EGU25-11587, updated on 03 Jul 2025
https://doi.org/10.5194/egusphere-egu25-11587
EGU General Assembly 2025

Solar Wind Speed Forecasting From Solar Images Using Distributional Regression

Daniel Collin[1,2], Yuri Shprits[1,3,4], Stefan Hofmeister[5], Stefano Bianco[1], Nadja Klein[6], and Guillermo Gallego[2,7]

[1]Space Physics and Space Weather, GFZ Helmholtz Centre for Geosciences, Potsdam, Germany (collin@gfz.de)
[2]Department of Electrical Engineering and Computer Science, Technical University of Berlin, Berlin, Germany
[3]Institute of Physics and Astronomy, University of Potsdam, Potsdam, Germany
[4]Department of Earth, Planetary, and Space Sciences, University of California Los Angeles, Los Angeles, USA
[5]Columbia Astrophysics Laboratory, Columbia University, New York, USA
[6]Scientific Computing Center, Karlsruhe Institute of Technology, Karlsruhe, Germany
[7]Einstein Center Digital Future, Berlin, Germany

The solar wind, a stream of charged particles originating from the Sun, poses significant risks to technology and astronauts. It is driven by large structures on the solar surface like coronal holes and active regions, which can be identified in extreme ultra-violet (EUV) solar images several days before they become geoeffective. In this work, we propose to use a distributional regression algorithm to forecast the solar wind speed at the Lagrange 1 point from solar images. Instead of predicting a single value, this method models the entire conditional distribution as a function of input features. It allows computing the uncertainty of predictions and specifying the probability of the solar wind speed exceeding certain thresholds, which is especially useful for extreme event predictions like coronal mass ejections and high-speed solar wind streams. We employ a convolutional neural network to encode solar images from multiple wavelength channels into unstructured low-dimensional representations. Using a semi-structured distributional regression approach, we couple the deep learning encoder with structured physical input parameters, such as past solar wind properties and solar cycle information. Thereby, we incorporate physical knowledge into the model and enhance explainability. We predict the solar wind speed distributions with a one-hour cadence four days in advance. We train and evaluate our method using cross-validation on 15 years of data and compare it to current state-of-the-art models. We find that it provides an accurate forecast and especially models the heavy-tailed solar wind speed distribution well. We further show the advantages over standard regression approaches and how to use the predicted conditional quantiles to improve extreme event predictions, highlighting the potential for operational space weather forecasts.