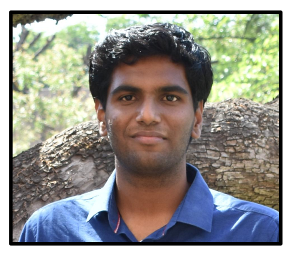
Solar wind prediction using Deep learning

Vishal Upendran[1,2], Mark C.M. Cheung[3,4], Shravan Hanasoge[5], Ganapathy Krishnamurthi[2]

[1]Inter-University Centre for Astronomy and Astrophysics, Post Bag-4, Ganeshkhind, Pune, India; [2]Department of Engineering Design, Indian Institute of Technology – Madras, Chennai, India; [3]Lockheed Martin Solar and Astrophysics Laboratory, Palo Alto, CA, USA; [4]Hansen Experimental Physics Laboratory, Stanford University, Stanford, CA, USA; [5]Department of Astronomy and Astrophysics, Tata Institute of Fundamental Research, Mumbai, India



Abstract: Emanating from the base of the Sun's corona, the solar wind fills the interplanetary medium with a magnetized stream of charged particles whose interaction with the Earth's magnetosphere has space weather consequences such as geomagnetic storms. Accurately predicting the solar wind through measurements of the spatio-temporally evolving conditions in the solar atmosphere is important but remains an unsolved problem in heliophysics and space-weather research. In this work, deep learning is used for the prediction of solar-wind speed. Extreme Ultraviolet (EUV) coronal data is used to predict solar wind speed measured at Lagrange point L1. The proposed model obtains a best fit correlation of **0.54 ± 0.04** with the wind speed . Visualization and investigation of activations of the model reveals higher activation at the coronal holes for fast wind prediction, and at the active regions for slow wind prediction, with the **higher activation at the coronal holes** being **3 to 4 days prior to prediction** for the fast wind. These trends bear an uncanny similarity to the influence of regions potentially being the sources of fast and slow wind, as reported in literature. This suggests that the our model was able to learn some of the **salient associations between coronal and solar wind structure without built-in physics knowledge**. Such an approach may help us discover hitherto unknown relationships in heliophysics data sets. (Work under review in AGU: Space Weather as Upendran et al.).

Problem

Solar wind (SW) speed prediction given Solar Coronal EUV imagery data.

Motivation: The best physics-based models give a 0.57 correlation between the hourly predicted and observed solar wind speed (Owens et al., 2008). 0.60 correlation between the prediction and observation has been obtained through linear regression fits between fraction Coronal Hole (CH) area from EUV data and solar wind speed (Rotter et al., 2015).

Idea: Use Deep learning (DL) to avoid using hand-engineered features, to see if an improvement in performance can be obtained, and if we can extract any existing/new physics out.

All about Data

Input Data: EUV 193 Å and 211 Å full disc images from the Machine learning dataset of Galvez et al., (2019), curated from the data of AIA onboard SDO (Lemen et al., 2012).

Prediction Data: Daily-averaged SW speed measured at L1. Measurements are retrieved from NASA OMNIWEB dataset.

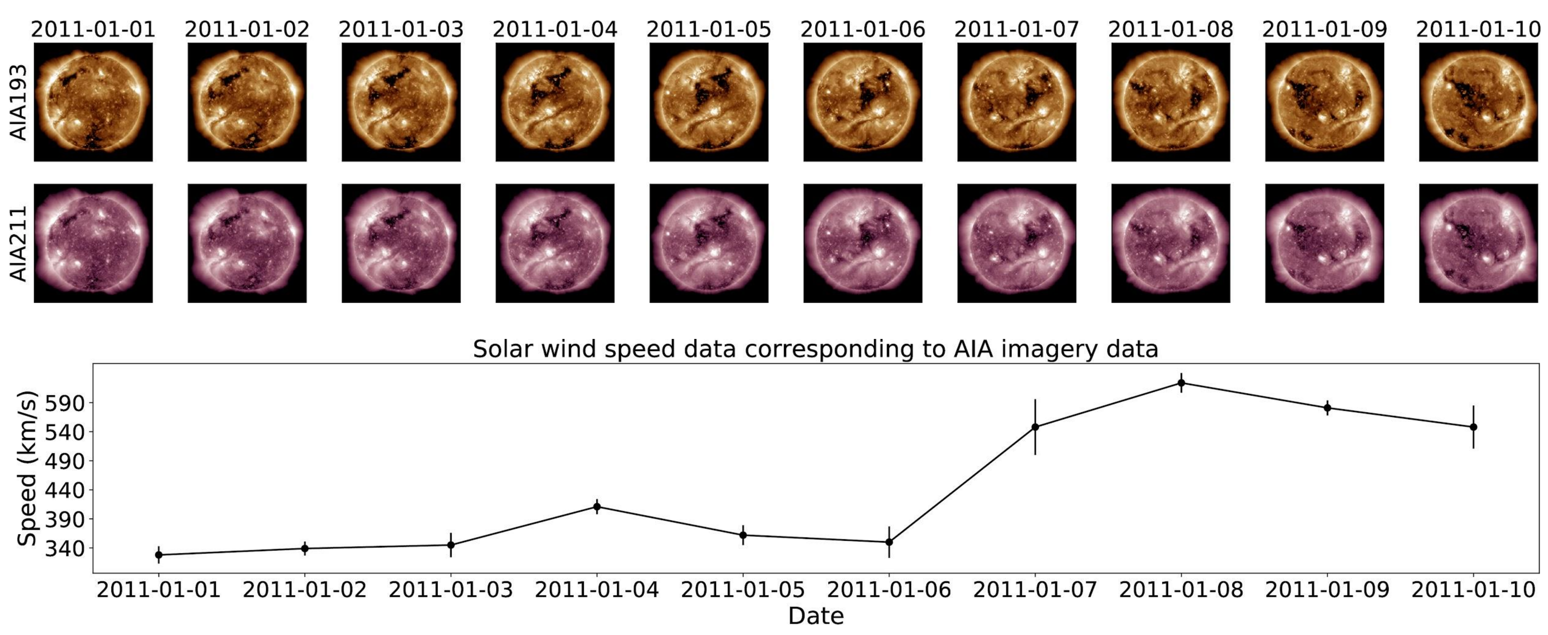


Fig 1. 10 days of AIA EUV data in 193 Å and 211 Å, and corresponding solar wind speed from OMNIWEB dataset.

Data preprocessing:

- EUV images log scaled, thresholded, resized to 224x224, and replicated into 3 channels.
- 5 fold cross validation performed.

Model architecture

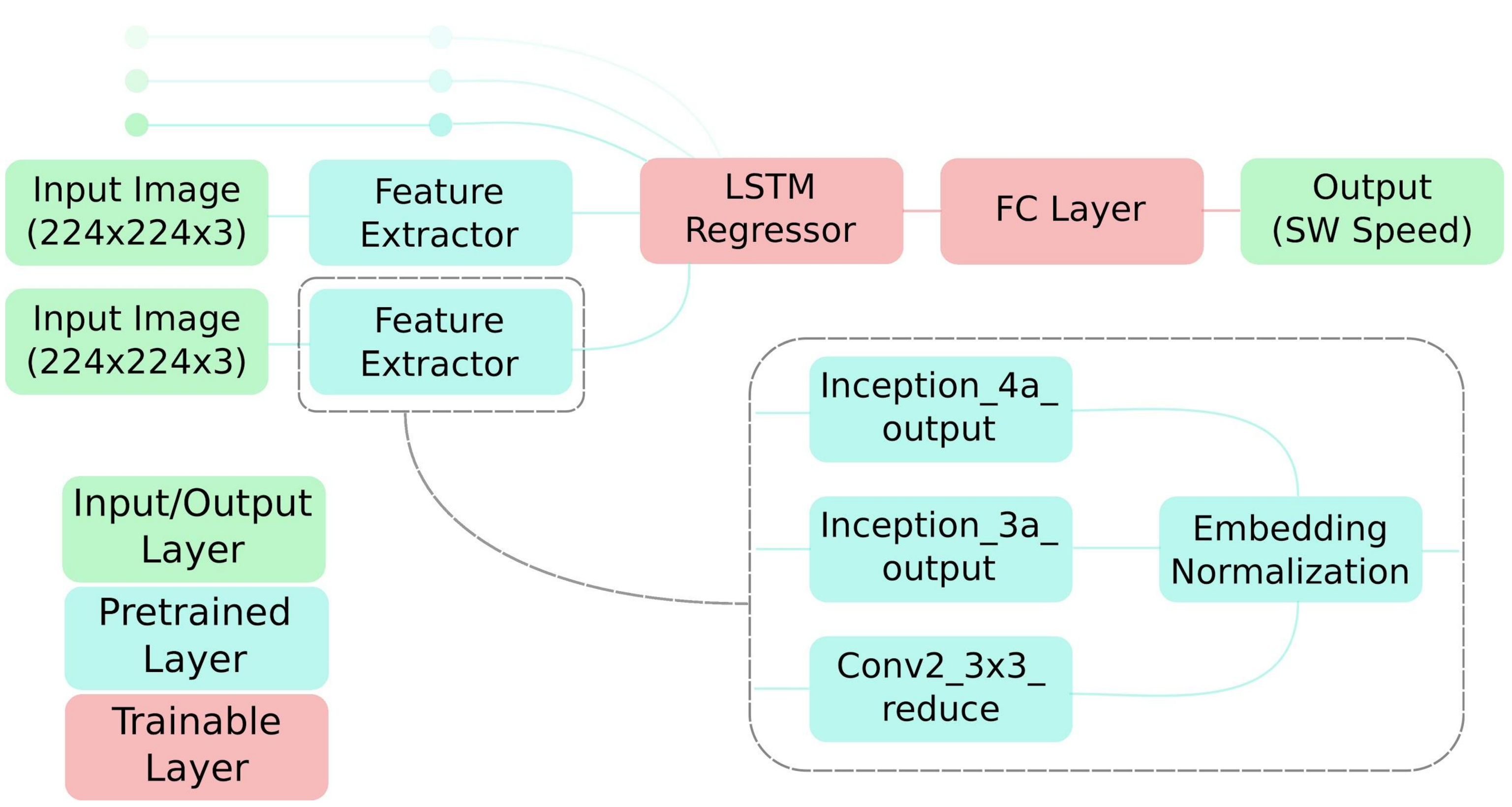


Fig 2. Model uses initial layers of a pretrained GoogleNet as feature extractor, and regresses these features against the SW speed using an LSTM.

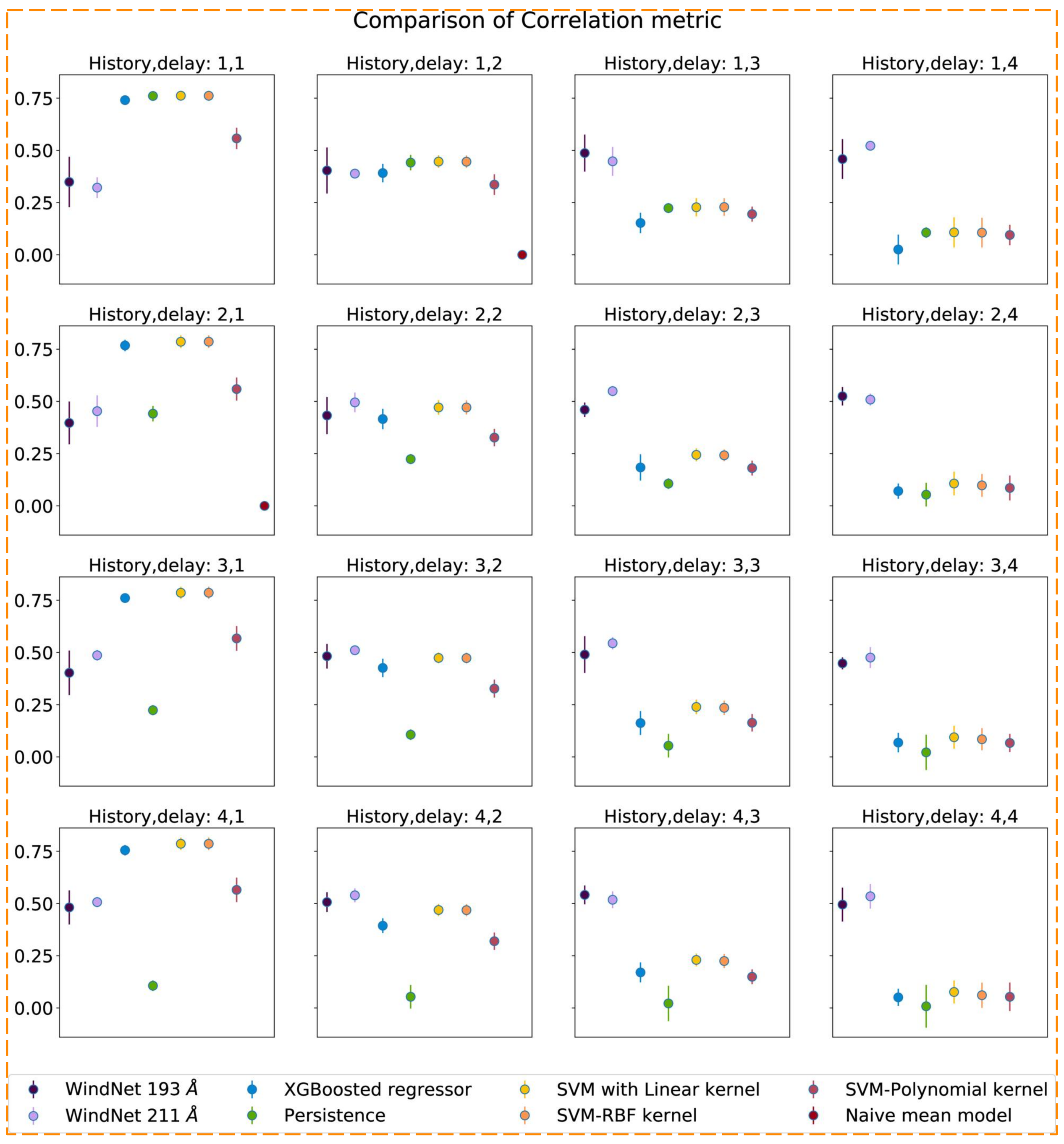
WindNet: The proposed model, called **WindNet**, consists of a Pre-trained **GoogleNet** (Szegedy et al., 2015) feature extractor with an **LSTM** (Hochreiter and Schmidhuber, 1997) regressor to SW speed. LSTM layer is trainable, and feature extractor is kept fixed.

Control parameters: History and Delay defined.

Model benchmarking: Benchmarking performed with autoregressive models using:

- XGBoost (Chen & Guestrin, 2016).
- Support Vector Regression (from Scikit-learn package, Pedregosa et al., 2011)
- Naive Mean model.
- Persistence model.

Results



- Best-fit correlation of **0.54 ± 0.04** for the 193 Å model, and **0.55 ± 0.02** for 211 Å model.
- WindNet outperforms benchmark models for delays more than 1.
- Grad-CAM (Selvaraju et al.,2017) used for Activation visualization.
- CH and Active Region (AR) segmentation using Otsu thresholding (Otsu, 1979) and Gaussian mixture model (Pedregosa et al.,2011) respectively.
- Segmentation maps with Grad-CAM maps give activation values.

Fig 3. A summary of predictive performance of proposed model WindNet against benchmarked autoregressive and naive models, in terms of Pearson correlation metric.

Activation Visualization

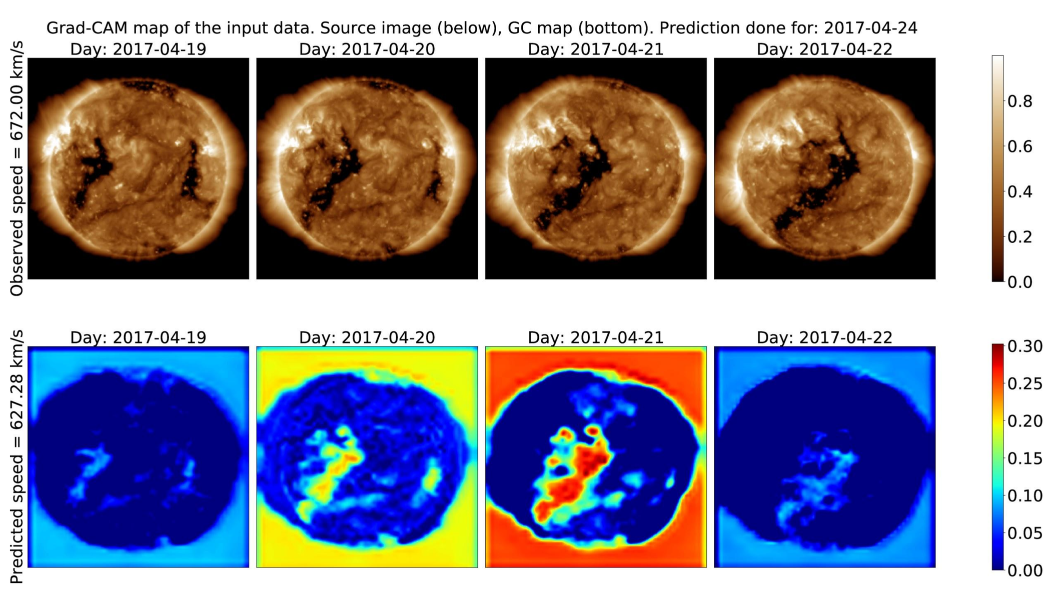


Fig 4. An example of activation of WindNet visualized using Grad-Cam map for a fast wind prediction.

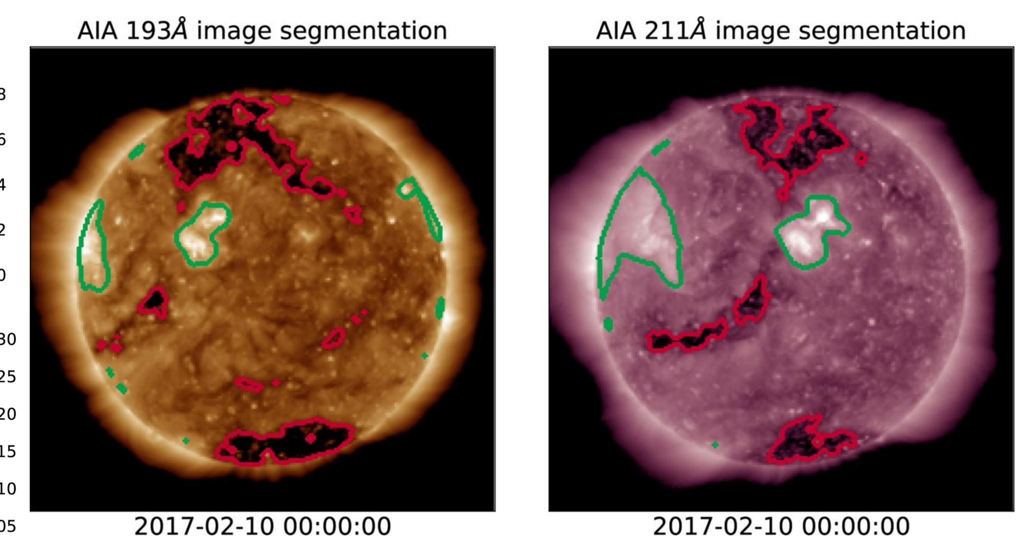


Fig 5. Example of segmented EUV images for generating AR and CH masks.

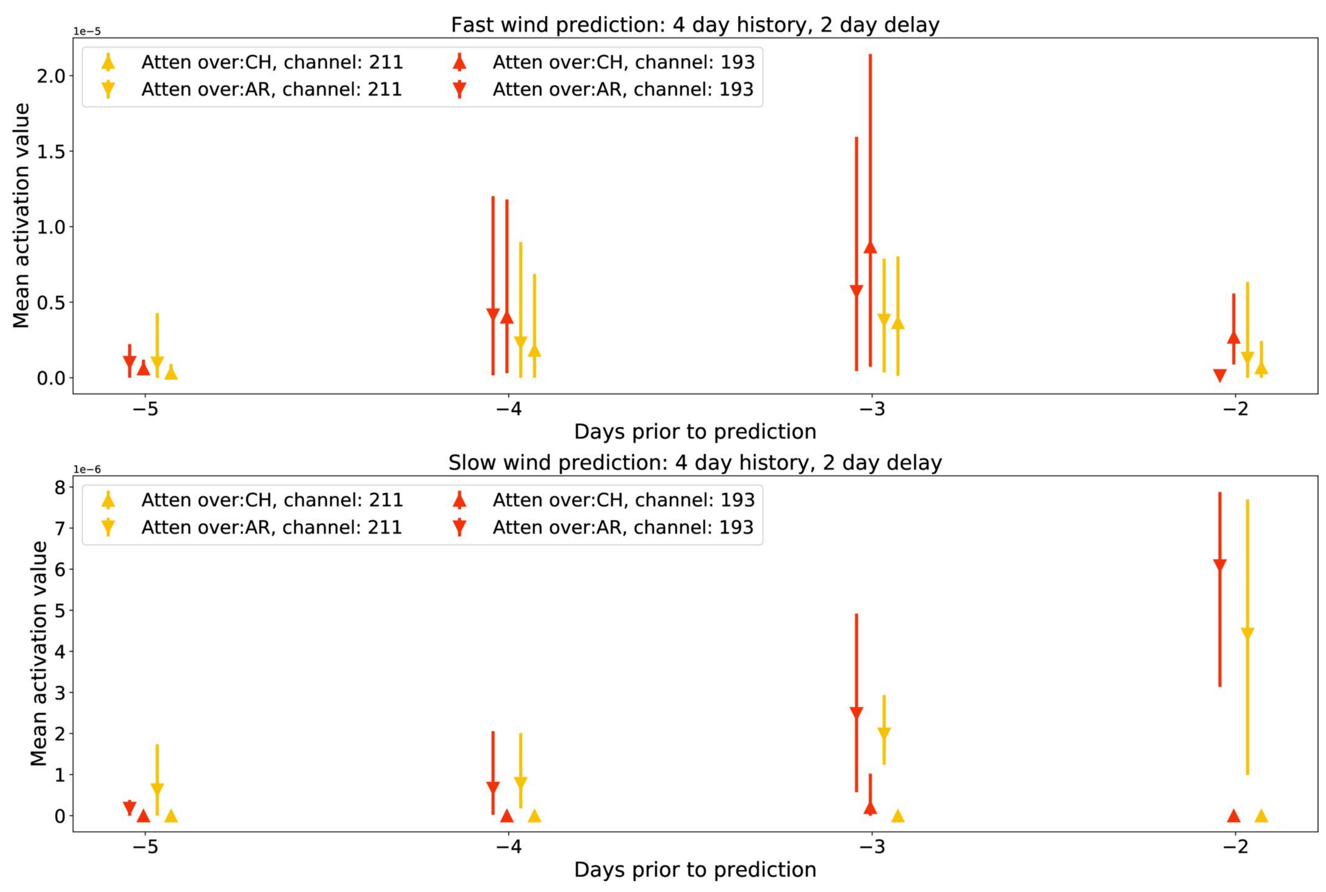


Fig 6. Mean activation value, normalized by mask area is shown with days prior to prediction. The **CHs show increased activation 3-4 days prior to prediction for a fast wind**. The bounds represent maximum and minimum activation values obtained over the cross validation dataset. The ARs alone show increased activation in case of a slow wind prediction - however, the increase occurs closer to prediction than further away.

References

- Owens, M. J. et al., *Space Weather*, 6, 2008.
- Rotter, T., Veronig, A. M., Temmer, M., & Vrˇsnak, B., Sol Phys, 290(5),1355, 2015.
- Galvez, R. et al., ApJ Supplement, 242 (1), 2019.
- Lemen, J.R. et al., Sol Phys, 275: 17, 2012.
- Szegedy, C. et al., CVPR, 2015, pp. 1-9.
- Hochreiter, S., & Schmidhuber, J., Neural Computation,9(8), 1735, 1997.
- Chen,T. and Guestrin, C. KDD16, 2016
- Pedregosa *et al.*, JMLR 12, pp. 2825-2830, 2011.
- Selvaraju, R.R. et al, ICCV 2017, 2017.
- Otsu, N., IEEE Transactions on Systems, Man, and Cybernetics, 9(1), 62, 1979.