PRediction Of Geospace Radiation Environment and Solar wind parameterS

Work Package 3
Forecast of the evolution of geomagnetic indices

Deliverable 3.4
Development of existing Kp and Dst models

P. Wintoft, M. Wik, S. Walker, H.-L. Wei
February 5, 2018

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 637302.

Document Change Record

Issue	Date	Author	Details
v1	Dec 12, 2016	P. Wintoft	First draft.
v2	Dec 19, 2016	P. Wintoft, S. Walker, H.-L. Wei	Second draft.
v3	Dec 22, 2016	P. Wintoft, S. Walker	Third draft.
1.0	Dec 23, 2016	P. Wintoft, M. Wik, S. Walker, H.-L. Wei	Final.
1.1	March 14, 2017	P. Wintoft, M. Wik, S. Walker, H.-L. Wei, V. Yatsenko	Updated according to review report.
1.2	Feb 5, 2018	P. Wintoft	Removed contribution from SRI as that will be covered in separate report.

Contents

1 Introduction **5**

2 Development of the IRF-Kp-2017 model **5**

3 Development of the UoS-Kp-NARX-2016 model **9**

4 Development of the IRF-Dst-2017 model **13**

5 Conclusions **18**

Summary

The overall aim of WP 3 concerns improvement and new development of models based on data driven modelling, such as neural networks and NARMAX. Existing models for *Dst* and *Kp* will be analysed and verified with the aim of finding weaknesses and to suggest improvements. Solar wind and geomagnetic indices shall also be analysed in order to develop models for the identification of features, such as (but not limited to) shocks, sudden commencements, and substorms. Such categorisation will aid the model development and verification, and can also serve as alternative approach to models providing numerical input-output mapping. In addition to the development of *Dst* and *Kp* models new models will be developed to forecast *AE*. The models will be implemented for real-time operation at IRF and data and plots will be provided on a web server.

This deliverable is targeted at the development of existing *Kp* and *Dst* forecast models that are driven by solar wind data. The models have been evaluated in this WP with the results described in D3.3. Further evaluation has been carried out during the development process. The existing forecast models from IRF are named IRF-Kp-2000 and IRF-Dst-2002 with the number indicating the year of their publication, and the updated models are named IRF-Kp-2017 and IRF-Dst-2017. Similarly the University of Sheffield models is named UoS-Kp-NARX-2016. The models by Space Research Institute in Ukraine (SRI) are covered in a separate report. This report summarised the development but leaves the detailed description to published or submitted manuscripts, or manuscripts under preparation.

Acronyms

ACE	Advanced Composition Explorer
DSCOVR	Deep Space Climate Observatory
GFZ	GeoForschungsZentrum
GSFC	Goddard Space Flight Center
IRF	Institutet för rymdfysik (Swedish Institute of Space Physics)
NASA	National Aeronautic and Space Administration
NCEI	National Centers for Environmental Information
NOAA	National Oceanographic and Atmospheric Administration
SWPC	Space Weather Prediction Center
SRI	Space Research Institute (Ukraine)
UoS	University of Sheffield
WDC	World Data Center
WP	Work Package

1 Introduction

A large number of algorithms and models exist for the prediction of *Kp* and *Dst* as described in the model overview in deliverable D3.1. The sections below describe the further development of existing models previously developed by the individual team members. For this purpose the data collected and described in D3.2 have been used. The evaluation of the models (D3.3) indicates weaknesses and provides a starting point for the further development.

2 Development of the IRF-Kp-2017 model

The model by Boberg et al. (2000) consists of two neural networks that predict *Kp* at lower and higher activity levels, respectively. The outputs from the two networks are combined through a weighting to produce a single output. The model is driven by real-time solar wind at L1 and has been in operation at IRF-Lund since then (see example in Figure 1). The model is driven by ACE real-time data but will be switched over to data from the DSCOVR spacecraft.

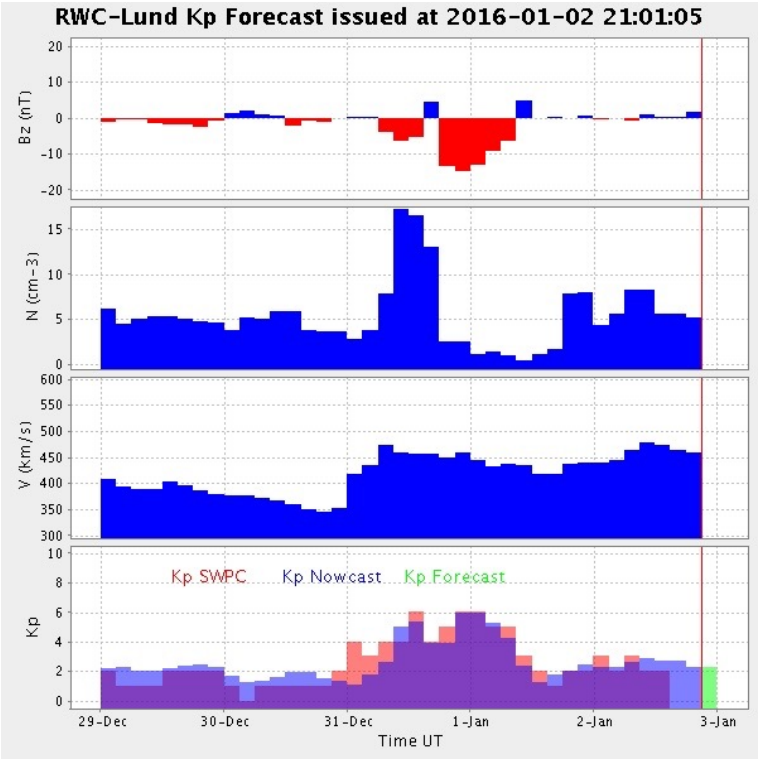

Figure 1: Example of real-time forecast currently published at IRF-Lund driven by ACE solar wind data.

In the following we refer to the model named LUND_NC in D3.3 as the IRF-Kp-2000 model, were 2000 refers to the year of its publication. There is also a LUND_FC model in D3.3 that we do not address any further after this paragraph. The difference between

LUND_NC and LUND_FC is that the NC model predicts Kp for the same timestamp as given by the latest solar wind, while the FC model predicts the next 3-hour interval. The FC model forecast works in a statistical sense, although storm onsets will in most cases be missed. The prediction lead-time for the NC model is governed by the solar wind propagation time from L1 to Earth. The verification of the IRF-Kp-2000 model (D3.3) showed that the mean absolute error (MAE) was 0.72 for the nowcast model and 0.88 for the forecast model, and with linear correlation coefficients (CORR) of 0.88 and 0.79, respectively. Note that Kp is an ordinal variable with 28 ordered categories $\{0o, 0+, 1-, 1o, 1+, \ldots, 9-, 9o\}$ that we transform to a real number in the interval $[0, 9]$.

In Wintoft et al. (2017) further evaluation of the IRF-Kp-2000 nowcast model were carried out during the development. It was noted that the MAE showed a large degree of variation with Kp level (see Figure 2).

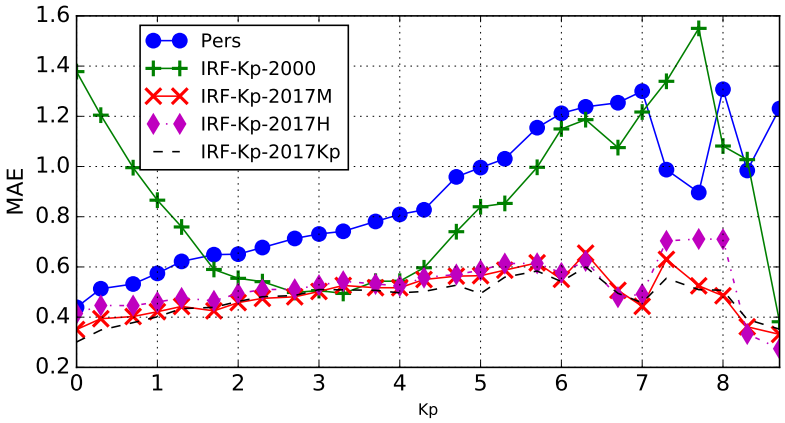

Figure 2: The variation of MAE as function of observed Kp for persistence (Pers) model and the different IRF-Kp models. Figure from Wintoft et al. (2017).

Another finding was that the onsets of some storms were missed. After further analysis, together with the fact that the Kp index is measuring the range of magnetic disturbances (Mayaud 1980), we concluded that the 3-hour averages of the solar wind used for inputs do not capture the physical processes due to too low temporal resolution.

To capture phenomena that occur on short timescales (minutes) but still have a 3-hour cadence in agreement with Kp we introduced 3-hour maxima and minima of the solar wind data. This new parametrisation and the use of more data since the publication of the IRF-Kp-2000 model made substantial improvements on the forecasts (Wintoft et al. 2017). The updated models MAE as function of Kp are shown in Figure 2 and are called IRF-Kp-2017. Two variants exist, one that is driven by one-minute resolution solar wind data (IRF-Kp-2017M) and one that is driven by hourly resolution data (IRF-Kp-2017H). We will come back to the two variants in the next paragraph. The high resolution ACE solar wind data used extend over the years 1998 to 2015. From this series we set aside data for the years 2001 and 2011 for testing. Data from the years 2000, 2006, and 2012 make up the validation set, while the training set consist of the remaining data. The overall MAE, using only data from the test set, dropped from 0.68 to 0.47–0.51 (Table 1).

Similar scores were obtained when evaluated on all available data.

	BIAS	MAE	RMSE	CORR	MSESS:PERS
Pers	-0.00	0.64	0.87	0.77	nan
IRF-Kp-2000	0.31	0.68	0.84	0.83	0.08
IRF-Kp-2017M	0.04	0.47	0.61	0.90	0.51
IRF-Kp-2017H	0.09	0.51	0.66	0.89	0.43
IRF-Kp-2017Kp	0.03	0.45	0.60	0.91	0.53

Table 1: Measures and scores for predicted *Kp* on the test set. The models are: persistence (Pers); Boberg et al. (2000) (IRF-Kp-2000); model with minute resolution solar wind (IRF-Kp-2017M); and model with hourly resolution solar wind (IRF-Kp-2017H). Table from Wintoft et al. (2017)

As the *Kp* forecast algorithm also should be able to use predicted solar wind, and because the temporal resolution of the predicted solar wind is lower, another model was developed that takes 1-hour averages as input (IRF-Kp-2017H). The same datasets and model structure was used for the development of the H model as for the M model. Naturally, we expect the H model to perform slightly poorer than the M model, and the MAE is slightly higher (0.51) and the correlation is slightly lower (0.89) as shown in Table 1. But the H model can be driven by the forecast solar wind delivered by WP2 and thereby completing the Sun-Earth chain.

An example with a detailed view of the prediction using 1-minute and 1-hour resolution solar wind data is shown in Figure 3. In both cases the solar wind is provided with a one-minute cadence where the hourly data are 1-hour running averages. The sudden change in solar wind conditions as seen in the velocity (top panel) is marked with a blue dot at the point in time when the velocity increases and with a red diamond when the jump has occurred. The blue dot and red diamond are placed on the same timestamps for the observed horizontal geomagnetic field (second panel). The jump in the geomagnetic field is marked with a green plus. The geomagnetic field variation is related to the *Kp* index as *Kp* is derived from geomagnetic data. The bottom panel shows the propagation time in minutes from L1 to Earth. Initially it is about 55 minutes but at the time of the velocity jump it decreases to about 40 minutes. At the timstamp marked with the blue dot in the solar wind the predicted *Kp* (third panel) from the IRF-Kp-2017M model will have a lead-time of ≈ 55 minutes, thus the predicted *Kp* value is marked with the corresponding blue dot. The red horizontal bar extending backwards from the blue dot indicates a 3-hour interval. A few minutes later the L1–Earth propagation lead time has decreased to ≈ 40 minutes, which means that the predicted *Kp* values may be located at timestamp for which prediction already exist, or even at earlier timstamps. The red curve joining the blue dot and red diamond in the third panel shows this: consecutive predictions are placed on the same timestamps or even earlier which means that the curve goes straight up or even backwards. In practice this means that the $Kp \approx 3$ forecast at the time of the blue dot becomes invalidated at the timestamp of the red diamond and replaced with the $Kp \approx 8$ forecast, and this is just a consequence of that the propagation time from L1–Earth decreases faster than the one-minute sampling rate. It should be noted that

the observed Kp value (blue step curve) from 00UT to 03UT can not be determined until after 03UT.

The green dashed curve in the third panel shows the predicted Kp from the H model, and naturally it shows smoother variation and also some lagging, however, that is not crucial as the intention is to drive it with predicted solar wind data which means that the total prediction lead time will be days.

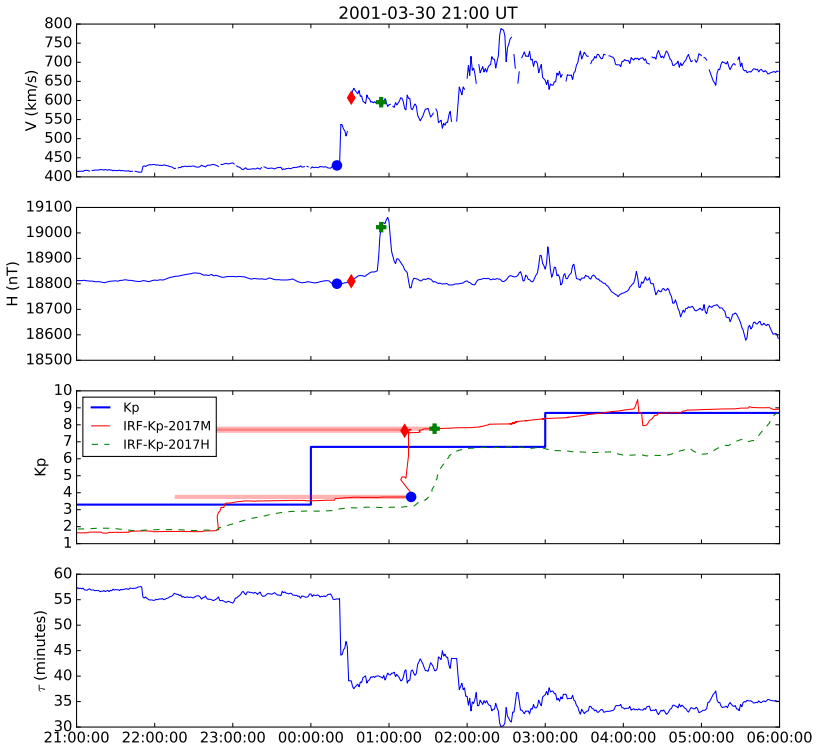

Figure 3: Predicted Kp at one minute cadence. From top to bottom: solar wind velocity at spacecraft location; horizontal geomagnetic field at NGK; final and predicted Kp; prediction lead time τ in minutes. In the third plot, the red solid curve is predicted Kp driven by minute resolution solar wind data, and green dashed curve using hourly resolution data. The blue circle, red diamond, and green plus mark three consecutive points in time. The three horizontal red lines indicates the corresponding three hour intervals that the forecasts represent. Figure from Wintoft et al. (2017).

In Figure 4 the observed and forecast Kp are shown over a longer period driven by hourly OMNI solar wind data from the test set. The event from Figure 3 is visible in the beginning of the period. It is seen that the predicted Kp follows closely the observed Kp over the full range from low to high Kp values. In many cases the onsets are timely predicted with one clear exception in 2001-04-04 where the prediction lags by one step

(i.e. 3 hours). During the strongest event the prediction goes off scale and reach a value close to $Kp = 12$.

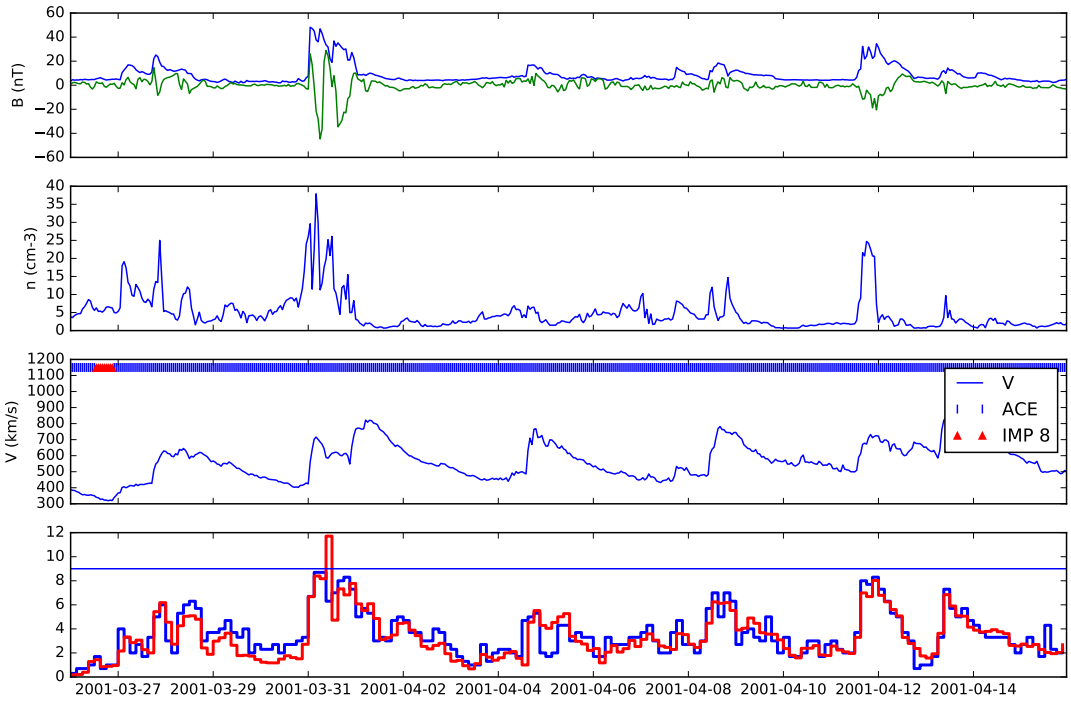

Figure 4: Prediction using the IRF-Kp-2017H model driven by OMNI hourly resolution data over the period 2001-03-26 to 2001-04-15 (from test set). From top to bottom: solar wind B and B_z; solar wind particle density n; solar wind velocity V with measuring spacecraft indicated; observed (blue) and predicted (red) Kp with blue line indicating the maximum $Kp = 9$.

The new models, IRF-Kp-2017Mand IRF-Kp-2017H, are available for implementation and further testing can be performed during WP7. It will be especially interesting to study those cases when the predictions fail and try to understand the cause and see if improvements can be made.

3 Development of the UoS-Kp-NARX-2016 model

The UoS-Kp-NARX-2016 model employs a data-based or system identification to develop transparent and parsimonious models for the prediction of the Kp index. The aim of the system identification modelling approach is to find a mathematical model that can be used to characterise the behaviour of a system based on discrete-time observational data of the input and output of the system (Billings 2013, Wei, Billings & Balikhin 2004,

Wei, Billings & Liu 2004, Wei & Billings 2008). Two different implementations of NARX models for the prediction of *Kp* have been developed and compared (Ayala Solares et al. 2016), namely a recursive sliding window model and a direct prediction model. These models are based on nonlinear system identification techniques and the NARX (nonlinear autoregressive with exogenous inputs) method. A major advantage of these models is that they provide a clear indication of the dependence of the *Kp* index on various solar wind parameters. While the role of solar wind speed and dynamic pressure as drivers of the *Kp* index has been confirmed, this new development provides an explicit formula linking the contribution of these drivers to changes in the *Kp* index. Thus, this technique provides scientific insight into the dependence of *Kp* on solar wind speed and dynamic pressure. It was found that the latter, direct prediction model had the best performance so is used as the basis for the UoS-Kp-NARX-2016 model.

The NARX methodology is used to determine the relationship between the input and output parameters of a backbox system. The various input-output parameters used in this development of the UoS-Kp-NARX-2016 model are listed in Table 2. The data for these parameters are taken from the low resolution OMNI dataset, which consists of hourly averaged near-Earth measurements of the solar wind magnetic field and particle environment from several spacecraft such as IMP 8, WIND, Geotail and ACE. The output parameter is the *Kp* index, which is measured every three hours. In order to match the time resolutions between the input and output signals, the observed *Kp* values are interpolated to 1-hour resolution by simply repeating the last measured value during the next two hours.

Table 2: Input and Output Parameters

Variable	Symbol	Description
Inputs	V	Solar wind speed [km/s]
	Bs	Southward interplanetary magnetic field [nT]
	VBs	Southward interplanetary magnetic field [VBs = VBs/1000]
	p	Solar wind pressure [nPa]
	Sp	Square root of solar wind pressure
Output	Kp	Index of interest

The NARX methodology (Billings 2013, Wei, Billings & Liu 2004, Wei & Billings 2008) was applied to the set of input/output parameters from the year 2000. The first half of this dataset was used for model training while the second half were reserved for model testing. A total of four models have been identified to provide forecasts with lead times of 3, 6, 12, and 24 hours. The model performances are summarised in Table 3. Figure 5 shows, from top to bottom, example output from the 3, 6, 12, and 24 hour ahead forecasts. The measured *Kp* is shown in black. It is seen that in general there is excellent agreement between the forecasts and the measured *Kp* except for a short period around July 15-16. This time corresponds to a period when the *Kp* index exhibited a large fluctuation that is not captured by the model. Reasons for this are still being investigated in a hope to

improve our model. Possible solutions include: i) to further optimise some key model parameters, for example, the maximum lags of the input and output variables, and the model complexity (e.g. the nonlinear degree of the polynomial model); these parameters are important; ii) the training data were chosen from the year 2000 in some arbitrary sense. These training data may not be sufficiently representative for the data dynamics of the other years. Future investigation would therefore consider finding a good common model structure that can accommodate severe nonlinearity and non-stationarity in Kp index, as well as in various solar wind parameters and geomagnetic field indices.

Table 3: 3, 6, 12 and 24 hour ahead predictions

Horizon	RMSE	Pearson Coefficient	PE
3	0.759	0.871	0.7585
6	0.833	0.842	0.710
12	0.862	0.831	0.690
24	0.872	0.827	0.682

Comparing the results obtained with those presented in Wing et al. (2005), the values of the two model performance metrics (i.e. prediction performance and correlation coefficient) calculated from our results are slightly lower. This may be explained from several factors: i) all the data for all input and output variables used for model estimation in this study are raw data sampled hourly; no pre-processing (e.g. smoothing, interpretation, etc.) was performed; ii) the model input variables used in this work are not exactly the same as those used in previous studies; iii) some coefficients required by the models, for example the maximum lags of the input and output variables, may need to be further optimised.

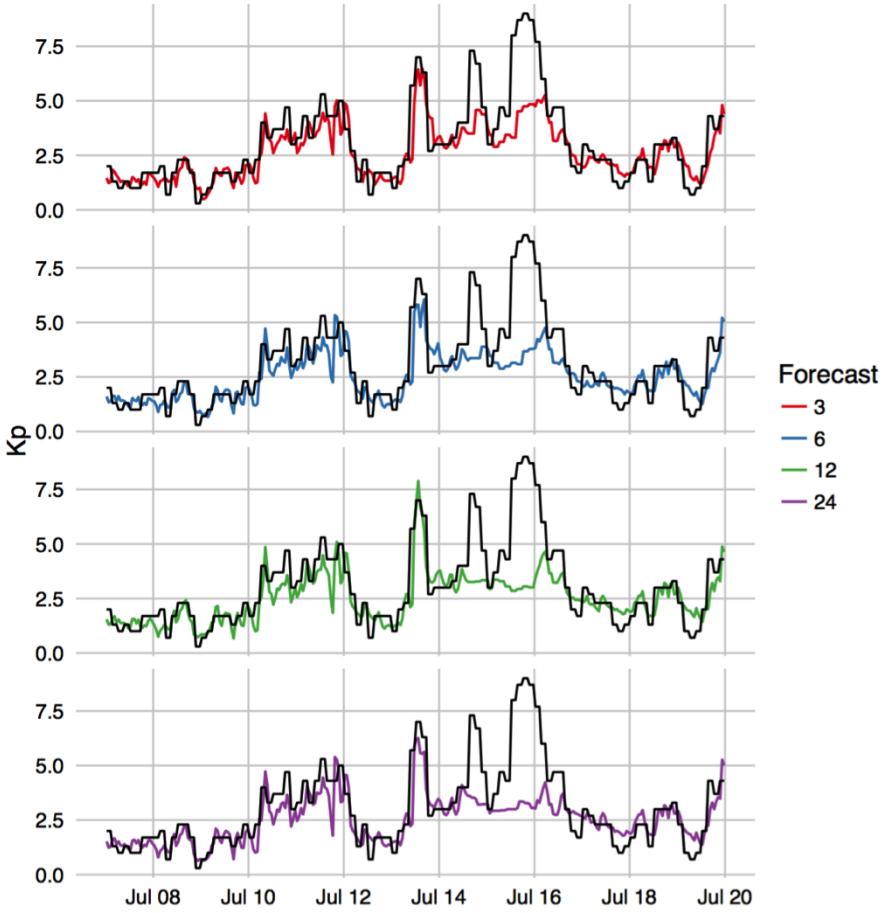

Figure 5: Predictions of the *Kp* index for the four horizons of interest during the middle of July 2000 using the direct approach. The black line corresponds to the measured *Kp* values.

4 Development of the IRF-Dst-2017 model

The *Dst* prediction model by Lundstedt et al. (2002) is a recurrent (time delayed outputs are fed back into the inputs) neural network known as the Elman network. The Elman network automatically finds the dynamics (memory) of the system during the training phase by adjusting the context weights (recurrent weights). The Elman network thus has the nice property that it can learn both the directly driven relation from the inputs and the dynamical component. The main disadvantage is that the data need to be contiguous and presented ordered in time which leads to a higher degree of complexity when selecting datasets and handling datagaps.

The IRF-Dst-2002 model has be operation for more than 15 years and the predictions have published in real-time at IRF-Lund. An example of predictions using the IRF-Dst-2002 model is shown in Figure 6.

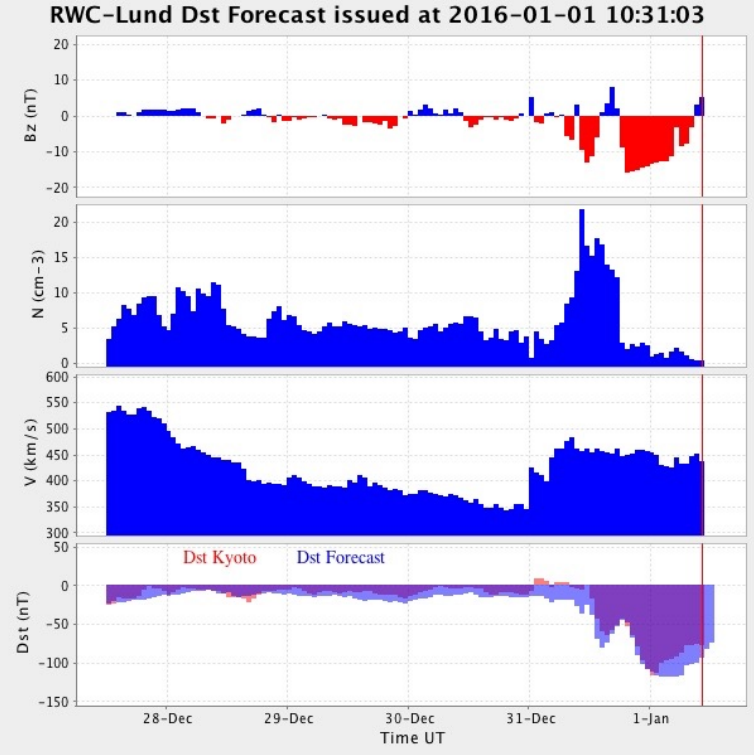

Figure 6: Example of real-time forecast currently published at IRF-Lund driven by ACE solar wind data.

The verification of the IRF-Dst-2002 model (D3.3) showed that the mean absolute error (MAE) was 10.6 nT and with linear correlation coefficients (CORR) of 0.85 based on OMNI solar wind data (D3.2) for the years 1998–2014. Overall this performance is very good, however, it was also noted that there was a saturation of the model so that predictions did not reach below about -230 nT. It should also be noted that the model was developed on data before about year 2002 (year of publication), thus today we have much more data with good coverage. Another goal with the model was to make it as

small as possible by having few processing units and thereby limiting the number of free parameters.

Within this project new *Dst* prediction models (IRF-Dst-2017) have been developed (**?**) and compared to the IRF-Dst-2002 model. To simplify the data selection procedure time-delay neural networks are applied instead of recurrent networks. The length of the time-delay line (the embedding dimension) needs to be defined. There is an unlimited number of ways that the embedding can be constructed. However, a typical isolated *Dst* storm has a recovery phase that lasts for several hours. In the early work by Burton et al. (1975) a constant decay term with decay time 7.72 hours was used. Later work (O'Brien & McPherron 2000) suggested a variable decay term with decay time between 2.4 and 19 hours. This means that it will take 44 hours to reach within 10% of the quiet *Dst* with a decay time of 19 hours. The driven *Dst* main phase is caused by changes in the solar wind (negative turning of B_z, jumps in n and V) while the recovery phase is decoupled from the solar wind and is dependent on the overall evolution. Therefore, we assume that features more distant in time do not require a detailed representation. Thus, instead of providing each hourly value 44 hours back in time we take temporal averages of increasing sizes when going back in time. The embedded velocity then becomes

$$\begin{aligned}\bar{V}(t) \;=\; & [V(t),\\ & \langle V(t-1), V(t-2)\rangle,\\ & \langle V(t-3), V(t-4), V(t-5), V(t-6)\rangle,\\ & \langle V(t-7), \ldots, V(t-14)\rangle,\\ & \langle V(t-15), \ldots, V(t-30)\rangle,\\ & \langle V(t-31), \ldots, V(t-62)\rangle] \end{aligned} \quad (1)$$

where $\langle\cdot\rangle$ is the average. Performing the same embedding on B, B_y, B_z and n results in $6 \times 5 = 30$ inputs.

Using the above embedding and removing all samples with data gaps in the solar wind results in the number of available samples and range of *Dst* as shown in Figure 7. From this set we extract the four years 1981, 1996, 2001, and 2008 which we call the Test set. The Test set is not used in the model development and optimisation. The remaining data are divided into two sets, A and B, each consisting of data from even and odd years, respectively.

A large set of different networks have been initialised with different weights and number of hidden units. Different types of learning algorithms and learning parameters have also been tested. The networks have been trained (weight adaption) on one set (A or B) and validated against the other set (B or A). The forecast lead-time was explored by allowing the lead-time to vary from 0 hours up to 3 hours. It should be noted that this lead time is in addition to the variable solar wind propagation lead time. The result is shown in Figure 8. In a statistical sense there is no real difference between the $\tau = 0$ and $\tau = 1$ hour models, which is expected considering the *Dst* storm build-up time.

The performance of the $\tau = 1$ hour model is compared with the IRF-Dst-2002 model with statistical measures summarised in Table 4. It is clear that the IRF-Dst-2017 model performs better than the IRF-Dst-2002 model. There are several reasons for this improvement. Firstly, we have more data to train and evaluate the model. Secondly, we allow the

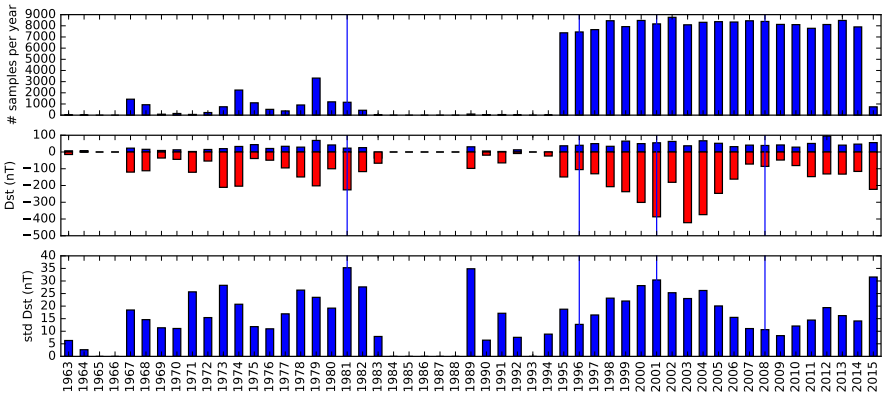

Figure 7: Number of samples per year (top panel), maximum and minimum *Dst* for each year (middle panel), and one standard deviation *Dst* (bottom panel).

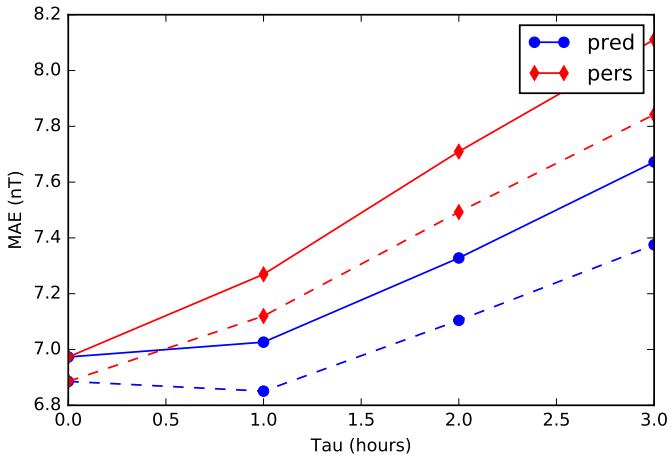

Figure 8: The MAE as function of lead-time τ in hours. Blue circles shows MAE for models with different τ. Red diamonds is the $\tau = 0$ model simply shifted for comparison. Solid lines are for all data and dashed lines for the Test set.

selection of a bigger neural network, in this case it has 10 hidden units as compared to 4 in the IRF-Dst-2002 model. Finally, we have also included the total solar wind magnetic field B and the Y-component B_y. The importance of the latter will be further explored in the future.

Model	Set	n	BIAS	MAE	RMSE	CORR
IRF-Dst-2016	All	177116	0.29	7.03	9.60	0.88
IRF-Dst-2002	All	177116	-6.09	9.65	12.40	0.85
IRF-Dst-2016	A	72511	0.02	7.21	9.58	0.90
IRF-Dst-2002	A	72511	-6.24	10.07	13.05	0.85
IRF-Dst-2016	B	79616	0.27	6.92	9.49	0.86
IRF-Dst-2002	B	79616	-6.17	9.27	11.64	0.85
IRF-Dst-2016	Test	24989	1.13	6.85	9.97	0.89
IRF-Dst-2002	Test	24989	-5.36	9.65	12.83	0.85

Table 4: The BIAS, MAE, RMSE, and CORR computed for the two models using all data (All) and using the subsets (A, B, Test). The number of samples (n) is also given.

Inevitably, the models will perform worse on the extremes of the *Dst* range due to the limited number of samples for strongly negative and positive *Dst* values. Therefore, we study the performance as a function of *Dst* (Figure 9). In the range $Dst \in [-100, 0]$ nT the models have comparable BIAS and MAE, while for the more extreme values the IRF-Dst-2017 model performs better.

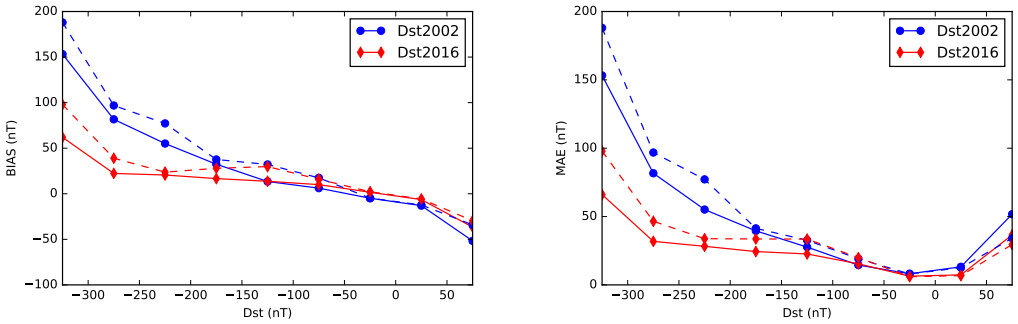

Figure 9: The BIAS and MAE as function of *Dst* for the IRF-Dst-2002 model (blue circles) and the IRF-Dst-2017 model (red diamonds) using all data (solid lines) and Test set (dashed lines). The markers are located at the centres of each *Dst* bin, except the $Dst = -325$ which includes all values with $Dst < -300$ nT.

The detailed performance of the models are illustrated with a few examples from the test set in year 2001 in Figure 10. The first event in the beginning av March 2001 reached $Dst \approx -70$ nT. The IRF-Dst-2002 and the nowcast IRF-Dst-2017 models comes very close, while the 1-hour forecast IRF-Dst-2017 reaches around -90 nT. However, looking carefully the onset of the storm the two new models follows *Dst* very well.

The second example at the end of March 2001 is stronger and reaches $Dst \approx -390$ nT. None of the models reach that value, this illustrates the great challenge to predict the rare but strong events. The 2017 models comes closest at predicted values of -300 nT, while the 2002 model saturates at slightly below -200 nT.

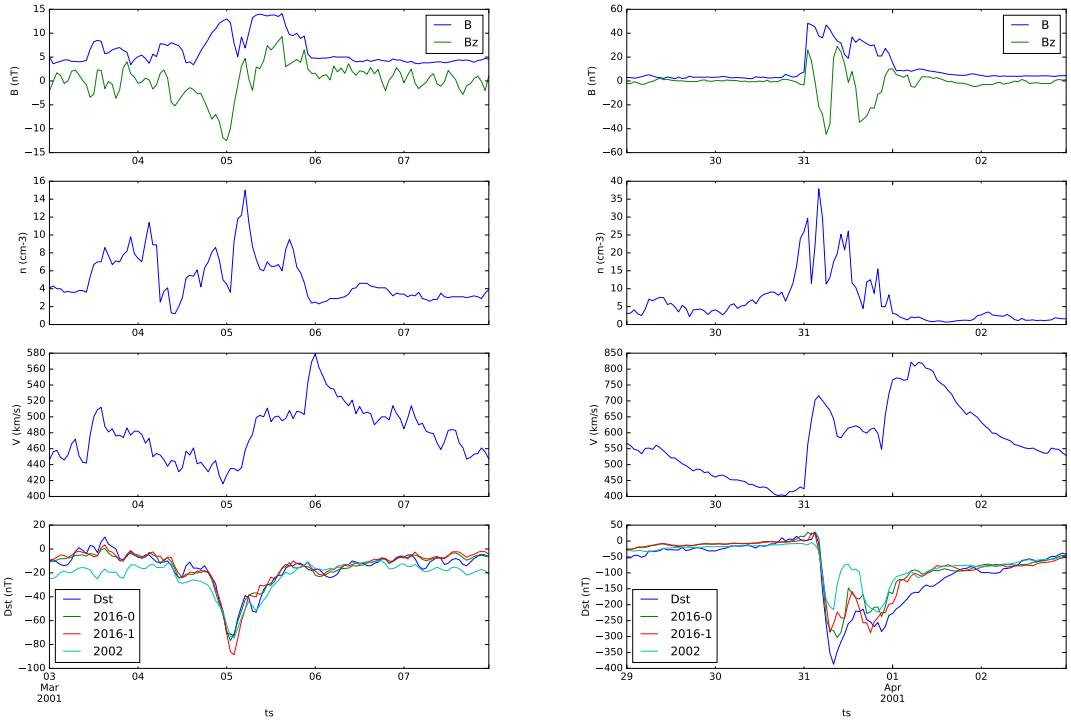

Figure 10: Forecast of *Dst* using the IRF-Dst-2002 model (2002), the 0-hour forecast (nowcast) IRF-Dst-2017 model (2017-0), and the 1-hour forecast IRF-Dst-2017 model (2017-1). The top three panels show solar wind magnetic field (B and B_z), density (n), and speed (V) from the OMNI datasets.

5 Conclusions

The forecasting algorithms for *Kp* and *Dst* have been updated and shows improved performance. As the algorithms are computationally lightweight it is straightforward to set them up for real-time operation. The algorithms can be driven by real-time measurement from the L1 solar wind location or from predicted solar wind provided from WP 2 *Propagation of the Solar Wind from the Sun to L1*. The algorithms are available for the implementation as part of WP7 *Fusion of forecasting tools*. The different forecast models can be run in parallel and their outputs can be compared against each other and against observed *Kp* and *Dst*.

During WP7 further testing and updates of the models can be performed. This will include the whole data processing pipeline from measured and predicted solar wind data, through the *Kp* and *Dst* forecast algorithms, to the final end products. Further error analysis on the forecasts can be carried out in order to better understand situations when the forecasts fails. The use of solar wind data from the new DSCOVR spacecraft will also be interesting to study.

References

Ayala Solares, J. R., Wei, H.-L., Boynton, R. J., Walker, S. N. & Billings, S. A. (2016), 'Modelling and prediction of global magnetic disturbance in near-earth space: A case study for kp index using narx models', *Space Weather* **14**(10), 899–916.

Billings, S. A. (2013), *Nonlinear System Identification: NARMAX Methods in the Time, Frequency,and Spatio-Temporal Domains*, Wiley.

Boberg, F., Wintoft, P. & Lundstedt, H. (2000), 'Real time Kp predictions from solar wind data using neural networks', *Physics and Chemistry of the Earth, Part C: Solar, Terrestrial & Planetary Science* **25**, 275–280.

Burton, R. K., McPherron, R. L. & Russell, C. T. (1975), 'An empirical relationship between interplanetary conditions and dst', *Journal of Geophysical Research* **80**(31), 4204–4214.

Lundstedt, H., Gleisner, H. & Wintoft, P. (2002), 'Operational forecasts of the geomagnetic *Dst* index', *Geophysical Research Letters* **29**(24), 34–1–4.

Mayaud, P. N. (1980), *Derivation, meaning, and use of geomagnetic indices*, Vol. 22 of *Geophysical monograph*, American Geophysical Union.

O'Brien, T. P. & McPherron, R. L. (2000), 'Forecasting the ring current dst in real time', *Journal of Atmospheric and Solar-Terrestrial Physics* **62**, 1295–1299.

Wei, H.-L. & Billings, S. A. (2008), 'Model structure selection using an integrated forward orthogonal search algorithm assisted by squared correlation and mutual information', *International Journal of Modelling, Identification and Control* (3(4)), 341–356.

Wei, H. L., Billings, S. A. & Balikhin, M. (2004), 'Prediction of the dst index using multiresolution wavelet models', *Journal of Geophysical Research* **109**, A07212.

Wei, H.-L., Billings, S. A. & Liu, J. (2004), 'Term and variable selection for non-linear system identification', *International Journal of Control* (77(1)), 86–110.

Wing, S., Johnson, J. R., Jen, J., Meng, C. I., Sibeck, D. G., Bechtold, K., Freeman, J., Costello, K., Balikhin, M. & Takashi, K. (2005), 'Kp forecast models', *Journal of Geophysical Research* **110**, A04203.

Wintoft, P., Wik, M., Matzka, J. & Shprits, Y. (2017), 'Forecasting kp from solar wind data: input parameter study using 3-hour averages and 3-hour range values', *Journal of Space Weather and Space Climate* **7**, A29.