

The Role of Empirical Space-Weather Models (in a World of Physics-Based Numerical Simulations)

Mathew J. Owens¹ Pete Riley² and Tim Horbury³

¹Space and Atmospheric Electricity Group, Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading RG6 6BB, UK

email: m.j.owens@reading.ac.uk

²Predictive Science Inc., 9990 Mesa Rim Rd, Suite 170, San Diego, CA 92121, USA

email: pete@predsci.com

³Blackett Laboratory, Imperial College London, London SW7 2BZ, UK

email: t.horbury@imperial.ac.uk

Abstract. Advanced forecasting of space weather requires prediction of near-Earth solar-wind conditions on the basis of remote solar observations. This is typically achieved using numerical magnetohydrodynamic models initiated by photospheric magnetic field observations. The accuracy of such forecasts is being continually improved through better numerics, better determination of the boundary conditions and better representation of the underlying physical processes. Thus it is not unreasonable to conclude that simple, empirical solar-wind forecasts have been rendered obsolete. However, empirical models arguably have more to contribute now than ever before. In addition to providing quick, cheap, independent forecasts, simple empirical models aid in numerical model validation and verification, and add value to numerical model forecasts through parameterization, uncertainty estimation and ‘downscaling’ of sub-grid processes.

Keywords. solar wind, solar-terrestrial relations

1. Introduction

State-of-the-art space-weather forecasting typically involves a chain of coupled numerical models with domain-specific approximations to the relevant physics (Merkin *et al.* 2006; Toth *et al.* 2005). Photospheric magnetic field observations are used to constrain a coronal model (e.g., Mikic *et al.* 1999), which then drives a magnetohydrodynamical (MHD) heliospheric model (e.g., Odstrcil, 2003; Riley *et al.* 2001), out to 1 AU. Note that even in full MHD schemes, semi-empirical methods are still necessary to relate coronal structure and solar wind speed (see section 4.1). Time-dependent phenomena, such as coronal mass ejections (CMEs), are incorporated in an ad-hoc manner (Odstrcil *et al.* 2004; Lionello *et al.* 2013). Forecast lead times are generally on the order of the solar wind Sun-Earth transit time, approximately 2-5 days. Advances in computing power and code efficiency have seen a steady increase in the spatial and temporal resolution of models. Similarly, better understanding of the physics of the corona and solar wind has led to better a representation of physical processes within models.

An alternative to the first-principles, physics-based approach is the use of statistical methods and empirical models. Despite their relative simplicity, such models can often approximate physical features and processes that the numerical simulations have difficulty with. These empirical models provide quick, cheap, independent forecasts and aid in state-of-the-art model validation and verification. Most importantly, they add value to numerical model forecasts through parameterization of missing or poorly represented

physical processes, uncertainty estimation and ‘downscaling’ of sub-grid and stochastic processes. Using examples from near-Earth solar wind forecasting, these concepts are illustrated in the following sections.

2. Providing independent forecasts

There are a number of near-Earth solar wind prediction methods that can provide independent forecasts to complement the state-of-the-art numerical forecasts. In-ecliptic solar wind is particularly difficult to forecast due to the dominance of slow solar wind, which is poorly understood relative to the fast polar wind (Cranmer *et al.* 2017).

The simplest possible forecast, if it can really be called that, is climatology, wherein the value of a given solar wind parameter at all times in the future is simply forecast to be the mean value of that parameter over some long-term period in the past (say, 1965-2016). Thus unlike other forecasts, the predictive capability of climatology is independent of forecast lead time. For near-Earth solar wind speed $[V_{SW}]$, the climatological mean (and therefore forecast) is 433 km s^{-1} , whereas for the out-of-ecliptic heliospheric magnetic field component $[B_Z]$, it is close to 0 nT . Using simple point-by-point metrics, such as correlation or mean-square error, to compute forecast ability relative to observations, climatology often compares favorably with state-of-the-art numerical forecasts (e.g., Owens *et al.* 2008). In reality, this is often the result of numerical models suffering ‘double penalties’, whereby small timing errors in the forecast of a large event results in both a forecast miss and a forecast false alarm, whereas climatology only records a miss (see Figure 4 of Owens *et al.* 2005). Clearly, climatology is not useful as an ‘actionable’ forecast. It can, however, be valuable for long-term planning, such as setting design requirements for space hardware.

A slightly more useful forecast is persistence, whereby the solar wind in the future is assumed to be the same as that at the forecast time. Clearly, the performance of a persistence forecast reduces with forecast lead time, but speed of this degradation varies greatly with different solar wind parameters. For V_{SW} , persistence provides a more accurate forecast than climatology out to a lead time of around a day. For B_Z , however, persistence falls behind climatology at forecast lead times beyond an hour (see Figure 3 of Owens *et al.* 2017). But perhaps the simplest useful forecast comes from exploiting the recurrent nature of the solar wind (Bartels, 1934). All solar wind parameters display autocorrelation at a lag of around 27 days, the synodic rotation period of the Sun, as well as its harmonics with decreasing amplitude (Owens *et al.* 2013). Thus a recurrent forecast (sometimes also called 27-day persistence) predicts the solar wind in 27 days time will be the same as that today. It outperforms climatology for V_{SW} , but not B_Z .

The recurrent forecast is valuable for a number of reasons. First, it is entirely independent of the state-of-the-art forecasts, providing forecast robustness when used in conjunction. Second, the observational input to recurrence is *in situ* solar wind observations, rather than photospheric magnetograms, providing forecast resilience to data availability. Third, as the forecast uses recent solar wind observations to predict future solar wind observations, there is almost zero forecast bias, such as, e.g., a systematic underestimate of the magnetic field strength. Fourth, recurrence offers forecast lead times well in excess of the 2-5 days provided by numerical simulations.

A further level of empirical forecast sophistication can be achieved using so-called ‘analogue forecasts’, whereby past solar wind observations are a good analogue of the future. Past analogues are defined as times in historic observations when the solar wind was similar to, say, the last 24 hours. The analogue periods can then be used to make a probabilistic forecast of the future (Riley *et al.* 2017; Owens *et al.* 2017). In addition

to the benefits of recurrence forecasts, analogue forecasts are also much more accurate, outperforming climatology for forecast lead times of many tens of days for V_{SW} , and even around 10 hours for B_Z . The same analogue forecasting method can be applied to centennial and millennial reconstructions of the heliospheric magnetic field, enabling forecasting of space climate decades into the future (Barnard *et al.* 2011).

3. Model validation and verification

It is vital for both model development and model selection that forecasts are accurately verified and validated. In isolation, simple metrics, like correlation or mean-square error, are rather meaningless; they are unable to distinguish between periods where there is a genuine problem with the forecast and periods where the solar wind is simply unpredictable. By comparing the performance with a baseline forecast, however, this distinction can be made. Climatology can provide a baseline, as was implicitly done in the discussion in the previous section, but recurrence or analogue forecasts provide a more appropriate comparison, as they contain realistic solar wind structure (Siscoe *et al.* 2004).

4. Adding value to state-of-the-art numerical forecasts

Perhaps most importantly, empirical models can be used to add value to state-of-the-art numerical forecasts. There are three broad ways in which this is achieved.

4.1. *Parameterization of missing physics*

The MHD approximation, computational limits in the spatial and temporal scales, and limitations in observational boundary conditions means some physical processes are not (well) captured by numerical simulations. Of particular relevance to near-Earth solar wind forecasting, the heating of the corona and subsequent solar wind acceleration does not produce sufficient contrast between fast and slow wind regimes (Riley *et al.* 2015). To enable correct inputs to solar wind models, this process is parameterized in coronal models using empirically-derived relations to coronal magnetic field topology (Wang & Sheeley, 1990; Riley *et al.* 2015). In addition to enabling more accurate solar wind forecasts, refinement of these empirical relations also aids scientific understanding of coronal heating and thus can potentially inform a first-principles model of solar wind acceleration.

4.2. *Uncertainty estimation*

Numerical models are typically used to provide a single, deterministic forecast of solar wind conditions in near-Earth space. Use of numerical model ensembles enables a measure of forecast uncertainty, but due to computational expense ensembles are typically limited to a few tens of members. Empirical methods for propagating the solar wind from Sun to Earth are not quite as accurate as the full 3-dimensional MHD solution (Riley & Lionello, 2011), but they are much less computationally expensive. Thus these empirical methods allow thousands of quasi-random samples of the near-Sun solar wind to be propagated to Earth. The variation in the resulting ensemble provides an accurate measure of uncertainty in the deterministic MHD forecast (Owens & Riley, 2017).

4.3. *“Downscaling” of sub-grid and stochastic processes*

Finally, empirical models are also used to compensate for limited spatial and temporal resolution of numerical models. The Earth’s magnetosphere responds to solar wind variations at time scales far below those currently reproduced by numerical solar wind models (Merkin *et al.* 2007). Future improvements to spatial and temporal resolution of

models will not resolve this issue, as much of the high frequency variations in the solar wind are the result of stochastic processes such as turbulence. Thus a stochastic solution is required. Using empirical relationships between solar wind high-frequency variations and bulk properties, solar wind model results can be “downscaled” to include noise in a statistical sense prior to their use in driving magnetospheric simulations. This provides a more realistic representation of the magnetosphere (Owens *et al.* 2014).

5. Summary

Physics-based numerical models and simple, empirical models are not an either/or proposition; they both have qualities which contribute to space-weather forecasting. As independent forecasts, empirical models add robustness to state-of-the-art methods, and can enable forecast lead times not achievable by numerical techniques. By relying on different observational inputs, empirical models produce resilience to interruptions in the observational data pipeline. Model validation/verification is aided by empirical models, as they provide a useful baseline against which numerical forecasts can be compared.

Perhaps most importantly, empirical models can add value to physics-based, numerical models in areas that are difficult to address from a first-principles approach. Empirical models can be used to parameterise physical processes that are not well represented in numerical models, to incorporate sub-grid and statistical processes, and to provide an accurate assessment of model uncertainty in a computationally inexpensive way.

6. Acknowledgements

M.O. is funded by Science and Technology Facilities Council (STFC) grant ST/M000885/1. We have benefited from discussions as part of the NASA LWS funded ProjectZed (NNX15AF39G).

References

- Barnard, L., *et al.* 2011, *Geophys. Res. Lett.*, 381, doi:10.1029/2011GL048489
- Bartels, J. 1934, *Terr. Magn. Atmos. Electr.*, 39, 201, doi:10.1029/TE039i003p00201
- Cranmer, S., Gibson, S. & Riley, P. 2017, *Space Sci. Rev.*, in press
- Lionello, R., *et al.* 2013, *Astrophys. J.*, 777, 1, 76
- Merkin, V., *et al.* 2007, *Space Weather*, 5, S12001, doi:10.1029/2007SW000335
- Mikic, Z., Linker, J. A., Schnack, D. D., Lionello, R. & Tarditi, A. 1999, *Phys. Plasma*, 6, 2217
- Odstrcil, D. 2003, *Adv. Space Res.*, 32, 497-506
- Odstrcil, D., Riley, P. & Zhao, X. P. 2004, *J. Geophys. Res.*, 109, doi:10.1029/2003JA010135
- Owens, M., *et al.* 2005, *J. Geophys. Res.*, 110, doi:10.1029/2005JA011343
- Owens, M., *et al.* 2008, *Space Weather*, 6, doi:10.1029/2007SW000380
- Owens, M., Challen, R., Methven, J., Henley, E. & Jackson, D. 2013, *Space Weather*, 11, 225-236, doi:10.1002/swe.20040
- Owens, M., *et al.* 2014, *Space Weather*, 12, 395-405, doi:10.1002/2014SW001064
- Owens, M., Riley, P. & Horbury, T. 2017, *Sol. Phys.*, 292, 5, 69, doi: 10.1007/s11207-017-1090-7
- Owens, M. & Riley 2017, *Space Weather*, submitted
- Riley, P., Linker, J. & Mikic, Z. 2001 *J. Geophys. Res.*, 106, A8, 15889-15901
- Riley, P. & Lionello, R. 2011, *Sol. Phys.*, 270, 575-592, doi:10.1007/s11207-011-9766-x
- Riley, P., Linker, J. A. & Arge, C. N. 2015, *Space Weather*, 13, 154-169, doi:10.1002/2014SW001144
- Riley, P., *et al.* 2017, *Space Weather*, 15, 3, 526-540, doi:10.1002/2016SW001589
- Siscoe, G., *et al.* 2004, *J. Atmos. Sol. Terr.*, 66, 1469-1480
- Toth, G., *et al.* 2005, *J. Geophys. Res.*, 110, A12226, doi:10.1029/2005JA011126
- Wang, Y. M. & Sheeley, N. R. Jr. 1990, *Astrophys. J.*, 335, 726