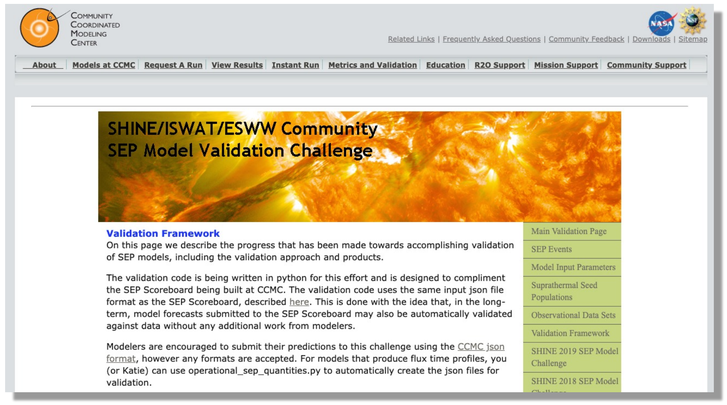
Community Effort towards Solar Energetic Particle Model Validation Motivated by Space Radiation Operations

Kathryn Whitman (NASA JSC SRAG/ University of Houston), Eddie Semones (NASA JSC SRAG), M. Leila Mays (NASA GSFC CCMC), Yaireska Collado-Vega (NASA GSFC M2M) on behalf of the ISEP collaboration; *corresponding author Kathryn.Whitman@nasa.gov*

Introduction

Solar Energetic Particle (SEP) events pose a radiation hazard for humans in space, particularly outside of Low Earth Orbit (LEO) away from the Earth's protective magnetosphere. A community effort to validate Solar Energetic Particle (SEP) models has been ongoing through the SHINE, ISWAT, and ESWW meetings in support of the Integrated Solar Energetic Proton Event Alert/Warning System (ISEP) project, a joint collaboration between the NASA Johnson Space Radiation Analysis Group (SRAG) , NASA Goddard's Community Coordinated Modeling Center (CCMC), and the new Moon to Mars Space Weather Analysis Office (M2M) at NASA Goddard. The ISEP project was established to improve SRAG's SEP forecasting capability for human exploration missions beyond LEO.

Methods to fairly, consistently, and quantitatively validate the diverse set of SEP models developed in the community must be established in order to assess performance, both from a scientific and operational standpoint. **We present an overview of the ongoing validation efforts resulting from discussions with the research community and direct participation of modelers.**



- **A community model validation effort is ongoing through SHINE, ISWAT and ESWW, lead by SRAG, CCMC, and SWPC.** Supporting information available at: https://ccmc.gsfc.nasa.gov/assessment/topics/SEP/campaign2020.php https://iswat-cospar.org/H3-01
- Website is updated as campaign progresses – check back regularly!

Methodology for Development of the SEP Model Validation Framework

1. Collect predictions from all SEP model types through a community effort (SHINE/ISWAT/ESWW)
2. Create a code (OpSEP) to derive equivalent quantities from observational data sets
3. Create a generalized and flexible framework to pair observations and predictions and validate for *science and operations*

Supporting information provided to the SEP Modeling Community:

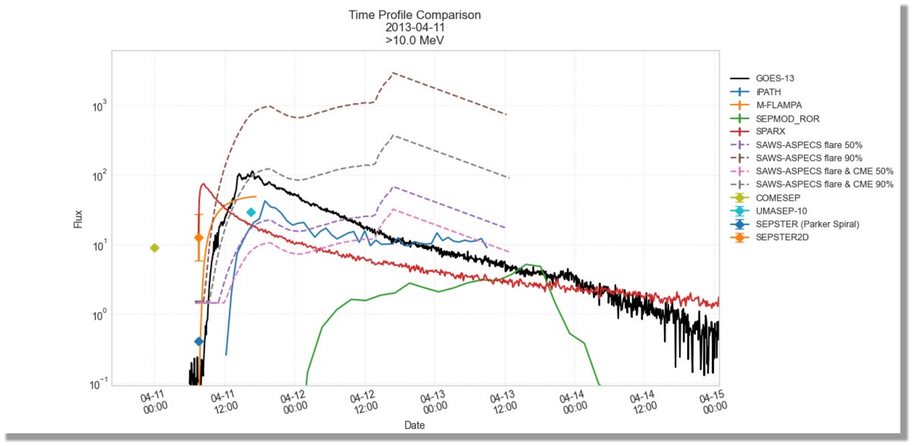
- **Selected 10 SEP events**
- **Standardized Inputs:**
  - Provided forecasting-quality flare and CME inputs for model triggers
  - Derived observed suprathermal spectra for each event to determine if useful for modelers
- **Observations:**
  - Developed set of recommendations for best use of available data sets for validation
- **Ease of participation:**
  - Created a validation framework that accepts the same JSON file format as the CCMC SEP Scoreboard https://ccmc.gsfc.nasa.gov/challenges/sep.php#format

1. Collect Predictions from Models

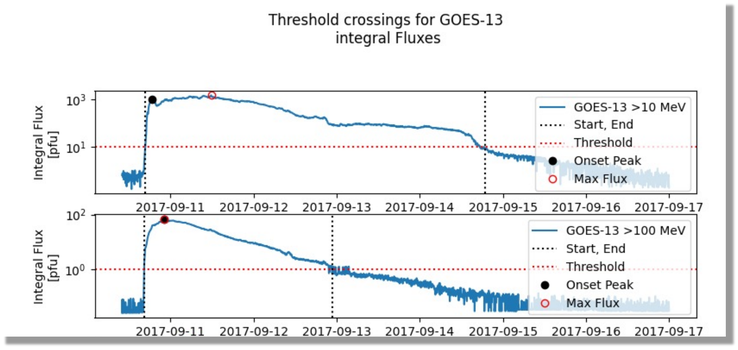
Modelers representing a wide variety of model types submitted forecasts for 1 or more of the challenge events. Forecast types include probability, all clear, event onset, peak flux, and partial and full time profiles. Modelers were encouraged to use inputs provided and to run models in a "forecast mode", however this is not required for participation. Continuously accepting submissions.

PARTICIPATING MODELS:

- **ASPECS** (Papaioannou et. al.)
- **COMESEP** (Dierckxsens et al.)
- **HESPERIA/REleASE** (Posner, Kuhl, Malandraki)
- **iPATH + ZEUS** (Li, Hu)
- **MAG4 SEP** (Falconer, Khazanov)
- **M-FLAMPA** (Sokolov, Zhao)
- **SEPMOD + ENLIL** (Luhmann)
- **SEPCaster** (iPATH + AWsoM) (Li, Jin)
- **SEPSTER** (Richardson)
- **SEPSTER2D** (Bruno)
- **SPARX** (Marsh, Dalla, Swalwell)
- **STAT** (MAS + EPREM) (Linker, Schwadron)
- **UMASEP** (Núñez)



2. Derive Matching Observed Values: OpSEP



The OpSEP code was created to derive observed SEP quantities for validation of models, including timing, onset peak flux, max flux, and event fluence, by applying a threshold to the flux time profile. The code always applies operational thresholds and allows users to specify additional thresholds. Users may also input modeled flux time profiles to extract quantities in the same way as observations.

- Operational thresholds >10 MeV exceeds 10 pfu, >100 MeV exceeds 1 pfu
- Includes native observational data sets and post-processing features
- https://github.com/ktindiana/operational-sep

3. Validate: Example Output from the Validation Framework

VALIDATION CODE CAPABILITIES:

1. Accepts a list of observations and a list of model forecast jsons
2. Automatically connect observations and predictions by date and energy channel + threshold
3. Calculate skill scores and metrics pertinent to each model's forecasts – bias, accuracy, distribution, skill scores

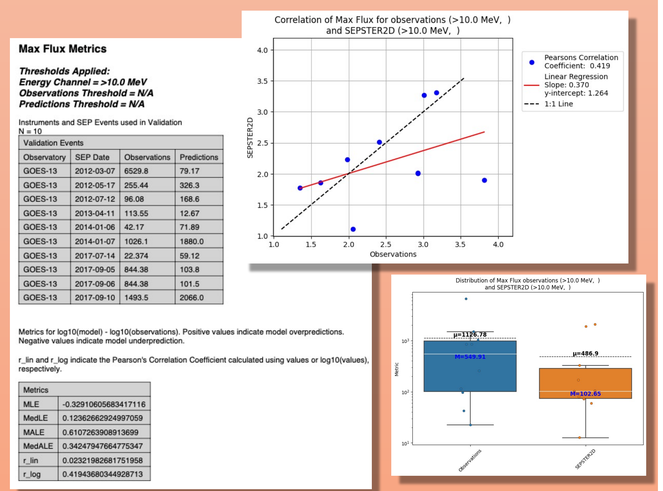
OUTPUTS:

- PDF report of the validation results for each individual model
- Produces a report showing all model results side-by-side

VALIDATED FORECAST QUANTITIES

- **Timing:** start time, end time, onset peak and max flux time,
- **Peak:** Onset peak (flux after initial increase - black points in 2.) and maximum flux between the start and end time
- **Probability of occurrence**
- **All Clear** (binary)
- **Event fluence**
- **Overall flux time profile**

Page of the report produced for SEPSTER2D validating max SEP flux with accompanying correlation and box plots



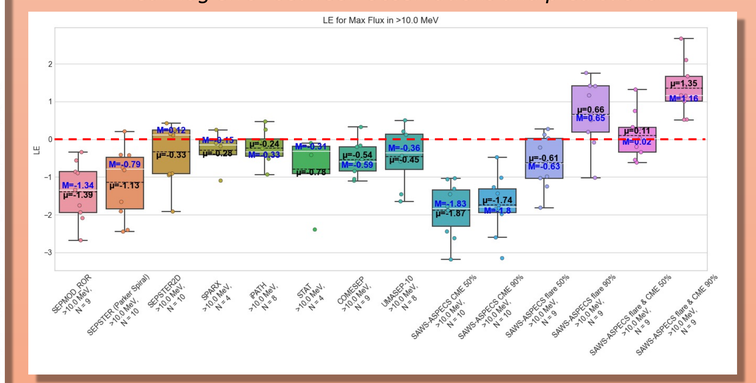
MAG4_HARP All Clear contingency table and skill scores for >10 MeV

Contingency Table

	Observed Yes	Observed No
Forecast Yes	6	0
Forecast No	1	0

SKILL SCORES	
Hits (TP)	6
Misses (FN)	1
False Alarms	0
Correct Negatives	0
Percent Correct	0.8571428571428571
Bias	0.8571428571428571
Hit Rate	0.8571428571428571
False Alarm Ratio	0.0
False Alarm Rate	nan
Frequency of Hits	1.0
Frequency of Misses	0.14285714285714285
Probability of Correct Negatives	nan
Detection Failure Ratio	1.0
Frequency of Correct Negatives	0.0
Threat Score	0.8571428571428571
Odds Ratio	nan
G Skill Score	0.0
True Skill Score	nan
Heidke Skill Score	0.0
Odds Ratio Skill Score	nan

Example comparison of model forecasts for >10 MeV maximum SEP flux. The box plots show the distribution of mean log error between observations and predictions.



➢ ***Example* metrics and plots derived by the validation code are shown here**
➢ These are **not representative validation results** due to small numbers and limited types of events, rather they are examples of the types of output and metrics produced by the validation code