Bayesian Forecasting

by David Will, Mentor: Josh Rigler

Abstract

Our current weather models are prone to error. Not only do they fail to accurately predict current weather conditions, they fail to even enclose the true conditions within the expected limits of their error. Since these models are statistical in nature, they are associated with a probability distribution, which, for simplicity, we assume to be normal. Likewise, since the observed data is experimental, it too has a probability distribution associated with it. Combining these distributions through Bayesian statistics, and after rescaling, a new distribution can be derived, which is expected to more accurately predict the true weather conditions. Using the Wang-Sheeley-Arge model, a deterministic model, and perturbing the input to it, we produce an ensemble of data, which can be resolved into a mean and standard deviation – a time-varying probability distribution. This first step toward a Bayesian model is discussed here.

The Wang-Sheeley-Arge Model

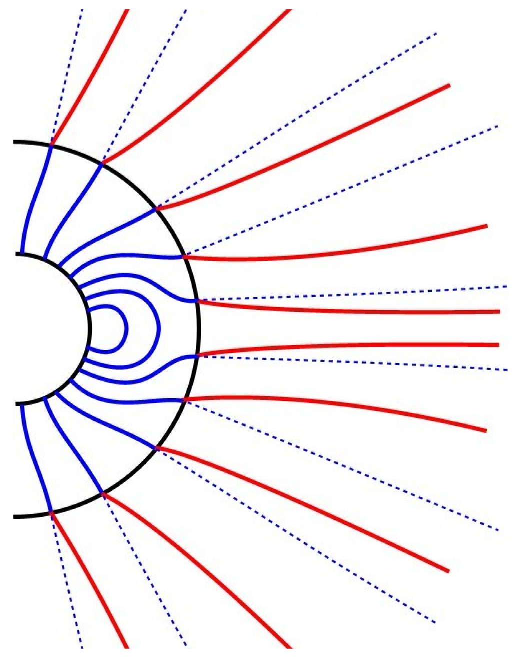


Figure 1

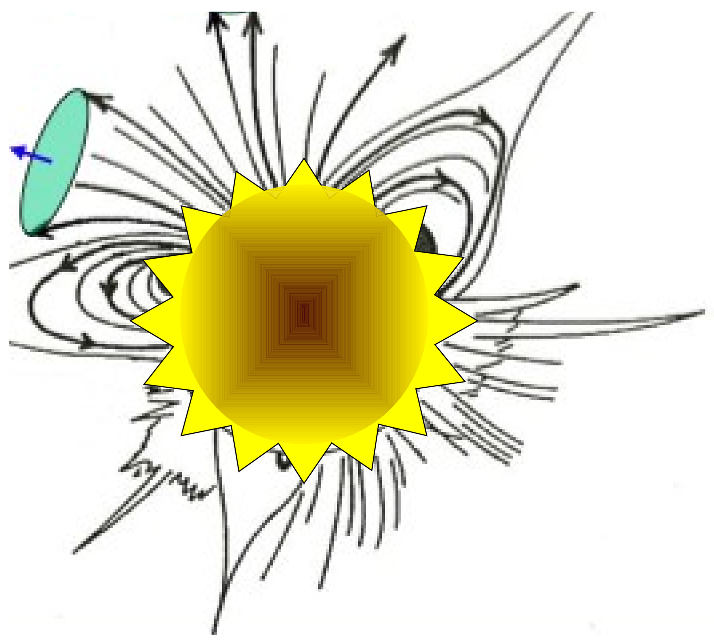


Figure 2

The WSA model takes as input magnetic field data from the sun and uses it to calculate a potential field. The sun's magnetic field behaves differently at different distances from the sun. To account for this, the WSA model calculates the magnetic field under different constraints for different regions. Briefly, the model uses one set of constraints to construct a potential field at distances less than 2.5 solar radii from the photosphere and another to calculate a potential field betweeen 2.5 and 5 solar radii. Near the sun's surface the model assumes that the magnetic field is predominantly radial and produces a field based on the magnetic input. The second region is a new improvement to the model which adjusts the field lines to give the magnetic field more of a current sheet topology. Figure 1 demonstrates the effect of this change. Beyond 5 solar radii the model doesn't affect the magnetic field, it simply continues on outward forever at whatever value it had when the model decided to stop fiddling with it.

Ensemble Wind Predictions

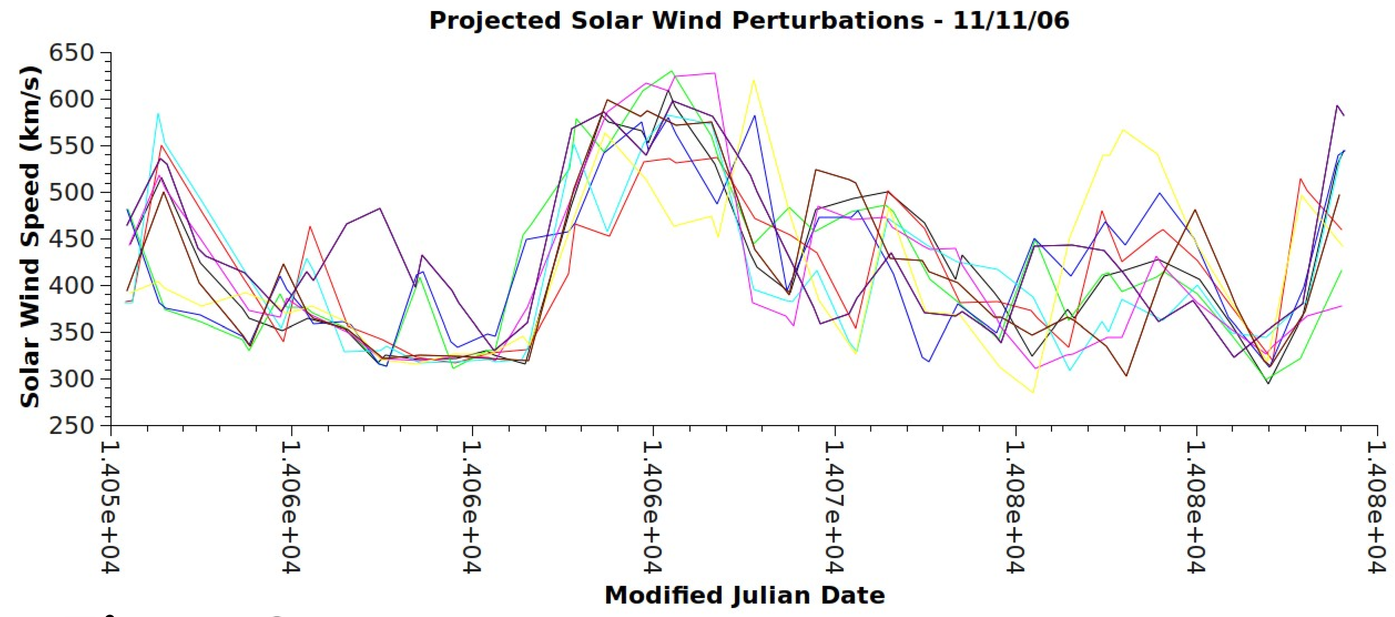


Figure 3

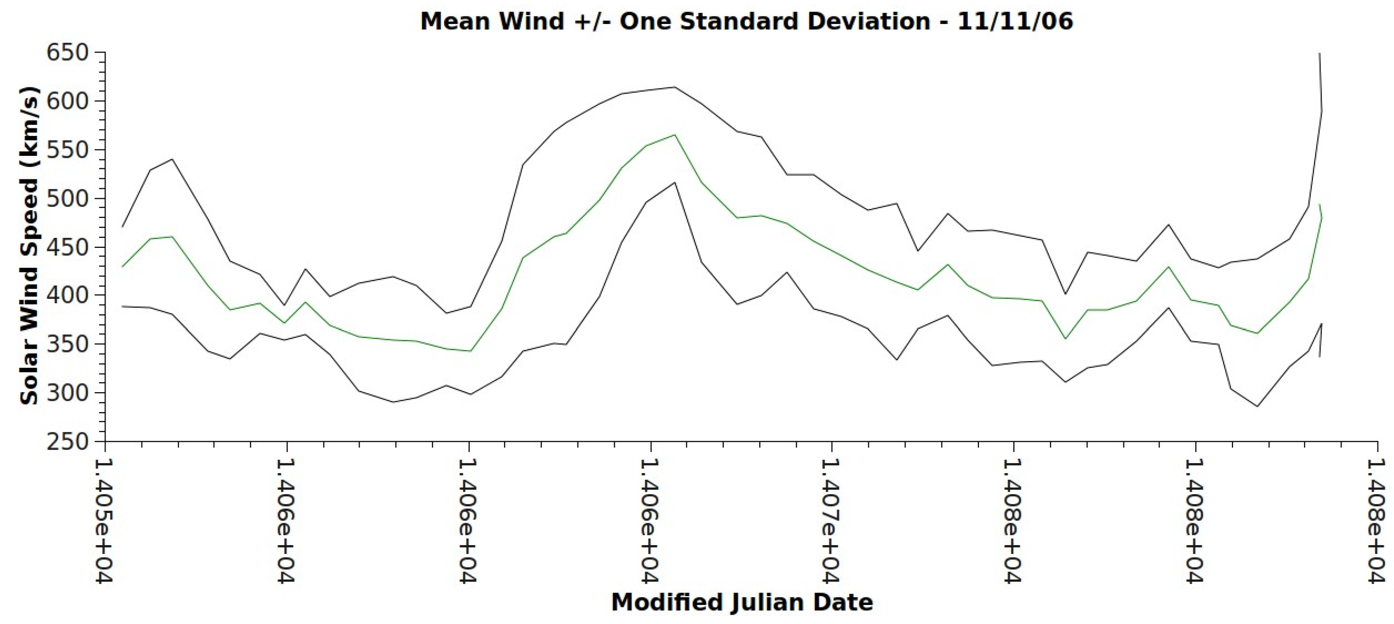


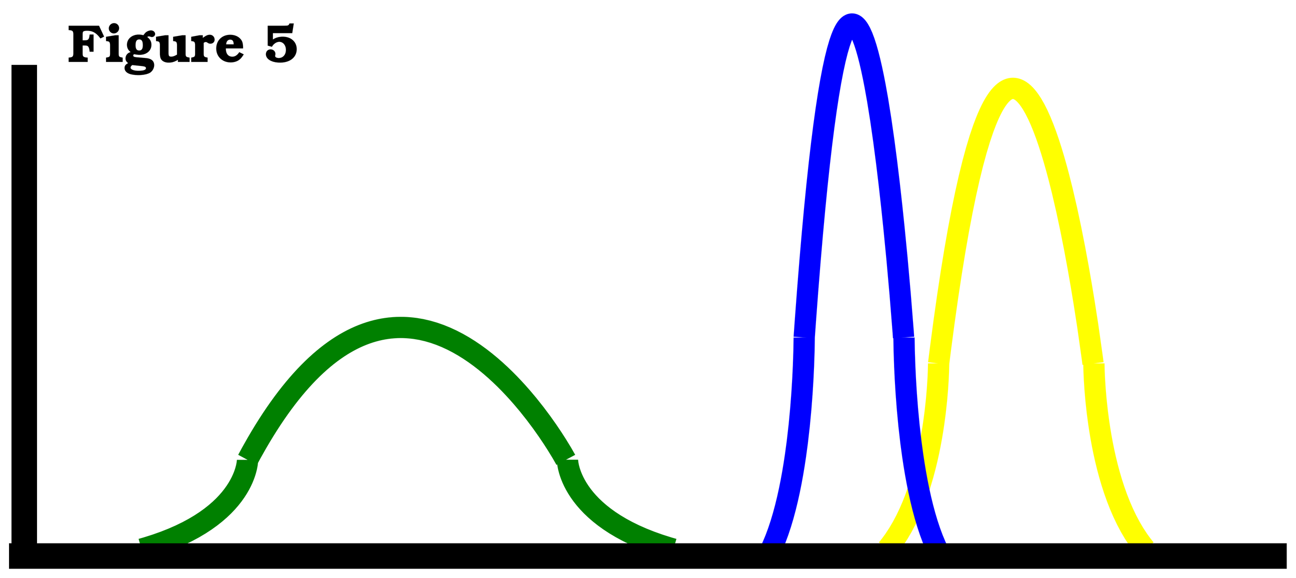
Figure 4

The wind predicting routine uses whatever data it receives to guide its predictions. So in our case we had data for each day from 11/11/06 to 12/08/06. If ever there is more than a day's lapse between input, the routine generates a guess for that day. If more than seven days pass without new input, the routine stops since the predictions start to become unreliable as divergence properties begin to dominate. Hence, these graphs show wind speed predictions at Earth between 11/11/06 and 12/15/06. Wind speed is plotted on the vertical, time (as a modified Julian date) on the horizontal.

Bayesian Statistics

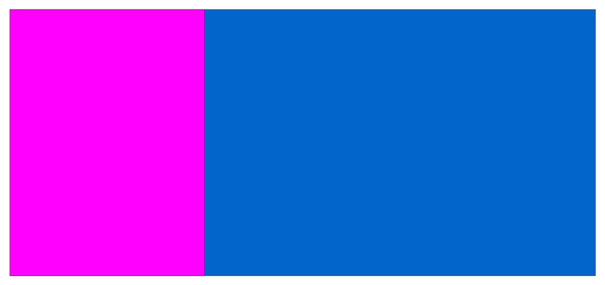
if we take a look at any snapshot in time of Figure 3, we would see a distribution for the model data, and since observations are experimental data, they have errors associated with them as well and so can also be expressed as a distribution. In Figure 5 below, the model is arbitrarily represented by the green curve, while the yellow curve is meant to represent some sort of actual observation. The blue curve is derived by applying Bayes rule to the green and yellow distributions and rescaling.

Figure 5

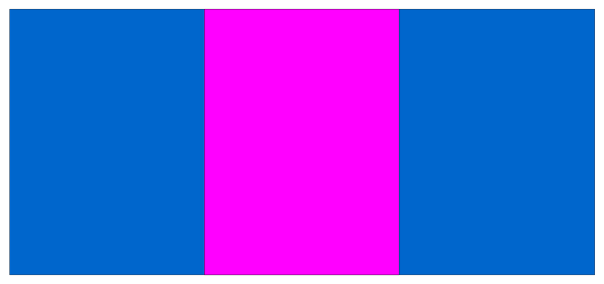


Perturbation Schemes

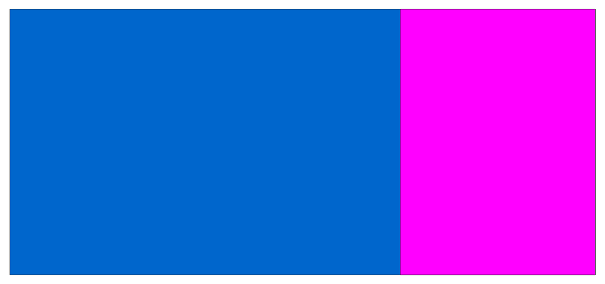
Original Magnetogram



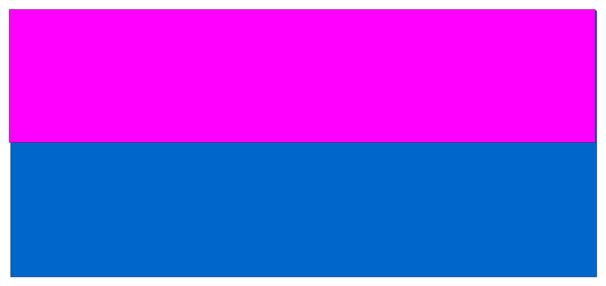
“Left” Longitude Perturbation



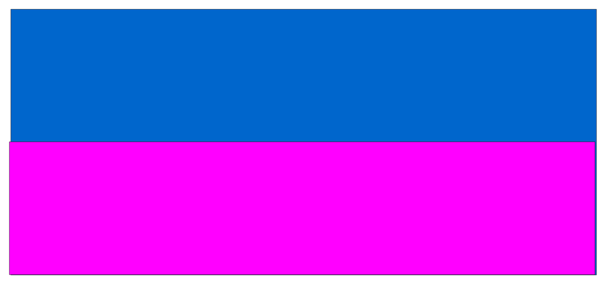
“Middle” Longitude Perturbation



“Right” Longitude Perturbation



Nothern Hemisphere Perturbation



Southern Hemisphere Perturbation

Each magnetogram contains magnetic field data for the sun. Longitude is plotted horizonally. Latitude is plotted vertically. In the diagrams below, a blue region indcates unchanged; that is, anything blue has not been modified from the original version of the magnetogram. The pink indicates a region that has either been increased to twice its original magnetic field strength or reduced to half its original magnetic field strength.

Future Work

The work up to this point comprises only the first step of the project. Now that we have a mean and standard deviation for the model data, it needs to be compared against observation. As more perturbation schemes are run, a greater knowledge of the model's covariance matrix can be achieved. As more obversed data is analyzed, its covariance matrix will also become more comprehensible. Finally, once all this is accomplished, a Bayesian model for predicting the solar wind can be produced.

Acknowledgements

My deepest thanks go out to Josh Rigler, who has mentored and worked tirelessly and passionately to guide this project, Leslie Mayer for her help with the WSA code, Nick Arge for graciously allowing us to use the WSA model, and to the National Science Foundation for funding this research.