

Predicting Time Lagged Effects of Solar Activity

Mandar Chandorkar^{1,2}, Cyril Furthlener², Bala Poduval³, Enrico Camporeale^{1,4,5}, Michele Sebag²

CWI

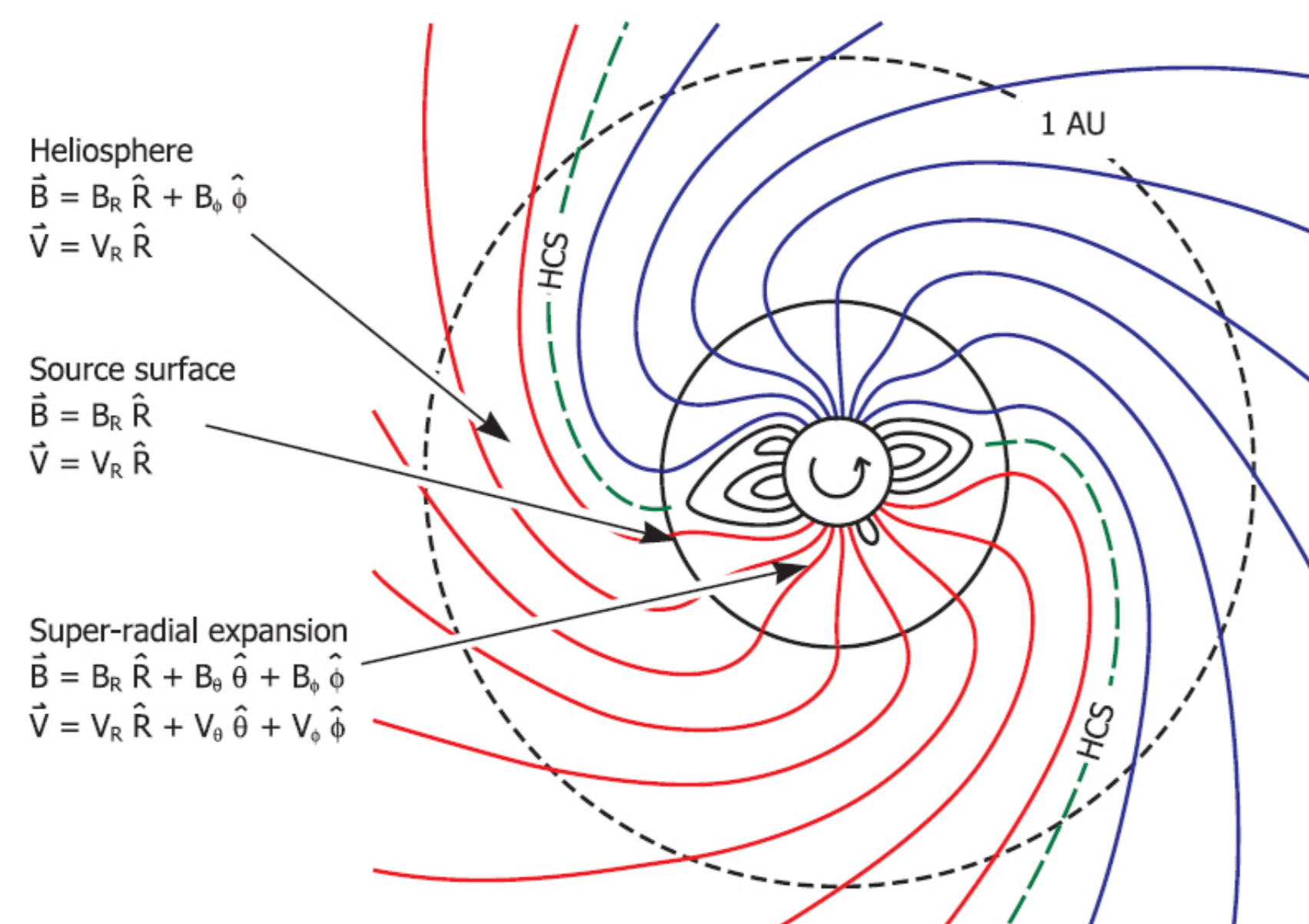
Centrum Wiskunde & Informatica

1 CWI, Amsterdam, 2 TAU, INRIA Paris-Saclay 3 UNH, New Hampshire 4 CIRES, CU, Boulder 5 NOAA, SWPC

www.mlspaceweather.org

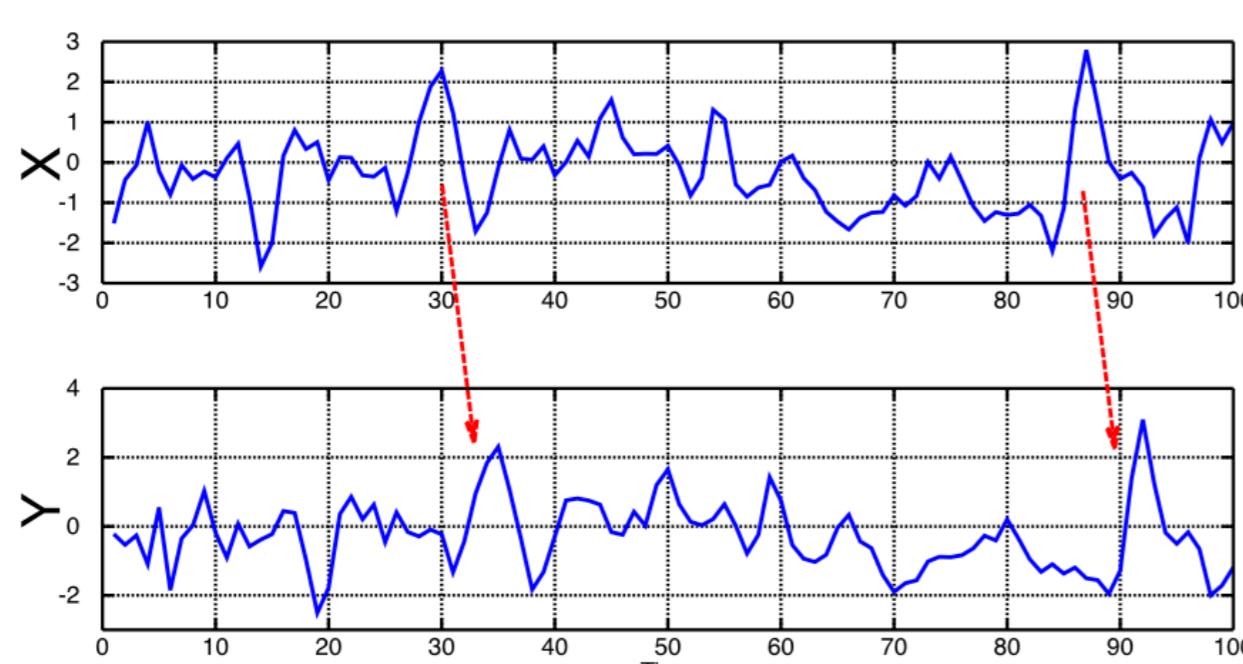
Inria
INVENTEURS DU MONDE NUMÉRIQUE

Origin of the Solar Wind



Supersonic expansion of the solar magnetic field leads to the formation of the *solar wind*. The charged particles follow the *Parker spiral* into the inter-planetary medium. Forecasting solar wind speed at 1 AU from solar observations is a challenging task because the propagation time of the solar wind is uncertain (2 to 5 days) and dynamic.

Dynamic Time-Lag Regression



Learning the time delay between causes and effects, from data, is an important and challenging problem. We call this problem *Dynamic Time Lag Regression* (DTLR) and formulate it as follows.

- Inputs:** Two time series, the causes $x(t)$ and the observed effects $y(t)$
- Outputs:** A function $f : \mathcal{X} \rightarrow \mathbb{R}$ which maps each input pattern $x(t_1)$ to an output $y(t_2)$, and $g : \mathcal{X} \rightarrow \mathbb{R}^+$ which maps the time delay between the input and output patterns $t_2 = t_1 + g(x(t_1))$.

$$y(t + \Delta(t)) = f[x(t)]$$

$$\Delta(t) = g[x(t)]$$

Proposed Solution

Define for every time step t , a time window $[t + \ell, t + \ell + h)$.

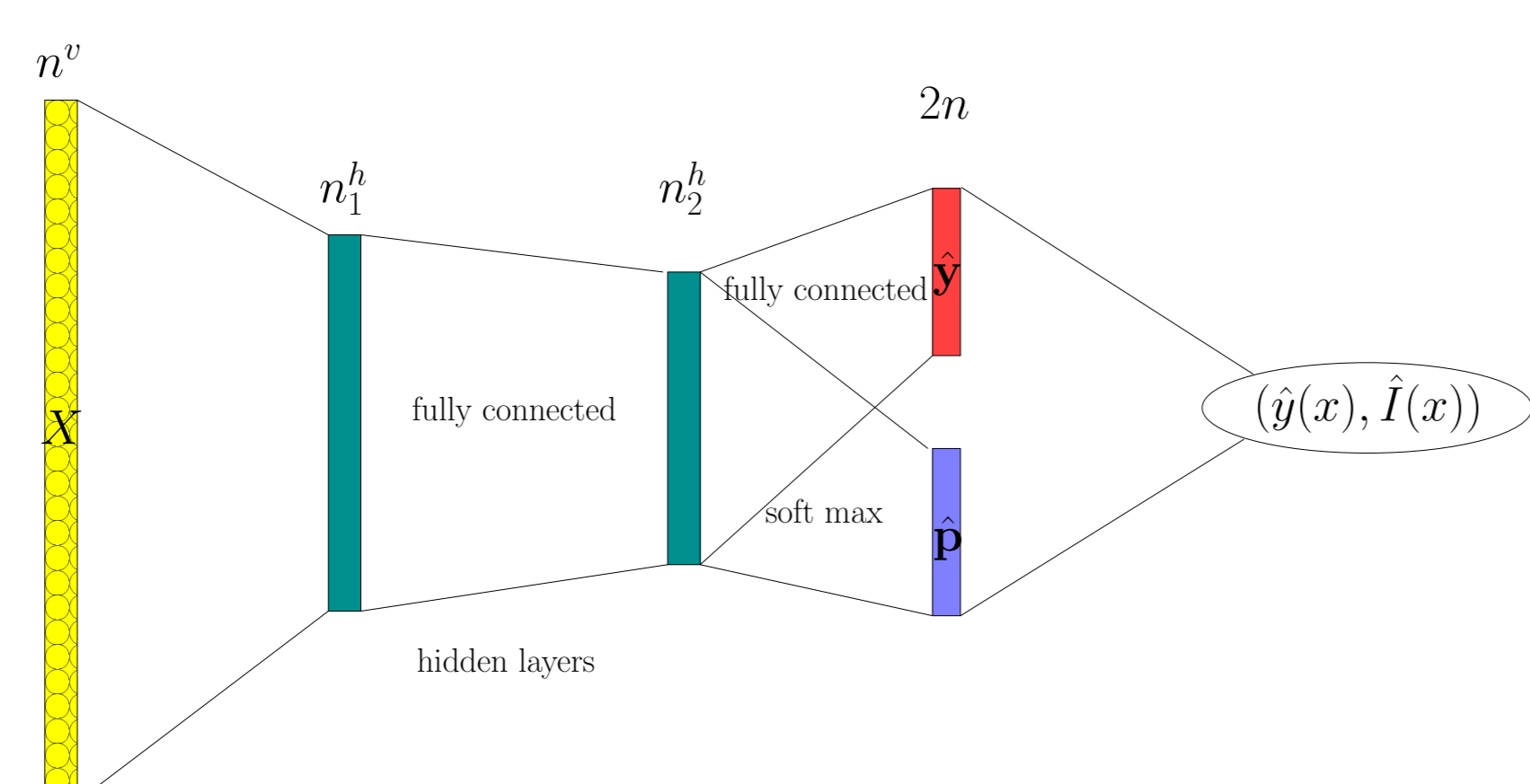
Inputs: Time series $x(t) \in \mathbb{R}^n$

Model Outputs:

- Targets $\hat{y} = [\hat{y}(t + \ell), \dots, \hat{y}(t + \ell + h - 1)]$
- Time Lag Probabilities $\hat{p} = [\hat{p}(t + \ell), \dots, \hat{p}(t + \ell + h - 1)]$, $\sum_{i=0}^{h-1} \hat{p}_i = 1$

Predictions: For an input $x(t)$, predict the target $\hat{y}(t + \ell + i^*)$ corresponding to the most probable propagation time $i^* = \arg \min_i \hat{p}_i(x(t))$

Architecture: The proposed model architecture takes the form of a neural network model as seen below.



Model Training: We need an objective/loss function which: 1. rewards accurate predictions for time window: \hat{y} and 2. penalises \hat{p} in a principled manner. After trying several

^aWang Sheeley
^bDistance from the Coronal Hole Boundary
^cWang Sheeley Arge

candidates, we settled on:

$$\mathcal{L}[\mathbf{x}, \mathbf{y}; \hat{\mathbf{y}}, \hat{\mathbf{p}}, \sigma, \alpha] = -|T| \log(\sigma) - \frac{1}{n_m} \sum_m \left[\sum_{i \in T} \frac{1}{2\sigma^2} (y_{m+i} - \hat{y}_i(x_m))^2 - \log(Z_m(\hat{\mathbf{y}}, \hat{\mathbf{p}}, \sigma, \alpha)) \right],$$

with n_m being the mini-batch size and Z_m is given as:

$$Z_m(\hat{\mathbf{y}}, \hat{\mathbf{p}}, \sigma, \alpha) = \sqrt{(1 + \alpha)} \sum_{i \in T} \hat{p}_i(x_m) \exp\left(-\frac{1}{2\sigma^2} \alpha (y_{m+i} - \hat{y}_i(x_m))^2\right).$$

After initialisation, the model parameters are updated in an alternating fashion. 1. Update the neural network parameters; keep α and σ fixed. 2. Update α and σ ; keep the network parameters fixed.

Solar Wind Forecasting

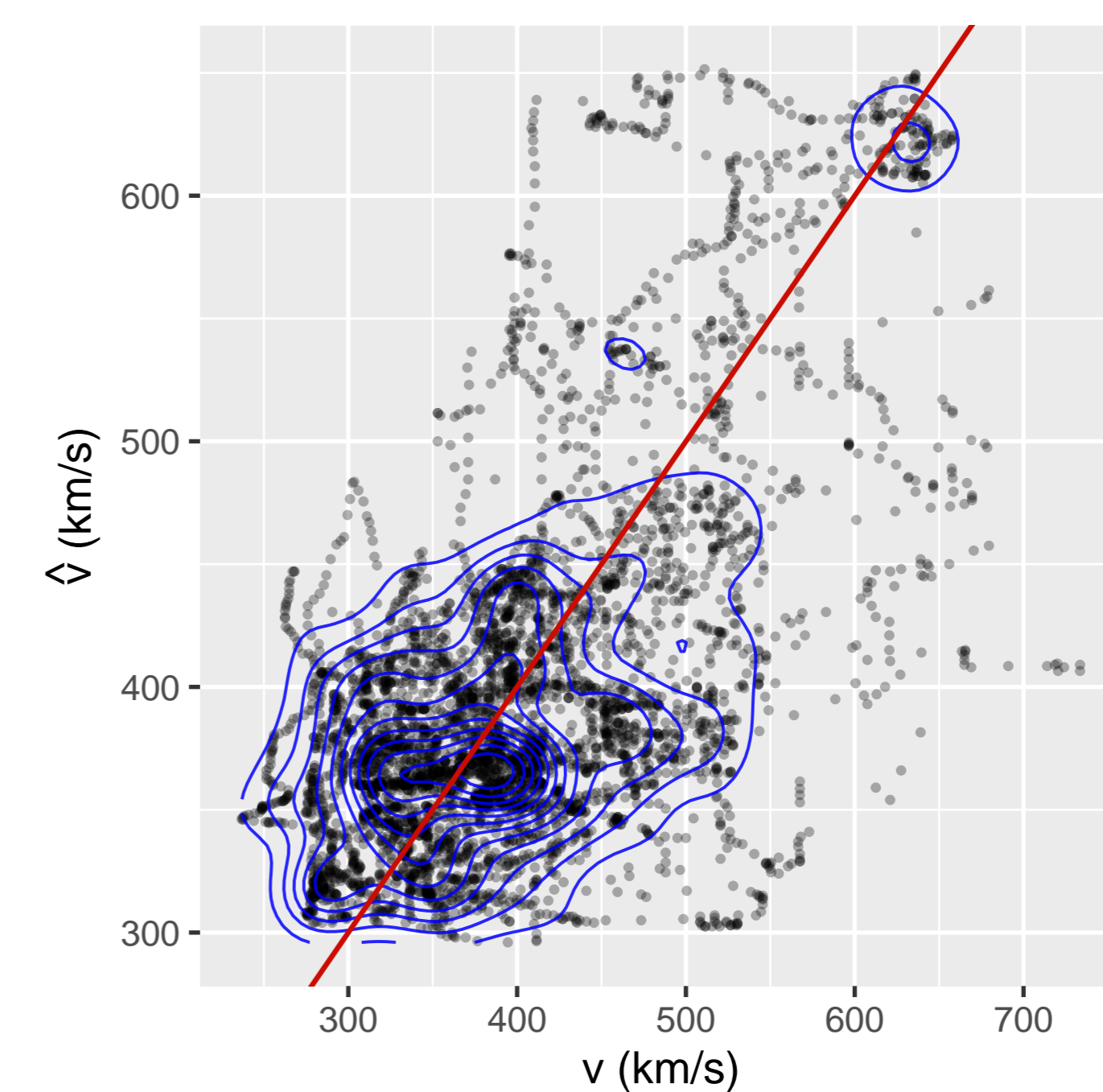
Input Data

- Magnetic flux-tube expansion $\log f_S$ and source surface field \mathbf{B}_{ss} . Both computed from GONG synoptic maps using the *Current-Sheet Source Surface* (CSSS) [2] model.
- Φ , the Carrington longitude of the computed of each time-stamped f_S .
- Sun spot number SSN and $F_{10.7}$ indices from OMNI.
- The solar wind speed 27 days before the data time-stamp, v_{27}

Validation: 9-fold cross-validation on Carrington rotations: 2077, 2090, 2104, 2117, 2130, 2143, 2157, 2171, and 2184.

Benchmarks: On Carrington rotation 2077, model benchmarks from Reiss et al. [1] were used. We also trained a fixed time-lag baseline model, which forecasts solar wind speed with a time horizon of $t + h/2$.

Results



Results of 9-fold cross validation.

Model	M.A.E	R.M.S.E
DTLR (this work)	54.41	66.22
Fixed Lag Baseline	67.33	80.39
WS ^a	74.09	85.27
DCHB ^b	83.83	103.43
WSA ^c	68.54	82.62
Ensemble Median (WS)	71.52	83.36
Ensemble Median (DCHB)	78.27	100.04
Ensemble Median (WSA)	62.24	74.86
Persistence (4 days)	130.48	161.99
Persistence (27 days)	66.54	78.86

Performance Comparison on CR 2077: DTLR, Fixed Lag Base Line vs Reiss et al. [1]

References

- Reiss Martin A., MacNeice Peter J., Mays Leila M., Arge Charles N., Möstl Christian, Nikolic Ljubomir, Amerstorfer Tanja. Forecasting the Ambient Solar Wind with Numerical Models. I. On the Implementation of an Operational Framework // The Astrophysical Journal Supplement Series. feb 2019. 240, 2. 35.
- Zhao Xuepu, Hoeksema J. Todd. Prediction of the interplanetary magnetic field strength // Journal of Geophysical Research: Space Physics. 1995. 100, A1. 19–33.