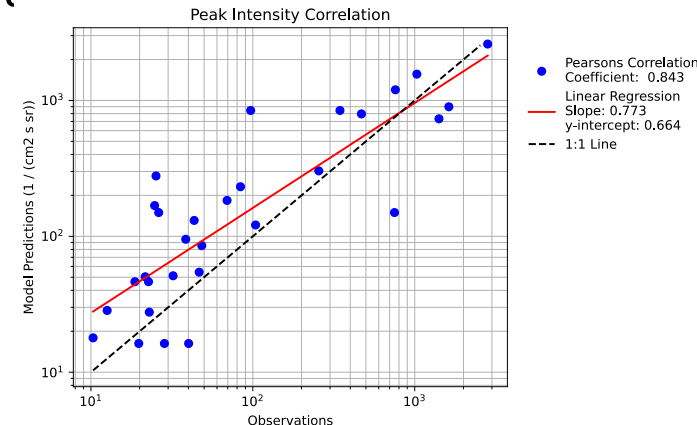
Solar Particles in the Heliosphere validation INfrastructure for SpWx (SPHINX)

SPHINX Goal: A generalized, automated tool that can validate any kind of forecasted quantity from any type of solar energetic particle (SEP) prediction model.

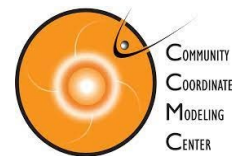
Developers: *Kathryn Whitman, Ricky Egeland, Luke Stegeman, Clayton Allison (NASA JSC SRAG)*

Contributors to the logical workflow: Collaborations with BIRA-IASB, NOAA SWPC, NASA CCMC, NASA M2M, and the participants of the SHINE, ISWAT, and SFPVAI community challenges.

Availability: Hosted at CCMC in the future.



VIVID Visualization Tool



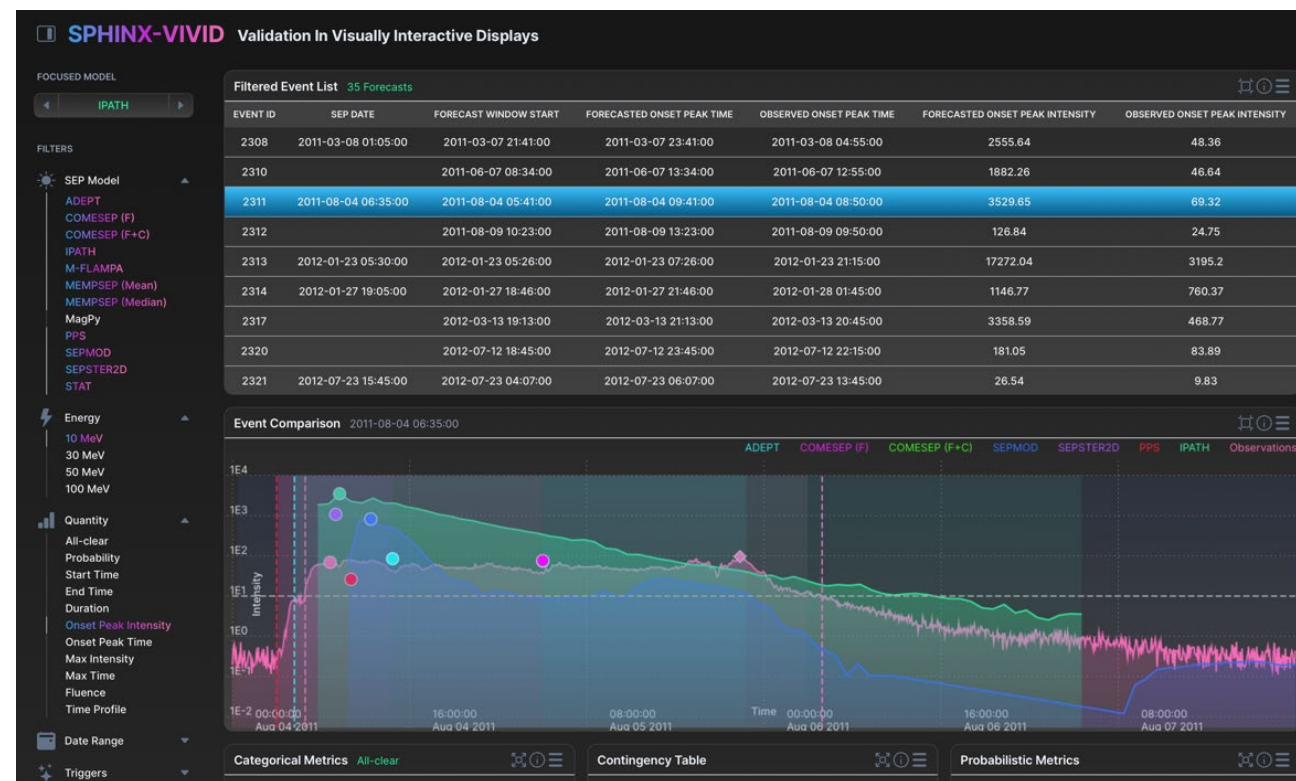
And the 17+
participating model
developers' institutions

Validation in Visually Interactive Displays (VIVID)

VIVID Goal: Web application for displaying the validation results of SPHINX in a dashboard of interactive plots and tables.

Developer: *Phil Quinn (NASA JSC SRAG)*

Availability: Hosted at CCMC in the future.



SPHINX Development

Historical validation and community challenges:

- Forecasts submitted to the community challenges have exercised SPHINX enabling the development of a generalized framework.

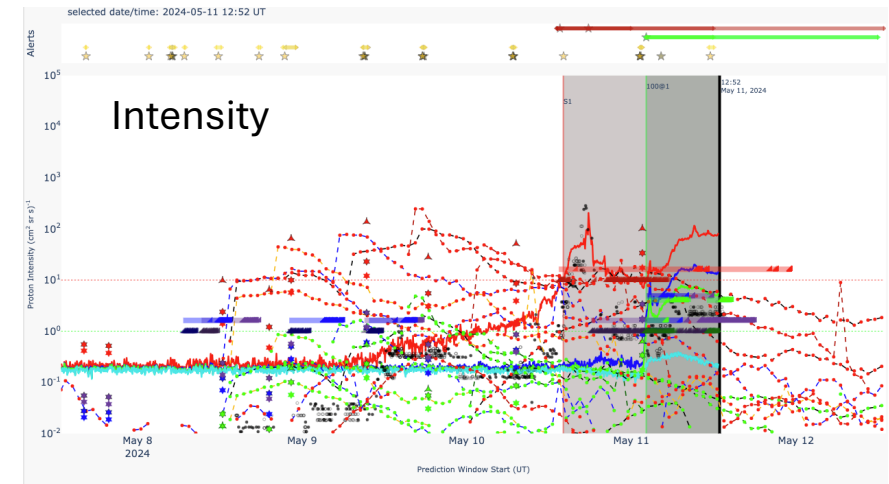
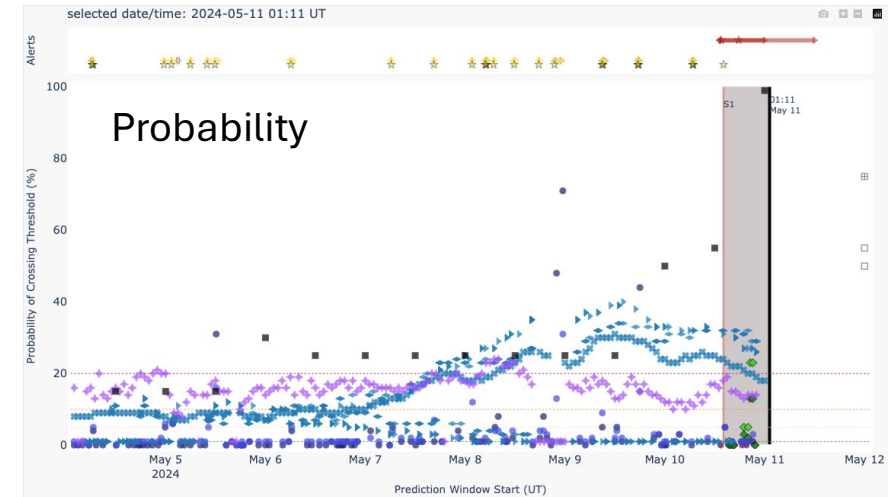
Validation of the SEP Scoreboards:

- Some models produce forecasts every few minutes, resulting in **millions of forecasts** over the lifespan of the Scoreboards.
- This drives the development of an automated methodology for the full validation process

SEP Scoreboards

May 11, 2024

Mother's Day/Gannon Memorial Storms



SEPVAL 2023 US and Europe Meetings

Following a multi-year validation effort through the SHINE, ISWAT, and ESWW workshops, we established the SEPVAL workshops which were focused on validation and SEP forecasting in ops.



SEPVAL organizers:

Provided a list of challenge time periods and triggers (flares, CMEs)

Defined rules of participation to encourage modelers to produce forecasts in a real time-like scenario

Performed the validation using SPHINX

Made the validation results available to attendees (R2O2R)

SEP model developers:

Provided forecasts and supplementary information

Followed the rules of participation

To provide feedback about the forecast/prediction process and the validation results

Activity of SEPVAL Community Effort

- **September and November 2023:** Received forecasts from SEPVAL participants (19 SEP models)
- **September and November 2023:** Preliminary “show and tell” of results at dedicated workshops, SEPVAL US & Europe
- **November 2023 – April 2024:** Improvements and additions to SPHINX in response to feedback and issues encountered during SEPVAL
- **April 2024:** SEPVAL results circulated to all model developers - individual model results and all models (anonymized)
- **April - May 2024:** One-on-one communication and meetings with developers (SPRINTS, UNSPELL, UMASEP, ASPECS)
- **May 2024:** Two virtual Q&A meetings to present specifics about SPHINX and clarify the outputs that were produced (26 developers and interested parties)
- **July & August 2024:** Presentation of results at COSPAR and SHINE

Forecasts Received for SEPVAL 2023 (19 Models)

United States
Europe

Model	Developer Point of Contact	Affiliation	Energy Channels	Forecasted Quantities	# Forecasts Submitted	# Forecasts Processed
ADEPT 1hr, 6hr	Stephen White	US Air Force	>10 MeV	Time Profile	25	25
COMESep flare, flare+CME	Mark Dierckxsens	BIRA	>10	Probability, Peak	63, 63	60, 63
cRT+AE10	Ming Zhang	Florida Institute of Technology	>10	Probability	63	63
ENLIL+SEPMOD	Janet Luhmann	UC Berkeley	>10, >30, >50, >100	Time Profile	63	63
Lavasa	Eleni Lavasa	National Observatory of Athens	>10	All Clear	58	58
MagPy	David Falconer, Tilaye Tedesse	UA Huntsville, NASA JSC SRAG	>10	Probability	2182	2182
MEMPSEP Mean, Median	Subhamoy Chatterjee	Southwest Research Institute	>10	Probability	60, 60	60, 60
MFLAMPA	Igor Sokolov	University of Michigan	>10, >30, >50, >100	Time Profile	9	8
PPS (SFS Update)	Stephen White	US Air Force	>10, >100	Peak Flux	61	61
SEPSAT	Ming Zhang	Florida Institute of Technology	>10, >100	Time Profile	64	64
SEPSTER	Ian Richardson	University of Maryland	>10, >30, >50, >100	Peak Flux	64	64
SEPSTER2D	Alessandro Bruno	NASA GSFC	>10, >30, >50, >100	Peak, Fluence	60	60
SPREAdFAST	Kamen Kozarev	Bulgarian Academy of Sciences	>10, >30, >50, >100	Time Profile	8	8
SPRINTS 0-24, 24-48, 48-72, 72-96	Alec Engell	NextGen	>10, >30, >50, >100	Probability, Peak	13134, 13134, 13134, 13134	13134, 13134, 13134, 13134
STAT	Jon Linker	Predictive Science, LLC	>10, >30, >50, >100	Time Profile	6	6
UMASEP-10, -100	Marlon Nunez	University of Malaga	>10, >30, >50, >100	Peak, Start	27572, 32240	27161, 32106
UNSPELL	Sigiava Alminalragia-Giamini	SPARC	>5	Probability	61	61
ZEUS+iPATH	Gang Li, Junxiang Hu	UA Hunstville, NASA GSFC	>10, >30, >50, >100	Time Profile	57	56

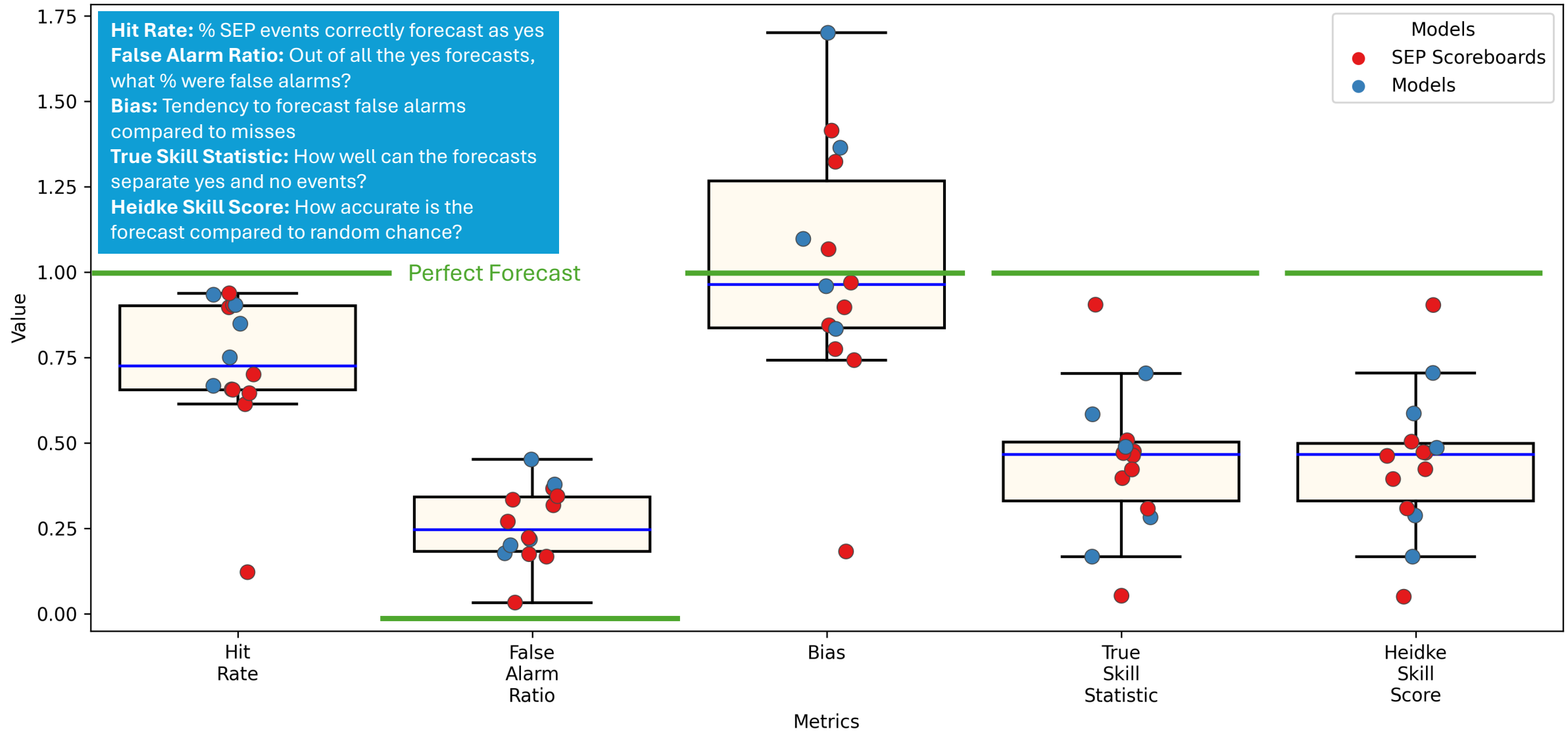
SAWS-ASPECS

Developer: Athanasios Papaioannou+, Affiliation: National Observatory of Athens, Energy Channels: >10, >100 MeV

Model SAWS-ASPECS	Forecasted Quantities	# Forecasts Submitted	# Forecasts Processed	Model SAWS-ASPECS	Forecasted Quantities	# Forecasts Submitted	# Forecasts Processed
CME (CACTus) 50%	Time Profile	60	60	CME (CACTus) electrons 50%	Time Profile	60	60
CME (CACTus) 90%	Time Profile	60	60	CME (CACTus) electrons 90%	Time Profile	60	60
CME (CACTus)	Probability/All Clear	60	60	CME (CACTus) electrons	Probability/All Clear	60	60
CME (CDAW) 50%	Time Profile	63	63	CME (CDAW) electrons 50%	Time Profile	63	63
CME (CDAW) 90%	Time Profile	63	63	CME (CDAW) electrons 90%	Time Profile	63	63
CME (CDAW)	Probability/All Clear	63	63	CME (CDAW) electrons	Probability/All Clear	63	63
Flare+CME (CACTus) 50%	Time Profile	57	57	Flare+CME (CACTus) electrons 50%	Time Profile	57	57
Flare+CME (CACTus) 90%	Time Profile	57	57	Flare+CME (CACTus) electrons 90%	Time Profile	57	57
Flare+CME (CACTus)	Probability/All Clear	57	57	Flare+CME (CACTus) electrons	Probability/All Clear	57	57
Flare+CME (CDAW) 50%	Time Profile	60	60	Flare+CME (CDAW) electrons 50%	Time Profile	60	60
Flare+CME (CDAW) 90%	Time Profile	60	60	Flare+CME (CDAW) electrons 90%	Time Profile	60	60
Flare+CME (CDAW)	Probability/All Clear	60	60	Flare+CME (CDAW) electrons	Probability/All Clear	60	60
Flare 50%	Time Profile	60	60	Flare electrons 50%	Time Profile	60	60
Flare 90%	Time Profile	60	60	Flare electrons 90%	Time Profile	60	60
Flare	Probability/All Clear	60	60	Flare electrons	Probability/All Clear	60	60

Performance of SRAG/CCMC SEP Scoreboard Models for SEPVAL Events

All Clear Group 7 (min.10.0.max.-1.0.units.MeV, threshold.10.0.units.1 / (cm² s sr))



The SEP Scoreboards contain models across all levels of performance, including top performers.

SEPVAL

Community Challenge



Designed, thorough, structured challenge gives an independent characterization of model performance



Provided an understanding of the state-of-the-art performance of SEP models across the research community (never been done before)



Receiving output from many types of models enabled the development of SPHINX as a generalized tool



SEPVAL contributors are now better-prepared to participate in the SEP Scoreboards



We can establish target benchmarks for model performance now that we have a clearer view of state-of-the-art

SPHINX

Validation Framework



We can identify which metrics are most meaningful for SRAG



Provide feedback to SEP model developers to inspire improvements in an R2O2R cycle



Identify which aspects of performance a model needs to improve to reach benchmark goals



Provide a tool to evaluate gap closure



Future: continuously updated assessment of SEP Scoreboard real-time model performance

Validation Effort R2O2R Outcomes

- Gain an understanding of the state-of-the-art of SEP forecasting
 - Along with Bain et al. (2021), get an understanding of how well we can predict SEPs
- SEPVAL effort resulted in a set of consistent, high-quality CME fits produced by the M2M office and cataloged on DONKI that can now be used by the research community
- Development of validation tools and standards for the community to enable an activity that isn't typically supported by research funding
- End user / developer interaction allows the developer to communicate the framework in which the model is intended to be used

Validation Effort R202R Outcomes

- **Specific model actions in response to SEPVAL:**
 - Developers identified bugs in their model that impacted performance and required fixes
 - Developers with research models thought through how to present their forecasts to operators for the first time – predicted value, timing, CCMC JSON schema
 - Some developers identified issues with the method/pipeline used to create the forecasts
 - Developers considered the need for recalibration after viewing validation results
- **Specific example - iPATH improvements based on SEPVAL/RT validation:**
 - Shock finding algorithm improvements
 - Extended transport runs to 4 days
 - Added projection of CME speed to equatorial plane (2D projection based on CME latitude)
 - Turbulence levels are calculated from ACE/DSCOVR
 - Modified density calculation for slow CMEs (help prevents false alarms found in initial SEPVAL results)

Effective R2O2R Summary

- One-on-one work between end-users and model developers
- Knowledgeable, expert end-users that can assist
- Direct communication to allow end-users to express needs
- Models forecasting in a real time environment with the support and attention of operators to provide feedback
- Concentrated validation effort to provide feedback about model performance – highlight strengths and aspects of models that need improvement

Back Up

Quantities Validated by SPHINX

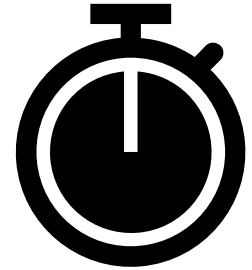
Will an SPE occur?

- All Clear (boolean threshold crossing)
- Probability of Occurrence



When? How long?

- Start Time
- Peak Time
- End Time
- Duration



How big?

- Peak Flux
- Fluence
- SEP Intensity Time Profile



How much warning?

- Advanced Warning Time

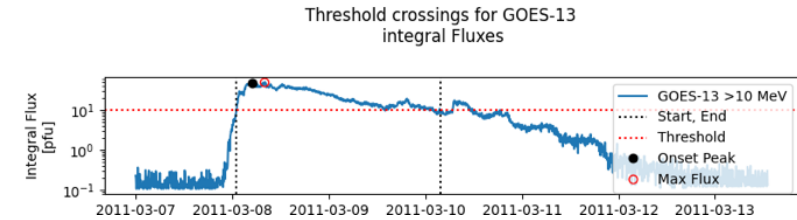


We validate many aspects of SEP events that are of interest during contingency operations for protecting astronaut health.

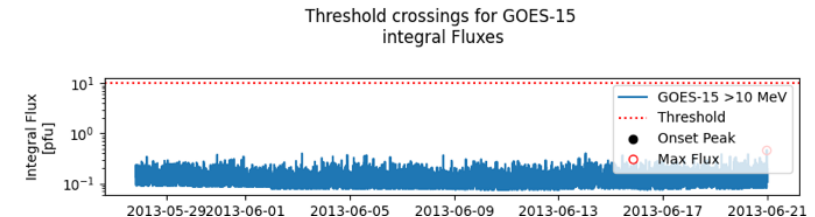
SEPVAL Challenge Events

- **33 SEP events** between 2011 – 2023 (Hits & Misses)
- **30 non-event** periods between 2012 – 2023 (False Alarms & Correct Negatives)
 - Eruption followed by no enhancement in proton flux
 - Eruption followed by a small or below threshold enhancement of proton flux
- **M2M Office remeasured all CMEs to ensure consistent measurements and give additional info**
- **Encouraged developers to run models in a real-time-like workflow**

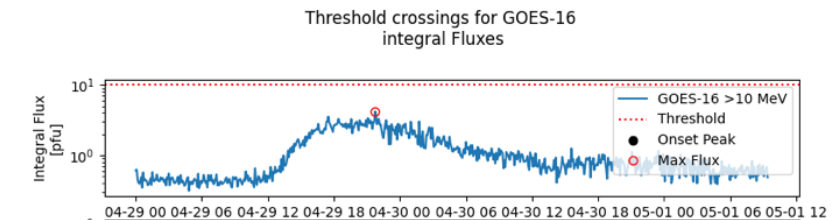
SEP Event 2011-03-08



Non-Event 2013-05-27



Non-Event 2022-04-29



Next Steps for SPHINX and SEPVAL

- **SPHINX will be automated to validate the SEP Scoreboards in real time**
 - Hosted and served at CCMC with a public facing interface
- **Reach out to additional SEP models**
 - Aware of 10+ SEP models that exist but have not yet/recently participated
 - Models beyond US and Europe
- **Prepare a journal paper with SEPVAL results** working directly with each participating SEP model developer
- **Establish a focus for the next phase of SEPVAL** at ESWW2024, ISWAT2025:
 - Use SEPVAL results to motivate discussions between model developers
 - Are there cases where all models missed or hit events?
 - Are there cases where a few models correctly predicted an event while others missed?
 - What do these outcomes say about the predictive physical characteristics being captured by the models and what is missing?