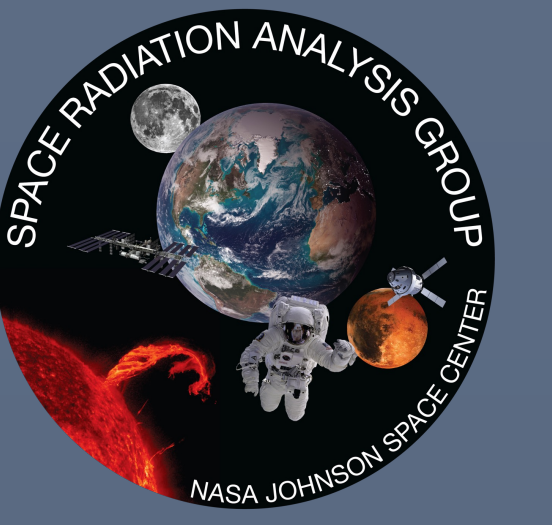


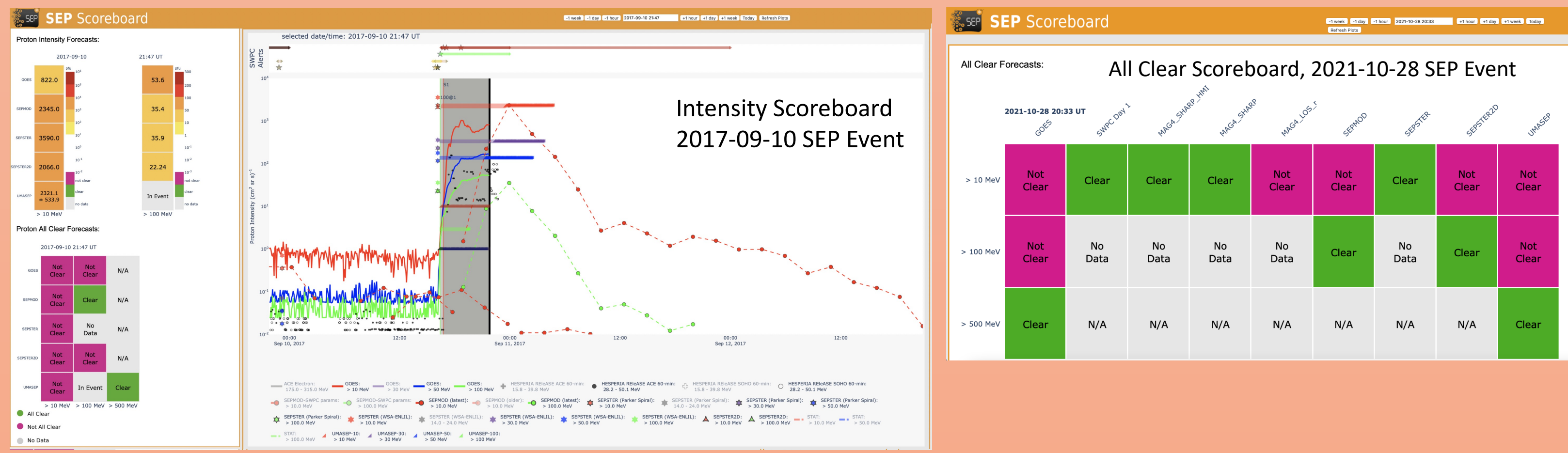
SPHINX: A GENERALIZED TOOL FOR SOLAR ENERGETIC PARTICLE MODEL VALIDATION



Kathryn Whitman (NASA JSC SRAG/KBR), Philip R. Quinn (NASA JSC SRAG/Leidos), Ricky Egeland (NASA JSC SRAG), Clayton Allison (NASA JSC SRAG/Leidos), Edward Semones (NASA JSC SRAG), Leila Mays (NASA GSFC), Yaireska Collado-Vega (NASA GSFC)

Abstract

The Integrated Solar Energetic Proton Alert/Warning (ISEP) project was established to aid in the transition of SEP models from the research realm to operational use. This project is a collaboration between the NASA Community Coordinated Model Center (CCMC), the NASA Moon to Mars Space Weather Analysis Office (M2M) and the NASA Space Radiation Analysis Group (SRAG). Through ISEP, the SEP Scoreboards (Intensity, Probability, and All Clear) were developed to visualize real time forecasts from SEP models for use in SRAG operations. To accomplish this, SRAG, CCMC, and M2M work closely with modelers on model development, CCMC works on Scoreboard implementation, and M2M provides human-in-the-loop space weather analysis and ensures that models run in real time.



Now that the SEP Scoreboards have been established and are running in real time, it is important to understand model performance to inform operators how to best interpret model forecasts. For this reason, the SPHINX validation tool is being developed at SRAG in collaboration with CCMC and M2M. SPHINX builds upon the community efforts carried out through the SHINE SEP model validation challenges in 2018, 2019, and 2022 and work accomplished through the ISWAT H3-01 section.

SHINE/ISWAT/ESWW SEP Model Validation Challenges

The SEP Model Validation Challenge past efforts and event lists are described on the website <https://ccmc.gsfc.nasa.gov/challenges/sep/>.

The current SEP challenge lists and Rules of Participation are available on the SEPVAL2023 website: <https://ccmc.gsfc.nasa.gov/community-workshops/ccmc-sepval-2023/>.

This ongoing validation effort has been carried out in three phases so far:

- **First Phase during SHINE 2019 & ISWAT 2022:** Collect forecasts for 10 Challenge SEP Events and encourage quantitative comparisons of model predictions to observations.
- **Second Phase during SHINE 2022:** Collect forecasts for 14 periods when no SEP was observed to test for false alarms & correct negatives.
- **Third and Current Phase during the upcoming SEPVAL 2023 working meetings:** Validation with statistically significant number of 30 SEP events and 33 non-events using standardized set of inputs and strict requirements to compute meaningful metrics and allow for cross-model comparisons.
 - The lists contain supporting flare and CME information. The CME 3D parameters have been reviewed and updated by M2M analysts to ensure measurements are of consistent quality throughout the event lists which extend from 2011 to present. M2M also added timing information to allow challenge participants to estimate simulated issued times to get a sense of advanced warning provided by the models.



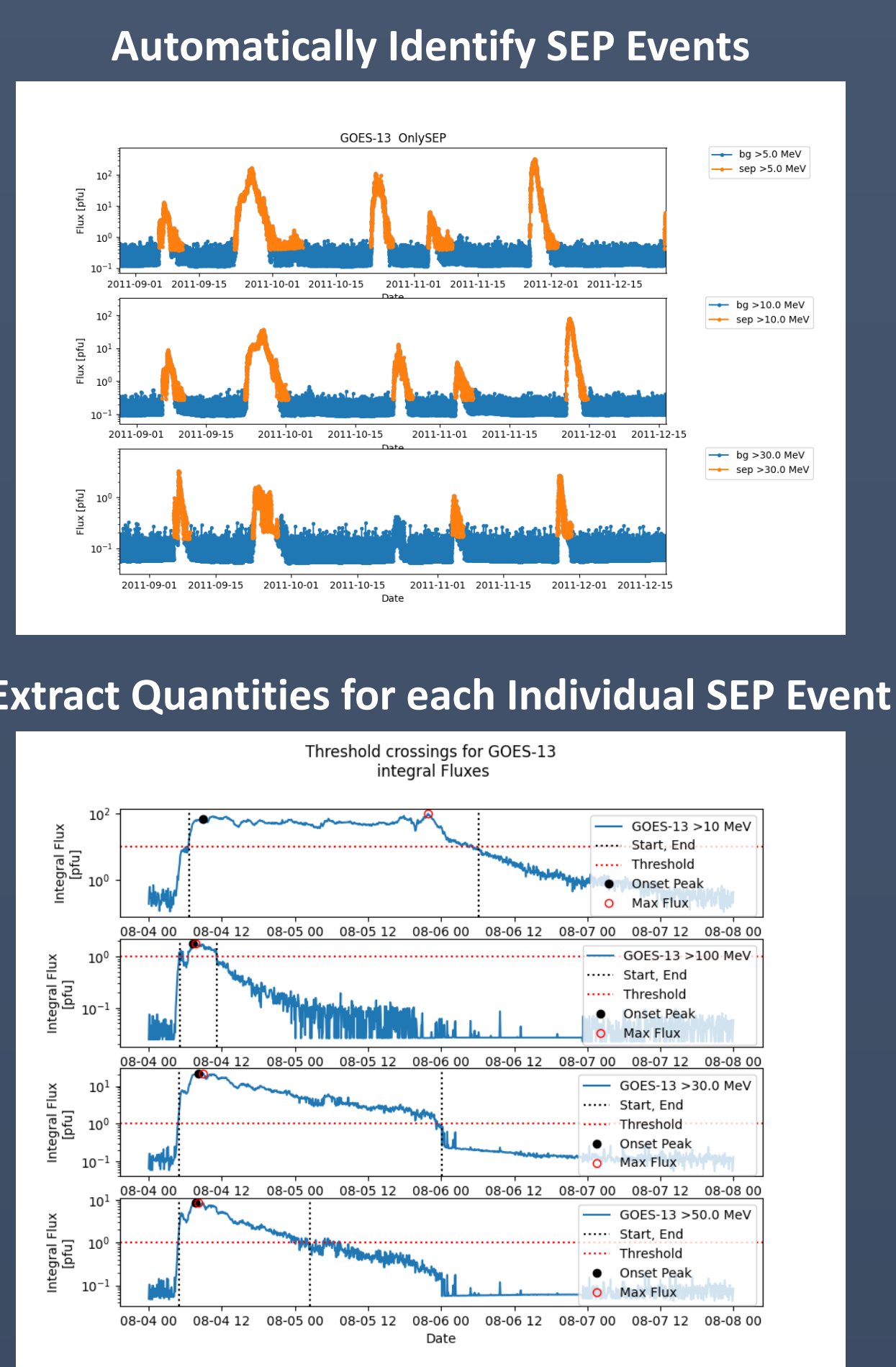
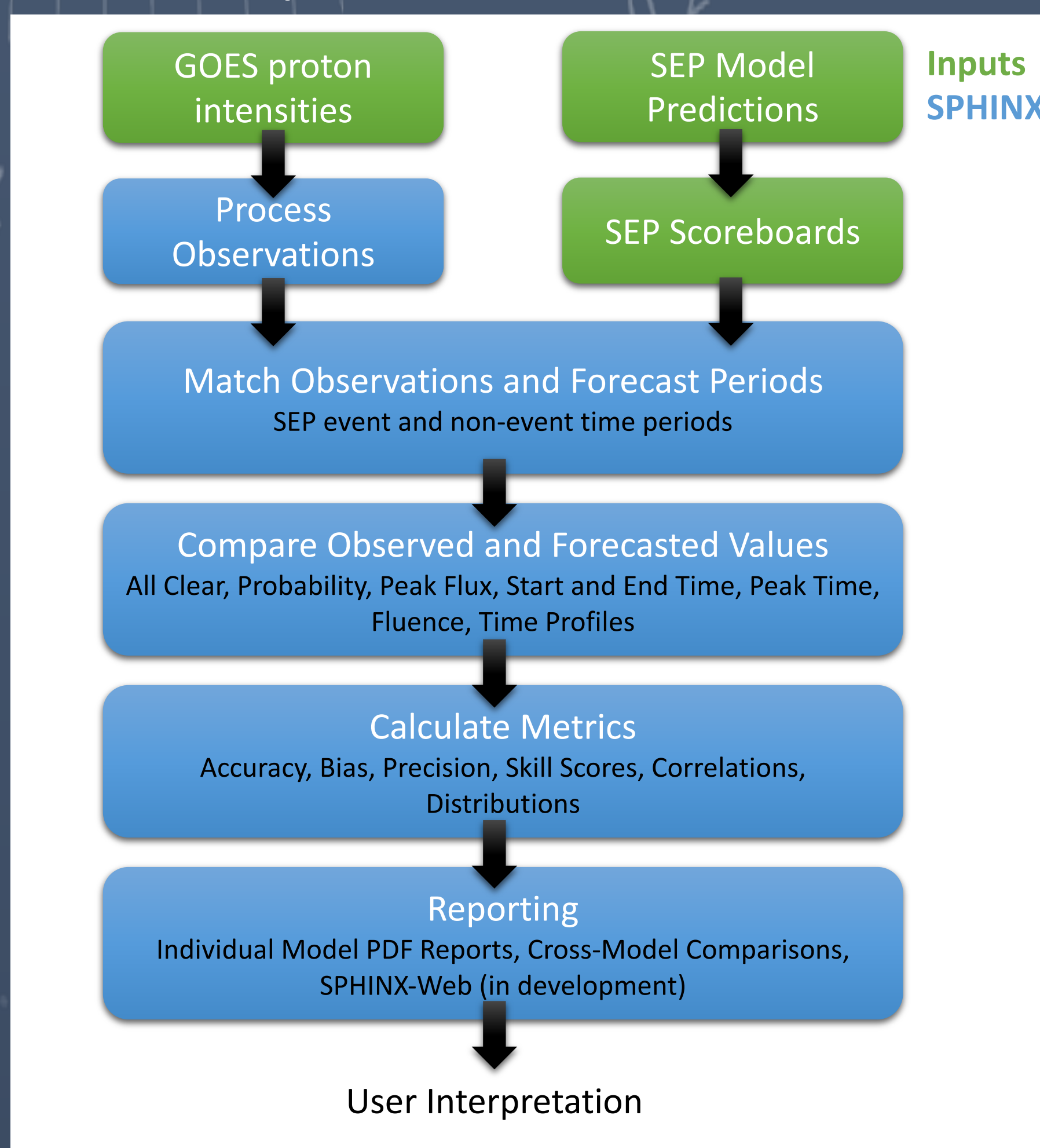
SPHINX

A gatekeeper that devours all who do not correctly answer her riddle.

Image credit: DALLE-2

Solar Particles in the Heliosphere validation INFrastructure for SpWx (SPHINX) is being developed to be a generalized, automated tool with the capability to validate any kind of forecasted quantity from any type of SEP prediction model to inform operations and science.

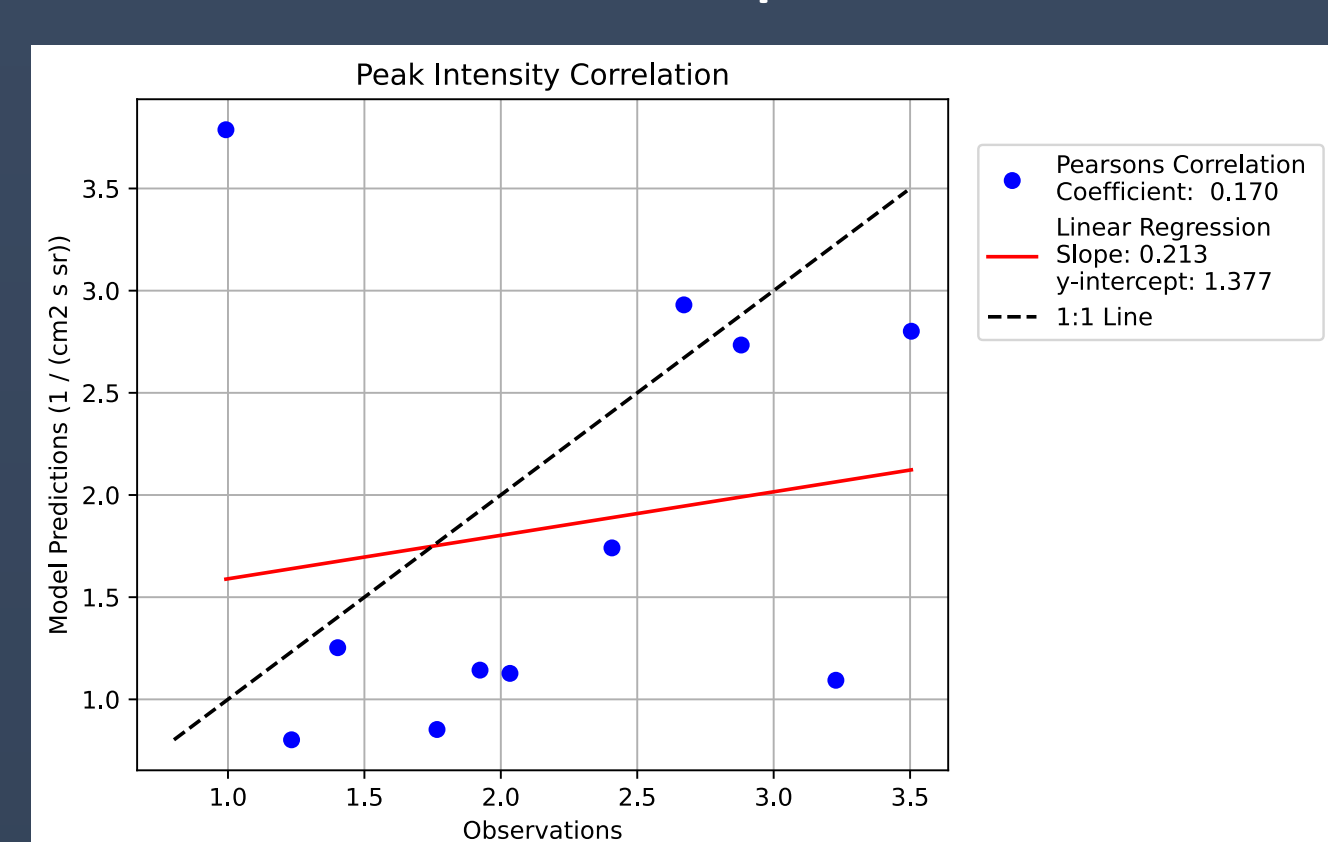
The main steps of SPHINX are:



Forecasts are matched to observed values using a completely automated and traceable process. Logic is applied using the timing of the model triggers and inputs (e.g. flare, CME, magnetogram, particle data) to associate forecasts to SEP events so that even complex sets of multiple forecasts are associated correctly. Metrics are calculated from the paired up forecasted and observed quantities.



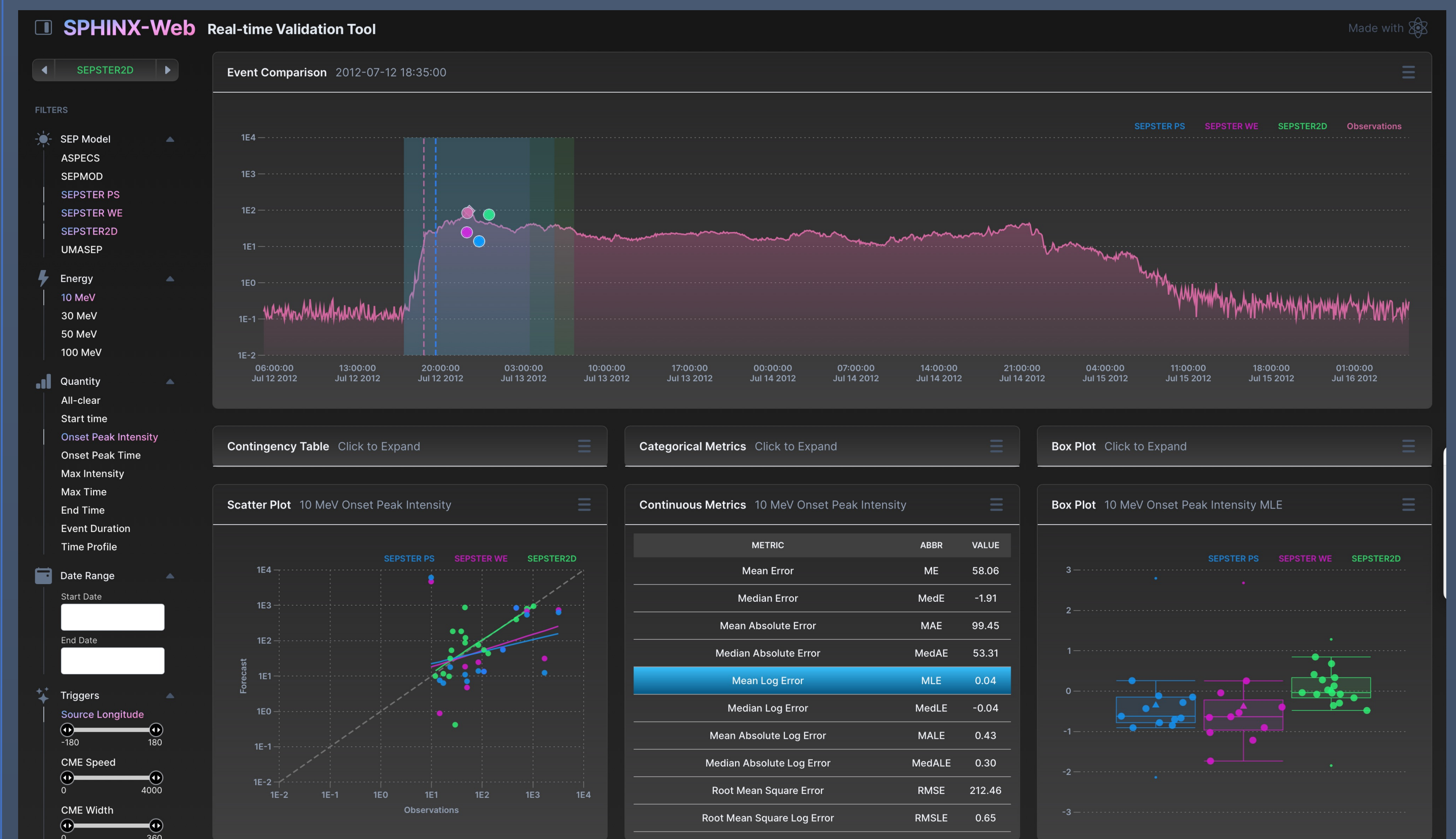
Automated matching appropriately associates forecasts to observations during complex time periods.



SPHINX generates plots and metrics for All Clear, Probability, Peak and Max Flux, Fluence, Fluence Spectrum, Timing, Time Profile, and Advanced Warning Time.

SPHINX-Web

An interactive tool to visualize validation results from SPHINX with the ability to apply filtering and recalculate metrics on the fly.



SPHINX-Web is the front-end to SPHINX and provides users with the following features:

- Visualize the validation results produced by SPHINX with interactive plots and tables
- Apply filters to the results like model, energy, quantity, date range, and model trigger
- Recalculate all metrics based on the selected filters to explore how models perform in specific scenarios
- Compare models by displaying results on the same plots
- Save plots and tables for use in publications
- (Coming soon) Run SPHINX from the browser
- (Coming soon) Load and visualize results from your own validation study
- (Coming soon) Learn additional info on how SPHINX performs validation and SPHINX-related publications

Example All Clear Metrics for SEPSTER (Parker Spiral) for 39 Forecasts in 2011 & 2012

Energy Channel: min.10.0.max.-1.0.units.MeV
Threshold: threshold.10.0.units.1 / (cm² s sr)

All Clear True Positives	All Clear False Positives	All Clear True Negatives	All Clear False Negatives	Percent Correct	Bias	Hit Rate	False Alarm Rate	Frequency of Misses	Frequency of Hits
9	2	26	2	0.90	1.00	0.82	0.07	0.18	0.82
Probability of Correct Negatives	Frequency of Correct Negatives	False Alarm Ratio	Detection Failure Ratio	Odds Threat Score Ratio	Gilbert Skill Score	True Skill Statistic	Heidke Skill Score	Odds Ratio Skill Score	Odds Ratio Skill Score
0.93	0.93	0.18	0.07	0.69	58.50	0.60	0.75	0.77	0.97

Example Metrics for All Clear Forecasts