Understanding the performance of neural network models for short-term predictions applied to geomagnetic indices

Swedish Institute of Space Physics
Peter Wintoft and Magnus Wik

The Origin of Magnetic Storms.

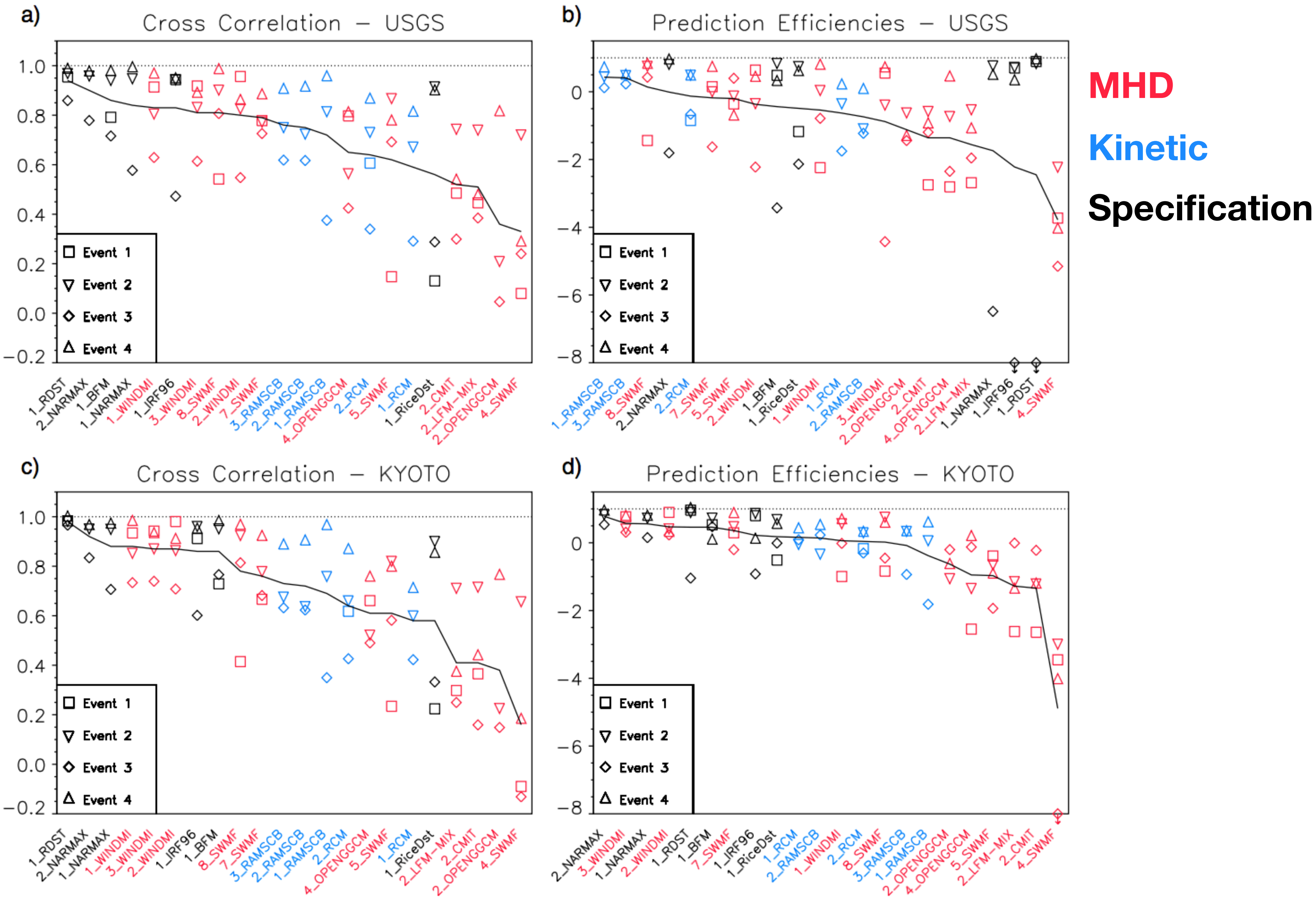
By ARTHUR SCHUSTER, F.R.S.

(Received January 9,—Read January 26, 1911.)

1. Lord Kelvin,* in discussing the origin of magnetic storms, came to the conclusion that they could not be due to a direct solar action on account of the enormous energy which would have to be supplied by the sun. This verdict was generally accepted until recently, when the theory of a direct solar action has been revived in a form, which is assumed to be free from the objection raised, the magnetic action being supposed to be due to a swarm of electrified corpuscles ejected by the sun. The question of energy has not, so far as I know, been discussed in this case, and it seems to be taken for granted that the total energy of the magnetic field due to a swarm of corpuscles is equal to the sum of the energies of each, calculated as if the others were not present. If the corpuscles are sufficiently far apart, this is approximately correct; but in that case the magnetic field itself would have to be negligible, except within molecular distance from each particle. How far we may go wrong by treating the energy as if it could be obtained by a process of addition becomes apparent when we consider that such treatment would lead to coefficients of self-induction which are proportional to the length of a circuit and independent of its shape.

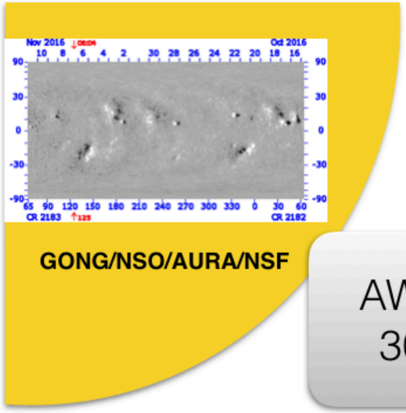
Proc. R. Soc. Lond. A 1911 85 44-50; DOI: 10.1098/rspa.1911.0019. Published 14 March 1911

- 30 model settings: 3D MHD models; kinetic models; specification models.
- 4 storms: Aug 2001 (-40 nT), Oct 2003 (-353 nT), Aug 2005 (-131 nT), Dec 2006 (-139 nT)



Rastätter et al., Geospace environment modeling 2008–2009 challenge: D_{st} index, SPACE WEATHER, VOL. 11, 187–205, doi:10.1002/swe.20036, 2013.

The PROGRESS project



AWSoM
30 Rs

SWIFT
L1

Neural Network
NARMAX

Kp
Dst
AU, AL, AE

Statistical
wave
models

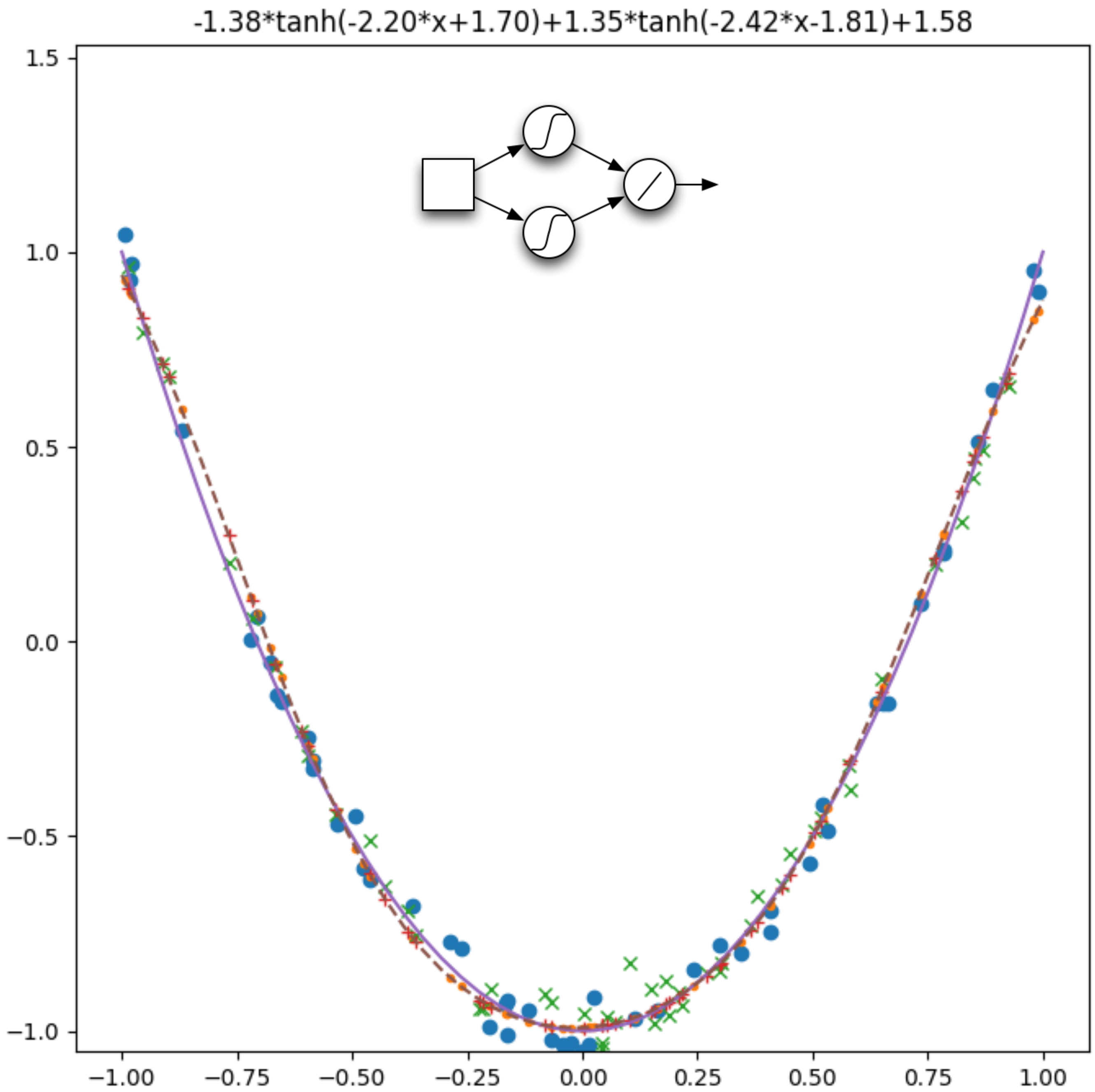
VERB
IMPTAM

- Ayala Solares, J. R., Wei, H.-L., Boynton, R. J., Walker, S. N. & Billings, S. A., 'Modelling and prediction of global magnetic disturbance in near-earth space: A case study for Kp index using narx models', *Space Weather, 2016*.
- Wintoft, P., M. Wik, J. Matzka and Y. Shprits, Forecasting Kp from solar wind data: input parameter study using 3-hour averages and 3-hour range values, J. Space Weather Space Clim. Volume 7, A29, 2017

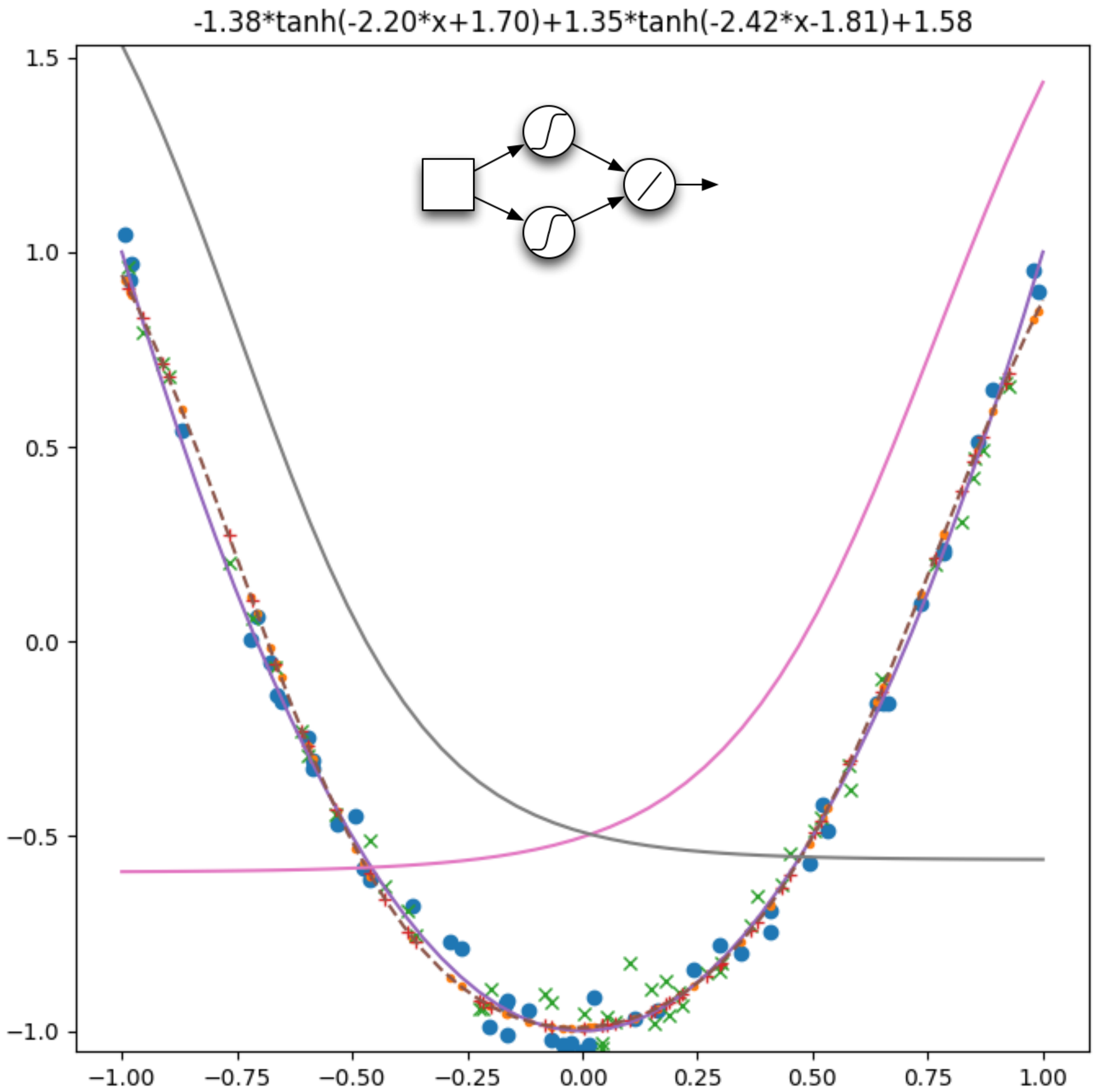
Neural networks

- Data …
- Data gaps (esp. hard for time-series).
- Division of datasets used for
  - training,
  - validation, and
  - testing.
- Training set => parameter estimation
- Validation set => hyper-parameter search
- Test set => final estimate of performance

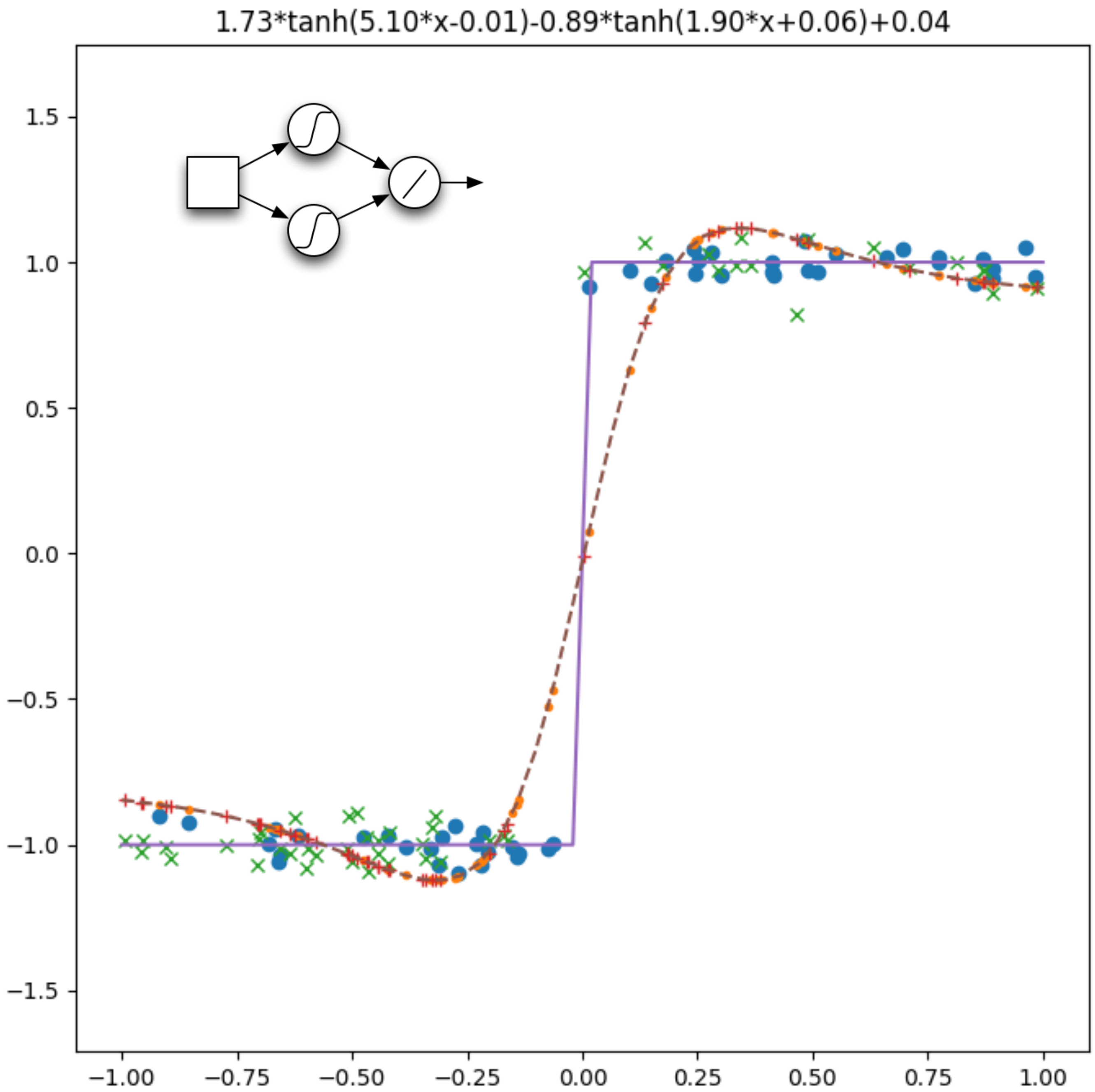
A lot of other considerations:
Learning algorithm
Input parameters
Transformation of inputs
Network architecture



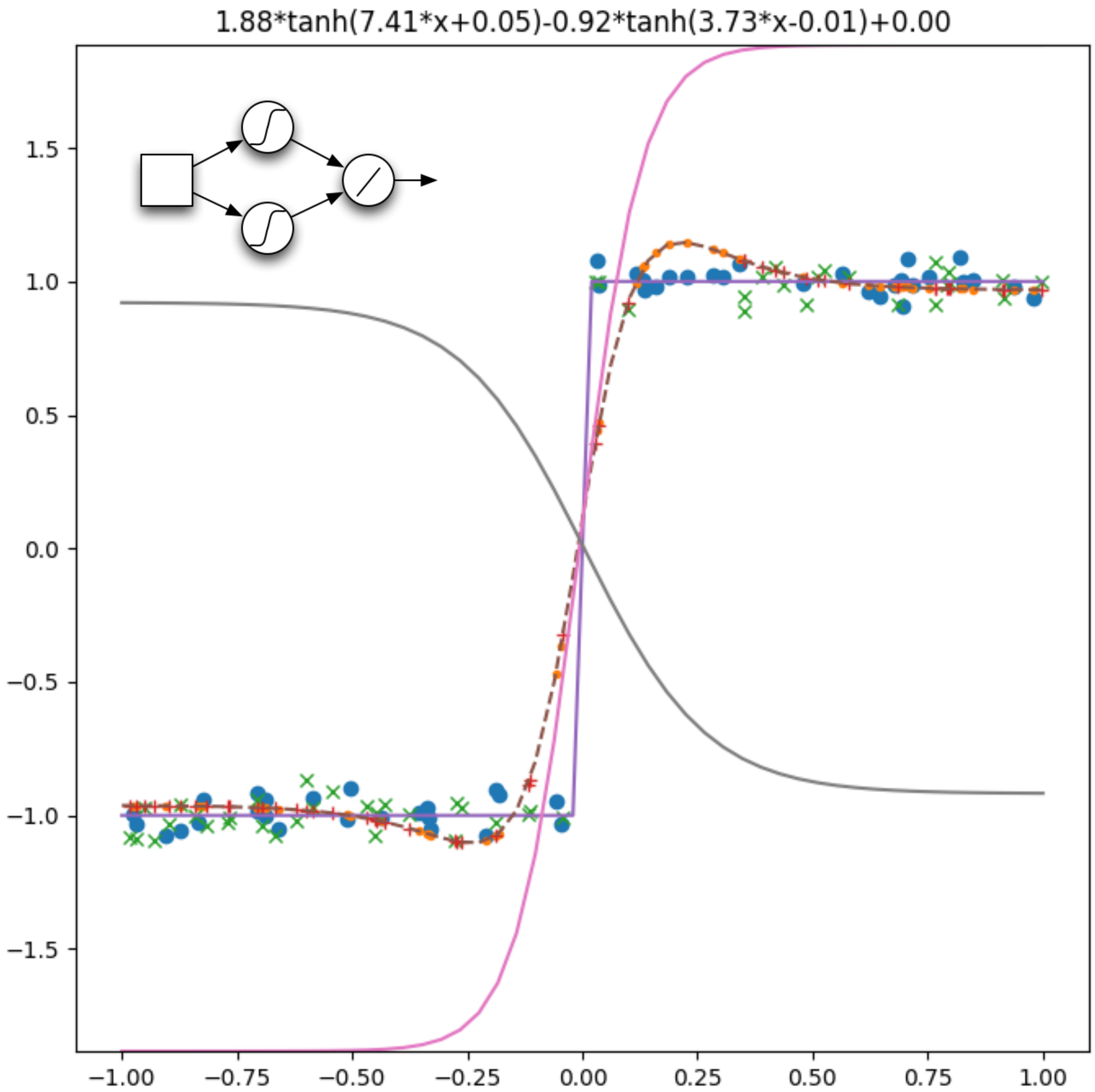
“Multilayer feedforward networks are universal approximators” (Hornik et al., 1989).



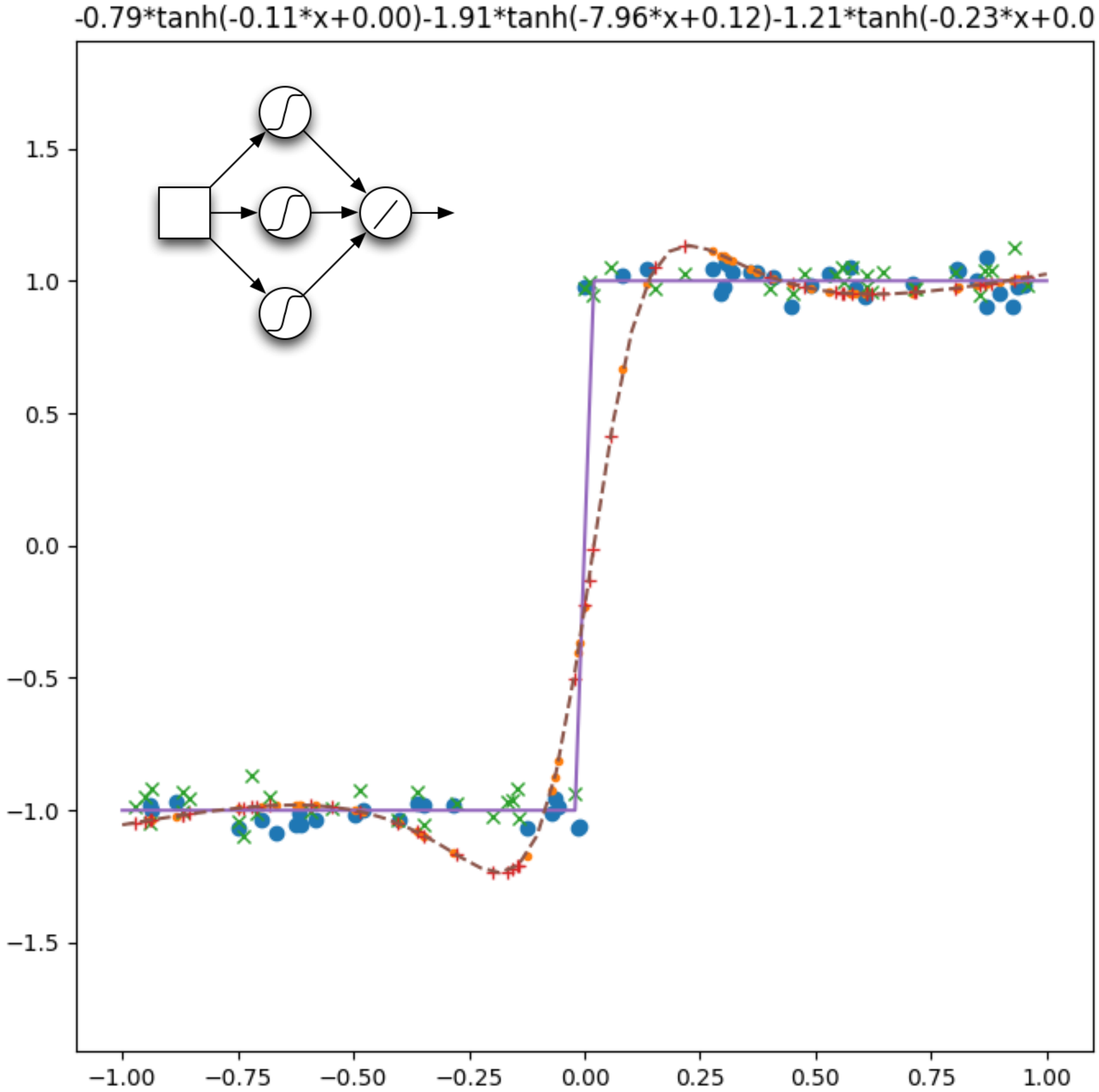
“Multilayer feedforward networks are universal approximators” (Hornik et al., 1989).



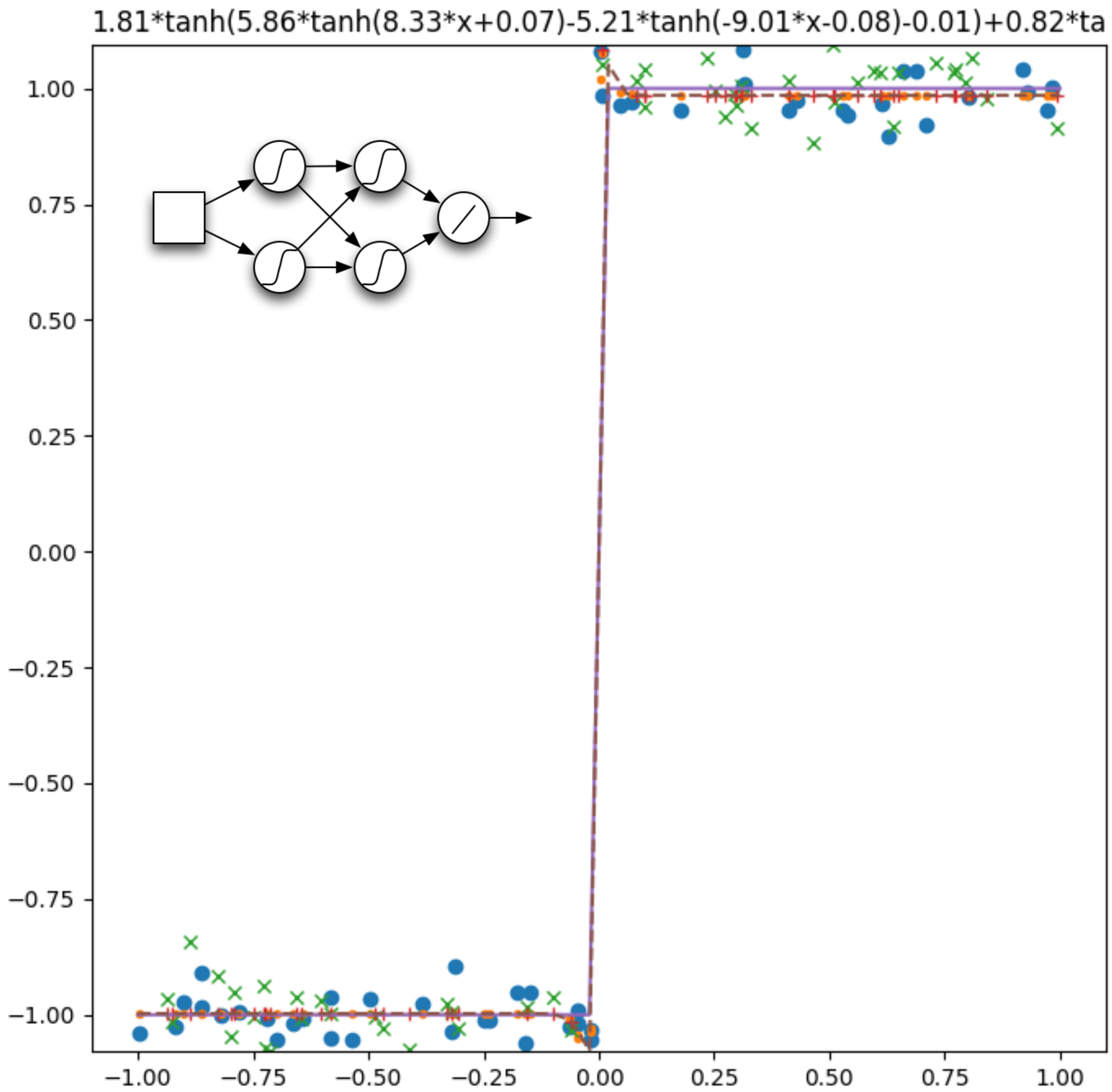
“Multilayer feedforward networks are universal approximators” (Hornik et al., 1989).



“Multilayer feedforward networks are universal approximators” (Hornik et al., 1989).

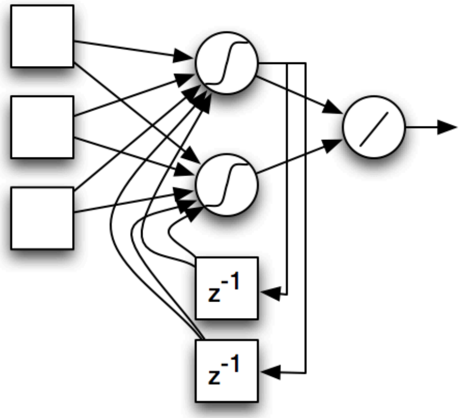


"Multilayer feedforward networks are universal approximators" (Hornik et al., 1989).



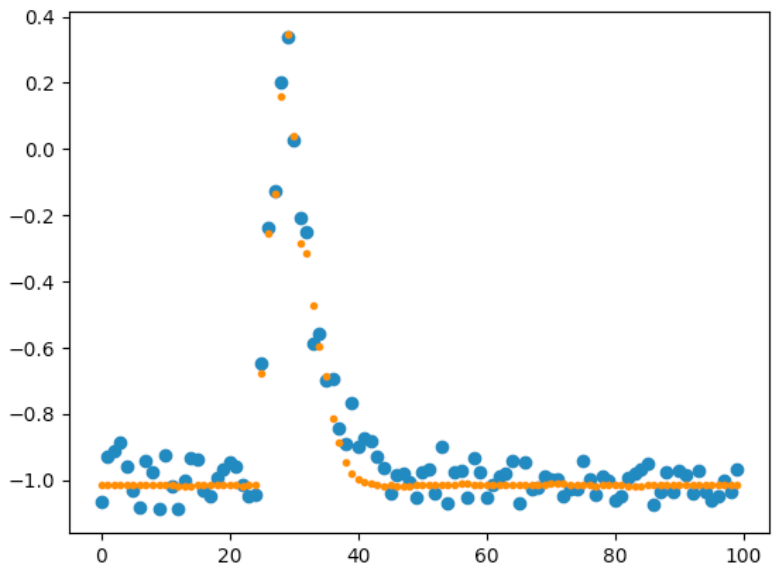
"Multilayer feedforward networks are universal approximators" (Hornik et al., 1989).

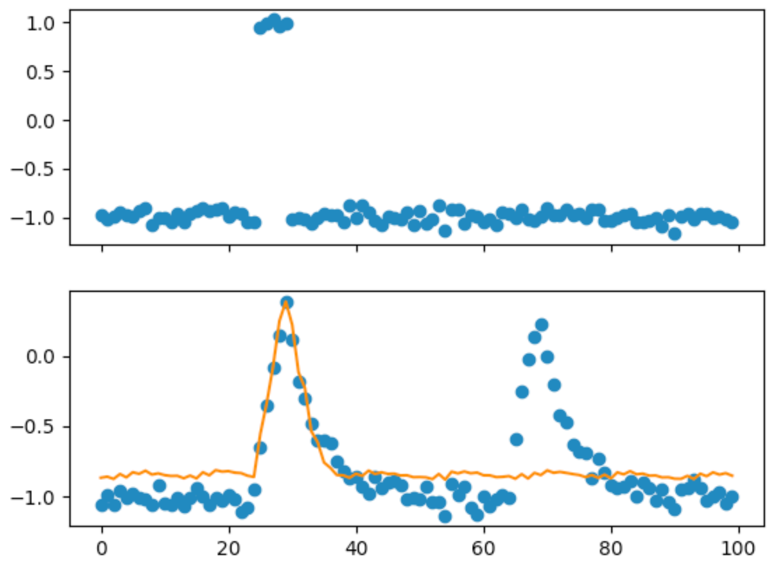
Dynamic NNs



- Time delays
- Recurrent connections
- More advanced: LSTM

$$Dst^*(t+1) = Dst^*(t) + [Q(t) - \lambda(t)Dst^*(t)]$$





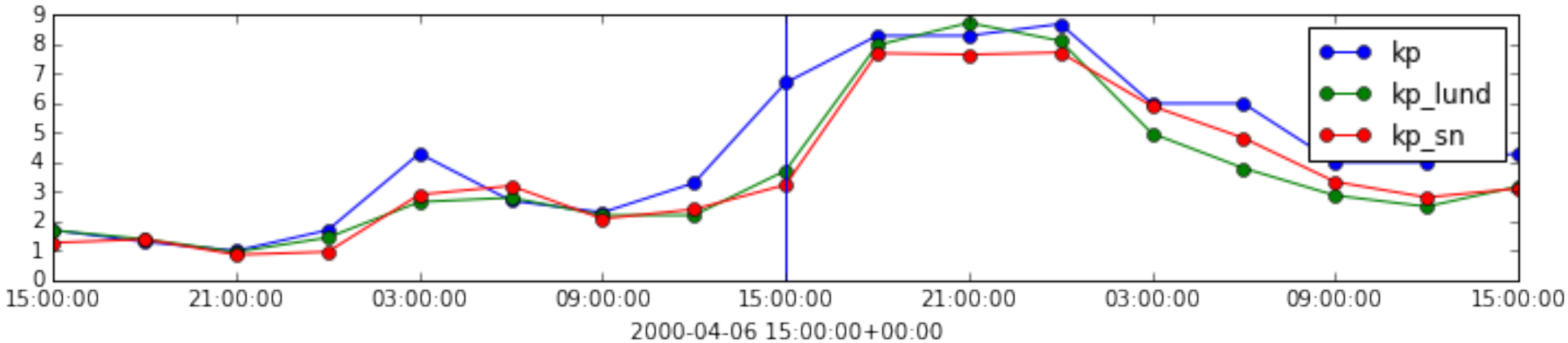
F. Takens, Detecting strange attractors in turbulence, Springer, 1981.

Recurrent Neural Networks are universal approximators, Schäfer and Zimmermann, Int J Neural Syst. 2007 Aug;17(4):253-63.

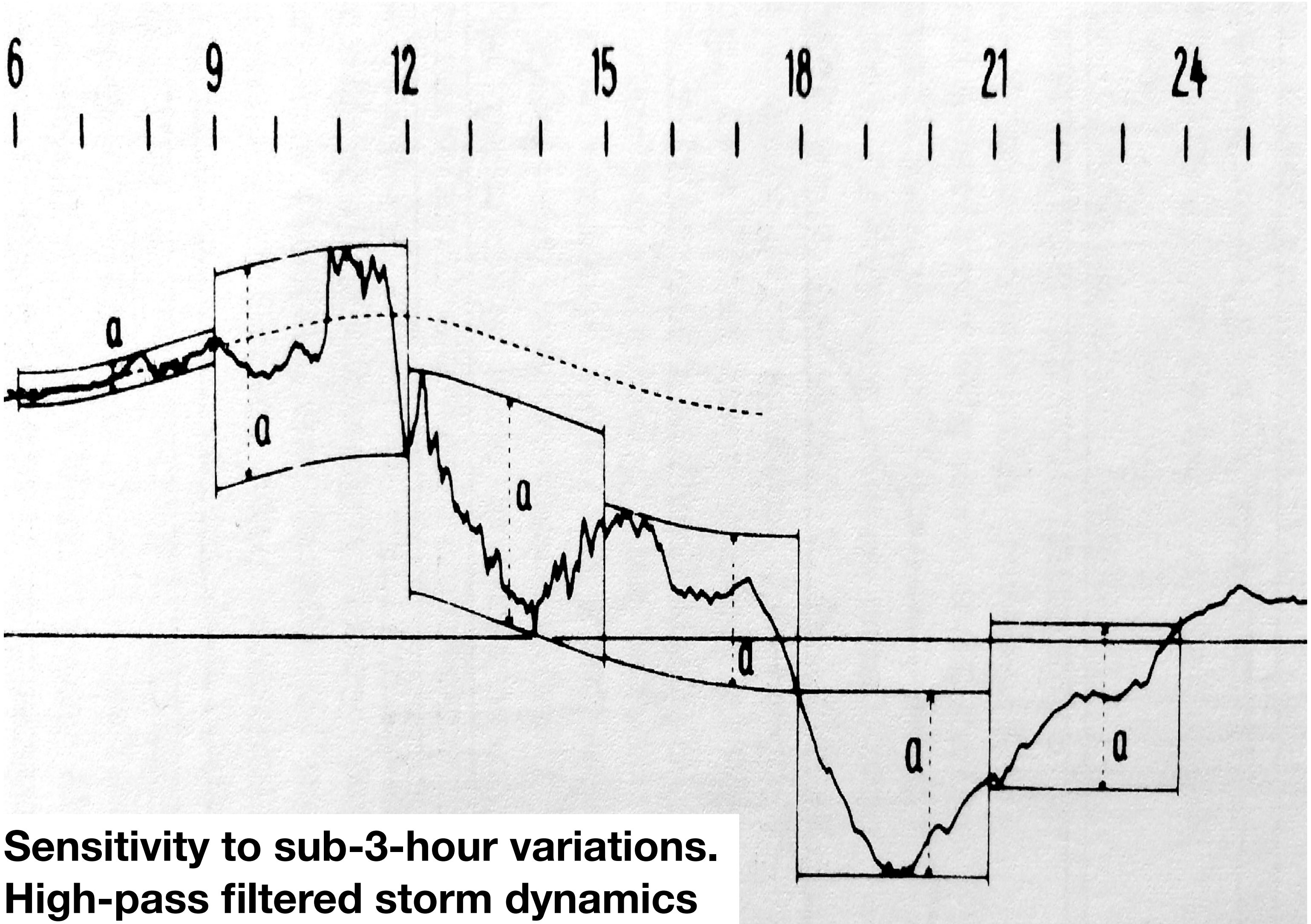
LSTM: A Search Space Odyssey, K. Greff and R. K. Srivastava and J. Koutník and B. R. Steunebrink and J. Schmidhuber, Transactions on neural networks and learning systems 28 2222-2232 (2017)

3-hour averages

Kp(t+3) = f(Bz(t), Bz(t-3), …, n(t), n(t-3), …, V(t), V(t-3), …)

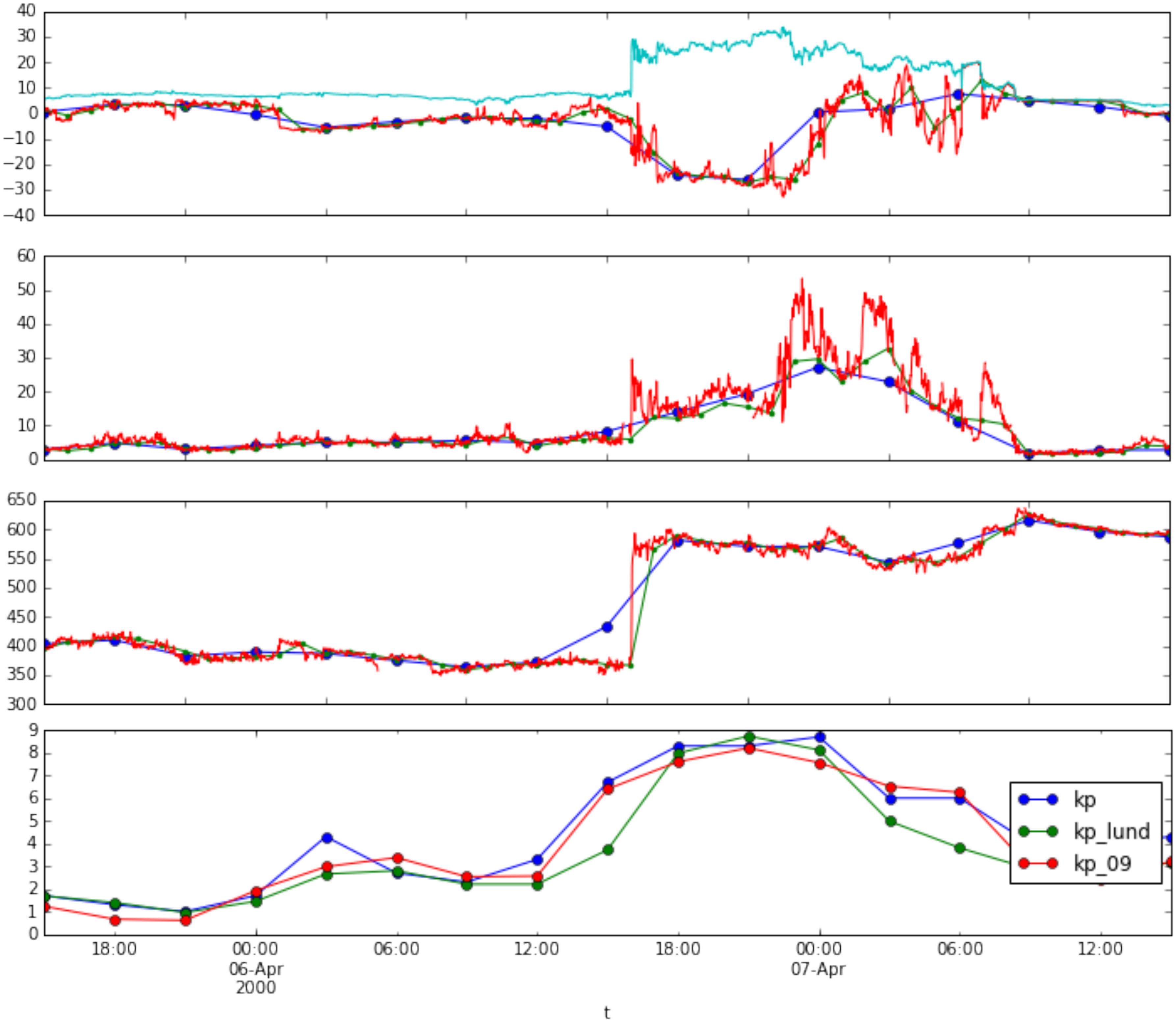


Derivation of Kp

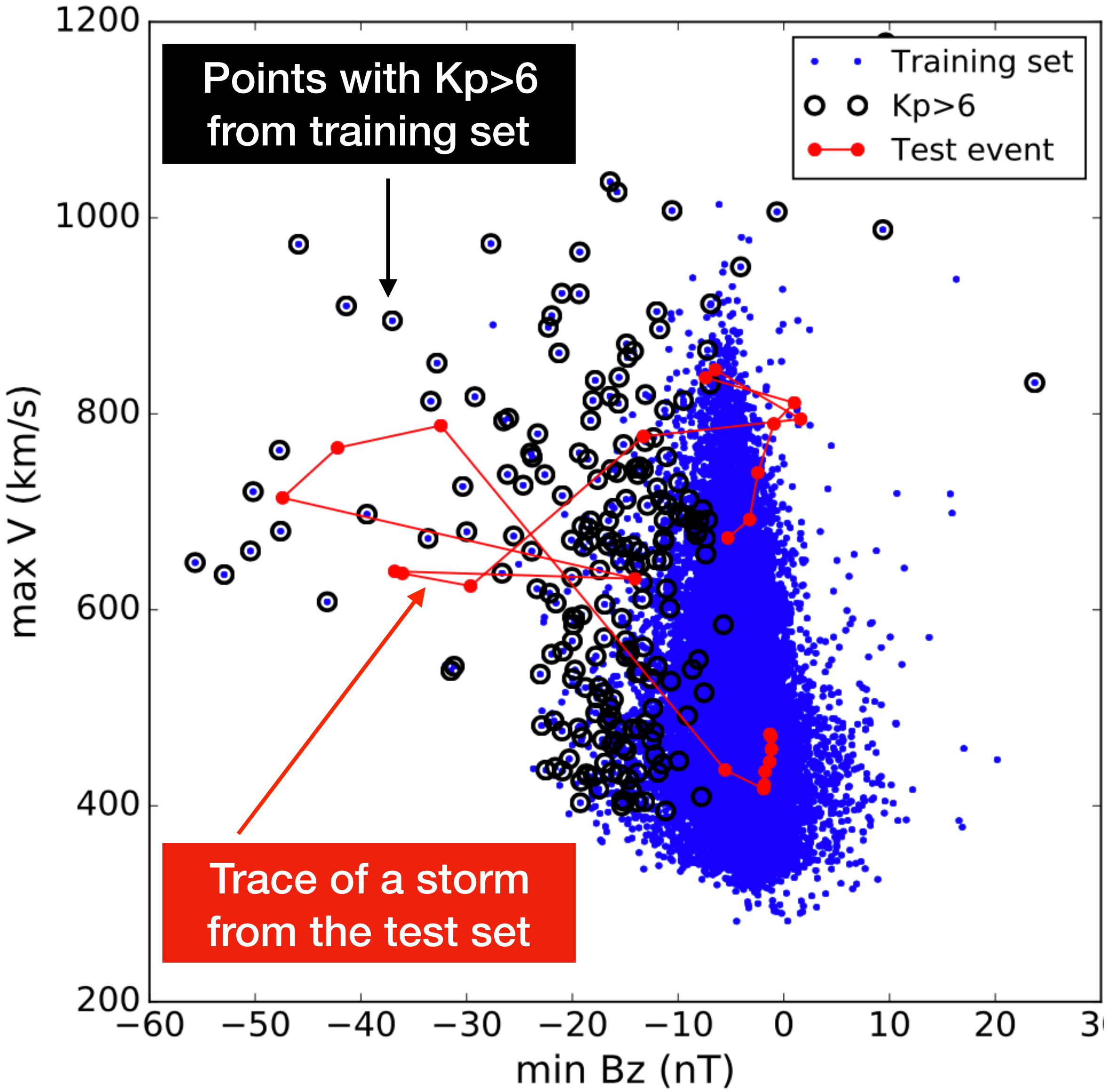


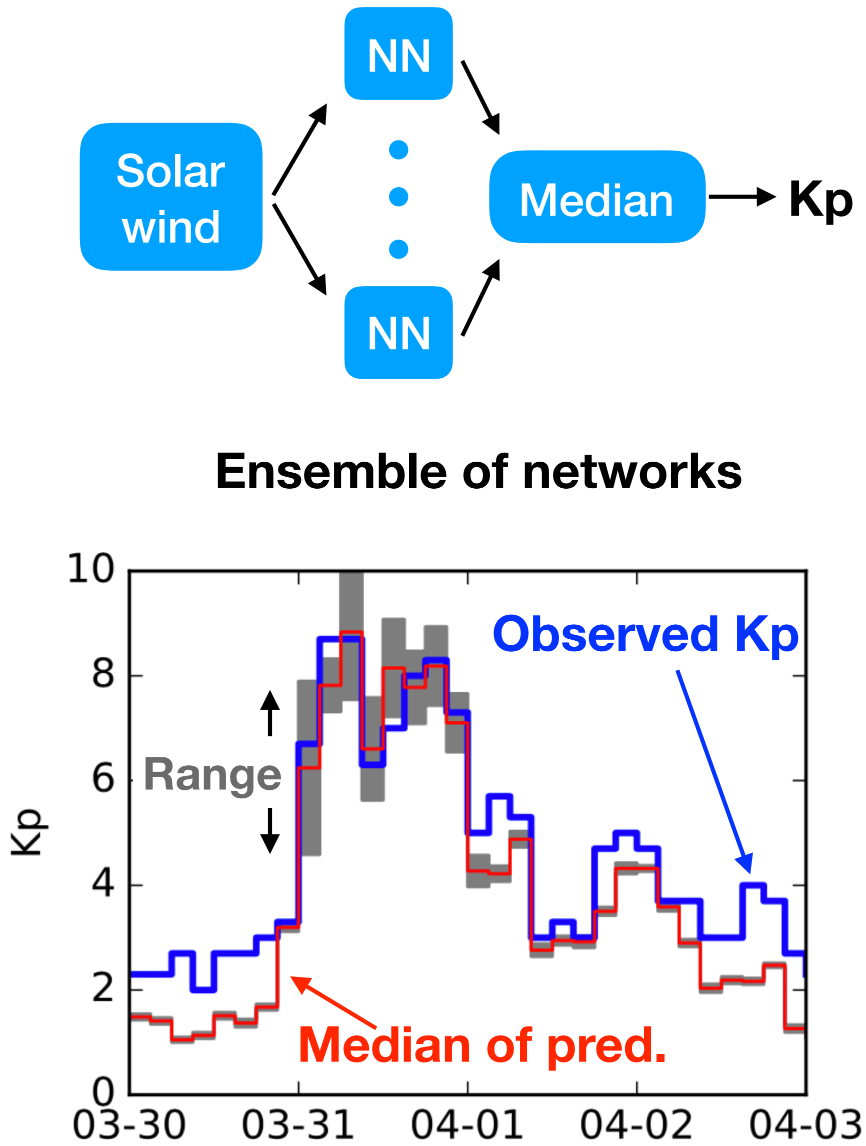
- **Sensitivity to sub-3-hour variations.**
- **High-pass filtered storm dynamics**

Derivation, meaning, and use of geomagnetic indices, P. N. Mayaud, AGU, 22, 1980.


40
30
20
10
0
−10
−20
−30
−40
60
50
40
30
20
10
0
650
600
550
500
450
400
350
300
9
8
7
6
5
4
3
2
1
0
18:00
00:00
06-Apr
2000
06:00
12:00
18:00
00:00
07-Apr
06:00
12:00
t
kp
kp_lund
kp_09

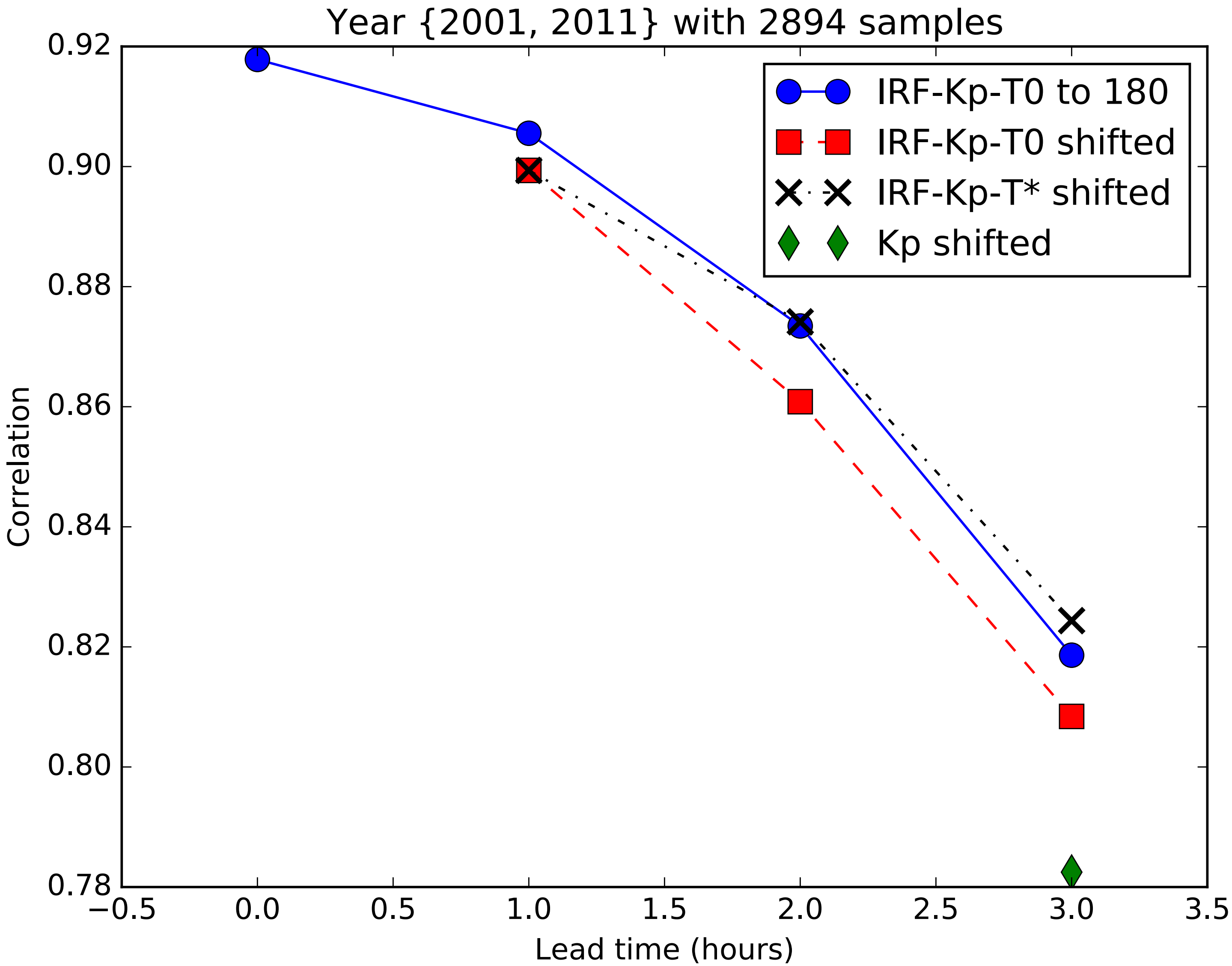
Performance in low-density state-space

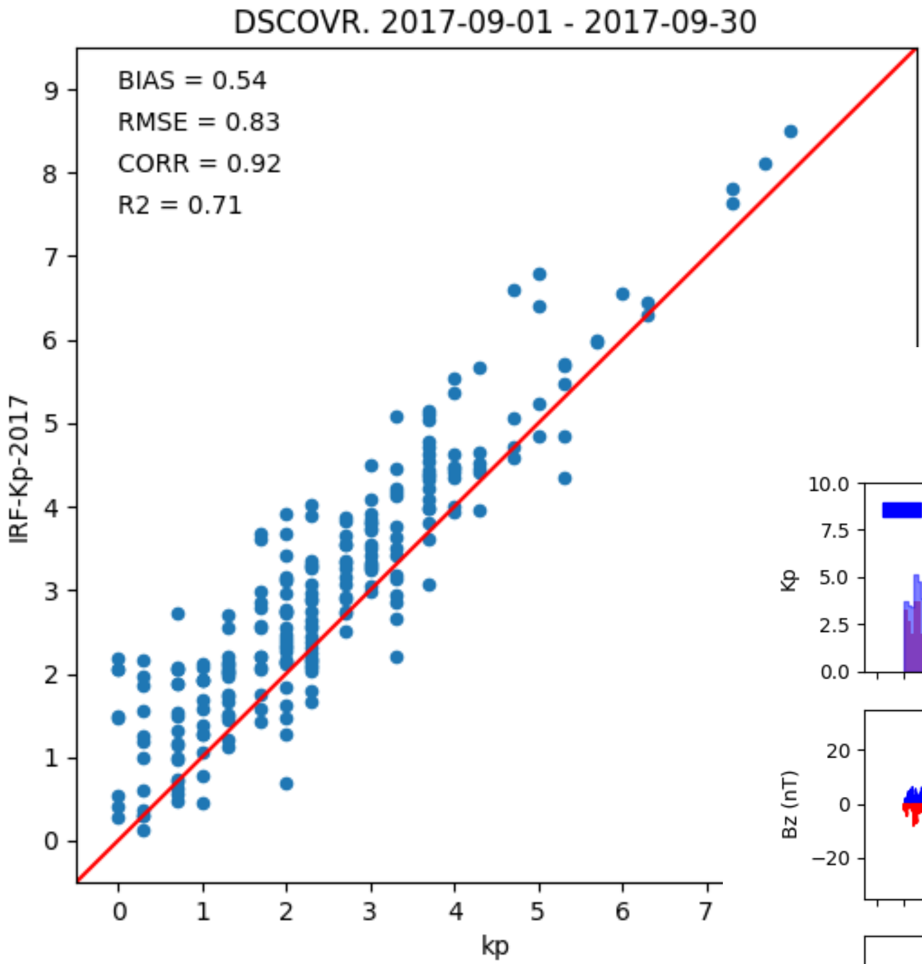




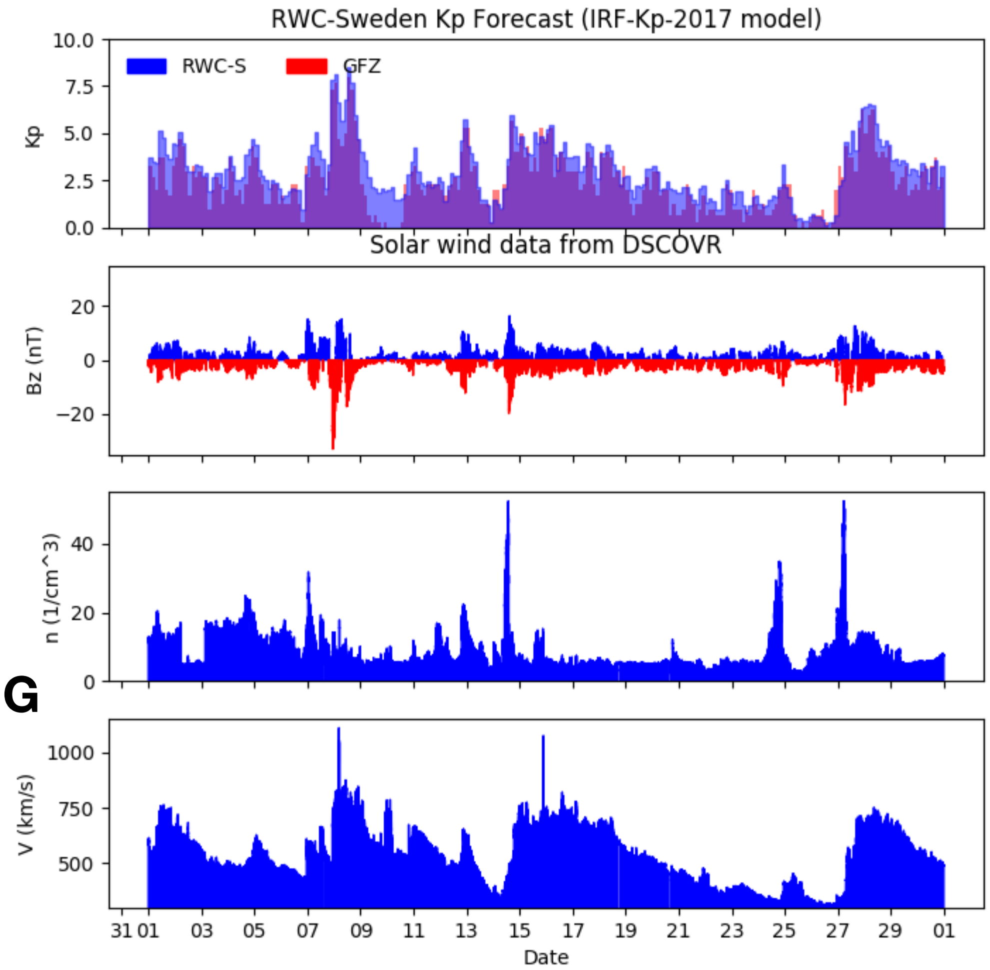
Wintoft P, Wik M, Matzka J, Shprits Y. 2017. Forecasting Kp from solar wind data: input parameter study using 3-hour averages and 3-hour range values. J. Space Weather Space Clim. 7: A29

Increasing Kp prediction lead time?



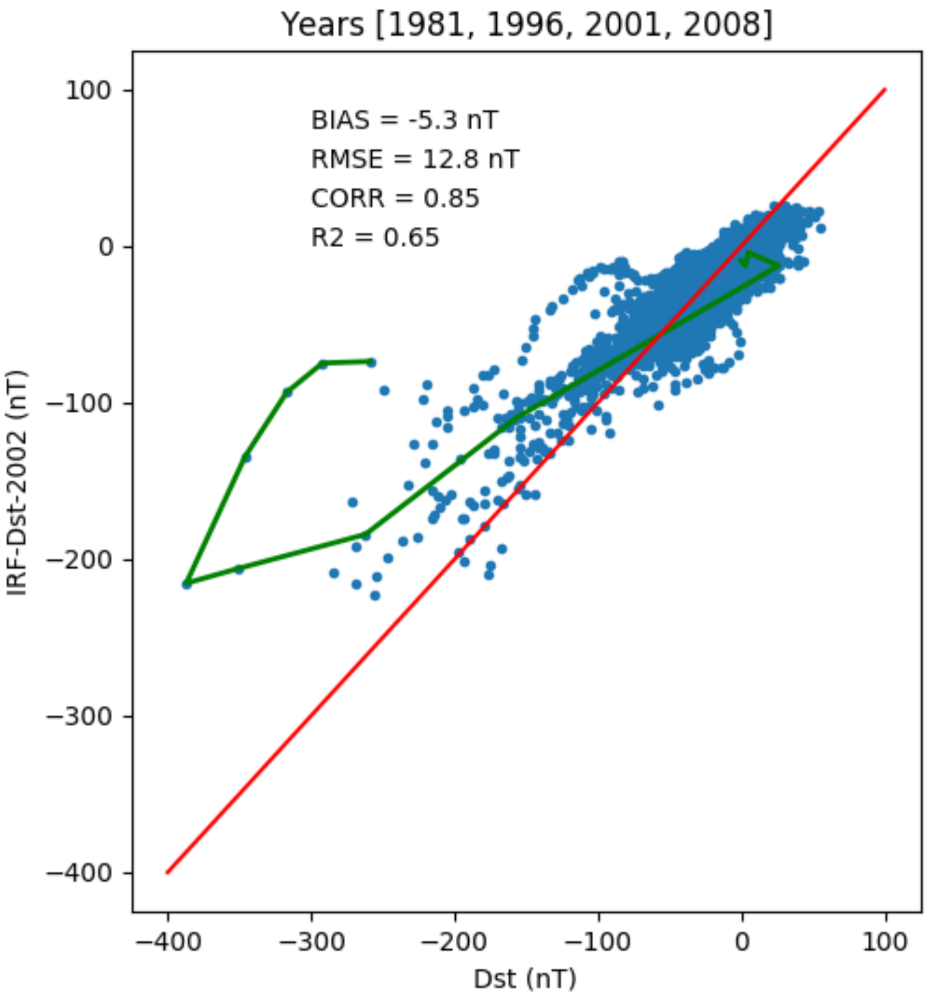


http://lund.irf.se/forecast/kp2017/

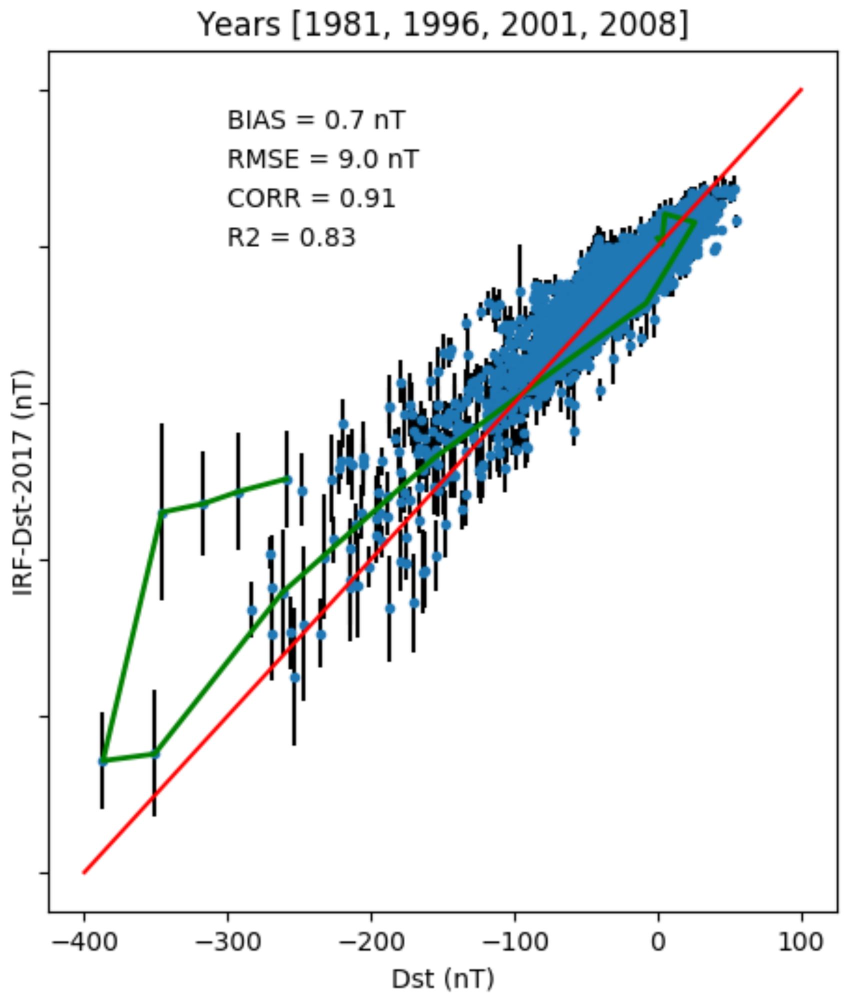


Will be included at ESA SSA ESC-G

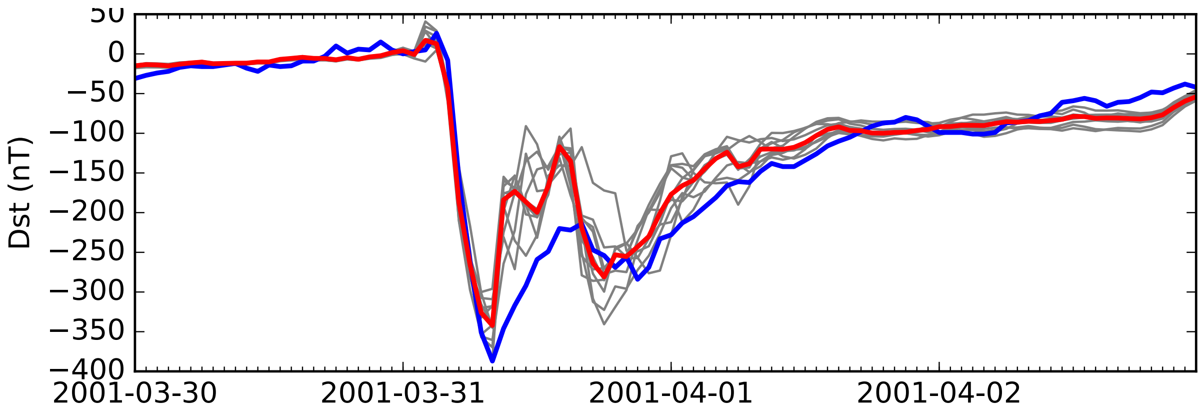
B, Bz, n, V



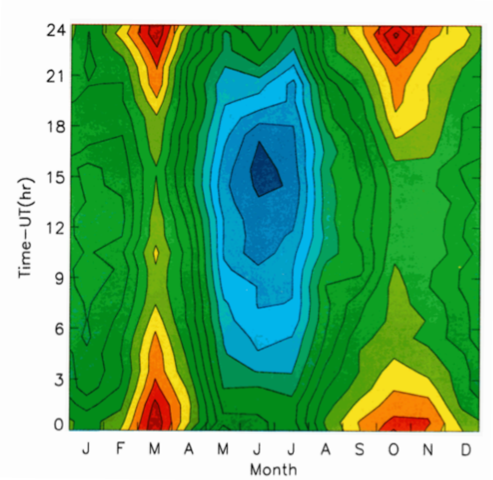
B, By, Bz, n, V, DOY, UT



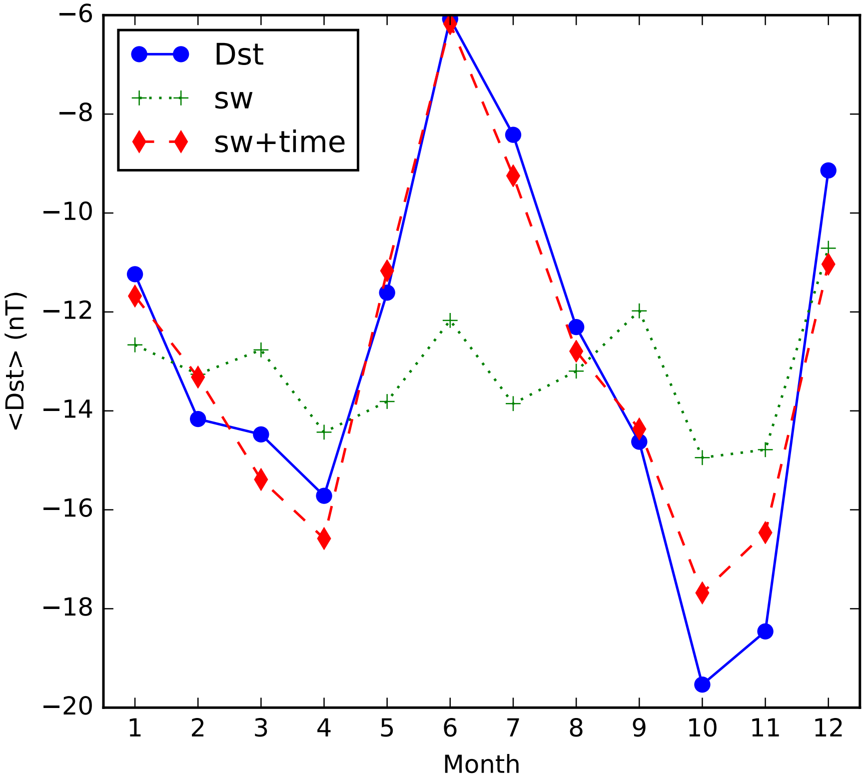
Operational forecasts of the geomagnetic Dst index, H. Lundstedt and H. Gleisner and P. Wintoft Geophysical Research Letters 29, 34-1–4, 2002.

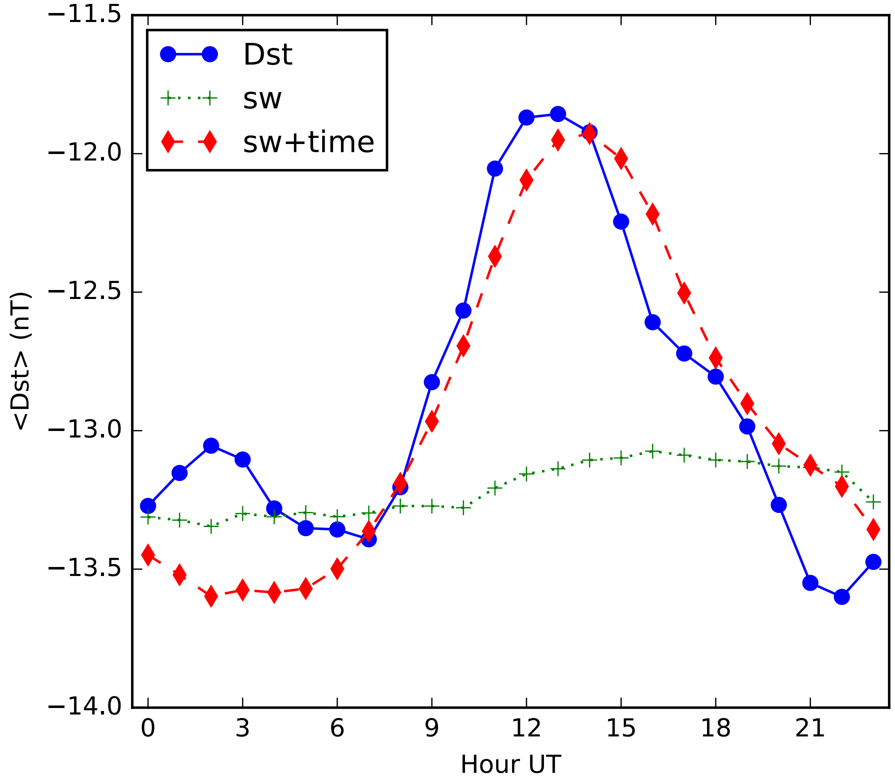


Semiannual variation of Dst

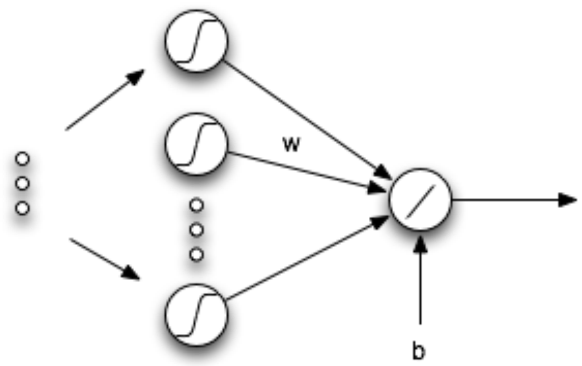


Semiannual variation of the geomagnetic Dst index: Evidence for a dominant nonstorm component, E. W. Cliver and Y. Kamide and A. G. Ling and N. Yokoyama, Journal of Geophysical Research 106 21,297-21,304 (2001)





Range of Dst from NN model



- **Outputs from final hidden layer are limited to [-1,+1].**
- **Sum |weights| + bias => possible range of Dst.**

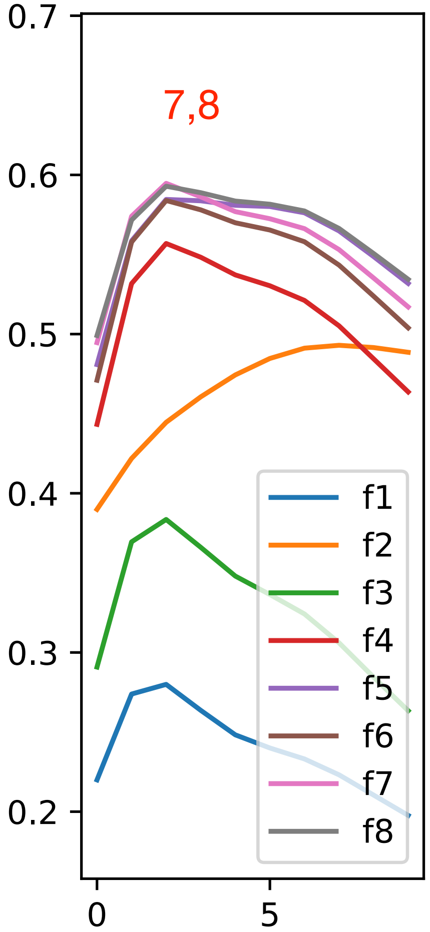
[-650, 190] nT

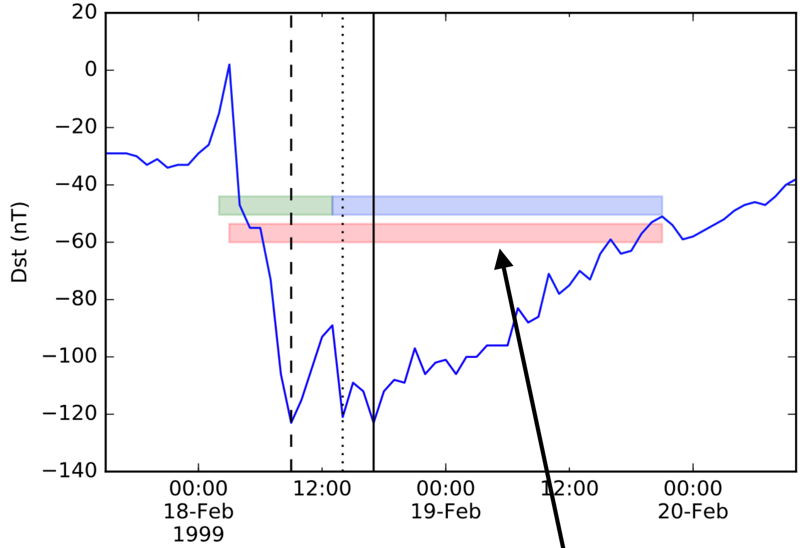
75%: [-500, 140] nT

$$R_{\mathrm{quick}} = 0.4\mu_o{}^{1/2} m_p{}^{1/2} \sin^2(\theta/2)\, C^{-1/2} n_{sw}{}^{1/2} v_{sw}{}^2 (1+\beta_s)^{-3/4}$$

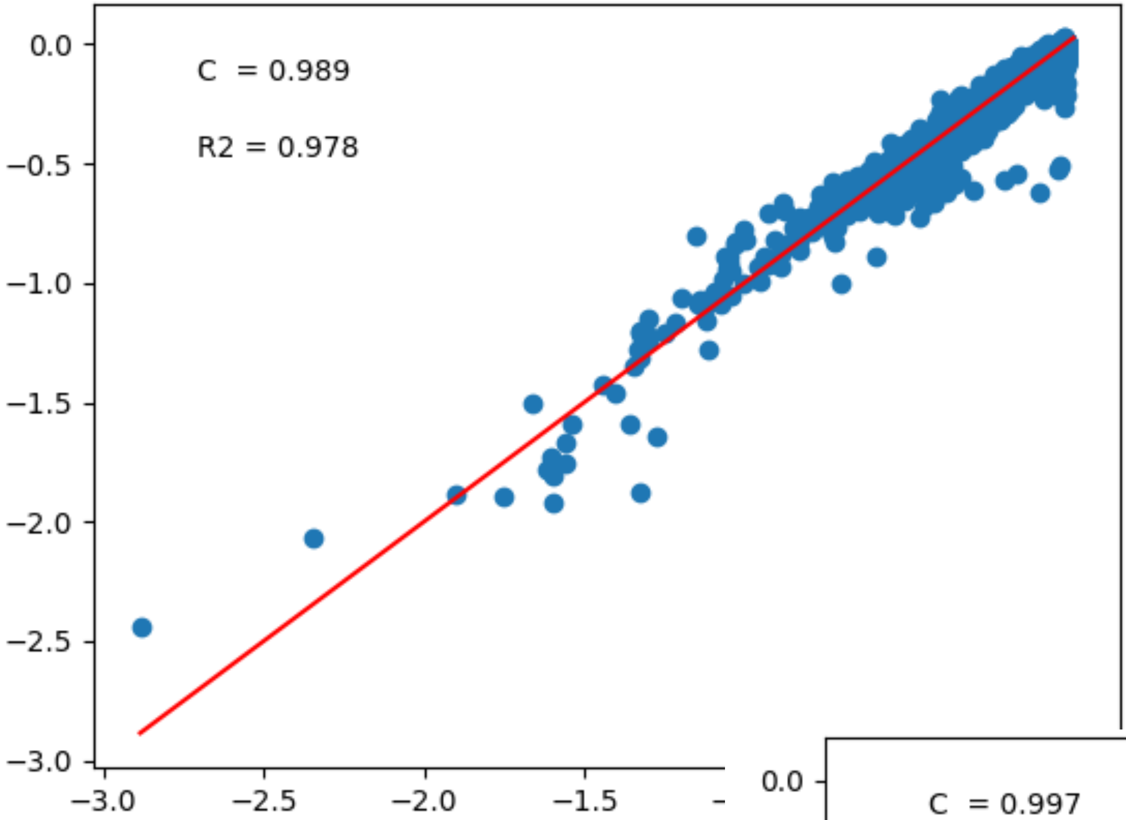
		$-Dst_2$
1	$\sin^2(\theta/2)$	0.279
2	$vB_\perp$	0.450
3	vB_z	0.384
4	vB_{south}	0.558
5	$vB_\perp \sin^2(\theta/2)$	0.587
6	$vB_\perp \sin^4(\theta/2)$	0.586
7	$v^{4/3} B_\perp{}^{2/3} \sin^{8/3}(\theta/2)$	0.596
8	R_{quick}	0.594

The solar wind electric field does not control the dayside reconnection rate, J. E. Borovsky and J. Birn, Journal of Geophysical Research: Space Physics 119 751--760 (2014)

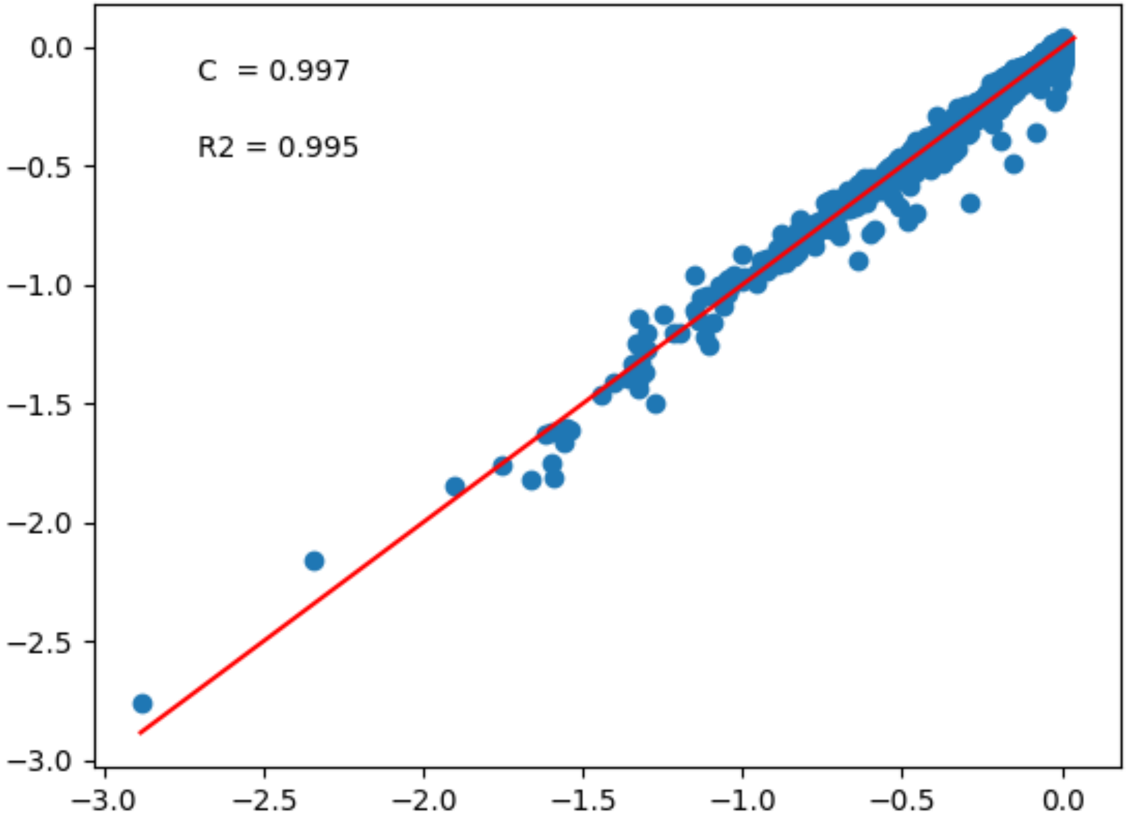




Echer, E., Gonzalez, W. D., Tsurutani, B. T. & Gonzalez, A. L. C. (2008), ‘Interplanetary conditions causing intense geomagnetic storms (Dst <-100 nT) during solar cycle 23 (1996–2006)’, Journal of Geophysical Research 113, A05221.



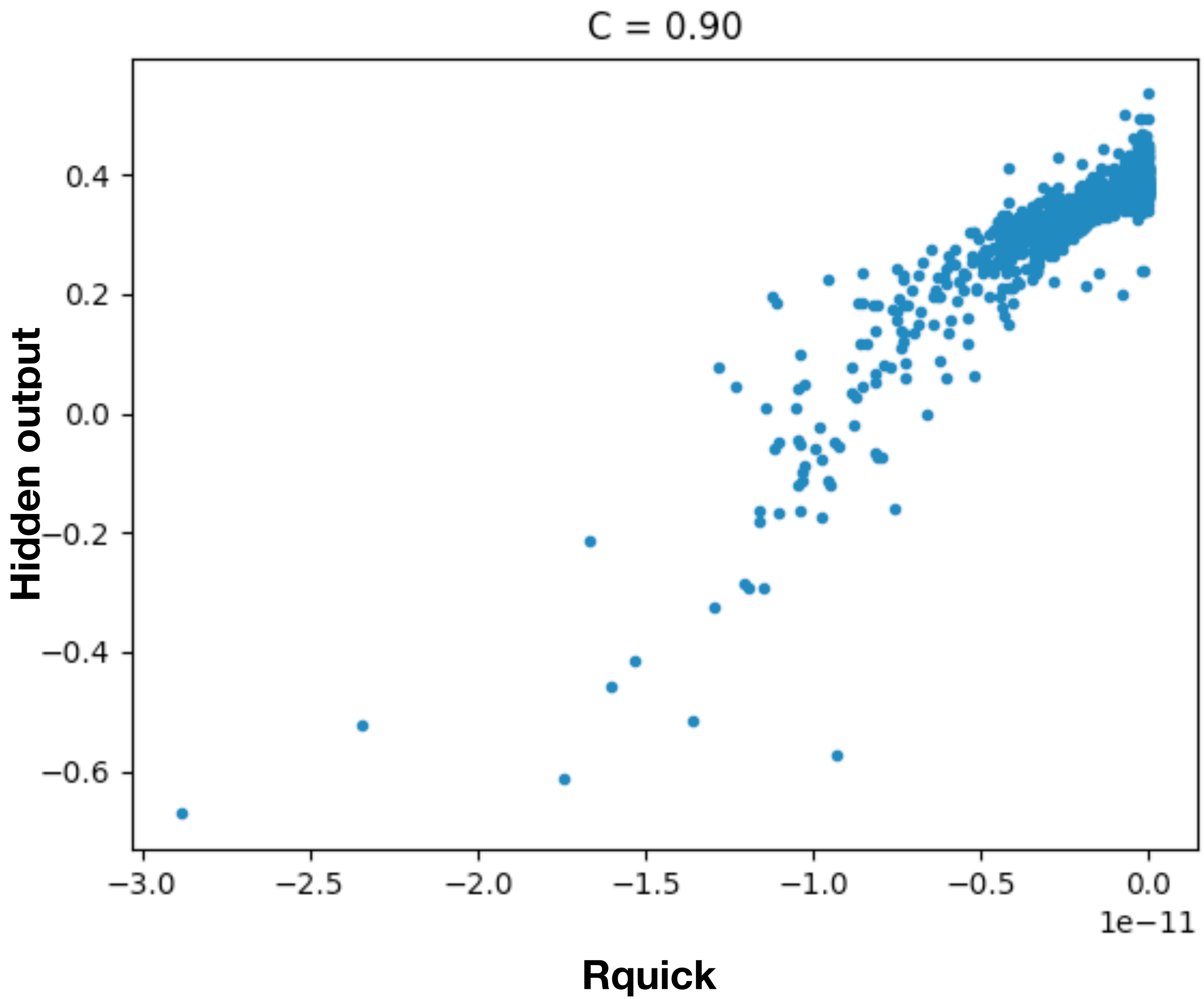
5,10,1 (71 weights)



5,10,10,1 (181 weights)

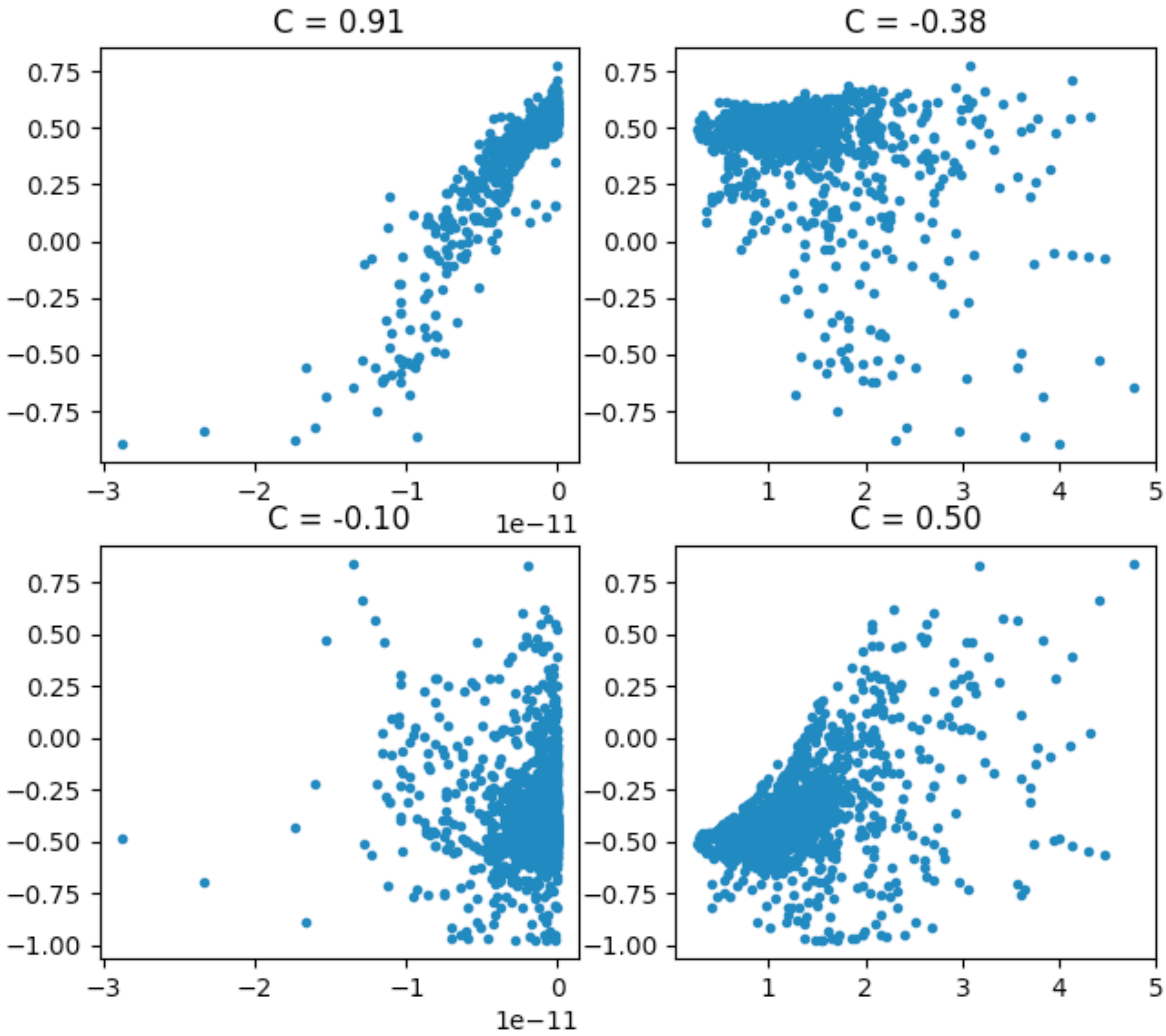
NN hidden output

5 inputs, 10 hidden, 1 hidden, 4 recurrent, 1 output



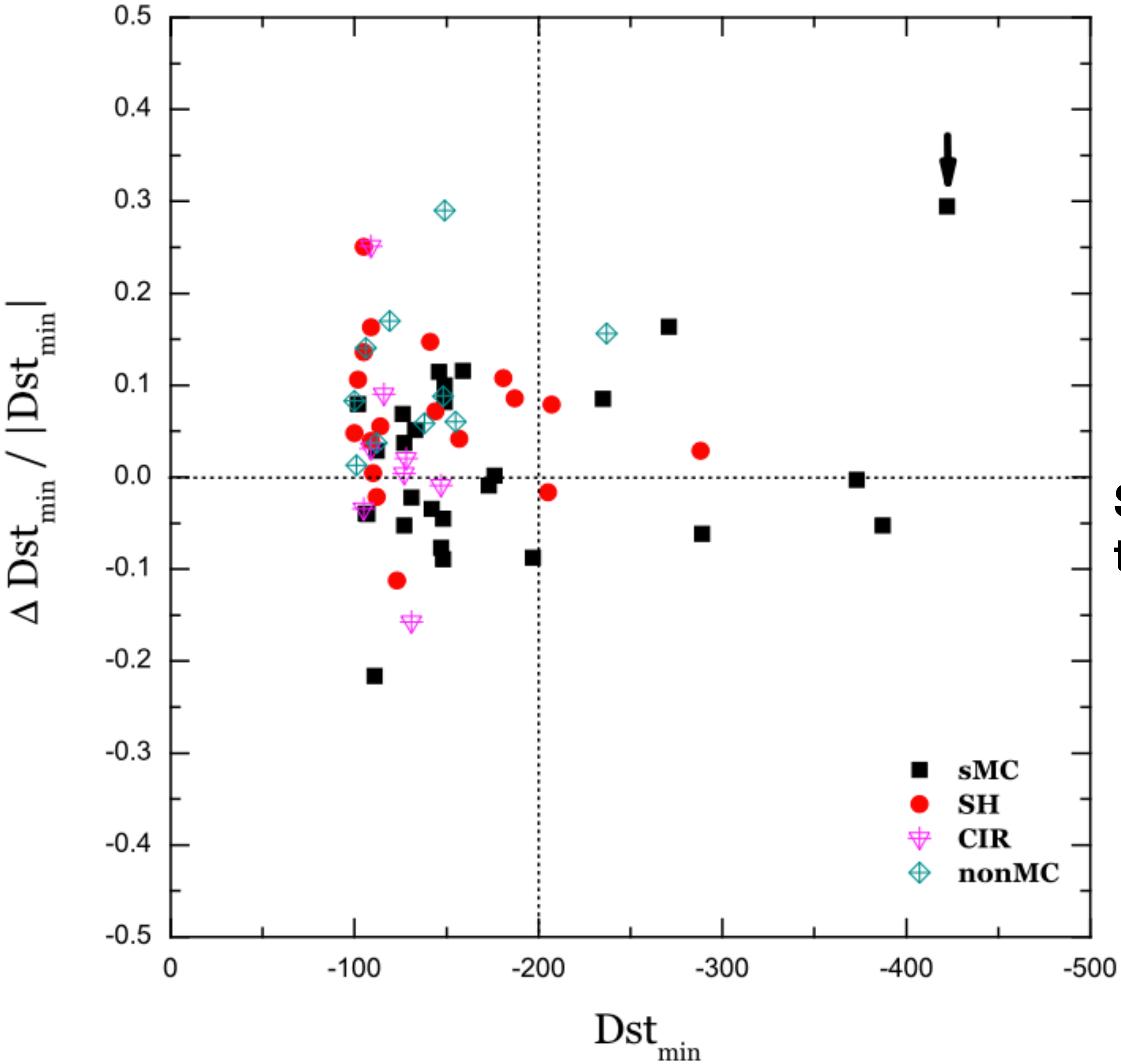
Node 1

Node 2



Rquick

sqrt(Pressure)



Some events belong to the "training set"!!

Figure 5. Relationship between $\Delta Dst_{\min}/|Dst_{\min}|$ and $|Dst_{\min}|$ for the TL model. Here the point symbols represent the four interplanetary structures. Super storms ($Dst \leq -200$ nT) are divided by the dotted vertical line.

Comparison of Dst forecast models for intense geomagnetic storms, E.-Y. Ji and Y.-J. Moon and N. Gopalswamy and D.-H. Lee, Journal of Geophysical Research 117 A03209 (2012)

Dst model for 1995–2002, M. Temerin and X. Li, Journal of Geophysical Research 111 A04221 (2006)

"Biological NN"

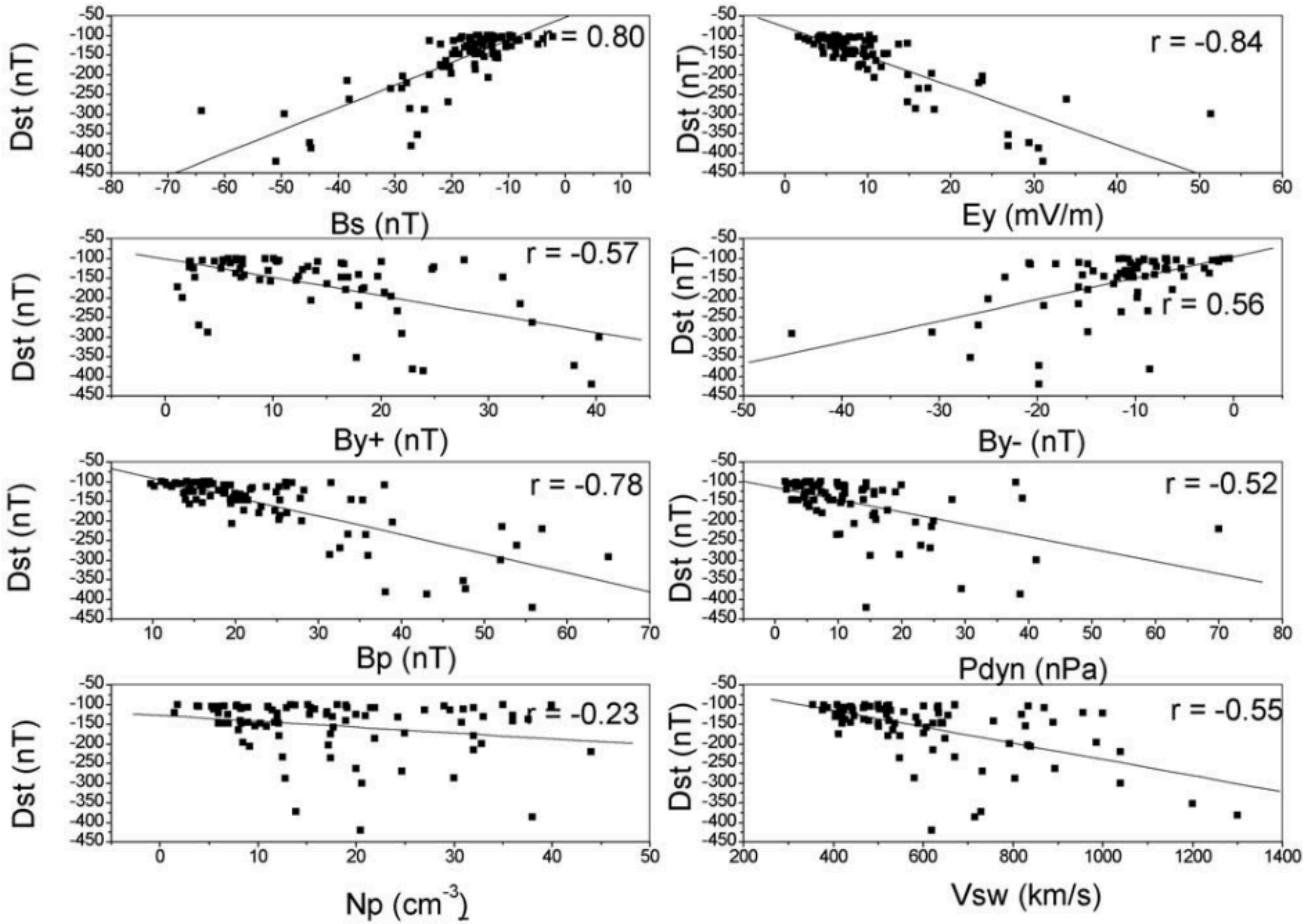
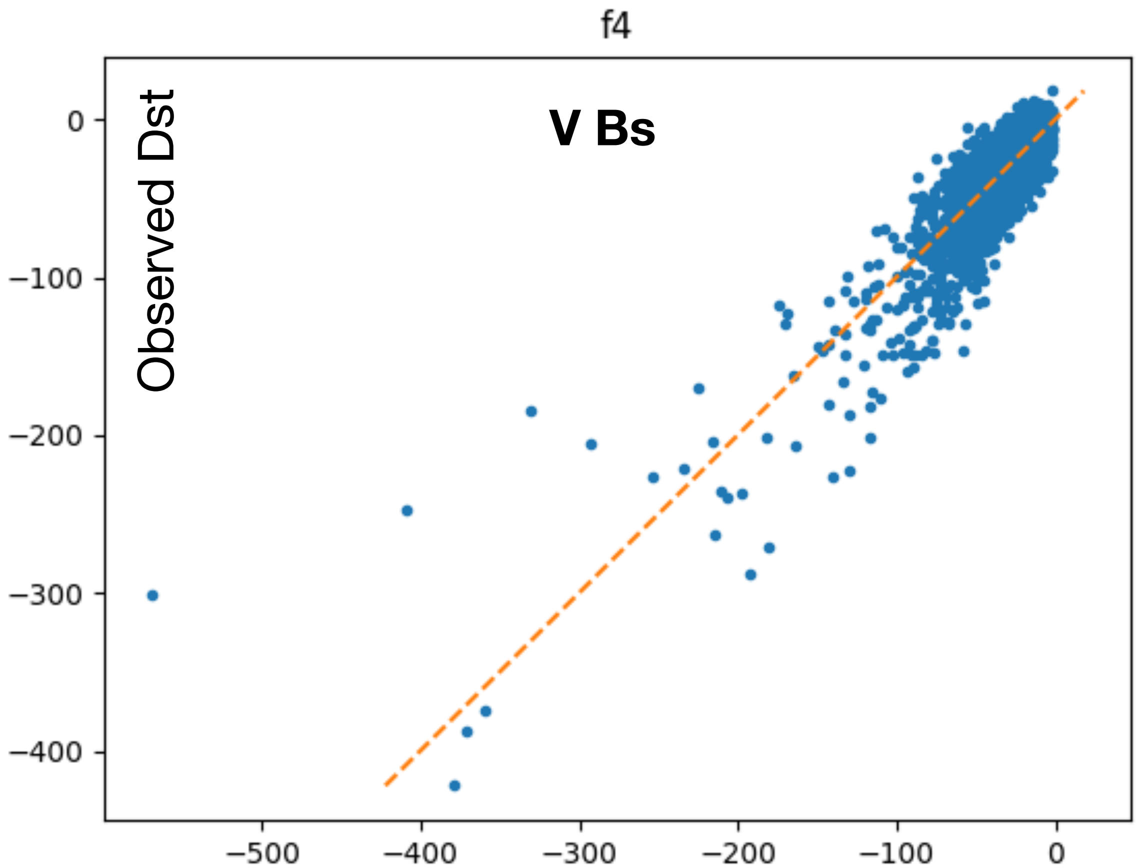
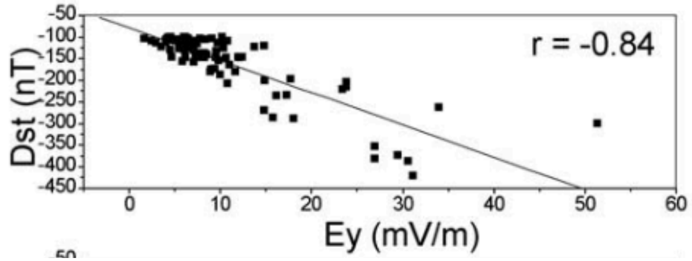
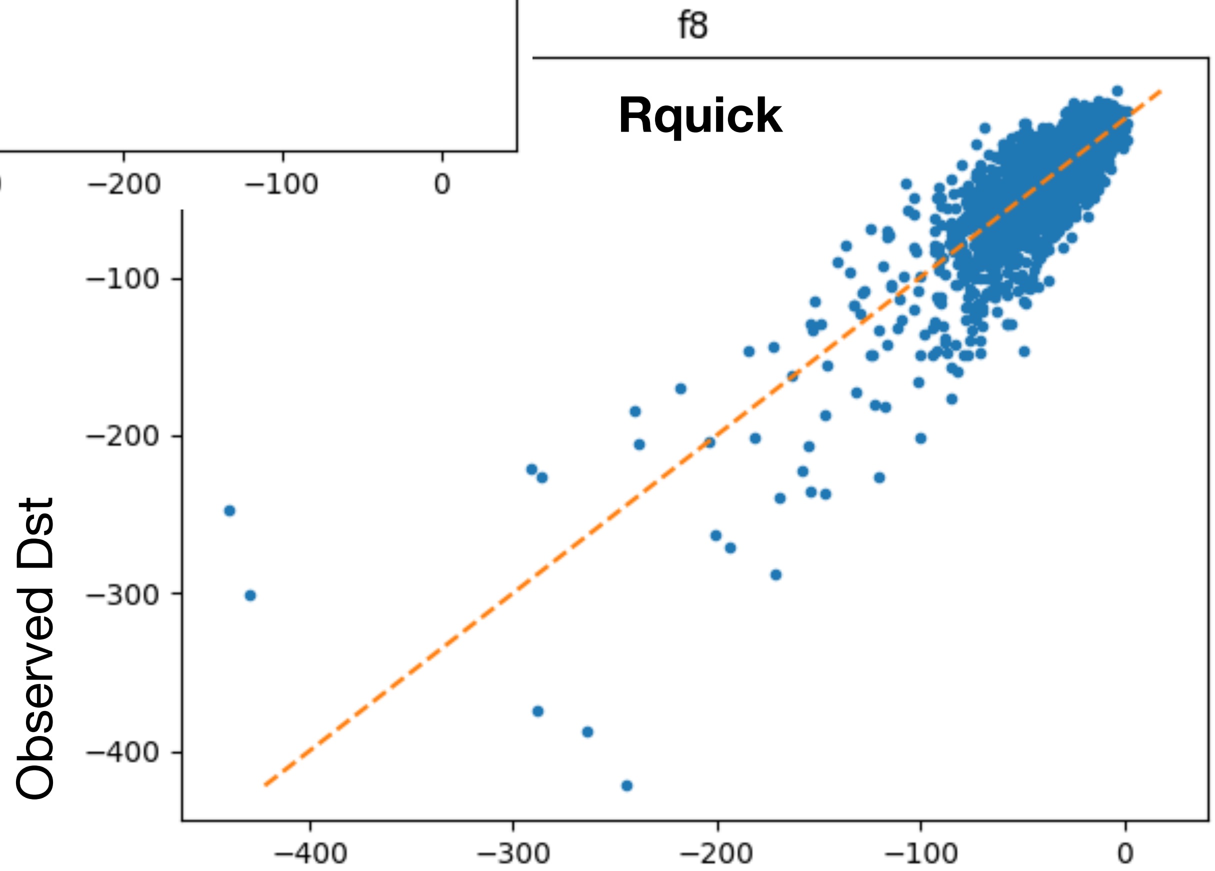


Figure 9. Correlation between peak Dst and interplanetary parameters: peak Bs, Ey, By+, By−, Bp, Pdyn, Np, and Vsw.

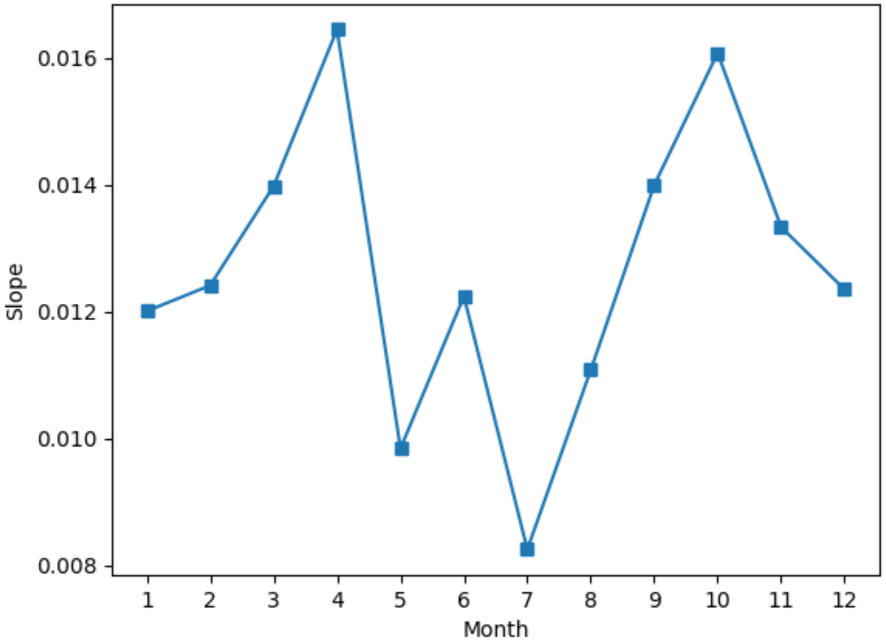
Interplanetary conditions causing intense geomagnetic storms (Dst $\leq$ -100 nT) during solar cycle 23 (1996–2006), E. Echer and W. D. Gonzalez and B. T. Tsurutani and A. L. C. Gonzalez, Journal of Geophysical Research 113 A05221 (2008)

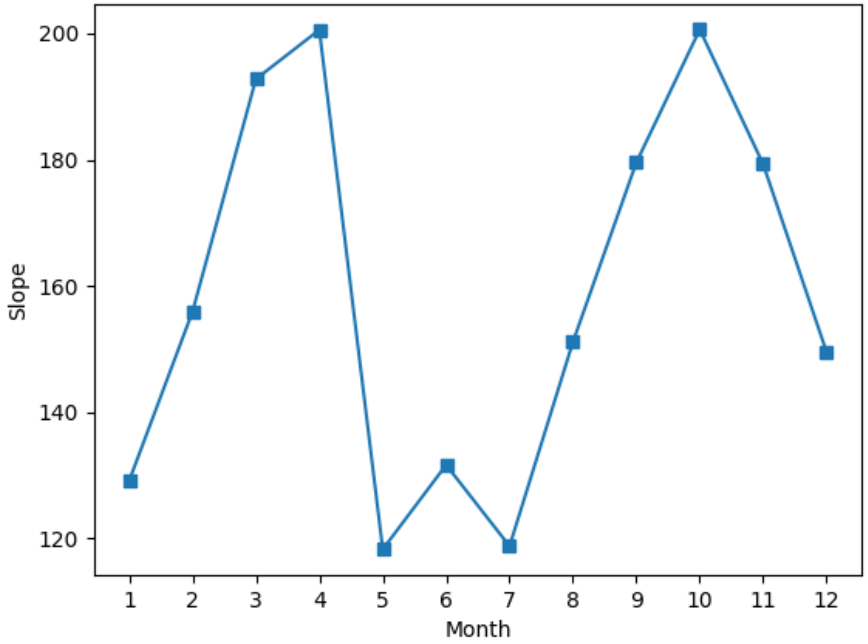




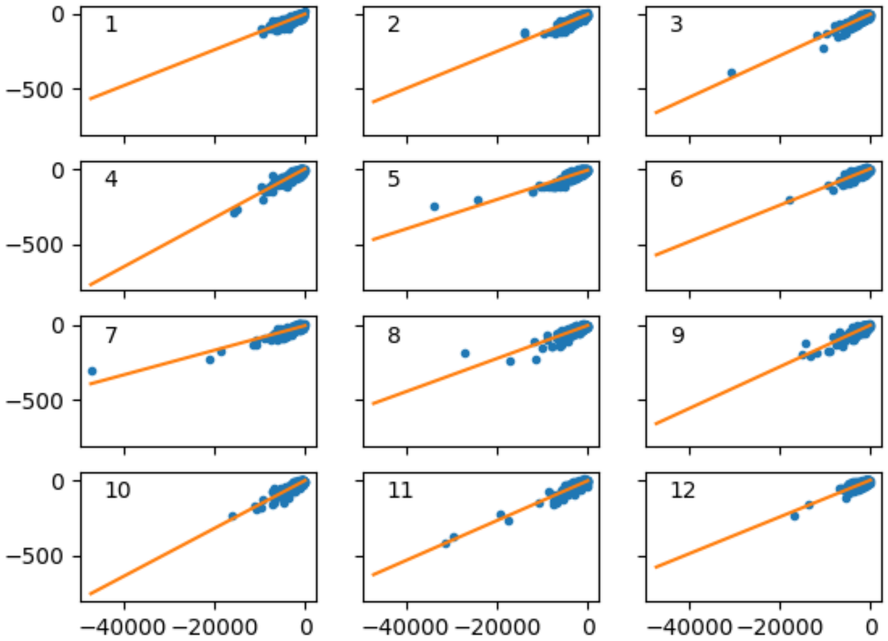


- **Storms identified using MODWT filtering.**
- **Determine storm minimum.**

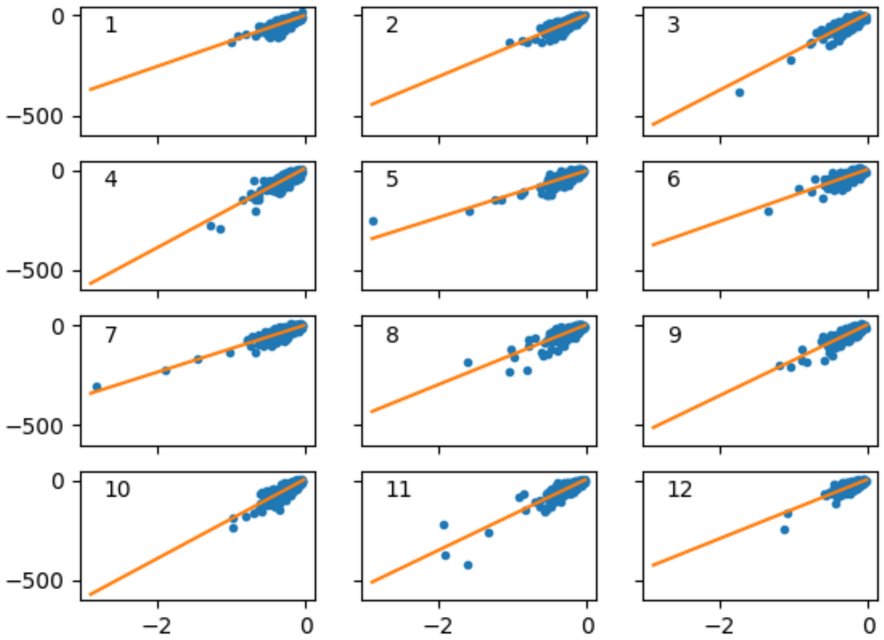


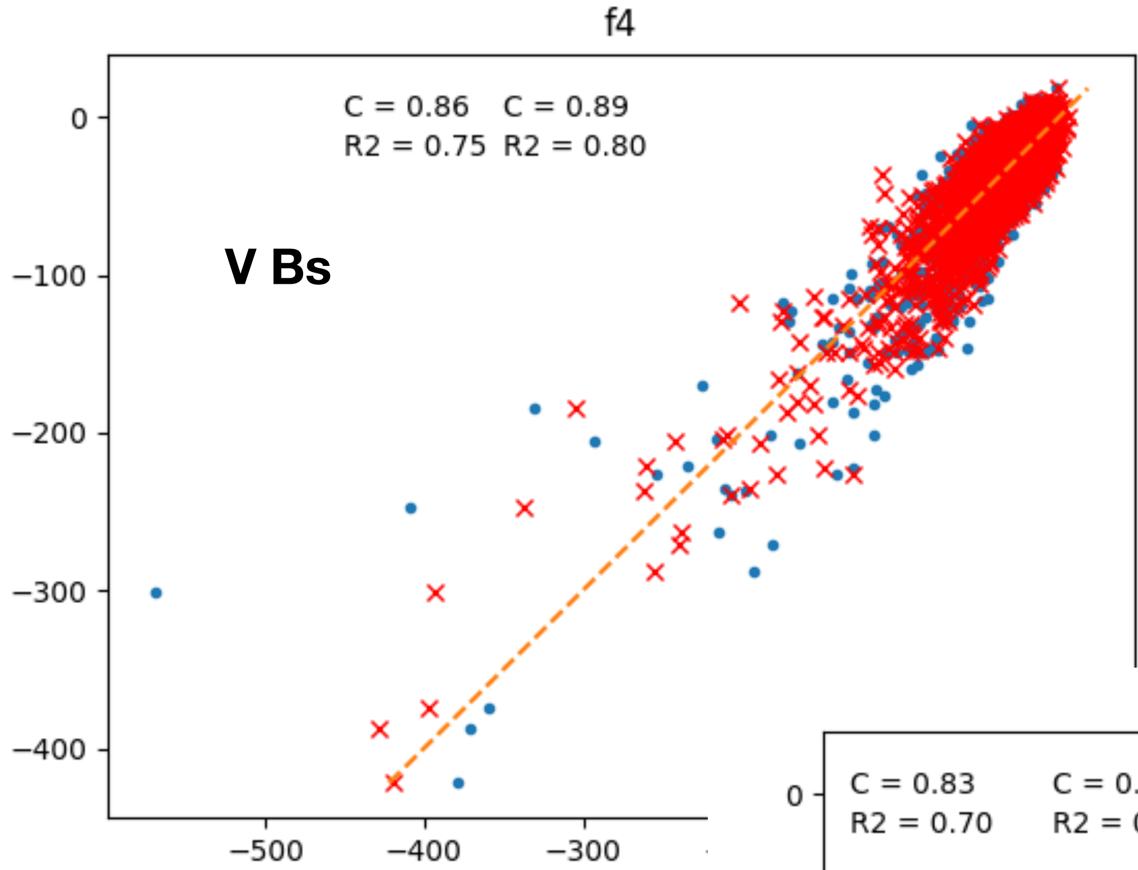


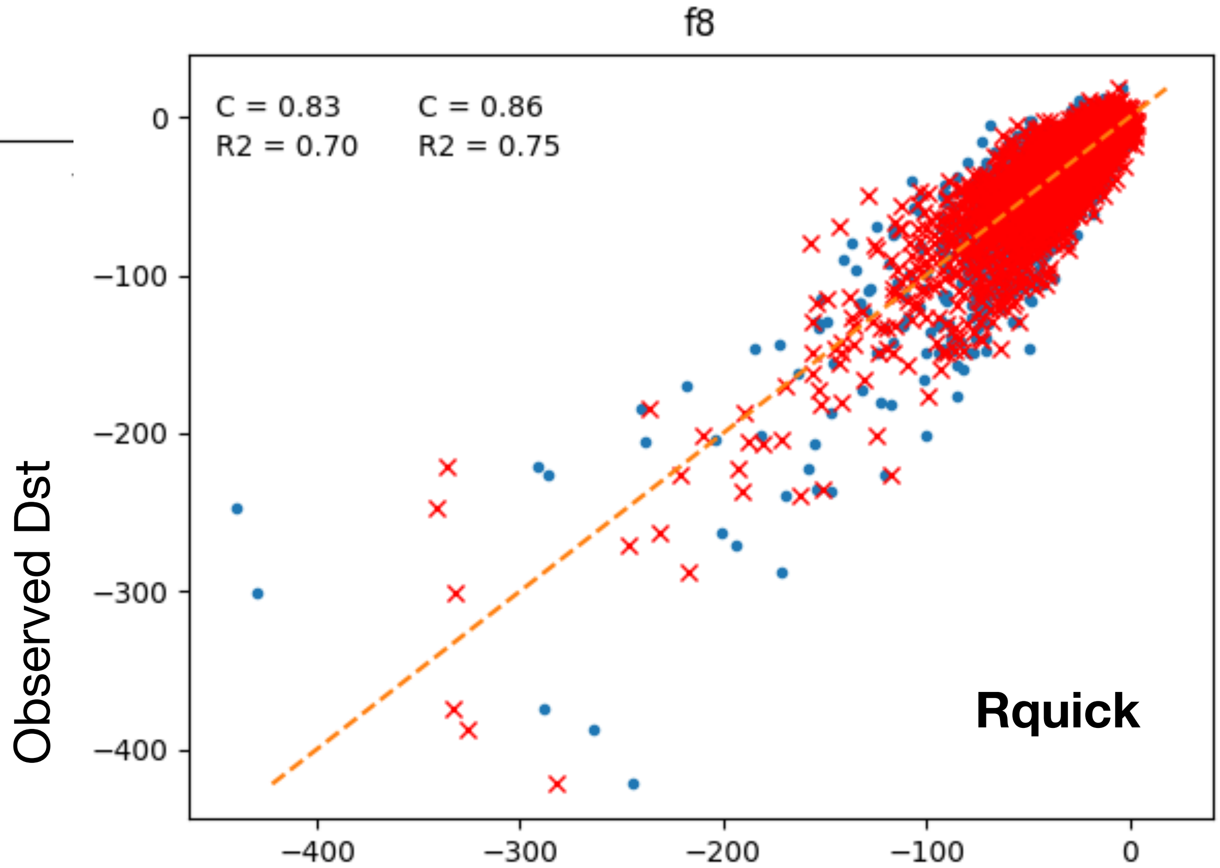
f4: V Bs



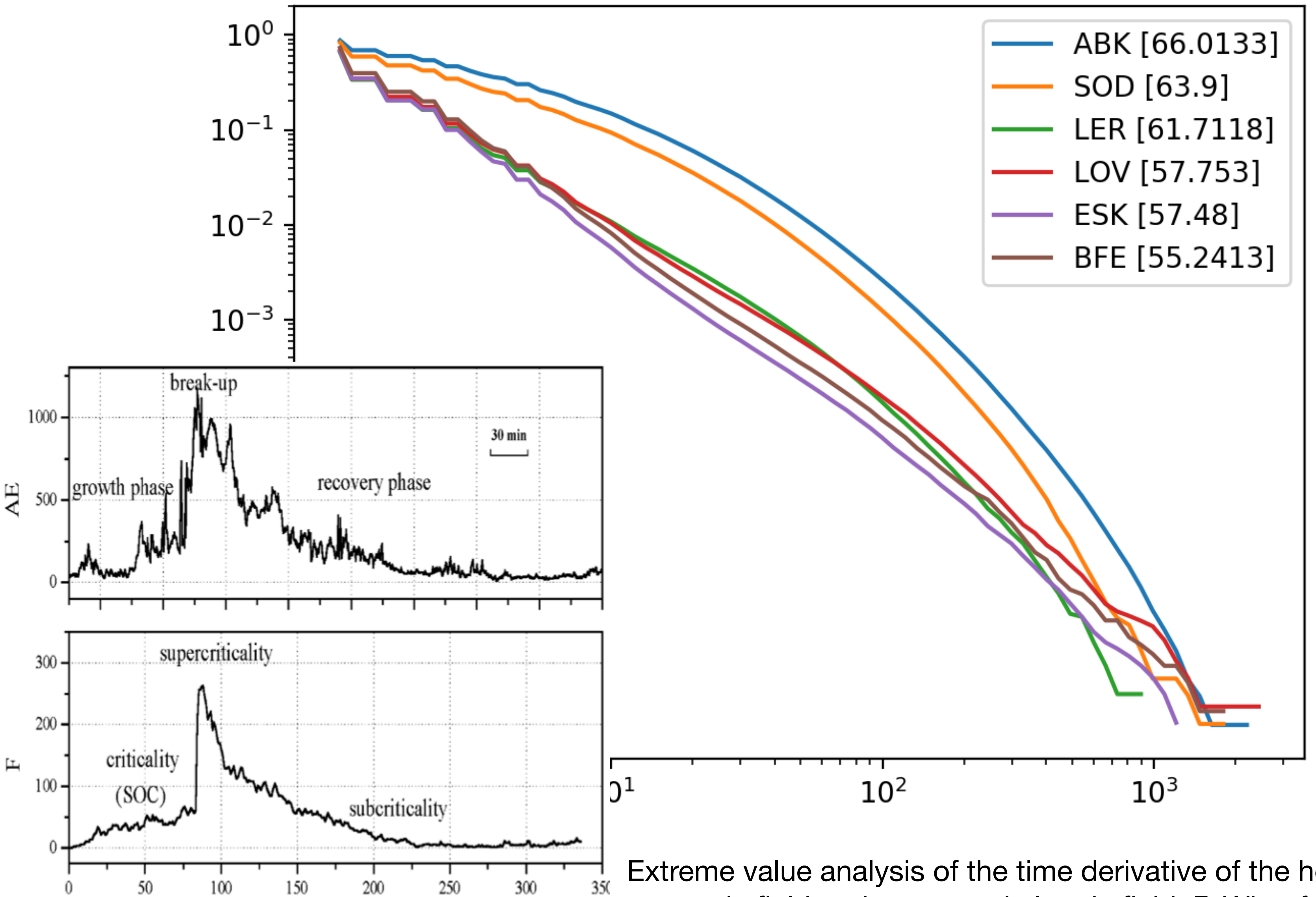
f8: Rquick




f4
C = 0.86 C = 0.89
R2 = 0.75 R2 = 0.80
V Bs
0
−100
−200
−300
−400
−500
−400
−300


f8
C = 0.83 C = 0.86
R2 = 0.70 R2 = 0.75
Observed Dst
Rquick
0
−100
−200
−300
−400
−400
−300
−200
−100
0

Solar wind driven empirical forecast models of the time derivative of the ground magnetic field, P. Wintoft and M. Wik and A. Viljanen, Journal of Space Weather and Space Climate 5 A7 (2015)



Geomagnetic substorms as perturbed self-organized critical dynamics of the magnetosphere, V. Uritsky and M. Pudovkin and A. Steen, Journal of Atmospheric and Solar-Terrestrial Physics 63 1415-1424 (2001)

Extreme value analysis of the time derivative of the horizontal magnetic field and computed electric field, P. Wintoft and A. Viljanen and M. Wik, Annales Geophysicae 34 485--491 (2016)

Quantifying extreme behavior in geomagnetic activity, A. Thomson and E. B. Dawson and S. J. Reay, Space Weather, 9, S100001 (2011)

Related

- Support Vector Machines:
  - Both classification and regression
  - Global minima
  - Support vectors drawn from training vectors
  - Nonlinearity through transformation to higher-dimensional space, but the space need not to be known only that there exist a dot-product => kernel trick (e.g. RBF corresponds to an infinite-dimensional space)
  - Example: Comparison of neural network and support vector machine methods for Kp forecasting, E.-Y. Ji and Y.-J. Moon and J. Park and J.-Y. Lee and D.-H. Lee, Journal of Geophysical Research 118 5109-5117 (2013)
- Nonlinear Auto-Regressive Moving Average with eXogenous inputs (NARMAX):
  - A "library" of terms (e.g. polynomials or other functions) are ranked and combined including time lags.
  - Will reveal a function that is more assessable for interpretation than SVM or NN.
  - Example: Data derived NARMAX Dst model, R. J. Boynton and M. A. Balikhin and S. A. Billings and A. S. Sharma and O. A. Amariutei, Annales Geophysicae 29 965-971 (2011).

Summary

- The selection of training, validation, and test data is important but can also be challenging:
  - Data gaps (esp. hard for time-series).
  - Division of data based on output distribution, but what about multidimensional inputs?
- Both NN and SVM will be limited by the range of the training data.
  - Use them in their valid regimes.
  - Is extrapolation improved by handling known non-linearities?
- Should simpler models be used at the extremes?
  - How to determine extreme?

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 637302, and the ESA SSA Space Weather ESC contract No 4000113185/15/D/MRP.