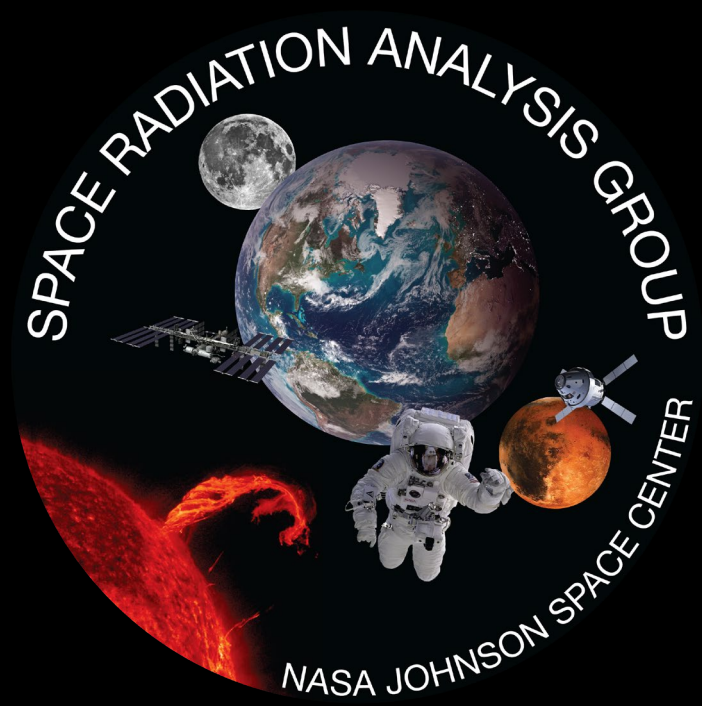
A Multi-Year Effort to Forward the Validation of Solar Energetic Particle Prediction Models

Kathryn Whitman, **Ricky Egeland**, Philip Quinn, Luke Stegeman, Clayton Allison

(NASA JSC SRAG)

With Mark Dierckxsens (BIRA-IASB), M. Leila Mays (NASA GSFC), Yaireska Collado-Vega (NASA GSFC), Hazel Bain (CU Boulder CIRES/NOAA SWPC)

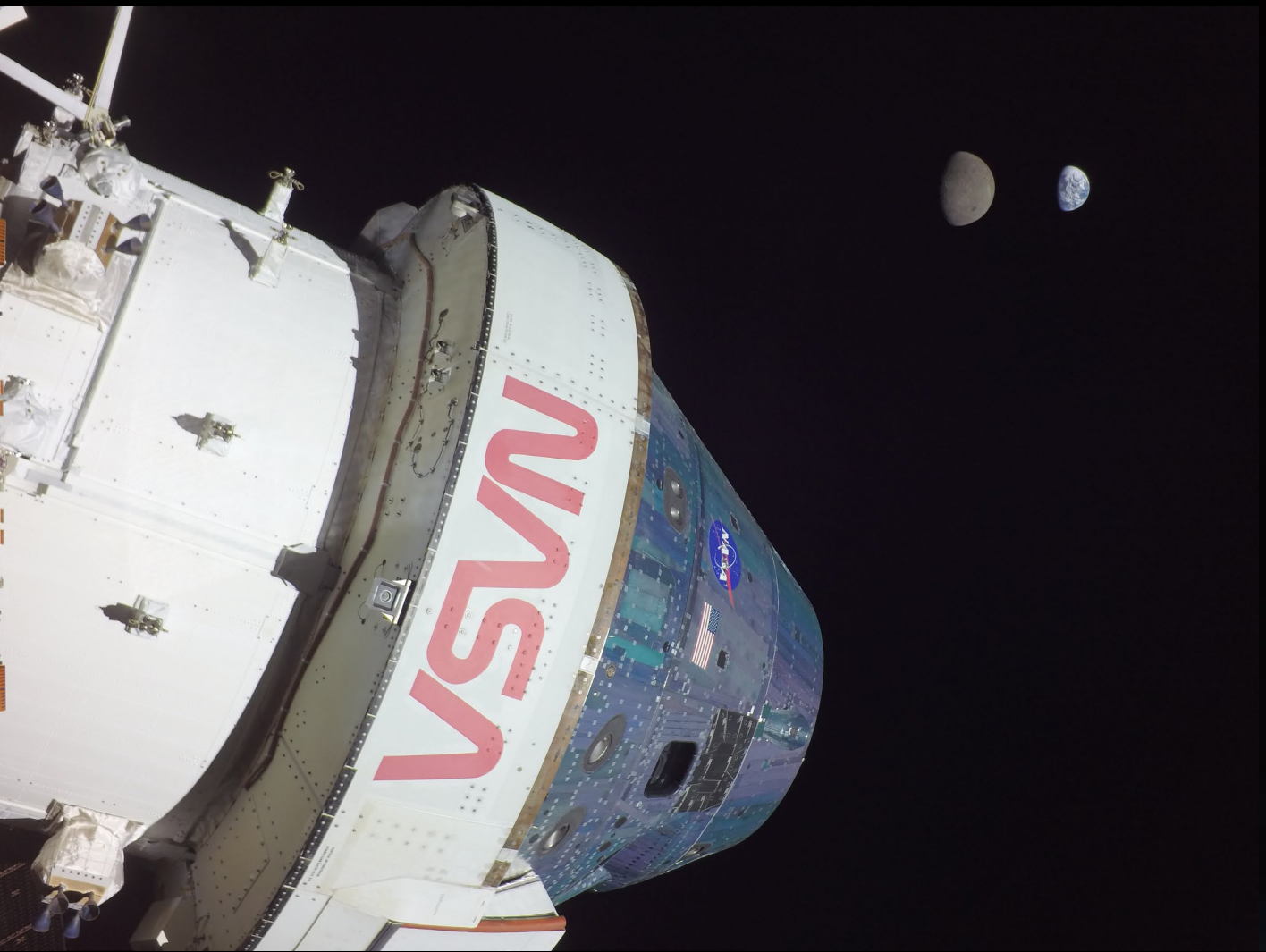
TESS, Dallas, TX
April 9, 2024



And the 17+ participating model developers' institutions

Artemis is Driving New SEP Forecasting Efforts

ISEP: Integrated Solar Energetic Proton Event Alert/Warning System



Overview of the ISEP Validation Effort

The ISEP validation effort began in 2018/2019

Prior to that, many developers of Solar Energetic Particle (SEP) models did not focus on validation and did not have a quantitative sense of how their models performed

NOAA SWPC is SRAG's operational space weather forecasting service and their human-in-the-loop performance was largely unquantified until Bain et al. 2021

Thus, at the start of ISEP, the state-of-the-art of SEP forecasting performance was unclear because it had not been assessed

Through ISEP efforts and a series of validation challenges in partnership with the research community, the SPHINX validation framework was developed

The latest, most comprehensive SEPVAL challenge, along with the maturation of the SPHINX tool, has now provided a quantitative understanding of the state-of-the-art of SEP model performance

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.

- Assess SEP model performance
- Establish standards
- Develop a generalized framework for model validation
- Operational use of SEP models
- NASA, NOAA, ESA infrastructure being developed for the R2O transitioning of models to ops

Organization of the SEPVAL Challenge

- **SEPVAL organizers:**
  - Provide a list of challenge time periods and triggers (flares, CMEs)
    - *M2M ensured quality CME measurements by checking all 3D CME parameters in DONKI provided for this challenge while also providing additional information*
  - Define rules of participation to encourage modelers to produce forecasts in a real time-like scenario
  - Perform the validation using SPHINX
  - Make the validation results available to attendees (R2O2R)
- **SEP model developers:**
  - Provide forecasts and supplementary information
  - Follow the rules of participation
  - Provide feedback about the forecast/prediction process and the validation results

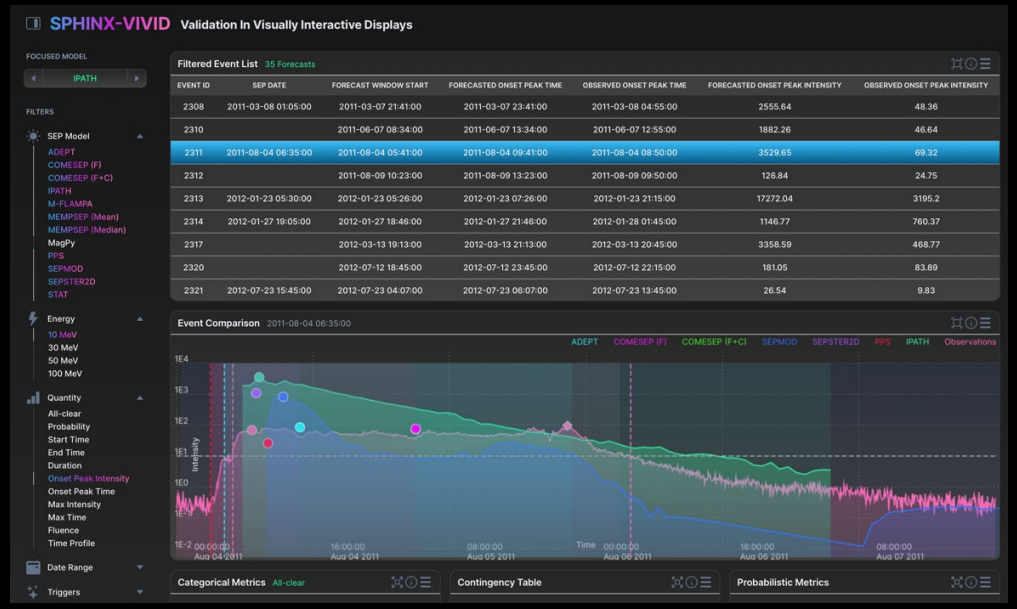
SPHINX Validation Code and VIVID Visualization Tool

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*

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*



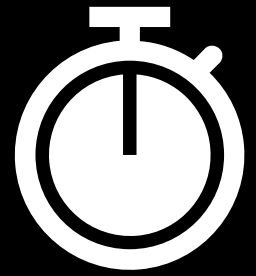
Quantities Validated by SPHINX

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.

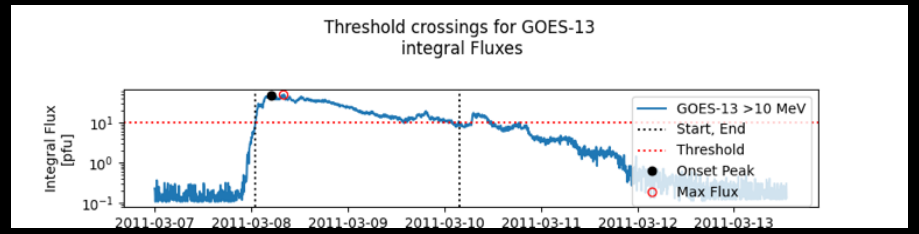
SEPVAL Challenge Events

- **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
- Events in both sets occurred throughout Solar Cycles 24 & 25

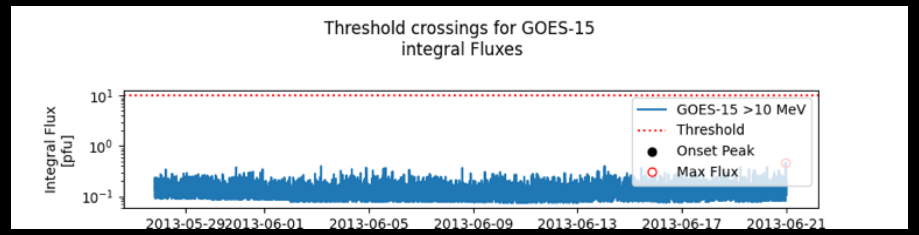
Event list:
https://docs.google.com/spreadsheets/d/11BiYbBALP-x0n4qxURo_78rCe0bDC6kKxuEmDO2noUk/edit?usp=sharing

Non-Event list:
https://docs.google.com/spreadsheets/d/1SPyMLBuopTp5IkMRjEGAMXVzGwcjO8sulpw7JHR78q0/edit?usp=sharing

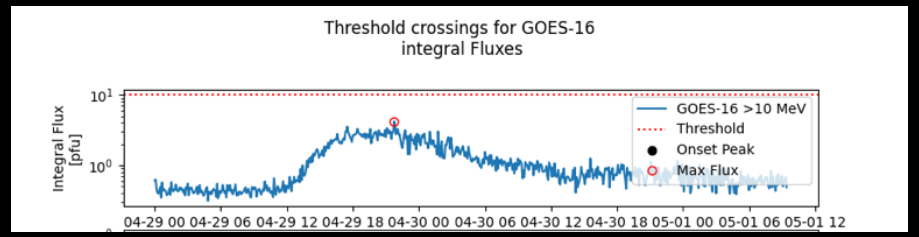
SEP Event 2011-03-08



Non-Event 2013-05-27

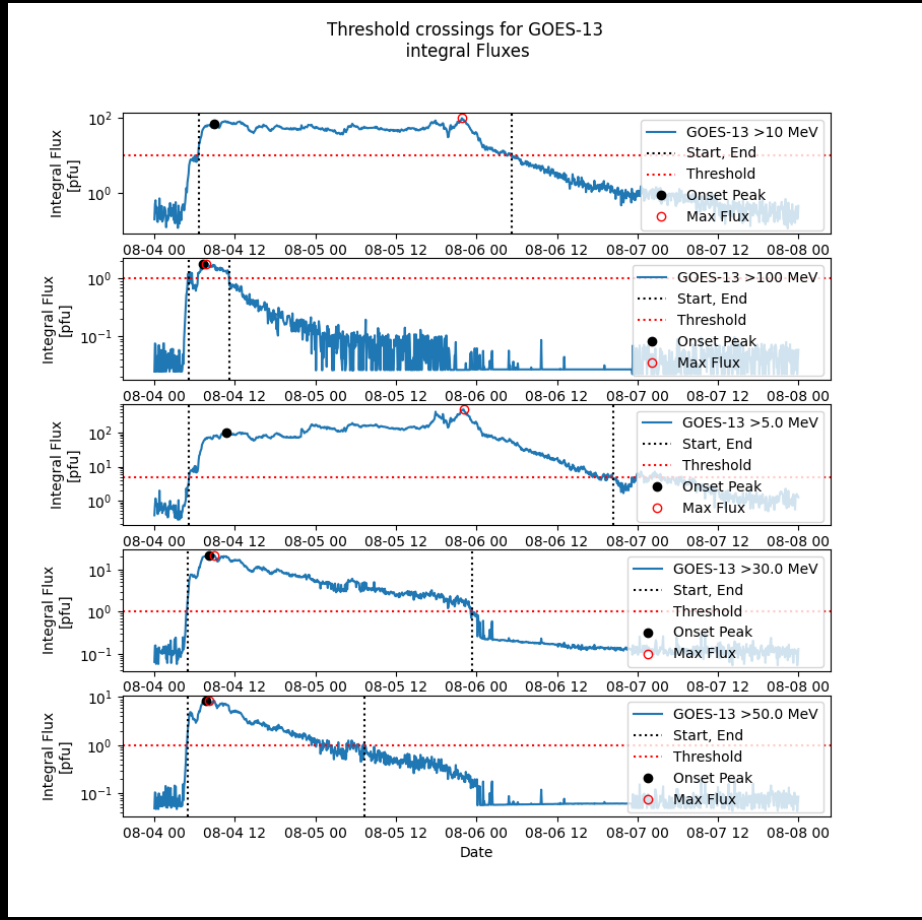


Non-Event 2022-04-29



Preparation of Observations for Validation

- Used the fetchsep package* developed by Katie Whitman to prepare the observational SEP event properties
- Energy channel and threshold combinations applied:
  - >5 MeV, 5 pfu
  - **>10 MeV, 10 pfu (NOAA, SRAG operational threshold)**
  - >30 MeV, 1 pfu
  - >50 MeV, 1 pfu
  - **>100 MeV, 1 pfu (SRAG operational threshold)**



* https://fetchsep.readthedocs.io

Forecasts Received for SEPVAL 2023

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	61	61
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

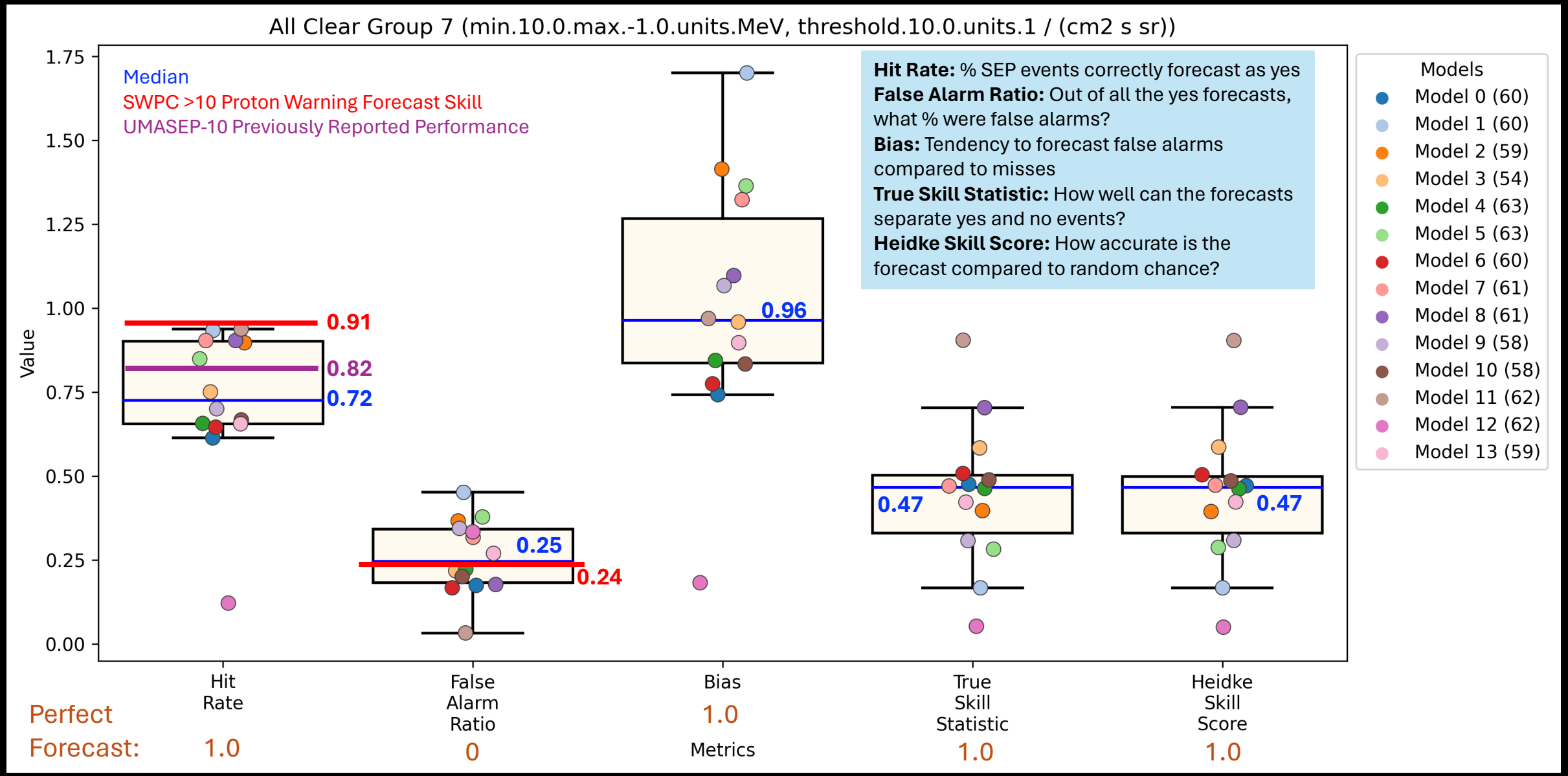
SAWS-ASPECS

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

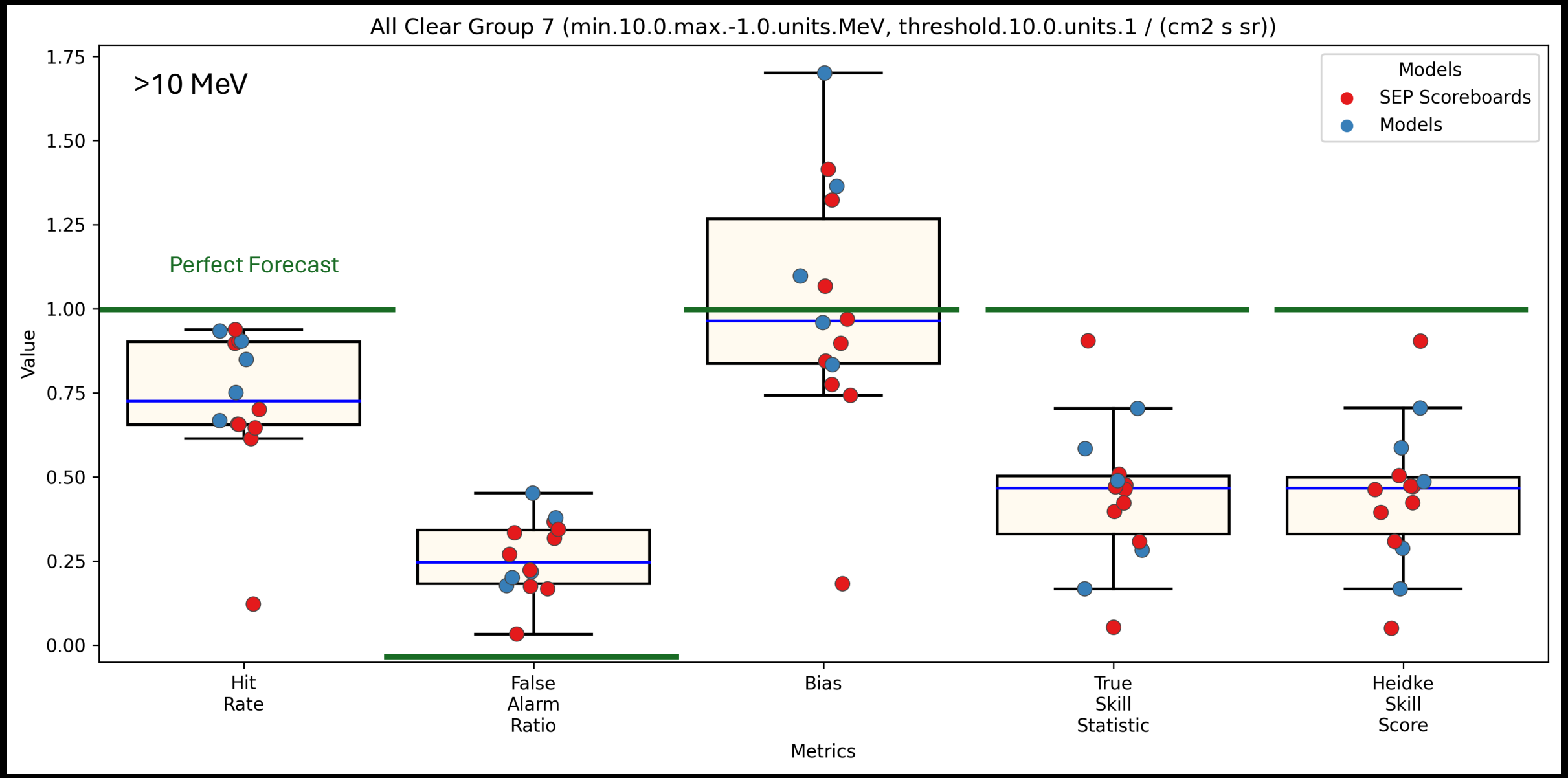
Model SAWS-ASPECS	Forecasted Quantities	# Forecasts Submitted	# Forecasts Processed
CME (CACTus) 50%	Time Profile	60	60
CME (CACTus) 90%	Time Profile	60	60
CME (CACTus)	Probability/All Clear	60	60
CME (CDAW) 50%	Time Profile	63	63
CME (CDAW) 90%	Time Profile	63	63
CME (CDAW)	Probability/All Clear	63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

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

All Clear Performance (>10 MeV)



Performance of SRAG/CCMC SEP Scoreboard Models



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

SEPVAL

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

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