



Northumbria
University
NEWCASTLE

A Threshold-Based Random Forest Forecasting Model For The Outer Radiation Belt

D.J.Weston¹, I.J.Rae¹, A.W.Smith¹, K.R.Murphy^{1,2}, C.E.J.Watt¹, F.X.Bocquet³, S.Bingham³, E.M.Henley³

¹Department of Mathematics, Physics and Electric Engineering, Northumbria University, Newcastle-Upon-Tyne, UK

²Department of Physics, Lakehead University, Thunder Bay, ON, Canada

³The Met Office, Exeter, UK

dylan.weston@northumbria.ac.uk

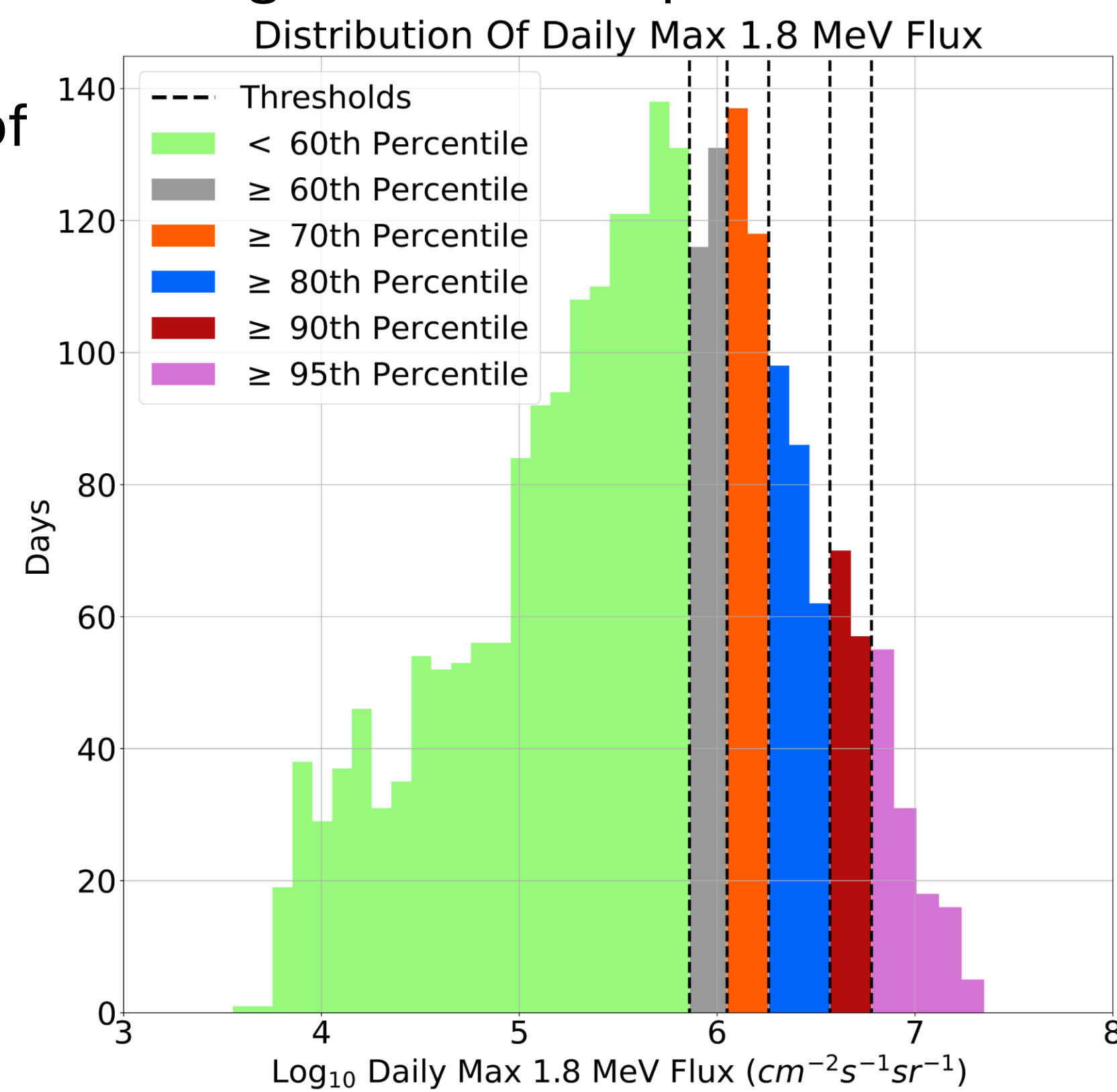
<https://github.com/D-J-Weston/RadiationBeltForecasting>



1. New Thresholds

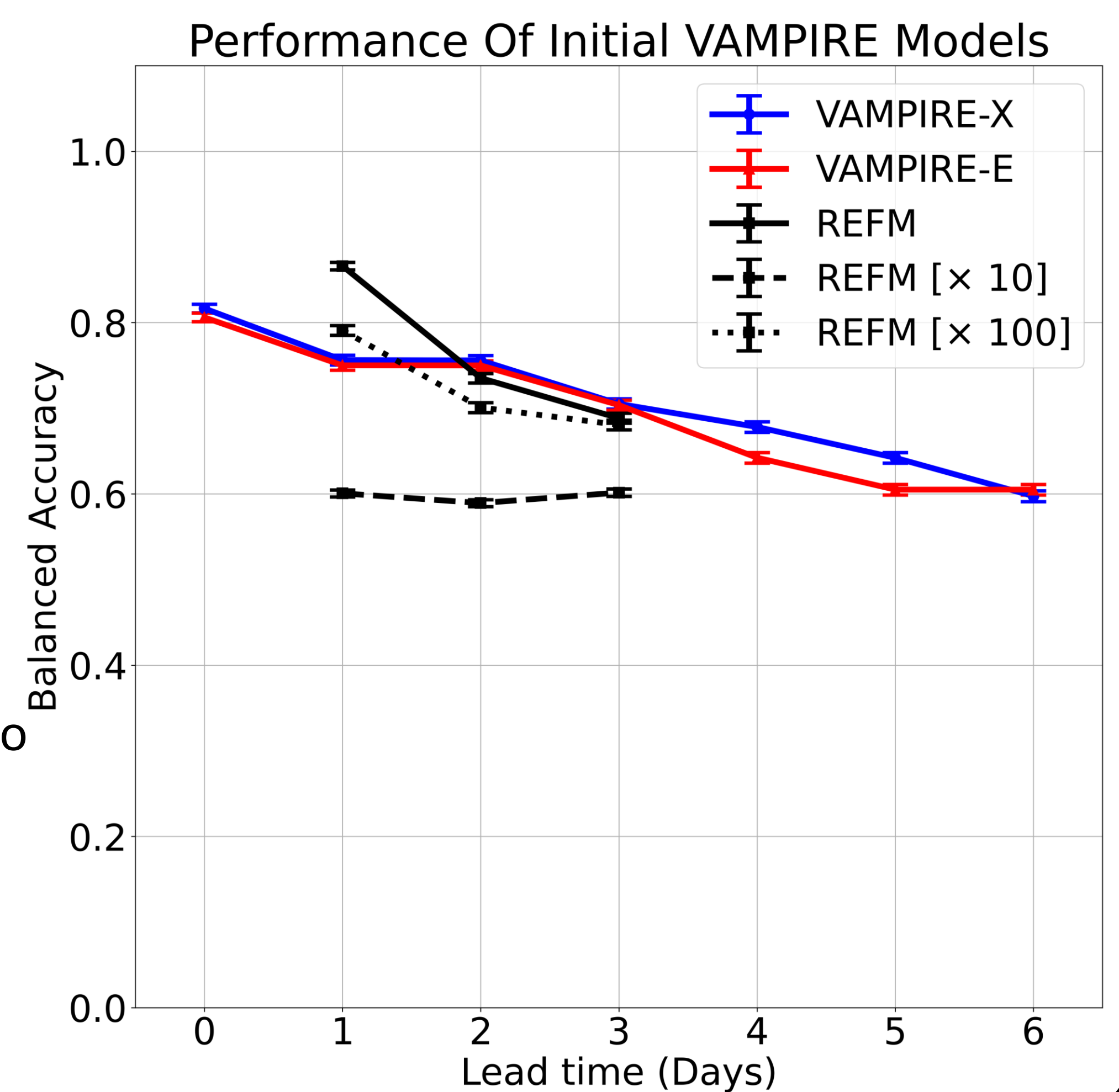
- MOSWOC and NOAA SWPC issue flux alerts when >2MeV flux at GOES exceeds 1000pfu (Forsyth et al., 2020)
- Outer Radiation Belt fluxes far higher than 1000pfu

- Aim to create a suite of models that nowcast and forecast the ORB with more realistic thresholds
- Assess model performance for
 - 60th Percentile
 - 70th Percentile
 - 80th Percentile
 - 90th Percentile
 - 95th Percentile

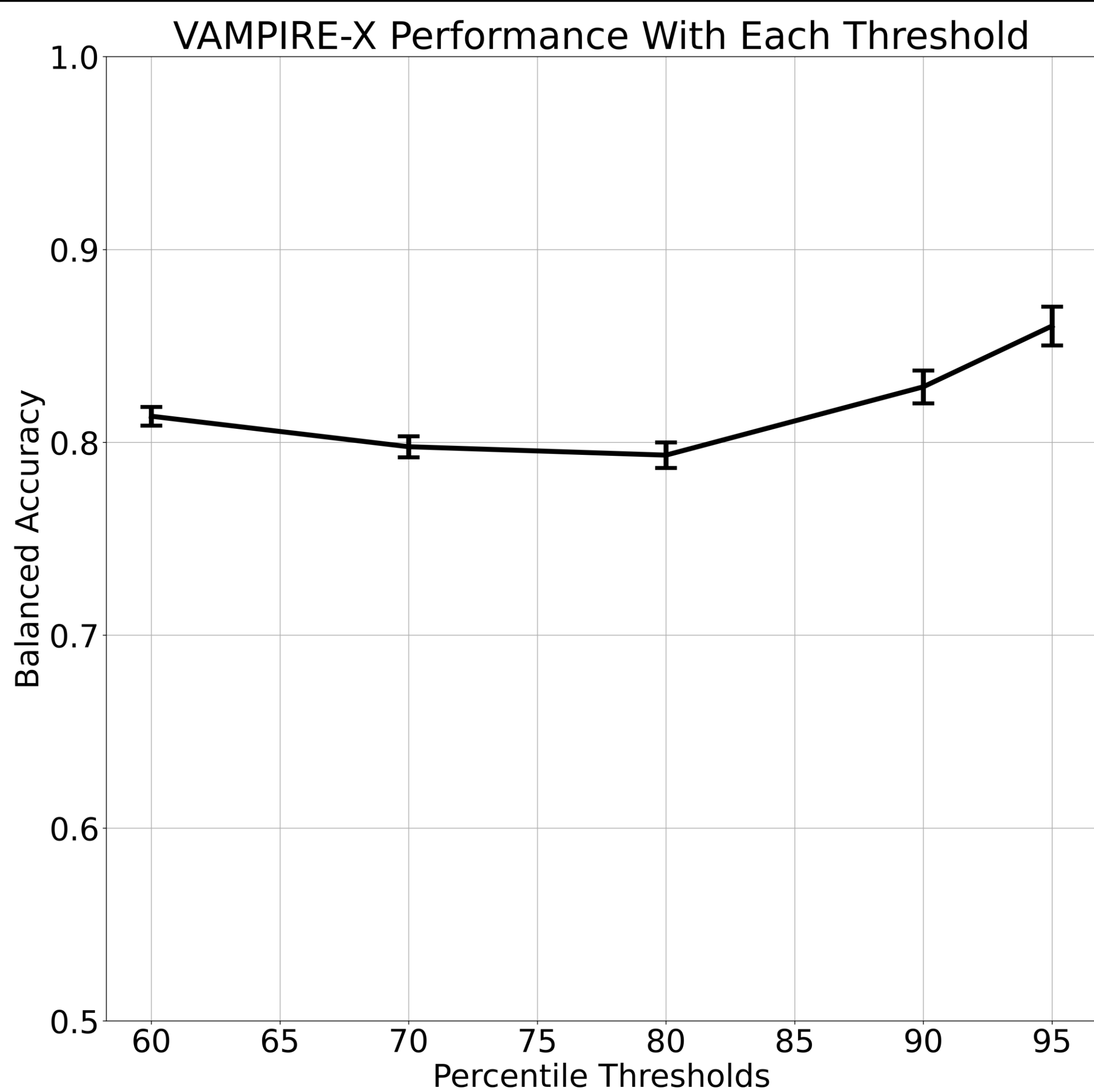


2. The VAMPIRE Model

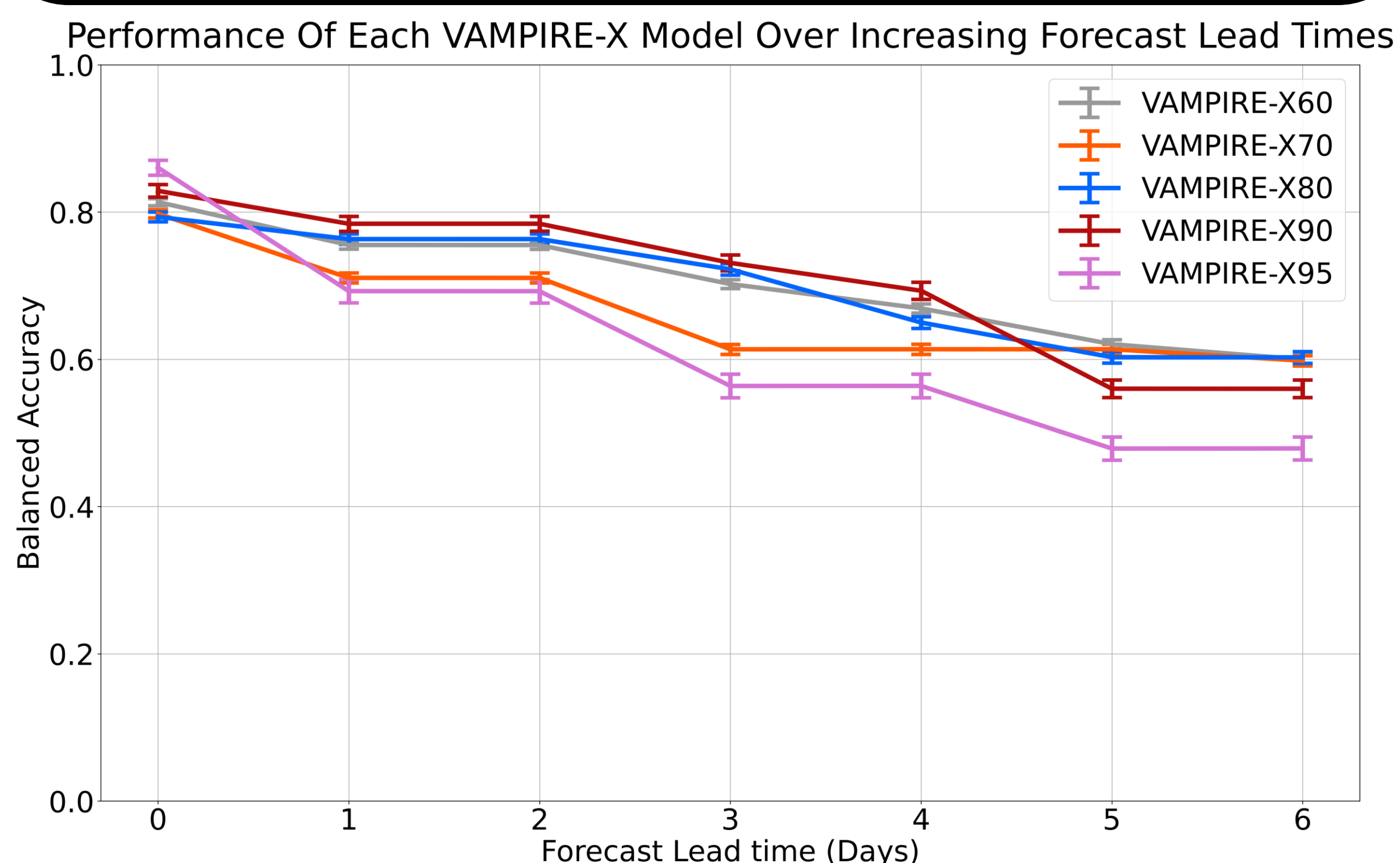
- V**an
Allen Belt
Multi-day
Predictions by
Implementing a
Random Forest for
Electrons
- Predict whether max daily flux will be above the **60th percentile** (Weston et al., 2025)
- Similar performance** to arbitrarily-scaled GEO operational REFM (Baker et al., 2019) but with **significantly extended lead times**



3. Nowcasting Performance



5. Forecasting longer lead times

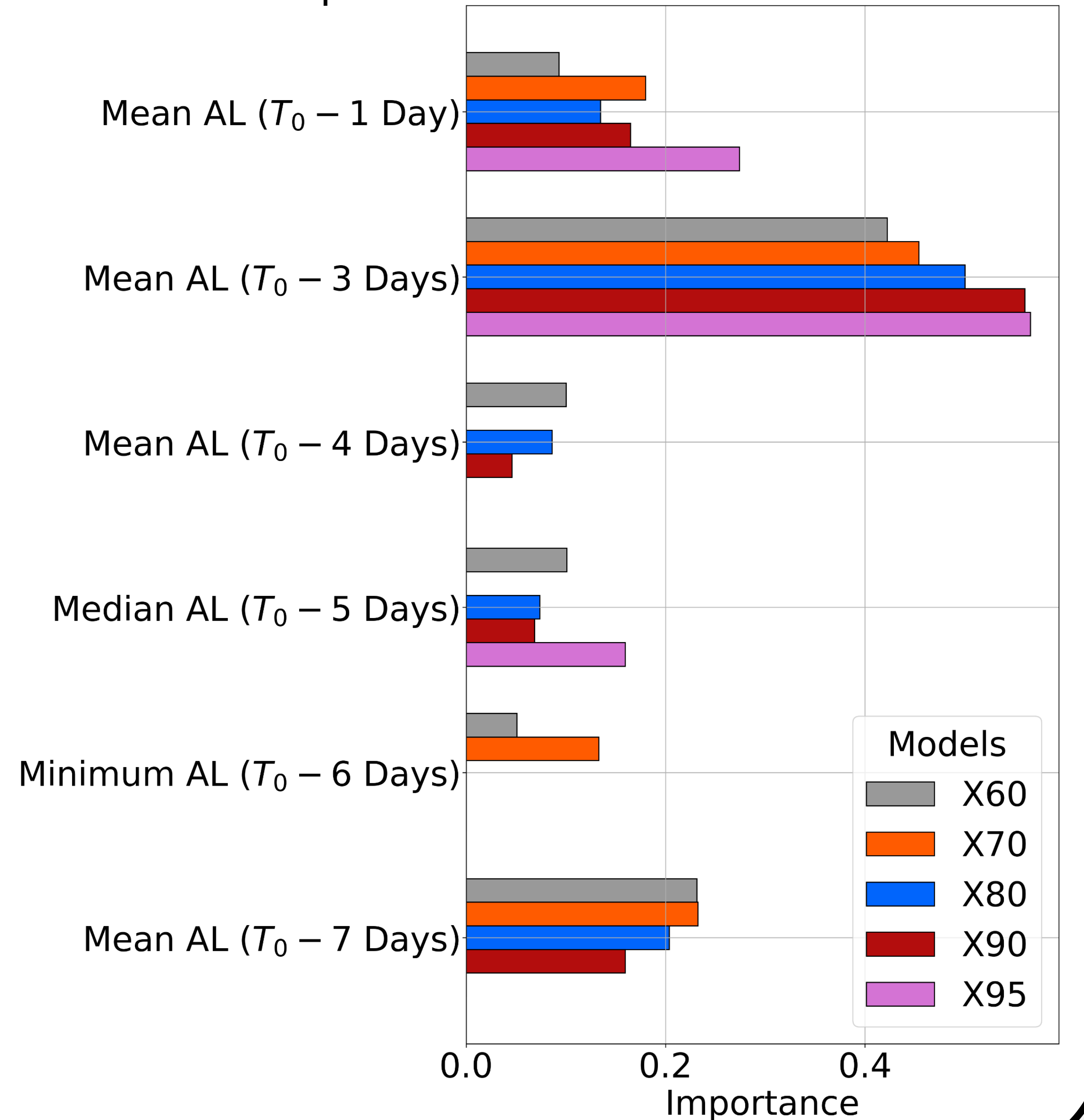


- All models display **excellent performance for ≤ 2 days** lead times
- Low errors (~0.02 or lower)
- X60, X70, X80, X90 all have some **skill for ≤ 6 days**
- X90 and X95 have lower and no skill for 5–6-day lead times, demonstrating that extreme flux spikes need a more recent time history

4. Model Inputs

- Low number of inputs** to each model (≤6)
- No solar wind or coupling functions selected; **Only AL needed** for all models
- Does not require the current state of the Belts to forecast
- AL from 3 days prior clearly the most important input
- Higher thresholds place more importance more recent information

Feature Importance For Each VAMPIRE-X Model



6. Conclusions

- VAMPIRE is a simple random forest model to forecast the state of the Outer Radiation Belt with **good levels of accuracy** over many different flux thresholds
- Each model requires **only AL as an input**, demonstrating that solar wind data is not required for accurate radiation belt forecasts
- All models show good skill up to 2 days in advance, and most up to 6 days in advance
- More extreme thresholds e.g., X95 show no skill at predicting flux for 5+ day lead times, suggesting that these thresholds are only crossed briefly and require a shorter lead time input

7. Future Work

- Split ORB forecasts by location as a function of satellite orbit or L* – currently issuing one forecast for the entire belt
- Aim to extend the lead time further than 6 days

References

PATEL, H. H. & PRAJAPATI, P. 2018. Study and Analysis of Decision Tree Based Classification Algorithms. *International Journal of Computer Sciences and Engineering*, 6, 74-78.
MAO, J. & ZHU, W. 2021. Precision-recall curve (PRC) classification trees. *Evolutionary Intelligence*, 15, 1545-1569.
FORSYTH, G., WATT, C. E. J., MOONEY, M. K., RAE, I. J., WALTON, S. D. & HORNE, R. B. 2020. Forecasting GOES 15-2 MeV Electron Fluxes From Solar Wind Data and Geomagnetic Indices. *Space Weather*, 18.
PEDREGOSA, F., GAEL, V., ALEXANDRE, G., VINCENT, M., BERTRAND, T., OLIVIER, G., MATHIEU, B., ANDREAS, M., JOEL, N., GILLES, L., PETER, P., RON, W., VINCENT, D., JAKE, V., ALEXANDRE, P., DAVID, C., MATHIEU, B. & BRUCHER, AND MATHIEU, P. & EDOUARD, D. 2018. Scikit-learn: Machine Learning in Python.
BREIMAN, L. (2001). Random Forests. *Machine Learning*, 45(1).
WESTON, D. J., RAE, I. J., SMITH, A. W., MURPHY, K. R., WATT, C. E. J., BOCQUET, F.-X., BINGHAM, S. & HENLEY, E. M. 2025. VAMPIRE: Using a Random Forest to Forecast Earth's Outer Van Allen Radiation Belt.

VAMPIRE
Pre-Print



DOI:
10.22541/essoar.175611399.91318861/v1