AGU100 ADVANCING EARTH AND SPACE SCIENCE

Space Weather

RESEARCH ARTICLE

10.1029/2018SW001955

Prediction of Solar Wind Speed at 1 AU Using an Artificial Neural Network

Yi Yang[1,2], Fang Shen[1,2,3], Zicai Yang[1,2], and Xueshang Feng[1,3]

[1]SIGMA Weather Group, State Key Laboratory of Space Weather, National Space Science Center, Chinese Academy of Sciences, Beijing, China, [2]College of Earth Sciences, University of Chinese Academy of Sciences, Beijing, China, [3]HIT Institute of Space Science and Applied Technology, Shenzhen, China

Abstract A hybrid intelligent source surface model applying the artificial neural network tactic for solar wind speed prediction is presented in this paper. The model is a hybrid system merging various observational and theoretical information as input. Different inputs are tested including individual parameters and their combinations in order to select an optimum. Then, the optimal model is implemented for prediction. The prediction is validated by both error analysis and event-based analysis from 2007 to 2016. The overall correlation coefficient is 0.74, and the root-mean-square error is 68 km/s. The probability for detecting a high-speed-event is 0.68, the positive predicted value is 0.73, and the threat score is 0.55.

Key Points:

- The prediction of solar wind speed with artificial neural network is presented
- Results are validated by model comparison, error analysis, and event-based analysis
- Our model generally performs as well as or better than the WSA model for predicting solar wind speed at 1 AU for a long period

Supporting Information:

- Supporting Information S1
- Table S1
- Table S2
- Table S3
- Table S4

Correspondence to:
F. Shen,
fshen@spaceweather.ac.cn

Citation:
Yang, Y., Shen, F., Yang, Z., & Feng, X. (2018). Prediction of solar wind speed at 1 AU using an artificial neural network. *Space Weather*, *16*. https://doi.org/10.1029/2018SW001955

Received 10 MAY 2018
Accepted 7 AUG 2018
Accepted article online 18 AUG 2018

1. Introduction

It is generally acknowledged that space weather effects have become a great threat to human society, because more and more modern technologies are exposed to the space environment. (Baker, 1998; Bothmer & Daglis, 2007; Feynman & Gabriel, 2000; Hapgood, 2011; Siscoe, 2000). One of the most damaging space weather effects is geomagnetic storms. Although most large geomagnetic storms are driven by coronal mass ejections (CMEs; Tsurutani et al., 1988), the ambient solar wind can also be very geoeffective, particularly in the corotating interaction regions (CIRs) where the fast and slow solar winds interact with each other (Richardson et al., 2002; Tsurutani et al., 2006). The high-speed streams (HSSs) and enhanced magnetic field of CIRs can cause long-living and recurrent geomagnetic storms, which may last for several days (Borovsky & Denton, 2006; Burlaga & Lepping, 1977; Gosling & Pizzo, 1999; Tsurutani et al., 1988, 1995, 2006; Zhang et al., 2007). Since there are fewer CMEs during the declining and minimum phases of a solar cycle (Alves et al., 2006), the overall contribution of CIRs to the geomagnetic activity is significant (Tsurutani et al., 2006). Thus, in order to mitigate the space weather effects, reliable forecasts of the near-Earth ambient solar wind conditions are required as well as the CMEs. Furthermore, the ambient solar wind parameters can provide background condition for kinematical and dynamical models of CME propagation (Gopalswamy et al., 2001; Lugaz et al., 2005; Shen et al., 2013, 2014; Vršnak & Gopalswamy, 2002) and interplanetary shocks (Dryer et al., 2004; Zhao & Dryer, 2014).

Currently, there are mainly two kinds of methods for predicting the ambient solar wind conditions. One kind is the physical-based magnetohydrodynamic (MHD) models, and the other is empirical or semiempirical models. As one of the most powerful tool to study and predict the space environment, a large number of MHD models have been developed and continuously improved by researchers all around the world. Examples of successful solar wind MHD models are the ENLIL model (Odstrcil, 2003; Odstrcil et al., 2004) used at the Space Weather Prediction Center of the National Oceanic and Atmospheric Administration (SWPC/NOAA), the Space Weather Modeling Framework (SWMF) at the Center for Space Environment Modeling (CSEM; Tóth et al., 2005), the CORona-HELiosphere (CORHEL) model developed by the Center for Integrated Space Weather Modeling (CISM; Riley et al., 2001), the Hybrid Heliospheric Modeling System (HHMS; Detman et al., 2006, 2011; Intriligator et al., 2012), a combining 3-D MHD model (Wu et al., 2007, 2011) with the Hakamada-Akasofu-Fry (HAF) code (Fry et al., 2001), the COrona-INterplanetary (COIN) model, the improved interplanetary total variation diminishing (IN-TVD) and SIP-AMR-CESE model developed by the SIGMA group at Chinese Academy of Sciences (Feng et al., 2007, 2010; Shen et al., 2009), the CRONOS model (Kissmann et al., 2008; Wiengarten et al., 2013), the space weather integrated forecasting framework (SWIFF; Lapenta et al., 2013), the Lyon-Fedder-Mobarry (LFM) heliospheric model (Merkin et al., 2011; Pahud et al., 2012), the MHD model by Usmanov and Goldstein (2003, 2006), the MHD model by

Hayashi (2005, 2012), and Hayashi, Tokumaru, et al. (2016), and other MHD models mentioned in the review of Wu and Dryer (2015).

Empirical or semiempirical models are also important research aspects for space weather, because of their efficiency, simplicity, and comparable forecasting skills to the more complicated and computational-expensive MHD models (Owens et al., 2008; Riley et al., 2006). One of the most commonly used semiempirical model is the Wang-Sheeley-Arge (WSA) model. It was first developed as a piecewise function by Wang and Sheeley (1990) based on the observed negative correlation between solar wind speed at 1 AU and the expansion factor (f_s) of coronal magnetic field at the source surface (2.5 solar radii [RS]). Arge and Pizzo (2000) improved the relation to a continuous function. Then, another parameter θ_b, which is the angular distance from the coronal hole boundary (Riley et al., 2001), was added by Arge et al. (2003) to give the current functional form. The Potential Field Source Surface (PFSS) model (Altschuler & Newkirk, 1969; Schatten et al., 1969) is usually used to calculate the coronal magnetic field and derive the f_s and θ_b parameters.

Another big family of empirical models relies on the relation between coronal hole area and solar wind speed, which was discovered many years ago (Levine et al., 1977; Nolte et al., 1976; Sheeley & Harvey, 1981). Researchers have improved this kind of models in various ways with the help of better quality observations such as SOHO and SDO (de Toma, 2011; Lowder et al., 2014; Luo et al., 2008; Obridko et al., 2009; Robbins et al., 2006; Verbanac et al., 2011; Vršnak et al., 2007). In addition to the solar wind speed, similar relations were discovered for other parameters (temperature, density, and magnetic field magnitude; Rotter et al., 2012, 2015; Vršnak et al., 2007).

Other techniques have been implemented in empirical models to predict solar wind parameters near the Earth. Owens et al. (2013) presented a simple way to make predictions of the solar wind parameters based on the 27-day periodicity, resulting from the 27-day rotation of the Sun. They made a comprehensive analysis of the model performance and pointed out that the method can represent a benchmark for other space weather forecast models. Innocenti et al. (2011) applied Kalman filters to implement data assimilation and made a significant improvement compared to the baseline model (Vršnak et al., 2007). Liu et al. (2011) predicted the solar wind speed by applying a support-vector-machine algorithm, but their predictions were only 3 hr ahead of real time. Bussy-Virat and Ridley (2014) developed the probability distribution function (PDF) model based solely on the 1-AU solar wind observations. They claimed that the PDF model can be better than the persistence model for a prediction horizon from 8 to 120 hr and better than the WSA model for prediction horizons shorter than 24 hr. The PDF model was then improved to predict high-speed events (Bussy-Virat & Ridley, 2016). Wintoft and Lundstedt (1997, 1999) developed a hybrid intelligent system combined the PFSS model and an artificial neural network (ANN) to give prediction of the daily-averaged solar wind speed.

The main advantages of using an ANN include the following: it can handle large data sets, it has the ability to capture nonlinear and complex underlying relationships of any physical process with plenty of data, it can be easily implemented in parallel computational architectures (such as multicore processors and graphic processing units) to reduce the processing time, and it has good fault tolerance, since the corruption of a few neurons does not prevent it from total functioning. ANN techniques have been applied to space weather for prediction of sunspot number and the solar cycle (Calvo et al., 1995; Conway et al., 1998; Fessant et al., 1996), prediction of CME transit time (Liu et al., 2018; Sudar et al., 2015), prediction of proton event (Gong et al., 2004), prediction of solar flare (Borda et al., 2002; Qahwaji & Colak, 2007; Wang et al., 2008), prediction of solar wind velocity (Wintoft & Lundstedt, 1999), prediction of electron flux in magnetosphere (Bortnik et al., 2016; Chu et al., 2017; Koons & Gorney, 1991; Ling et al., 2010; Wu & Lundstedt, 1997; Zhelavskaya et al., 2017), and prediction of geomagnetic activity (Boberg et al., 2000; Cai et al., 2010; Lazzús et al., 2017; Macpherson et al., 1995; Uwamahoro et al., 2012; Valach et al., 2009; Wintoft et al., 2017). A detailed introduction can be found in the review of Lundstedt (2005). Although the knowledge of solar wind have greatly improved from both observations and theories, the accuracy of ANN techniques in modeling solar wind conditions at 1 AU has not improved in the last two decades (Wintoft & Lundstedt, 1997, 1999). The recent improvements in space and ground-based instruments have enabled solar and heliospheric observations with much higher temporal and spatial resolution, which can improve the performance of data-driven models. For example, the accuracy of the PFSS model in modeling the coronal magnetic field depends on the quality of the magnetogram that was used as input of the model (Liu et al., 2012; Hayashi, Yang, et al., 2016; Riley et al., 2014). The PFSS modeling accuracy also depends on some other factors, such as the magnetogram

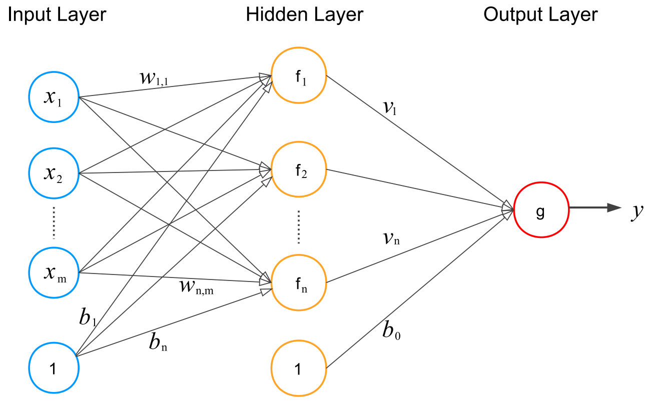

Figure 1. The architecture of a three-layer feedforward artificial neural network (ANN) used in this study. Each circular node represents an artificial neuron, and each arrow represents a direction to transmit signals. The white arrows can be considered as weights and biases of inputs from its left side.

smoothing procedures (Hayashi, Yang, et al., 2016), the hypothesis of source-surface radius (Lee et al., 2011), and the spherical polynomial calculation method (Tóth et al., 2011). We also noticed that theoretical studies may be useful to improve the solar wind prediction by ANN. Studies have linked the photo magnetic field strength to the solar wind speed observed at 1 AU (Fujiki et al., 2015; Suzuki, 2006; Suzuki & Inutsuka, 2005). Riley et al. (2015) demonstrated that the angular distance from the coronal hole boundary (θ_b) played an important role as a solar wind speed indicator. The θ_b has been routinely used with f_s in the current WSA model at SWPC/NOAA. Recent research studies have proved that global coronal magnetic field models can be improved by using image-based information (Conlon & Gallagher, 2010; Jones et al., 2016, 2017). Thus, the white light brightness observation from LASCO/SOHO or SECCHI/STEREO may also be useful to improve a solar wind prediction model.

Consequently, we aim to develop a hybrid ANN model to predict the solar wind parameters at 1 AU, using knowledge from both theory and observations. In this paper, we only focus on the prediction of the solar wind speed. Prediction of other solar wind parameters is left for future work. In section 2, we describe the ANN algorithm and the hybrid model used in our study. In section 3, we introduce the data source used for this study, and the data set prepared for ANN training. All the modeling results are presented in section 4. In section 4.1, we show the training performances of different ANNs with different input in order to select the best one. In section 4.2, we show the training results of the optimal ANN we discovered. Then, results of implementing the optimal ANN for prediction are presented in section 4.3. The conclusions are made in section 5.

2. Methods

2.1. ANN

ANN is an intelligent computing system inspired by the biological neural networks like human brains. ANN can progressively improve performance on a task by analyzing a large number of examples, like the learning process of humans.

An ANN is constructed by a group of interconnected nodes called artificial neurons. The artificial neurons can transmit a signal from one to another by their connection. A neuron calculates a weighted sum of its inputs and applies a transfer function to obtain the signal that will be transmitted to a next neuron. Artificial neurons are usually organized in layers. Different layers may perform different kinds of transformations on their inputs. ANN adjusts the weights to minimize the error between input and output as a learning process (Haykin, 2004).

Various types of ANNs have been established through these years, different from architecture of node connections and learning algorithms. The ANN we apply in this study is a widely used feedforward neural network with one hidden layer, similar to that used in Boberg et al. (2000) and Wintoft et al. (2017). The commonly used Levenberg-Marquardt backpropagation algorithm is implemented in the learning process. Figure 1 shows the architecture of this three-layer feedforward ANN. Each circular node represents an artificial neuron, and each arrow represents a connection from the output of one neuron to the input of another neuron. The blue, yellow, and red color of circles are used to distinguish different layers. The white arrows can be considered as weights of the input from its left side layer. The input layer contains m nodes, each for a kind of input characteristic parameter that may associate with the output. The hidden layer contains n nodes, and the output layer has only one node for a certain solar wind parameter.

The output of this ANN can be calculated as follows:

$$y = g\left(b_0 + \sum_{j=1}^{n} v_j\, f_j\left(b_j + \sum_{i=1}^{m} w_{ji} x_i\right)\right) \tag{1}$$

where x_i is the normalized input vector, w_{ji} is the weights between node i of input layer and node j of hidden layer, v_j are the weights between hidden layer and the output, b_j and b_0 represent the biases, and y is the final output. Bipolar sigmoid nonlinear activation function f_j is applied as the hidden layer transfer function, while a linear function g is applied for the output layer. The f function is calculated as

$$f(x) = \frac{2}{1 + e^{-2x}} - 1 \tag{2}$$

After a certain ANN architecture is determined, it has to be trained with an input-output data set to adjust its weights in order to generate the output as similar as possible to the true output. Before training, the full data set is split randomly into three sets called training, validation, and test set. Then, we iteratively use the training set to minimize the mean-squared errors (*MSEs*) between the network output and the target output. For the input data with a total number of Q, MSE is calculated as follows:

$$\mathrm{MSE} = \frac{1}{Q}\Sigma_{q=1}^{Q}\left(t_q - y_q\right)^2 \tag{3}$$

where y is the ANN output value and t is the observed target value. For each iteration in the training phase, the MSE is also computed by using validation set. The number of iterations of training phase is experimentally determined by assuring the MSE decreases. This procedure can help to avoid the training results biasing toward a particular section of the database. Finally, the specific weights that minimize validation MSE give the optimal network. The test set is left only for a performance evaluation.

The iteration step updates weights and biases using the Levenberg-Marquardt backpropagation algorithm (LMBP). We can define $\boldsymbol{W}$ as the vector of network weights and biases:

$$\boldsymbol{W} = [W_1\ W_2 \ldots W_N]^T = \left[w_{1,1}\ w_{1,2} \ldots w_{ji}\ v_1\ v_2 \ldots v_j\ b_1\ b_2 \ldots b_j\ b_0\right]^T, \tag{4}$$

where $N = n(m + 1) + (n + 1)$ is the total number of weights and biases. The error vector $\boldsymbol{e}$ is defined by

$$\boldsymbol{e} = [e_1\ e_2 \ldots e_Q]^T \tag{5}$$

$$e_q(\boldsymbol{W}) = t_q - y_q \tag{6}$$

where e_q is the error for the qth sample and Q is the total sample number. Then the training evaluation parameter MSE can be expressed as

$$\mathrm{MSE} = \frac{1}{Q}\Sigma_{q=1}^{Q}(e_q)^2 \tag{7}$$

And the Jacobian matrix for the network training can be expressed as

$$\boldsymbol{J}(\boldsymbol{W}) = \begin{bmatrix} \frac{\partial e_1(\boldsymbol{W})}{\partial W_1} & \frac{\partial e_1(\boldsymbol{W})}{\partial W_2} & \cdots & \frac{\partial e_1(\boldsymbol{W})}{\partial W_N} \\ \frac{\partial e_2(\boldsymbol{W})}{\partial W_1} & \frac{\partial e_2(\boldsymbol{W})}{\partial W_2} & & \frac{\partial e_2(\boldsymbol{W})}{\partial W_N} \\ \vdots & & \ddots & \vdots \\ \frac{\partial e_Q(\boldsymbol{W})}{\partial W_N} & \frac{\partial e_Q(\boldsymbol{W})}{\partial W_N} & \cdots & \frac{\partial e_Q(\boldsymbol{W})}{\partial W_N} \end{bmatrix} \tag{8}$$

From iteration step k to $k + 1$, the weights and biases can be updated by the following equations:

$$\boldsymbol{W}^{k+1} = \boldsymbol{W}^k + \Delta\boldsymbol{W}^k \tag{9}$$

$$\Delta\boldsymbol{W} = \left[\boldsymbol{J}^T(\boldsymbol{W})\boldsymbol{J}(\boldsymbol{W}) + \mu\boldsymbol{I}\right]^{-1}\boldsymbol{J}^T(\boldsymbol{W})\boldsymbol{e}(\boldsymbol{W}) \tag{10}$$

where $\boldsymbol{I}$ is the identity matrix and μ is the learning rate. A detailed derivation of the equations above and the whole LMBP algorithm theory was introduced by Hagan and Menhaj (1994) and Hagan et al. (1996).

After an ANN has been trained, the forecast results implementing the ANN need more quantitative evaluation to show the model performance. Besides the MSE, the following statistical parameters will also be used in this paper: (a) the root-mean-square errors (*RMSE*), (b) the standard deviation (*STD*), and (c) the correlation coefficient (*R*).

For a set of input data with a total number *Q*, the RMSE, STD, and *R* are calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{Q}\Sigma_{q=1}^{Q}\left(t_q - y_q\right)^2} \tag{11}$$

$$\mathrm{STD} = \sqrt{\frac{1}{Q}\Sigma_{q=1}^{Q}\left(y_q - \overline{y}\right)^2} \tag{12}$$

$$R = \frac{\Sigma_{q=1}^{Q}\left(t_{\mathrm{q}} - \overline{t}\right)\left(y_q - \overline{y}\right)}{\sqrt{\Sigma_{q=1}^{Q}\left(t_q - \overline{t}\right)^2}\sqrt{\Sigma_{q=1}^{Q}\left(y_q - \overline{y}\right)^2}} \tag{13}$$

where *y* is the ANN output value and *t* is the observed target value. While the RMSE and the *R* value directly compare the ANN output with the observations, the STD only characterizes the variability of one or the other (ANN or observations).

2.2. Hybrid Intelligent Source Surface (HISS) Model

In this paper, we develop an ANN using characteristic parameters obtained from both theoretical model and observational information as input to predict ambient solar wind condition at 1 AU. We call this combined system the HISS model, similar as the concept proposed by Wintoft and Lundstedt (1997).

PFSS (Altschuler & Newkirk, 1969, Schatten et al., 1969) is a commonly used physical model to extrapolate the coronal magnetic field from observed photospheric magnetograms. After the coronal magnetic field is determined, we can obtain the f_s and θ_b by tracing field lines. The f_s is calculated by the function:

$$f_s = \left(\frac{B_0}{B_{ss}}\right)\left(\frac{R_0}{R_{ss}}\right)^2 \tag{14}$$

where B_{ss} is the magnetic field strength at the source surface, B_0 is the magnetic field strength at the photosphere, R_0 is the solar radius, and R_{ss} is the radius of the source surface (we define $R_{ss} = 2.5R_0$ in this paper). After tracing field lines, we can know the location of open magnetic field, that is, the coronal holes, then θ_b is obtained by calculating the angular distance to the nearest coronal hole boundary. The field line foot-point location at the photosphere can also be determined. Then we note the heliolatitude of the foot-point as L_0. The latitude difference between a point and the heliospheric current sheet can be understood as magnetic latitude, noted L_m. Thus, we could obtain six possible input parameters from a PFSS model, namely, f_s, θ_b, B_{ss}, B_0, L_0, and L_m.

The white light brightness observed by coronagraph can be used to optimize the extrapolated coronal magnetic field as mentioned in section 1. As a derivative, coronal density can be inversed from observed polarized brightness (*pB*) data (Van de Hulst, 1950; Wang & Davila, 2014). The inversion process uses IDL routines associated with *pb_inverter.pro* from SolarSoftWare (SSW, http://www.lmsal.com/solarsoft/) library to calculate the field-of-view density map. Then we extract the density at the source surface to construct the density synoptic maps, similar to the magnetic synoptic maps. We note the density at the source surface N_{ss} and consider it as a possible character for the ANN input.

Most data-driven solar wind models only use solar observations as inputs, as described in section 1. However, the 1 AU in situ observations can also be used for a model as data assimilation technique, like a few works have done (Bussy-Virat & Ridley, 2014, 2016; Innocenti et al., 2011). We can use the value of a solar wind parameter 27 days ago as another input characteristic parameter of the ANN. The solar wind speed 27 days ago is further noted as PS_{27}. The persistence technique is implemented in the PDF model as well (Bussy-Virat & Ridley, 2014, 2016).

All in all, the HISS model presented in this paper is a three-layer feedforward ANN, hybridized with a PFSS physical model and merged information from both solar and heliosphere observations.

3. Data

3.1. Data Source

Global Oscillation Network Group (GONG) magnetograms of the observed photospheric magnetic field are provided by National Solar Observatory (NSO, http://gong.nso.edu/). GONG data are available since late 2006. We use GONG data as input data of the PFSS model in this study.

OMNI data set is a compilation of 1963-to-current near-Earth solar wind magnetic field and plasma data from several spacecraft in geocentric orbits or at the L1 Lagrange point. This study uses the hourly averaged solar wind data downloaded from OMNI web (http://omniweb.gsfc.nasa.gov).

The polarized brightness (*pB*) data we used in this study was derived from white light coronagraph observation of LASCO/SOHO, and it can be downloaded from the LASCO instrument website at NRL (http://lasco-www.nrl.navy.mil/content/retrieve/polarize/). The *pB* data are available since the SOHO launch in 1996.

3.2. Training Data Set Preparation

An input-output data set must be prepared before training the ANN. The selection of representative data sets and input characters is very important to the ANN performance. The possible input characters we prepared include f_s, θ_b, B_0, B_{ss}, L_0, L_m, N_{ss}, and PS_{27}, as presented in section 2. The output is the corresponding solar wind observation at 1 AU including velocity, density, temperature, and magnetic field strength. To associate the input data at the source surface and the output data at 1 AU, we use an average solar wind propagating time of four days. The input parameters at the time *t0* are used to predict the solar wind parameters up to t0 + four days. We have also tested to map back the solar wind parameters assuming a constant solar wind speed or considering the CIR effects using the method developed by Arge and Pizzo (2000). However, the ANN performance by using these methods would be slightly worse, and with an RMSE about 10 km/s higher than that by using the method presented in this paper.

We use a data scope from year 2007 to 2016. This period covers all the four phases (descending, minimum, ascending, and maximum) of a solar cycle, so that the ANN could learn the properties of all these phases. We divide the data of year 2007–2015 randomly into three sets named training set, validation set, and test set, as described in section 2. The percentage of the training set in the whole data set is 70%, while the validation set and the test set are both 15%. The data of year 2016 is left for a final validation of the model since this continuous period is totally not involved in training procedure.

Because the input and output variables have much different ranges, they have to be normalized before training. We normalize the data to the range from -1 to 1. The values will be unnormalized to physical values after training.

4. Results

In this section, we present a detailed study for the prediction of solar wind speed (V). Performances of different ANNs using different input parameters and hidden node numbers are introduced in section 4.1. Then, training results of the optimal ANN are shown in section 4.2. Finally, the optimal model is implemented for prediction and is validated by comparing with observations and other models in section 4.3.

4.1. Selection of ANN

Choosing the proper input parameters is the most important process in developing an ANN model. The hidden node number of an ANN architecture also affects the ANN performance. The best input parameters and the proper hidden node number could only be determined experimentally.

First, we tested the ANNs using different individual parameters as input. The tested parameters were f_s, θ_b, B_0, B_{ss}, L_0, L_m, N_{ss}, and PS_{27} respectively. Meanwhile, we also tested how the performances varied with the change of ANN hidden node number. The correlation coefficient (R) and RMSE are shown in Figure 2.

As shown in Figure 2, the correlation from the highest to the lowest is respectively PS_{27}, θ_b, f_s, B_{ss}, L_m, L_0, N_{ss}, and B_0. The RMSE from the lowest to the highest is PS_{27}, θ_b, f_s, B_{ss}, L_m, L_0, N_{ss}, and B_0 respectively. The R and

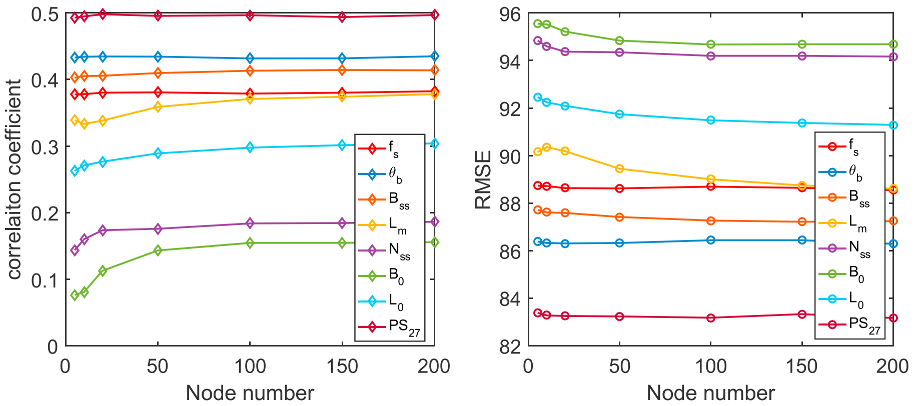

Figure 2. The performances of artificial neural networks (ANNs) with different individual input parameters and different hidden node number. The left panel shows the correlation coefficients *R*, and the right panel shows root-mean-square error (RMSE).

RMSE results accord with each other, as the higher *R* and the lower RMSE both indicate the better model performance. The θ_b and f_s have relatively good correlation with velocity as expected, since they have been successfully used in the WSA model for many years. The ANN using only f_s as input gives a *R* of 0.35 and a RMSE of 89 km/s, which is consistent with the results of Wintoft and Lundstedt (1999). Finally, the increase of hidden node number does not evidently improve the performances of all ANNs using individual parameters.

Since each parameter may provide unique useful constraining information on solar wind prediction, we discuss the effect of different combinations of these parameters, which can be easily realized in ANN. We start from only one parameter, then incrementally add one more parameter in each additional test. The sequence of addition is from the parameter of the best correlation to the parameter of less correlation. Although the L_0 has a better correlation, it is considered to be added only after B_0, because it is a parameter from the photosphere as B_0, and we assume parameters on the source surface to be more reliable for prediction. PS_{27} is added last in spite of its high correlation, since it is a near-Earth observational data while other parameters are derived from coronal observations. PS_{27} could be considered as a data assimilation parameter, different from all other parameters. For a future study, we could change PS_{27} to other near-Earth observation for different applications. For example, we could change it to one-day persistent data for a one-day-forward prediction. Overall, the combinations used here are the following ones: (θ_b, f_s); (θ_b, f_s, B_{ss}); (θ_b, f_s, B_{ss}, L_m); (θ_b, f_s, B_{ss}, L_m,

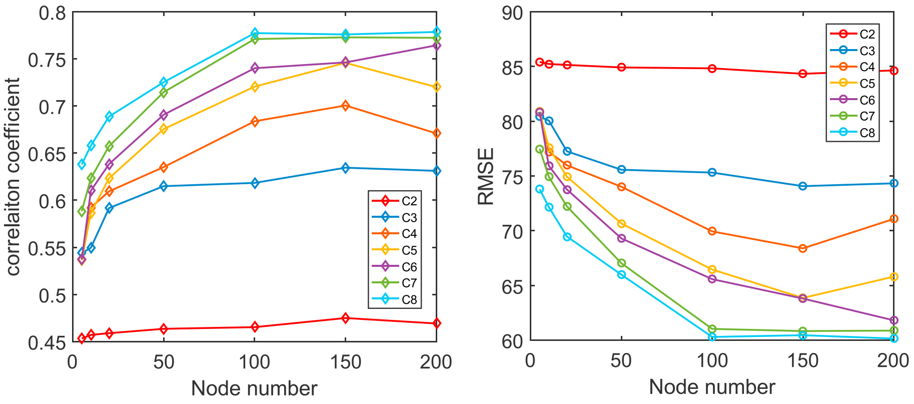

Figure 3. The performance of artificial neural networks (ANNs) using different combination of parameters as input and different hidden node number. The left panel shows the correlation coefficients *R*, and the right panel shows root-mean-square error (RMSE).

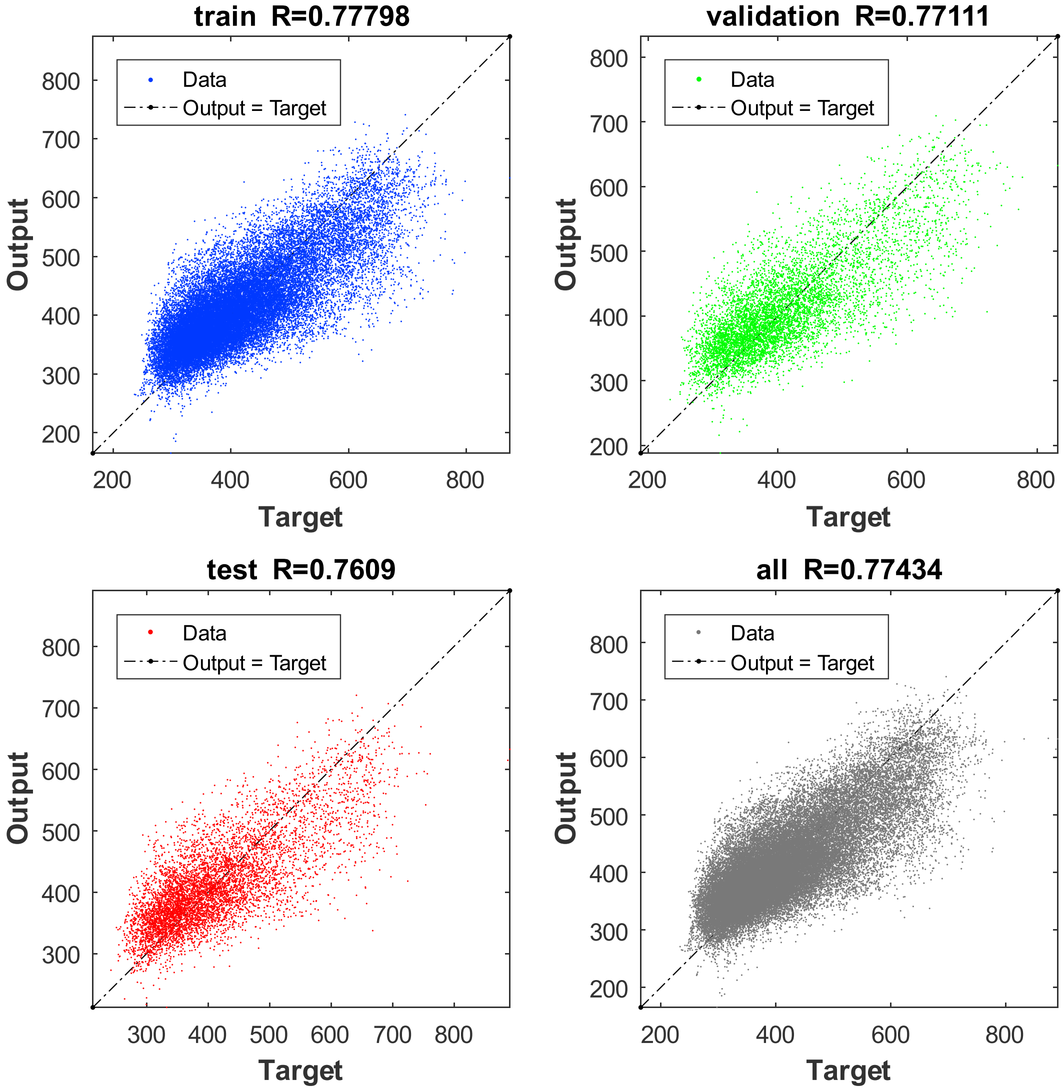

Figure 4. Regression results of the optimal artificial neural network (ANN). The blue, green, red, and grey dots represent the data points in training set, validation set, test set, and all data set, respectively. The horizontal axis is the observational value, and the vertical axis is the output of ANN.

N_{ss}); (θ_b, f_s, B_{ss}, L_m, N_{ss}, B_0); (θ_b, f_s, B_{ss}, L_m, N_{ss}, B_0, L_0); and (θ_b, f_s, B_{ss}, L_m, N_{ss}, B_0, L_0, PS_{27}), noted as C2, C3, C4, C5, C6, C7, and C8 for short, respectively.

As shown in Figure 3, using combined different parameters as input can give much better fit than any individual parameter. The performance of the correlation in the left panel and the RMSE in the right panel accord with each other, similar as in Figure 2.

Different combinations and node number have very different performances. Using the same number of hidden node, ANN performance gets better with more parameters included. Using all parameters as input gives the best fit. We can also see that the performances are better with more hidden nodes within 150, but it might go down when the node numbers are bigger than 150, depending on the inputs used. We could conclude from the figure that the more parameters are input, the bigger hidden node number should be used. But big hidden node number is not necessary for small input parameter number, since the slope is higher for models with a large number of inputs than that with a small number of inputs. For example, the *R* value of C2 only increased from 0.45 to 0.47, but the *R* value of C8 increased greatly from 0.64 to 0.77 when the node number varied from 5 to 200. This is also consistent with Figure 2, which shows that the performance of each model with a single input does not depend on the number of nodes.

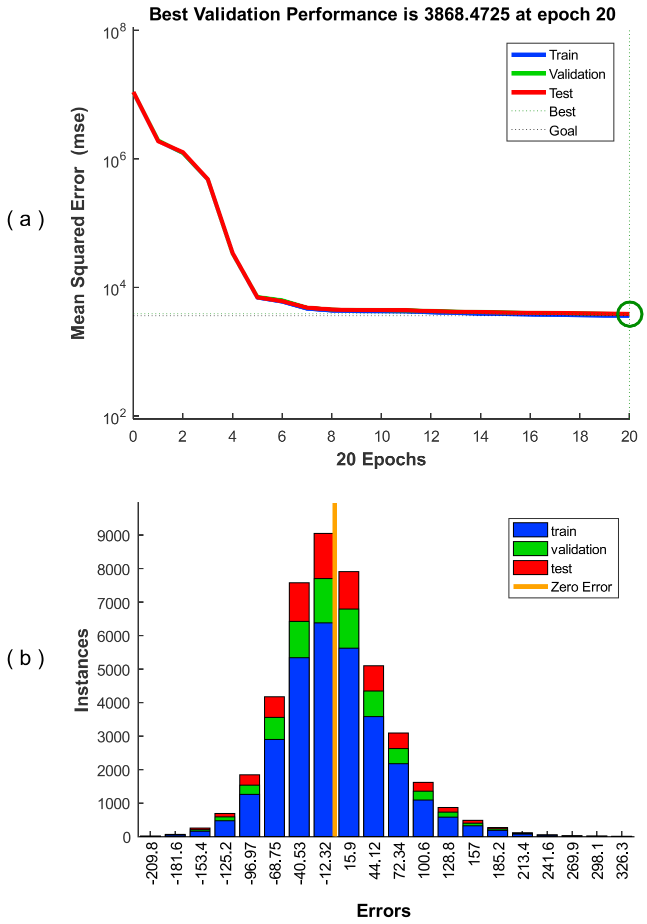

Figure 5. (a) The variation of the artificial neural network (ANN) performance (mean-squared error [MSE]) during the training procedure. The blue, green, and red lines represent the train, validation, and test sets, respectively. The green dotted line indicates the best validation performance, and the black dotted line indicates our training goal. (b) Distribution diagram of the ANN prediction error for our whole data set from 2007 to 2015. The horizontal axis is the error value, and the vertical axis is the instance of data point. The blue, green, and red color represent the training set, validation set and test set respectively.

Finally, the optimal ANN is the one with 150 hidden nodes and using all of θ_b, f_s, B_{ss}, L_m, N_{ss}, B_0, L_0, and PS_{27} as input parameters, with a correlation of ~0.78 and a RMSE of ~60 km/s. The weights and biases of the optimal ANN are recorded in the supporting information.

4.2. Training Performance of the Optimal ANN

Before applying the optimal ANN to a solar wind speed prediction, we should examine its training performance carefully. The regression results of the optimal ANN are presented in Figure 4. The training set, validation set, test set, and the whole data set are shown in different panels with different color. The correlation coefficients of different sets are noted in the title of each panel. We can see that all the sets performed nicely, with *R* values above 0.76. The training set is a little better with an *R* value of about 0.78, which could be anticipated.

Figure 5a shows the variation of ANN performance during the training procedure. We can see that the MSE decreases quickly in the first five epochs (iterations) and continues to decrease slowly until meet our training goal of 3,600 (km/s)^2 at the 20th epoch. The performance of the training, validation, and test sets are very close to each other during all the epochs.

The error distribution histogram of the optimal ANN is presented in Figure 5b. The training set, validation set, and test set are shown in blue, green, and red color, respectively. We can see that most errors are in a range of -69 to 72 (km/s), which is consistent with the RMSE results. However, there are still a few large errors, limited in a range of -210 to 326 (km/s). The large errors may be caused by some influence factors rather than the input characters we used, such as the influence of ICMEs, which are not excluded from our data set. It is also noticed that instances of different sets are proportional, which is consistent with the result shown in Figure 4.

From the training performance exhibited above, we can infer that this ANN may have good prediction capability and nice generalization ability.

4.3. Implementation of the Optimal ANN for Prediction

In the following, we make further validation of our model by applying it for solar wind speed prediction and comparing our results with the observation data and other models. It should be noted that our model is trained for the four-day advanced prediction, because four days is the average duration for solar wind to propagate from the source surface to the Earth. The WSA model presented in this paper is also used for the four-day advanced prediction.

Figure 6 presents a prediction result for 10 years from 2007 to 2016. The year 2016 is a totally out-of-sample period, which could further demonstrate the generalization ability of our model. As shown in this figure, our ANN prediction is fairly coincident with the OMNI observation. The WSA modeling result is also shown as a standard model, which performs well as expected.

The WSA model function used in this study is the following one:

$$V = V_s + \frac{V_f}{(1+f_s)^{a_1}}\left[1 - 0.8\exp\left(-\left(\frac{\theta_b}{a_2}\right)^{a_3}\right)\right]^{a_4} \tag{15}$$

where $V_s = 250\,\frac{\text{km}}{\text{s}}$, $V_f = 675$ km/s, $a_1 = 2/9$, $a_2 = 2$, and $a_3 = 1$, $a_4 = 1$. It should be noted that these values are determined according to the GONG magnetogram and PFSS model we have used in this paper. Different values may be used by other studies (Arge et al., 2003; MacNeice, 2009a; Riley et al., 2015).

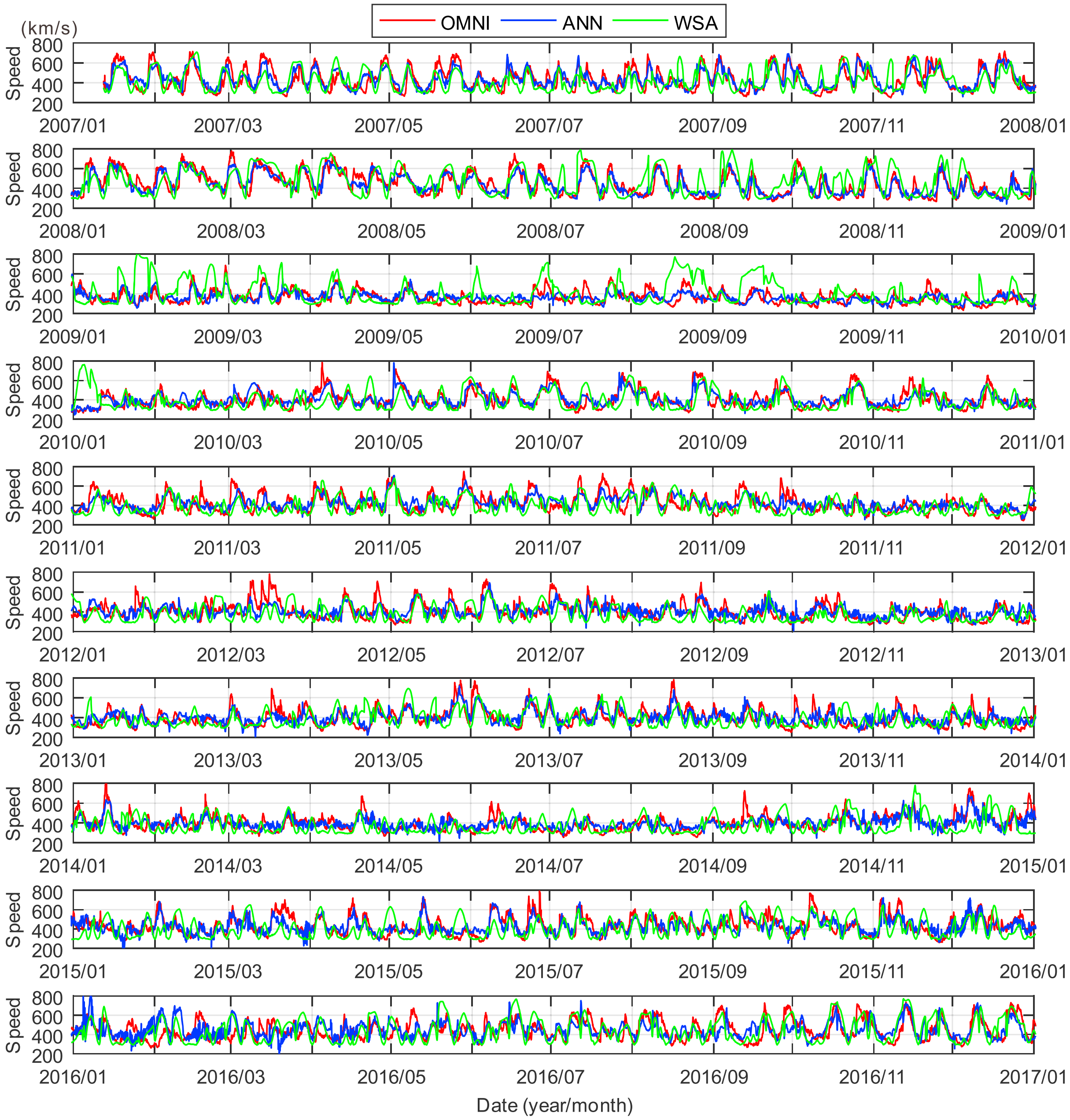

Figure 6. Solar wind speed predictions from our model and the Wang-Sheeley-Arge (WSA) model from 2007 to 2016. The red, blue, and green lines represent the OMNI observational data, artificial neural network (ANN) modeling result and WSA modeling result.

When we zoom into the picture and look the curves more carefully, we can find some better predicted peak values of the ANN result. For example, there are a lot HSSs faster than 600 km/s predicted by the WSA model from July 2008 to March 2009, but these HSSs are not really observed as in our model predictions. Figure 7 presents the zoomed-in pictures for four Carrington Rotations (CRs), respectively, from the four phases of the 24th solar cycle, namely, CR2077 from the minimum phase, CR2110 from the ascending phase, CR2140 from the maximum phase, and CR2178 from the descending phase. We could easily see that our ANN results are much closer to the observations than the WSA results for these four CRs. These results make sense, because the WSA model is a fixed function, which might not be suitable for some periods in the whole time we examined. However, our ANN may have learned how to predict these periods, for it was trained with 85% of the big data set from 2007 to 2015.

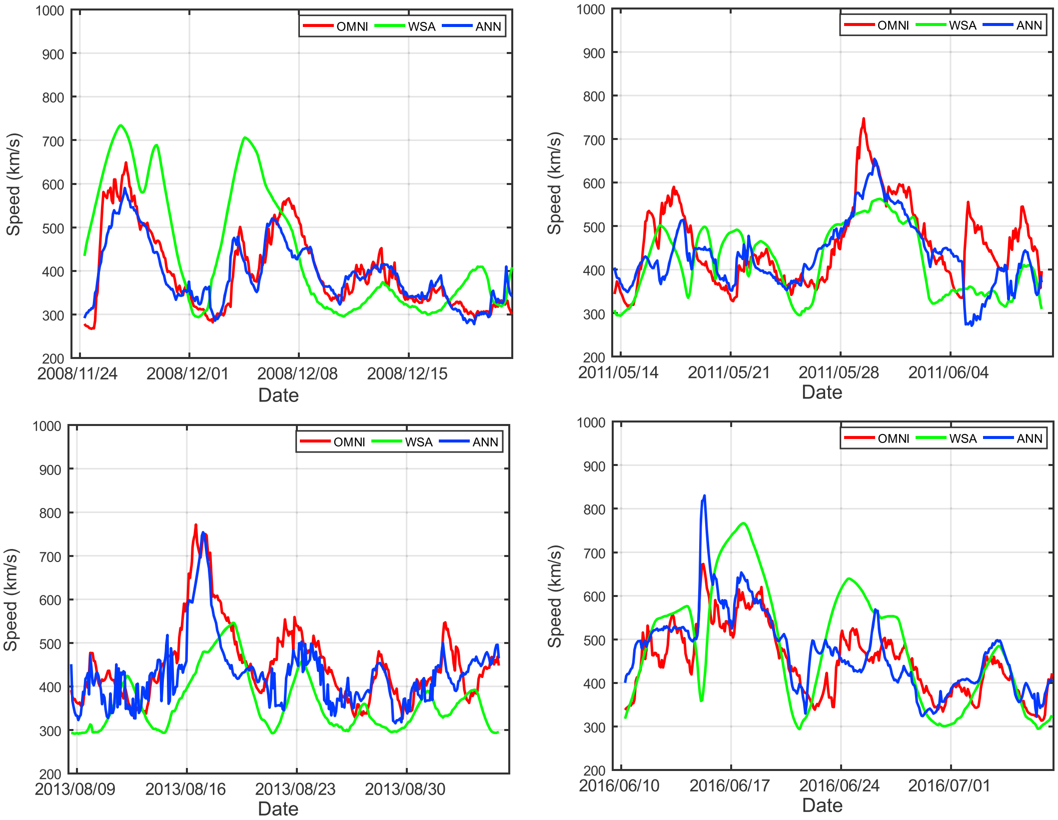

Figure 7. Solar wind speed predictions from our model and the Wang-Sheeley-Arge (WSA) model for four Carrington Rotations (CRs), respectively from the minimum, ascending, maximum, and descending phases of the 24th solar cycle, namely, CR2077 (top left), CR2110 (top right), CR2140 (bottom left), and CR2178 (bottom right). The red, blue, and green lines, respectively, represent the OMNI observational data, artificial neural network (ANN) modeling result, and WSA modeling result.

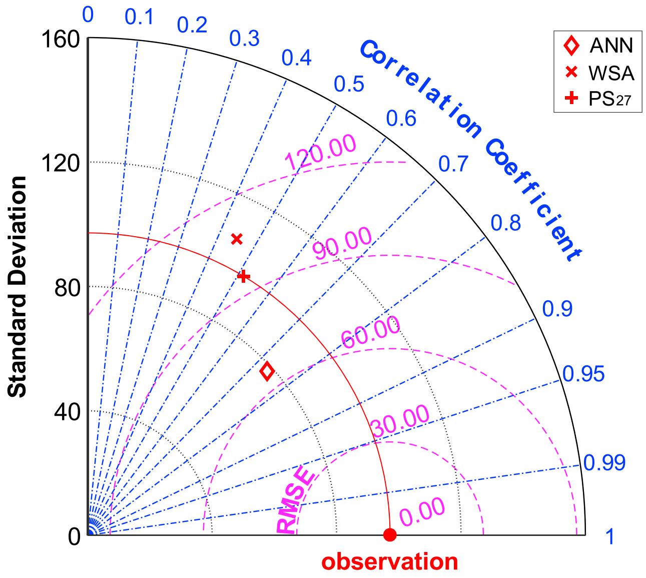

Figure 8. Taylor diagram for comparison of three solar wind speed prediction models. The results of artificial neural network (ANN) model, Wang-Sheeley-Arge (WSA) model, and a 27-day persistent model are represented by a diamond, multiple sign, and plus sign, respectively.

Moreover, the accuracy of the optimal ANN can be checked and compared with other models by a Taylor diagram (Taylor, 2001). The Taylor diagram contains three performance parameters (RMSE, STD, and *R*) in one figure and can present results of different models simultaneously by showing them with different markers. Figure 8 is a Taylor diagram for comparison of the ANN model, WSA model, and a 27-day persistent model. The black, blue, and magenta axes are stand for STD, *R*, and RMSE, respectively. The red round dot is the location of real observation. The red line represents the observed STD, which is about 95 km/s. In order to calculate the *R* and RMSE numbers in Figure 8, we took the *R* and RMSE between the red and blue lines (for ANN) and between the red and green lines (for WSA) from Figure 6. We can learn from the diagram that the ANN model has a bigger correlation coefficient and a smaller RMSE than the other two models. The overall correlation coefficients for the ANN, WSA, and PS_{27} model are 0.74, 0.45, and 0.51 respectively. The corresponding RMSEs are about 68, 106, and 95 km/s. However, due to the mathematical nature of a regression method, the ANN model has a smaller STD of 79 km/s compared to the observed STD of 95 km/s. The WSA STD (~106 km/s) is slightly closer to the STD of the observations than the ANN, which means that the WSA model does a better job as estimating the variability of the solar wind speed than the ANN.

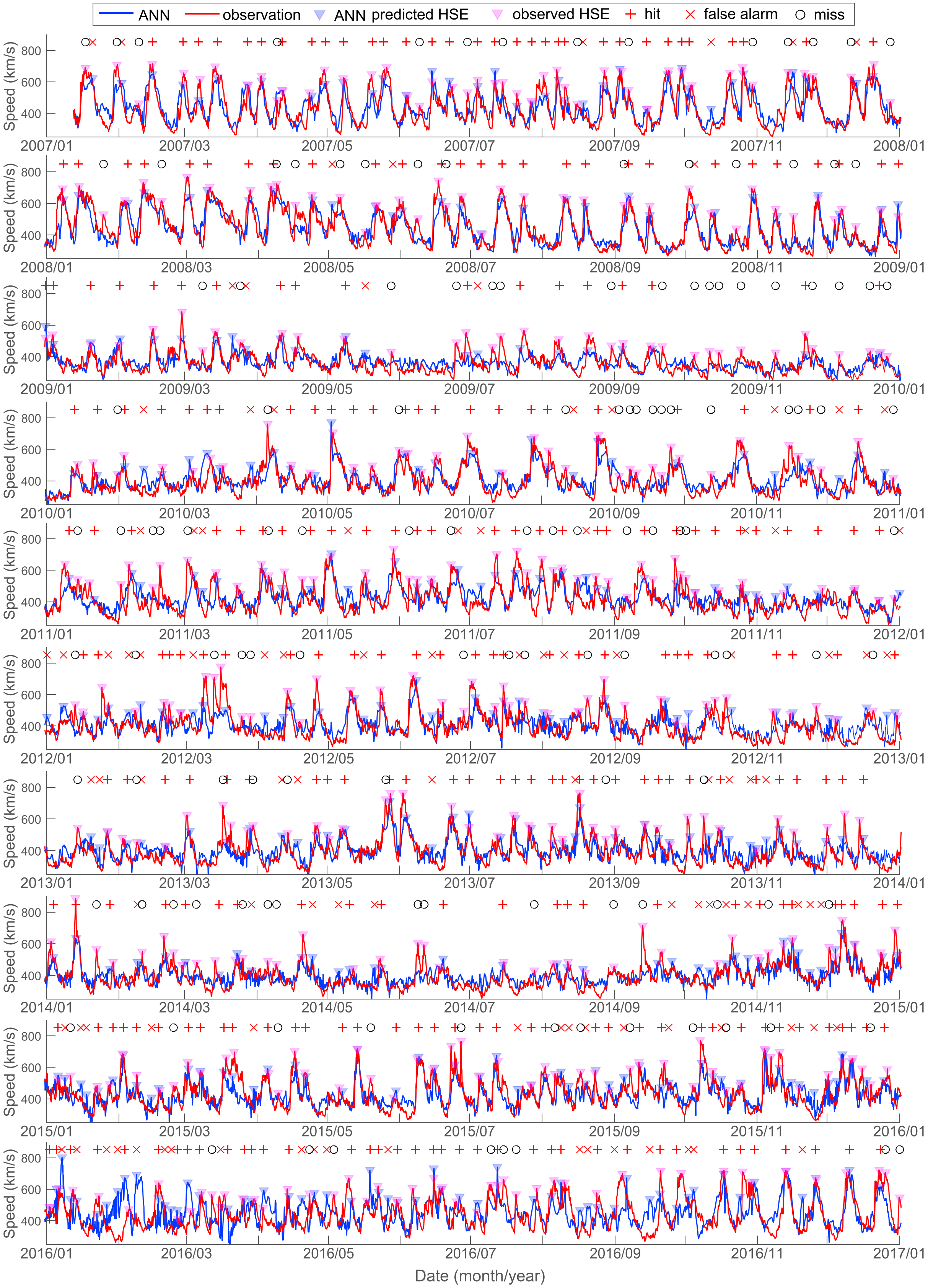

Figure 9. High-speed-enhancement (HSE) validation of our artificial neural network (ANN) model for the years from 2007 to 2016.

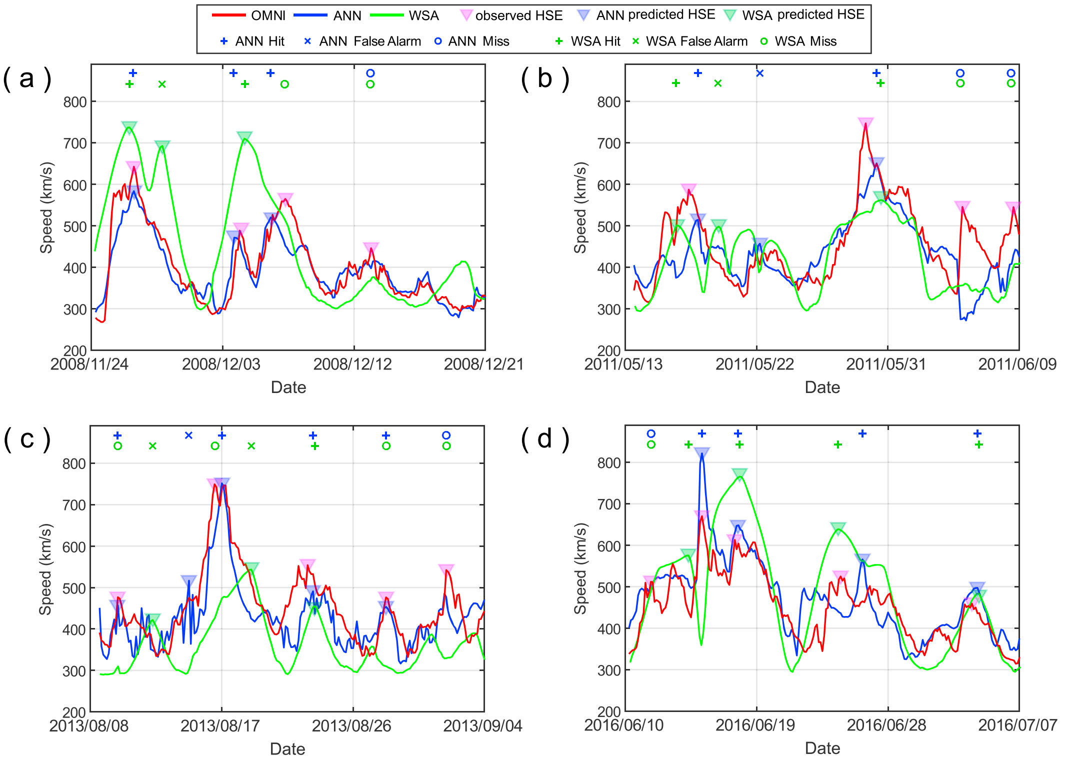

Figure 10. High-speed-enhancement (HSE) validation of our artificial neural network (ANN) model and Wang-Sheeley-Arge (WSA) model for four Carrington Rotations (CRs): (a) CR2077, (b) CR2110, (c) CR2140, and (d) CR2178.

Our results for WSA and PS27 models are reasonably consistent with studies made by others, considering the differences in these studies, such as the magnetogram sources, the time periods, the time resolutions, and the empirical formulae. For the WSA model, an RMSE of 94.9 km/s, an STD of 84.3 km/s, and an *R* of 0.51 were reported by Owens et al. (2008) for the period from 1995 to 2002; an RMSE of 99.8 km/s and an STD of 69.9 km/s were found by MacNeice (2009b) for the years 1976–2008; the RMSEs of 105 and 91 km/s were found for the years 2008 and 2011, respectively, by Bussy-Virat and Ridley (2014); and an RMSE of 99.5 km/s, an STD of 70.3 km/s, and an *R* of 0.35 were reported by Reiss et al. (2016) for 2011 to 2014. For the PS27 model, an RMSE of 75.4 km/s, an STD of 98.5 km/s, and an *R* of 0.47 were reported by Owens et al. (2013) for the period 1995–2003; the RMSEs of 163 and 128 km/s were found for the years 2008 and 2011, respectively by Bussy-Virat and Ridley (2014); and an RMSE of 100.4 km/s, an STD of 83.5 km/s, and an *R* of 0.28 were reported by Reiss et al. (2016) for the years 2011 to 2014.

Table 1
Comparison of the ANN Predicted and the Observed HSE From 2007 to 2016

Year	ANN predicted HSE	Observed HSE	Hit	False alarm	Miss	PD	PPV	TS
2007	38	46	32	6	14	0.70	0.84	0.62
2008	30	41	27	3	14	0.66	0.90	0.61
2009	22	35	18	4	17	0.51	0.82	0.46
2010	32	39	24	8	15	0.62	0.75	0.51
2011	39	46	29	10	17	0.63	0.74	0.52
2012	45	42	27	18	15	0.64	0.60	0.45
2013	43	40	32	11	8	0.80	0.74	0.63
2014	38	40	25	13	15	0.63	0.66	0.47
2015	53	50	38	15	12	0.76	0.72	0.58
2016	57	46	38	19	8	0.83	0.67	0.58
Total	397	425	290	107	135	0.68	0.73	0.55

For the validation of a solar wind speed prediction model, only error analysis is not enough. Furthermore, the prediction ability of our ANN model is examined by the high-speed-enhancement (HSE) method (e.g., MacNeice, 2009a; Owens et al., 2005; Reiss et al., 2016). In this paper, an HSE is defined as a speed increase larger than 100 km/s with a minimum peak speed of 400 km/s. The minimum distance of two HSE peaks is set as two days. Our HSE validation uses the same method described in detail by Reiss et al. (2016), except that we replaced the speed increase of 60 km/s in their HSE definition by 100 km/s in ours. The HSE validation is made by using the data set of years 2007 to 2016. Our validation results are presented as large-scale view in Figure 9 for the years from 2007 to 2016 and as zoomed-in view in Figure 10 for the same four CRs as that in Figure 7. The prediction horizon is four days, as in Figure 6. The performance results are shown in Table 1 for the ANN model and in Table 2 for the WSA model.

Table 2
Comparison of the WSA Predicted and the Observed HSE From 2007 to 2016

Year	WSA predicted HSE	Observed HSE	Hit	False alarm	Miss	PD	PPV	TS
2007	45	46	32	13	14	0.70	0.71	0.54
2008	45	41	27	18	14	0.66	0.60	0.46
2009	34	35	16	18	19	0.46	0.47	0.30
2010	39	39	21	18	18	0.54	0.54	0.37
2011	54	46	32	22	14	0.70	0.59	0.47
2012	43	42	23	20	19	0.55	0.54	0.37
2013	44	40	20	24	20	0.50	0.46	0.31
2014	53	40	22	31	18	0.55	0.42	0.31
2015	52	50	36	16	14	0.72	0.69	0.55
2016	57	46	37	20	9	0.80	0.65	0.56
Total	466	425	266	200	159	0.63	0.57	0.43

Every triangle in Figures 9 and 10 represents an HSE detected from observation or predicted by the models. In Figure 9, the red color presents the results from observation and the blue color represents the results from the ANN model. In Figure 10, the red and blue colors are consistent with Figure 9 and the green color represents the results from WSA model. In our HSE analysis, a *hit* means the predicted HSE meet the observation, a *false alarm* means the predicted HSE is not observed, and a *miss* means the observed HSE has not been predicted. The probability of HSE detection (*PD*) is usually defined as

$$\mathrm{PD} = \frac{\mathrm{hits}}{\mathrm{hits} + \mathrm{miss}} \tag{16}$$

The PD value of 1 means that all observed peaks are predicted, while the *PD* value of 0 means that all peaks are missed. A model could have a PD of 1 if it predicts a peak every time, so we introduce another metric named the positive predicted value (PPV) to monitor false alarms, which is defined as

$$\mathrm{PPV} = \frac{\mathrm{hits}}{\mathrm{hits} + \mathrm{false\ alarm}} \tag{17}$$

For example, a PPV of 0.8 means that 80% of the peaks predicted were observed, while 20% were false alarms. Another parameter for HSE analysis, named the threat score (*TS*), is a statistical measure of the overall model performance (Owens et al., 2005). The TS ranges from 0 to 1, which notes a totally no-skill with 0 and a perfect skill with 1. The TS can be calculated by

$$\mathrm{TS} = \frac{\mathrm{hits}}{\mathrm{hits} + \mathrm{miss} + \mathrm{false\ alarm}} \tag{18}$$

The statistical results of HSE for every year from 2007 to 2016 are shown in Table 1 for the ANN model and in Table 2 for the WSA model. For our ANN model, the PD ranges from 0.51 to 0.83 with an average of 0.68, the PPV ranges from 0.6 to 0.9 with an average of 0.73, and the TS varies from 0.45 to 0.63 with an average of 0.55. The largest PD is from 2016, and the biggest TS is from 2013. The lowest PD and TS are both from 2009, which is the year of solar minimum. This may be because the input parameters that we used to train the ANN are not very suitable for the solar minimum, and it is hard for the ANN to learn how to predict the speed in this period properly. The PPV of 2009 is higher than most years though, which demonstrates that the ANN has less false alarms near the solar minimum. However, the model has good results for other years, with PD larger than 0.6. The results from the descending phase of a solar cycle are better than other phases, such as 2007, 2008, and 2016. This may be because the input characters are more reliable in these periods.

For the WSA model shown in Table 2, the PD ranges from 0.46 to 0.8 with an average of 0.63, the PPV ranges from 0.42 to 0.71 with an average of 0.57, and the TS varies from 0.3 to 0.56 with an average of 0.43. Comparing these numbers with the corresponding ones in Table 1, we can conclude that our ANN model performs better for predicting HSEs than the WSA model. In previous studies, Owens et al. (2008) reported a PD of 0.59, a PPV of 0.84, and a TS of 0.51 for the period from 1995 to 2002; MacNeice (2009a) gave a PD of 0.4 and a PPV of 0.61 for the years 1976–2008; and Reiss et al. (2016) reported a PD of 0.37, a PPV of 0.51, and a TS of 0.28 for the years from 2011 to 2014. The differences of the definition for HSEs may cause the differences in the performance results.

5. Conclusions

In this paper, we present a HISS model, which applied an ANN with observational and theoretical input, to predict solar wind speed at 1 AU. The modeling results are validated by both error analysis and event-based analysis from 2007 to 2016. Our results demonstrate that the HISS model can predict the large-scale solar wind speed at 1 AU well, with an RMSE of 68 km/s, and a correlation coefficient of 0.74 for year 2007 to 2016. The HSE analysis shows that our model can predict 68.2% of the observed HSEs.

The comparison with WSA model demonstrates that our model generally performs better than the WSA model for predicting solar wind speed at 1 AU for a long period. This may illustrate that the input parameters more than f_s and θ_b in HISS model can be useful to predict solar wind speed. However, which parameter can cause a better prediction at a specific time needs further study.

The method used in this work is flexible, and it can be adjusted for specific purposes. For example, it can be applied separately to each of the four phases of the solar cycle, to examine whether or not some distinct differences among the phases exist or not. Since the influences of ICMEs might affect the prediction to some extent, a future work may train the ANN with a data set excluding ICMEs. Furthermore, the method in this paper can also be used to train ANNs for other parameters, such as temperature (T), density (N), and magnetic strength (B). Different solar wind parameters at 1 AU have correlation with each other (Yang et al., 2018), so we may use the same input characters such as f_s and θ_b to predict T, N, and B.

The artificial intelligence techniques have been greatly improved in the last few years, with the explosion of deep learning. It can be expected that the solar wind prediction will be greatly improved with more advanced machine learning techniques and more observational data collected. Therefore, a deep neural network with many hidden layers may further improve the performance of our work.

Acknowledgments
The data for this work are available at the official websites of NSO/GONG, SOHO/LASCO, and OMNI. We acknowledge the use of them. Global Oscillation Network Group (GONG) magnetograms are downloaded from the website of National Solar Observatory (NSO, http://gong.nso.edu/). SOHO is a mission of international cooperation between ESA and NASA. The SOHO/LASCO data used here are produced by a consortium of the Naval Research Laboratory (USA), Max-Planck-Institut fuer Aeronomie (Germany), Laboratoire d'Astronomie (France), and the University of Birmingham (UK). The polarized brightness (pB) data can be downloaded from the LASCO instrument website at NRL (http://lasco-www.nrl.navy.mil/content/retrieve/polarize/). The OMNI Web interface is provided by the Space Physics Data Facility (SPDF) of NASA/Goddard Space Flight Center (GSFC). The hourly-averaged OMNI data are obtained from the OMNI Web interface http://omniweb.gsfc.nasa.gov. This work is jointly supported by grants from the National Natural Science Foundation of China (41774184, 41474152, 41731067, and 41531073) and the Specialized Research Fund for State Key Laboratories. F. S. is also supported by the National Program for Support of Top-notch Young Professionals.

References

Altschuler, M. D., & Newkirk, G. Jr. (1969). Magnetic fields and the structure of the solar corona. *Solar Physics*, *9*(1), 131–149.

Alves, M., Echer, E., & Gonzalez, W. (2006). Geoeffectiveness of corotating interaction regions as measured by Dst index. *Journal of Geophysical Research*, *111*, A07S05. https://doi.org/10.1029/2005JA011379

Arge, C. N., Odstrcil, D., Pizzo, V. J., & Mayer, L. R. (2003). Improved method for specifying solar wind speed near the Sun. *AIP Conference Proceedings*, 679, 190–193. https://doi.org/10.1063/1.1618574

Arge, C. N., & Pizzo, V. J. (2000). Improvement in the prediction of solar wind conditions using near-real time solar magnetic field updates. *Journal of Geophysical Research*, *105*(A5), 10,465–10,479. https://doi.org/10.1029/1999JA000262

Baker, D. (1998). What is space weather? *Advances in Space Research*, *22*(1), 7–16.

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*(4), 275–280.

Borda, R. A. F., Mininni, P. D., Mandrini, C. H., Gómez, D. O., Bauer, O. H., & Rovira, M. G. (2002). Automatic solar flare detection using neural network techniques. *Solar Physics*, *206*(2), 347–357.

Borovsky, J. E., & Denton, M. H. (2006). Differences between CME-driven storms and CIR-driven storms. *Journal of Geophysical Research*, *111*, A07S08. https://doi.org/10.1029/2005JA011447

Bortnik, J., Li, W., Thorne, R. M., & Angelopoulos, V. (2016). A unified approach to inner magnetospheric state prediction. *Journal of Geophysical Research: Space Physics*, *121*, 2423–2430. https://doi.org/10.1002/2015JA021733

Bothmer, V., & Daglis, I. A. (2007). *Space weather: Physics and effects*. Berlin: Springer Science & Business Media. https://doi.org/10.1007/978-3-540-34578-7

Burlaga, L., & Lepping, R. (1977). The causes of recurrent geomagnetic storms. *Planetary and Space Science*, *25*(12), 1151–1160.

Bussy-Virat, C. D., & Ridley, A. J. (2014). Predictions of the solar wind speed by the probability distribution function model. *Space Weather*, *12*, 337–353. https://doi.org/10.1002/2014SW001051

Bussy-Virat, C. D., & Ridley, A. J. (2016). Twenty-four hour predictions of the solar wind speed peaks by the probability distribution function model. *Space Weather*, *14*, 861–873. https://doi.org/10.1002/2016SW001437

Cai, L., Ma, S. Y., & Zhou, Y. L. (2010). Prediction of SYM-H index during large storms by NARX neural network from IMF and solar wind data. *Annales Geophysicae*, *28*(2), 381–393. https://doi.org/10.5194/angeo-28-381-2010

Calvo, R., Ceccato, H., & Piacentini, R. (1995). Neural network prediction of solar activity. *The Astrophysical Journal*, *444*, 916–921.

Chu, X., Bortnik, J., Li, W., Ma, Q., Denton, R., Yue, C., et al. (2017). A neural network model of three-dimensional dynamic electron density in the inner magnetosphere. *Journal of Geophysical Research: Space Physics*, *122*, 9183–9197. https://doi.org/10.1002/2017JA024464

Conlon, P. A., & Gallagher, P. T. (2010). Constraining three-dimensional magnetic field extrapolations using the twin perspectives of STEREO. *The Astrophysical Journal*, *715*(1), 59–65.

Conway, A., Macpherson, K., Blacklaw, G., & Brown, J. (1998). A neural network prediction of solar cycle 23. *Journal of Geophysical Research*, *103*(A12), 29,733–29,742.

de Toma, G. (2011). Evolution of coronal holes and implications for high-speed solar wind during the minimum between cycles 23 and 24. *Solar Physics*, *274*(1–2), 195–217.

Detman, T., Intriligator, D., Dryer, M., Sun, W., Deehr, C., & Intriligator, J. (2011). The influence of pickup protons, from interstellar neutral hydrogen, on the propagation of interplanetary shocks from the Halloween 2003 solar events to ACE and Ulysses: A 3-D MHD modeling study. *Journal of Geophysical Research*, *116*, A03105. https://doi.org/10.1029/2010JA015803

Detman, T., Smith, Z., Dryer, M., Fry, C. D., Arge, C. N., & Pizzo, V. (2006). A hybrid heliospheric modeling system: Background solar wind. *Journal of Geophysical Research*, *111*, A07102. https://doi.org/10.1029/2005JA011430

Dryer, M., Smith, Z., Fry, C., Sun, W., Deehr, C., & Akasofu, S. I. (2004). Real-time shock arrival predictions during the “Halloween 2003 epoch”. *Space Weather*, *2*, S09001. https://doi.org/10.1029/2004SW000087

Feng, X., Yang, L., Xiang, C., Wu, S., Zhou, Y., & Zhong, D. (2010). Three-dimensional solar wind modeling from the Sun to Earth by a SIP-CESE MHD model with a six-component grid. *The Astrophysical Journal*, *723*(1), 300–319.

Feng, X., Zhou, Y., & Wu, S. (2007). A novel numerical implementation for solar wind modeling by the modified conservation element/solution element method. *The Astrophysical Journal*, *655*(2), 1110–1126.

Fessant, F., Bengio, S., & Collobert, D. (1996). On the prediction of solar activity using different neural network models, paper presented at Annales Geophysicae, Citeseer.

Feynman, J., & Gabriel, S. (2000). On space weather consequences and predictions. *Journal of Geophysical Research*, *105*(A5), 10,543–10,564.

Fry, C., Sun, W., Deehr, C., Dryer, M., Smith, Z., Akasofu, S. I., et al. (2001). Improvements to the HAF solar wind model for space weather predictions. *Journal of Geophysical Research*, *106*(A10), 20,985–21,001.

Fujiki, K. i., Tokumaru, M., Iju, T., Hakamada, K., & Kojima, M. (2015). Relationship between solar-wind speed and coronal magnetic-field properties. *Solar Physics*, *290*(9), 2491–2505.

Gong, J.-c., Xue, B.-s., Liu, S.-q., Zou, Z.-m., Miao, J., & Wang, J.-l. (2004). Short-term prediction of solar proton events by neural network method. *Chinese Astronomy and Astrophysics*, *28*(2), 174–182.

Gopalswamy, N., Lara, A., Yashiro, S., Kaiser, M. L., & Howard, R. A. (2001). Predicting the 1-AU arrival times of coronal mass ejections. *Journal of Geophysical Research*, *106*(A12), 29,207–29,217.

Gosling, J., & Pizzo, V. (1999). Formation and evolution of corotating interaction regions and their three dimensional structure. In A. Balogh, J. T. Gosling, J. R. Jokipii, R. Kallenbach, & H. Kunow (Eds.), *Corotating interaction regions* (Vol. 7, pp. 21–52). Dordrecht: Springer.

Hagan, M. T., Demuth, H. B., Beale, M. H., & De Jesús, O. (1996). *Neural network design*. Boston: Pws Pub.

Hagan, M. T., & Menhaj, M. B. (1994). Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*, *5*(6), 989–993.

Hapgood, M. (2011). Towards a scientific understanding of the risk from extreme space weather. *Advances in Space Research*, *47*(12), 2059–2072.

Hayashi, K. (2005). Magnetohydrodynamic simulations of the solar corona and solar wind using a boundary treatment to limit solar wind mass flux. *The Astrophysical Journal Supplement Series*, *161*(2), 480–494.

Hayashi, K. (2012). An MHD simulation model of time-dependent co-rotating solar wind. *Journal of Geophysical Research*, *117*, A08105. https://doi.org/10.1029/2011JA017490

Hayashi, K., Tokumaru, M., & Fujiki, K. I. (2016). MHD-IPS analysis of relationship among solar wind density, temperature, and flow speed. *Journal of Geophysical Research: Space Physics*, *121*, 7367–7384. https://doi.org/10.1002/2016JA022750

Hayashi, K., Yang, S., & Deng, Y. (2016). Comparison of potential field solutions for Carrington Rotation 2144. *Journal of Geophysical Research: Space Physics*, *121*, 1046–1061. https://doi.org/10.1002/2015JA021757

Haykin, S. (2004). A comprehensive foundation. *Neural Networks*, *2*, 41.

Innocenti, M. E., Lapenta, G., Vršnak, B., Crespon, F., Skandrani, C., Temmer, M., et al. (2011). Improved forecasts of solar wind parameters using the Kalman filter. *Space Weather*, *9*, S10005. https://doi.org/10.1029/2011SW000659

Intriligator, D. S., Detman, T., Gloecker, G., Gloeckler, C., Dryer, M., Sun, W., et al. (2012). Pickup protons: Comparisons using the three-dimensional MHD HHMS-PI model and Ulysses SWICS measurements. *Journal of Geophysical Research*, *117*, A06104. https://doi.org/10.1029/2011JA017424

Jones, S. I., Davila, J. M., & Uritsky, V. (2016). Optimizing global coronal magnetic field models using image-based constraints. *The Astrophysical Journal*, *820*(2), 113.

Jones, S. I., Uritsky, V., & Davila, J. M. (2017). Image-optimized coronal magnetic field models. *The Astrophysical Journal*, *844*(2), 93.

Kissmann, R., Kleimann, J., Fichtner, H., & Grauer, R. (2008). Local turbulence simulations for the multiphase ISM. *Monthly Notices of the Royal Astronomical Society*, *391*(4), 1577–1588.

Koons, H. C., & Gorney, D. J. (1991). A neural network model of the relativistic electron flux at geosynchronous orbit. *Journal of Geophysical Research*, *96*(A4), 5549. https://doi.org/10.1029/90JA02380

Lapenta, G., Pierrard, V., Keppens, R., Markidis, S., Poedts, S., Šebek, O., et al. (2013). SWIFF: Space weather integrated forecasting framework. *Journal of Space Weather and Space Climate*, *3*, A05.

Lazzús, J. A., Vega, P., Rojas, P., & Salfate, I. (2017). Forecasting the Dst index using a swarm-optimized neural network. *Space Weather*, *15*, 1068–1089. https://doi.org/10.1002/2017SW001608

Lee, C., Luhmann, J., Hoeksema, J., Sun, X., Arge, C., & de Pater, I. (2011). Coronal field opens at lower height during the solar cycles 22 and 23 minimum periods: IMF comparison suggests the source surface should be lowered. *Solar Physics*, *269*(2), 367–388.

Levine, R. H., Altschuler, M. D., & Harvey, J. W. (1977). Solar sources of the interplanetary magnetic field and solar wind. *Journal of Geophysical Research*, *82*(7), 1061–1065.

Ling, A., Ginet, G., Hilmer, R., & Perry, K. (2010). A neural network–based geosynchronous relativistic electron flux forecasting model. *Space Weather*, *8*, S09003. https://doi.org/10.1029/2010SW000576

Liu, D. D., Huang, C., Lu, J. Y., & Wang, J. S. (2011). The hourly average solar wind velocity prediction based on support vector regression method. *Monthly Notices of the Royal Astronomical Society*, *413*(4), 2877–2882. https://doi.org/10.1111/j.1365-2966.2011.18359.x

Liu, J., Ye, Y., Shen, C., Wang, Y., & Erdélyi, R. (2018). A new tool for CME arrival time prediction using machine learning algorithms: CAT-PUMA. *The Astrophysical Journal*, *855*(2). https://doi.org/10.3847/1538-4357/aaae69

Liu, Y., Hoeksema, J., Scherrer, P., Schou, J., Couvidat, S., Bush, R., et al. (2012). Comparison of line-of-sight magnetograms taken by the solar dynamics observatory/helioseismic and magnetic imager and solar and heliospheric observatory/Michelson Doppler imager. *Solar Physics*, *279*(1), 295–316.

Lowder, C., Qiu, J., Leamon, R., & Liu, Y. (2014). Measurements of EUV coronal holes and open magnetic flux. *The Astrophysical Journal*, *783*(2), 142.

Lugaz, N., Manchester, W. IV, & Gombosi, T. (2005). Numerical simulation of the interaction of two coronal mass ejections from Sun to Earth. *The Astrophysical Journal*, *634*(1), 651–662.

Lundstedt, H. (2005). Progress in space weather predictions and applications. *Advances in Space Research*, *36*(12), 2516–2523. https://doi.org/10.1016/j.asr.2003.09.072

Luo, B., Zhong, Q., Liu, S., & Gong, J. (2008). A new forecasting index for solar wind velocity based on EIT 284 Å observations. *Solar Physics*, *250*(1), 159–170. https://doi.org/10.1007/s11207-008-9198-4

MacNeice, P. (2009a). Validation of community models: Identifying events in space weather model timelines. *Space Weather*, *7*, S06004. https://doi.org/10.1029/2009SW000463

MacNeice, P. (2009b). Validation of community models: 2. Development of a baseline using the Wang-Sheeley-Arge model. *Space Weather*, *7*, S12002. https://doi.org/10.1029/2009SW000489

Macpherson, K. P., Conway, A. J., & Brown, J. C. (1995). Prediction of solar and geomagnetic activity data using neural networks. *Journal of Geophysical Research*, *100*(A11), 21,735–21,744. https://doi.org/10.1029/95JA02283

Merkin, V., Lyon, J., McGregor, S., & Pahud, D. (2011). Disruption of a heliospheric current sheet fold. *Geophysical Research Letters*, *38*, L14107. https://doi.org/10.1029/2011GL047822

Nolte, J., Krieger, A., Timothy, A., Gold, R., Roelof, E., Vaiana, G., et al. (1976). Coronal holes as sources of solar wind. *Solar Physics*, *46*(2), 303–322.

Obridko, V., Shelting, B., Livshits, I., & Asgarov, A. (2009). Contrast of coronal holes and parameters of associated solar wind streams. *Solar Physics*, *260*(1), 191–206.

Odstrcil, D. (2003). Modeling 3-D solar wind structure. *Advances in Space Research*, *32*(4), 497–506.

Odstrcil, D., Pizzo, V. J., Linker, J. A., Riley, P., Lionello, R., & Mikic, Z. (2004). Initial coupling of coronal and heliospheric numerical magnetohydrodynamic codes. *Journal of Atmospheric and Solar-Terrestrial Physics*, *66*(15–16), 1311–1320.

Owens, M. J., Arge, C. N., Spence, H. E., & Pembroke, A. (2005). An event-based approach to validating solar wind speed predictions: High-speed enhancements in the Wang-Sheeley-Arge model. *Journal of Geophysical Research*, *110*, A12105. https://doi.org/10.1029/2005JA011343

Owens, M. J., Challen, R., Methven, J., Henley, E., & Jackson, D. R. (2013). A 27 day persistence model of near-Earth solar wind conditions: A long lead-time forecast and a benchmark for dynamical models. *Space Weather*, *11*, 225–236. https://doi.org/10.1002/swe.20040

Owens, M. J., Spence, H. E., McGregor, S., Hughes, W. J., Quinn, J. M., Arge, C. N., et al. (2008). Metrics for solar wind prediction models: Comparison of empirical, hybrid, and physics-based schemes with 8 years of L1 observations. *Space Weather*, *6*, S08001. https://doi.org/10.1029/2007SW000380

Pahud, D., Merkin, V., Arge, C., Hughes, W., & McGregor, S. (2012). An MHD simulation of the inner heliosphere during Carrington rotations 2060 and 2068: Comparison with MESSENGER and ACE spacecraft observations. *Journal of Atmospheric and Solar-Terrestrial Physics*, *83*, 32–38.

Qahwaji, R., & Colak, T. (2007). Automatic short-term solar flare prediction using machine learning and sunspot associations. *Solar Physics*, *241*(1), 195–211.

Reiss, M. A., Temmer, M., Veronig, A. M., Nikolic, L., Vennerstrom, S., Schöngassner, F., & Hofmeister, S. J. (2016). Verification of high-speed solar wind stream forecasts using operational solar wind models. *Space Weather*, *14*, 495–510. https://doi.org/10.1002/2016SW001390

Richardson, I., Cane, H., & Cliver, E. (2002). Sources of geomagnetic activity during nearly three solar cycles (1972–2000). *Journal of Geophysical Research*, *107*(A8), 1187. https://doi.org/10.1029/2001JA000504

Riley, P., Ben-Nun, M., Linker, J., Mikic, Z., Svalgaard, L., Harvey, J., et al. (2014). A multi-observatory inter-comparison of line-of-sight synoptic solar magnetograms. *Solar Physics*, *289*(3), 769–792.

Riley, P., Linker, J., & Mikić, Z. (2001). An empirically-driven global MHD model of the solar corona and inner heliosphere. *Journal of Geophysical Research*, *106*(A8), 15,889–15,901.

Riley, P., Linker, J., Mikić, Z., Lionello, R., Ledvina, S., & Luhmann, J. (2006). A comparison between global solar magnetohydrodynamic and Potential Field Source Surface model results. *The Astrophysical Journal*, *653*(2), 1510–1516.

Riley, P., Linker, J. A., & Arge, C. N. (2015). On the role played by magnetic expansion factor in the prediction of solar wind speed. *Space Weather*, *13*, 154–169. https://doi.org/10.1002/2014SW001144

Robbins, S., Henney, C. J., & Harvey, J. (2006). Solar wind forecasting with coronal holes. *Solar Physics*, *233*(2), 265–276.

Rotter, T., Veronig, A. M., Temmer, M., & Vršnak, B. (2012). Relation between coronal hole areas on the Sun and the solar wind parameters at 1 AU. *Solar Physics*, *281*(2), 793–813. https://doi.org/10.1007/s11207-012-0101-y

Rotter, T., Veronig, A. M., Temmer, M., & Vršnak, B. (2015). Real-time solar wind prediction based on SDO/AIA coronal hole data. *Solar Physics*, *290*(5), 1355–1370. https://doi.org/10.1007/s11207-015-0680-5

Schatten, K. H., Wilcox, J. M., & Ness, N. F. (1969). A model of interplanetary and coronal magnetic fields. *Solar Physics*, *6*(3), 442–455.

Sheeley, N., & Harvey, J. (1981). Coronal holes, solar wind streams, and geomagnetic disturbances during 1978 and 1979. *Solar Physics*, *70*(2), 237–249.

Shen, F., Feng, X., & Song, W. (2009). An asynchronous and parallel time-marching method: Application to three-dimensional MHD simulation of solar wind. *Science in China Series E: Technological Sciences*, *52*(10), 2895–2902.

Shen, F., Shen, C., Wang, Y., Feng, X., & Xiang, C. (2013). Could the collision of CMEs in the heliosphere be super–elastic? Validation through three–dimensional simulations. *Geophysical Research Letters*, *40*, 1457–1461. https://doi.org/10.1002/grl.50336

Shen, F., Shen, C., Zhang, J., Hess, P., Wang, Y., Feng, X., et al. (2014). Evolution of the 12 July 2012 CME from the Sun to the Earth: Data-constrained three-dimensional MHD simulations. *Journal of Geophysical Research: Space Physics*, *119*, 7128–7141. https://doi.org/10.1002/2014JA020365

Siscoe, G. (2000). The space-weather enterprise: Past, present, and future. *Journal of Atmospheric and Solar-Terrestrial Physics*, *62*(14), 1223–1232.

Sudar, D., Vršnak, B., & Dumbović, M. (2015). Predicting coronal mass ejections transit times to Earth with neural network. *Monthly Notices of the Royal Astronomical Society*, *456*(2), 1542–1548.

Suzuki, T. K. (2006). Forecasting solar wind speeds. *The Astrophysical Journal Letters*, *640*(1), L75–L78.

Suzuki, T. K., & Inutsuka, S.-i. (2005). Making the corona and the fast solar wind: a self-consistent simulation for the low-frequency Alfvén waves from the photosphere to 0.3 AU. *The Astrophysical Journal Letters*, *632*(1), L49.

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*(D7), 7183–7192.

Tóth, G., Sokolov, I. V., Gombosi, T. I., Chesney, D. R., Clauer, C. R., De Zeeuw, D. L., et al. (2005). Space Weather Modeling Framework: A new tool for the space science community. *Journal of Geophysical Research*, *110*, A12226. https://doi.org/10.1029/2005JA011126

Tóth, G., van der Holst, B., & Huang, Z. (2011). Obtaining potential field solutions with spherical harmonics and finite differences. *The Astrophysical Journal*, *732*(2), 102. https://doi.org/10.1088/0004-637x/732/2/102

Tsurutani, B. T., Gonzalez, W. D., Gonzalez, A. L., Guarnieri, F. L., Gopalswamy, N., Grande, M., et al. (2006). Corotating solar wind streams and recurrent geomagnetic activity: A review. *Journal of Geophysical Research*, *111*, A07S01. https://doi.org/10.1029/2005JA011273

Tsurutani, B. T., Gonzalez, W. D., Gonzalez, A. L., Tang, F., Arballo, J. K., & Okada, M. (1995). Interplanetary origin of geomagnetic activity in the declining phase of the solar cycle. *Journal of Geophysical Research*, *100*(A11), 21,717–21,733.

Tsurutani, B. T., Gonzalez, W. D., Tang, F., Akasofu, S. I., & Smith, E. J. (1988). Origin of interplanetary southward magnetic fields responsible for major magnetic storms near solar maximum (1978–1979). *Journal of Geophysical Research*, *93*(A8), 8519–8531.

Usmanov, A., & Goldstein, M. (2003). A tilted-dipole MHD model of the solar corona and solar wind. *Journal of Geophysical Research*, *108*(A9), 1354. https://doi.org/10.1029/2002JA009777

Usmanov, A., & Goldstein, M. (2006). A three-dimensional MHD solar wind model with pickup protons. *Journal of Geophysical Research*, *111*, A07101. https://doi.org/10.1029/2005JA011533

Uwamahoro, J., McKinnell, L., & Habarulema, J. (2012). Estimating the geoeffectiveness of halo CMEs from associated solar and IP parameters using neural networks, paper presented at Annales Geophysicae, Copernicus GmbH.

Valach, F., Revallo, M., Bochníček, J., & Hejda, P. (2009). Solar energetic particle flux enhancement as a predictor of geomagnetic activity in a neural network-based model. *Space Weather*, *7*, S04004. https://doi.org/10.1029/2008SW000421

Van de Hulst, H. (1950). The electron density of the solar corona. *Bulletin of the Astronomical Institutes of the Netherlands*, *11*, 135–150.

Verbanac, G., Vršnak, B., Veronig, A., & Temmer, M. (2011). Equatorial coronal holes, solar wind high-speed streams, and their geoeffectiveness. *Astronomy & Astrophysics*, *526*, A20.

Vršnak, B., & Gopalswamy, N. (2002). Influence of the aerodynamic drag on the motion of interplanetary ejecta. *Journal of Geophysical Research*, *107*(A2), 1019. https://doi.org/10.1029/2001JA000120

Vršnak, B., Temmer, M., & Veronig, A. M. (2007). Coronal holes and solar wind high-speed streams: I. Forecasting the Solar Wind Parameters. *Solar Physics*, *240*(2), 315–330. https://doi.org/10.1007/s11207-007-0285-8

Wang, H., Cui, Y., Li, R., Zhang, L., & Han, H. (2008). Solar flare forecasting model supported with artificial neural network techniques. *Advances in Space Research*, *42*(9), 1464–1468.

Wang, T., & Davila, J. M. (2014). Validation of spherically symmetric inversion by use of a tomographically reconstructed three-dimensional Electron density of the solar corona. *Solar Physics*, *289*(10), 3723–3745. https://doi.org/10.1007/s11207-014-0556-0

Wang, Y.-M., & Sheeley, N. Jr. (1990). Solar wind speed and coronal flux-tube expansion. *The Astrophysical Journal*, *355*, 726–732.

Wiengarten, T., Kleimann, J., Fichtner, H., Cameron, R., Jiang, J., Kissmann, R., & Scherer, K. (2013). MHD simulation of the inner-heliospheric magnetic field. *Journal of Geophysical Research: Space Physics*, *118*, 29–44. https://doi.org/10.1029/2012JA018089

Wintoft, P., & Lundstedt, H. (1997). Prediction of daily average solar wind velocity from solar magnetic field observations using hybrid intelligent systems. *Physics and Chemistry of the Earth*, *22*(7), 617–622.

Wintoft, P., & Lundstedt, H. (1999). A neural network study of the mapping from solar magnetic fields to the daily average solar wind velocity. *Journal of Geophysical Research*, *104*(A4), 6729–6736.

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*. https://doi.org/10.1051/swsc/2017027

Wu, C. C., Dryer, M., Wu, S., Wood, B. E., Fry, C. D., Liou, K., & Plunkett, S. (2011). Global three-dimensional simulation of the interplanetary evolution of the observed geoeffective coronal mass ejection during the epoch 1–4 August 2010. *Journal of Geophysical Research*, *116*, A12103. https://doi.org/10.1029/2011JA016947

Wu, C. C., Fry, C., Wu, S., Dryer, M., & Liou, K. (2007). Three-dimensional global simulation of interplanetary coronal mass ejection propagation from the Sun to the heliosphere: Solar event of 12 May 1997. *Journal of Geophysical Research*, *112*, A09104. https://doi.org/10.1029/2006JA012211

Wu, J. G., & Lundstedt, H. (1997). Neural network modeling of solar wind-magnetosphere interaction. *Journal of Geophysical Research*, *102*(A7), 14,457–14,466.

Wu, S. T., & Dryer, M. (2015). Comparative analyses of current three-dimensional numerical solar wind models. *Science China Earth Sciences*, *58*(6), 839–858. https://doi.org/10.1007/s11430-015-5062-1

Yang, Z., Shen, F., Zhang, J., Yang, Y., Feng, X., & Richardson, I. G. (2018). Correlation between the magnetic field and plasma parameters at 1 AU. *Solar Physics*, *293*(2), 24.

Zhang, J., Richardson, I., Webb, D., Gopalswamy, N., Huttunen, E., Kasper, J., et al. (2007). Solar and interplanetary sources of major geomagnetic storms (Dst$\leq-$ 100 nT) during 1996–2005. *Journal of Geophysical Research*, *112*, A07S09. https://doi.org/10.1029/2005JA011476

Zhao, X., & Dryer, M. (2014). Current status of CME/shock arrival time prediction. *Space Weather*, *12*, 448–469. https://doi.org/10.1002/2014SW001060

Zhelavskaya, I. S., Shprits, Y. Y., & Spasojević, M. (2017). Empirical modeling of the Plasmasphere dynamics using neural networks. *Journal of Geophysical Research: Space Physics*, *122*, 11,227–211,244. https://doi.org/10.1002/2017JA024406