

JGR Space Physics



RESEARCH ARTICLE

10.1029/2018JA025957

Special Section:

The Earth's Magnetosphere:
New Tools, New Thinking,
New Results

System Identification and Data-Driven Forecasting of AE Index and Prediction Uncertainty Analysis Using a New Cloud-NARX Model

Y. Gu¹ , Hua-Liang Wei¹ , Richard J. Boynton¹ , Simon N. Walker¹ , and Michael A. Balikhin¹¹Department of Automatic Control and Systems Engineering, The University of Sheffield, Sheffield, UK

Key Points:

- A novel cloud-NARX model is presented for modeling and forecasting of AE index with solar wind parameters as input variables
- The cloud-NARX model generates 1-hr-ahead prediction of AE index with a confidence interval
- The uncertainty of model structure, parameters, and predictions can be quantified by the cloud-NARX model, and the model robustness can be improved

Correspondence to:

H.-L. Wei,
w.hualiang@sheffield.ac.uk

Citation:

Gu, Y., Wei, H.-L., Boynton, R. J., Walker, S. N., & Balikhin, M. A. (2019). System identification and data-driven forecasting of AE index and prediction uncertainty analysis using a new cloud-NARX model. *Journal of Geophysical Research: Space Physics*, 124, 248–263. <https://doi.org/10.1029/2018JA025957>

Received 1 AUG 2018

Accepted 20 NOV 2018

Accepted article online 13 DEC 2018

Published online 22 JAN 2019

Abstract Severe geomagnetic storms caused by the solar wind disturbances have harmful influences on the operation of modern equipment and systems. The modeling and forecasting of AE index are extremely useful to understand the geomagnetic substorms. This study presents a novel cloud-nonlinear autoregressive with exogenous input (NARX) model to predict AE index 1 hr ahead. The cloud-NARX model provides AE index forecasting results, with a correlation coefficient of 0.87 on the data of whole year 2015. The benchmarks on the data of the two interested periods of 17–21 March 2015 and 22–26 June 2015 are presented. The presented model uses uncertainty “cloud” model and cloud transformation to quantify the uncertainty throughout the structure detection, parameter estimation, and model prediction. The new predicted band can be generated to forecast AE index with confidence interval. The proposed method provides a new way to evaluate the model based on uncertainty analysis, revealing the reliability of model, and visualize the bias of model prediction.

1. Introduction

Many modern technological systems are sensitive to space weather disturbances, such as geomagnetic storms and substorms and ionosphere variability (Knipp, 2012; MacAlester & Murtagh, 2014; Schrijver, 2015). The severe situations of space weather can have harmful effect on power grid, navigation systems, and satellite system. Thus, it is extremely important to forecast the space weather disturbances to avoid damages and losses. In addition to the traditional first principle and statistical approaches for understanding the interactions between the solar wind and magnetosphere (e.g., Ala-Lahti et al., 2018, and the references therein), application of data-based methods and in particular techniques based on machine learning to the prediction of various geomagnetic indices resulted in many innovative forecasting models (e.g., Camporeale et al., 2018; Chandorkar et al., 2017; Wintoft & Cander, 2000).

The AE index, along with the Al and AU indices, was introduced by Davis and Sugiura (1966) as a measurement of global auroral electrojet activity (Mayaud, 1980). Changes in AE are driven by variations in the solar wind convection electric field produced by fluctuations in the solar wind velocity and interplanetary magnetic field (IMF). These two factors govern the efficiency of the coupling between the solar wind and terrestrial magnetosphere with the dominant role being played by a southward directed IMF. In this coupling process, the energy associated with the solar wind flow is converted into magnetic energy, which is transferred into the magnetosphere via reconnection processes on the dayside and is stored in the magnetotail. This energy is eventually released, energizing the plasma sheet, ring current, and ionosphere.

Three classes of interactions have been identified, depending upon the southward turnings of the IMF (see, e.g., Gonzalez et al., 1994). Short-lived southward turnings of the IMF with modest ($B_z \approx 3\text{ nT}$) give rise to minor intensifications of the ring current, yielding substorm events. Repeated southward turnings, referred to as high intensity, long duration, continuous AE activity events arise due to the occurrence of interplanetary Alfvén wave train embedded within the solar wind flow (Tsurutani & Gonzalez, 1987). These events result in a continued period of AE activity. Finally, coronal mass ejections (CMEs) or magnetic clouds exhibit extended periods in which a strong B_z is observed. This coupling, between the CME and terrestrial magnetosphere, results in a major intensification of the ring current, and large deviations in both AE and Dst, and is referred to as an intense magnetic storm.

©2018. The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Previous studies of substorm using AE index have provided accumulated evidence that the magnetosphere behaves as a nonlinear dynamic system, and it can be described by a small number of variables (Kamide et al., 1998). There are plenty of studies aiming to forecast AE index from solar wind measurements. Among the many approaches of modeling and forecast, neural network (NN) is a commonly used method. Early in 1997, NN models were constructed to study prediction of the AE index (Takalo & Timonen, 1997). Later, an ANN algorithm based at IMF measured on Lagrangian point L1 and plasma measurements was introduced in 2008 to predict AE index from 5 to 60 min ahead (Pallochia et al., 2008). The ANN models were further improved to achieve a correlation coefficient of 0.83 for 1-hr-ahead forecast and 0.80 for 3-hr-ahead forecast (Bala & Reiff, 2012). In addition, some other approaches are also applied for the analysis; for example, a correlation analysis with a technique of wavelet decomposition and selective reconstruction was applied to analyze the relationship between AE index and solar wind variables (Guarnieri et al., 2018). The advantage of NNs and its variants is that it can provide an efficient nonlinear representation to generate good model predictions. However, the identification process of NNs often involves a large number of variables, so that the model structure of NNs can be very complicated. From such model structure, it is quite difficult to further understand the nonlinear dynamic of the system, for example, which model term/variables are superior for describing the index and which model terms/variables are redundant. Nevertheless, it is obvious that such a model cannot provide a model structure that is simple and easy for understanding.

Another widely used approach for the modeling and forecast of magnetosphere is the nonlinear autoregressive with exogenous input (NARX) model. The NARX model is developed for the nonlinear system identification and can detect an appropriate model structure by selecting the most important model terms from a dictionary consisting of a huge number of candidate model terms (Billings, 2013). Thus, it is very efficient method for the space weather forecast due to the fact that the magnetosphere is a nonlinear process (Kamide et al., 1998). The NARX model have successfully solved the modeling and prediction of many magnetic indices, for example, the Dst index (Balikhin et al., 2011; Boynton et al., 2011; Wei et al., 2004) and the Kp index (Ayala Solares et al., 2016). Comparing to the NNs, the NARX method only uses a small number of effective model terms to describe the system, so that the system can be represented a linear-in-the-parameter form, which is parsimonious and transparent. It is achieved by a model selection algorithm called orthogonal forward regression (OFR), which was initially developed as subset selection method for nonlinear modeling problem with unknown structure (Chen et al., 1989). In recent years, several variants have been introduced to improve the performance of NARX model and OFR algorithm, for example, the wavelet NARX model (Billings & Wei, 2005; Wei & Billings, 2004), the iterative search algorithm with mutual information (Wei & Billings, 2008), and the common/robust model structure selection (RMSS) method (Gu & Wei, 2018; Li et al., 2016).

Under the assumption that the identified individual model structure elements can perfectly describe the true system components, most of the models are capable to provide accurate representations of the system. However, in many practical scenarios, this assumption is usually invalid due to the existence of uncertainty. Generally speaking, there are several types of uncertainty that may cause deleterious effect on the system modeling process. First, the uncertainty in data collection, for example, the experimental uncertainty and epistemic uncertainty, may lead to incomplete and inaccurate information in the data set. If the data samples are insufficient or some important variables are missing in the data set, it would be extremely difficult to find a model. Second, model type or model structure uncertainty directly affects the model performance. It is known that models are usually designed to represent some specific system features; that is, there are no single model type or structure that can perfectly describe all the true systems. Thus, it is essential to choose a correct model type to represent the system, and an inappropriate model type can largely reduce the model performance. Third, noise/disturbance is one of the main sources of uncertainty. The noise can be brought to the data through many ways, for example, measurement error from physical equipment, and external disturbances. The existence of noise could lead to biased parameter estimation and uncorrected selected model terms. Therefore, quantifying uncertainty is essential for establishing the significance of findings and making predictions with known confidence.

From the literature, it is known that estimating the true uncertainty remains as an exclusive goal for AE index forecast. Under the effect of uncertainty, the identified model usually cannot perfectly represent the system but only approximately describe the system behaviors. In these situations, a single model may not always work well on future new data, as there might be a risk on trusting and relying on a single model

for future system behavior forecasting. Thus, the model robustness becomes extremely important. Given the above concerns, this study focuses on the modeling and forecasting of AE index using NARX model structure with new uncertainty analysis method. In this study, a novel cloud-NARX model is presented to predict AE index 1 hr ahead and to quantify the uncertainty in the system modeling process. The reliability of the model can be quantified by the proposed uncertainty analysis, which makes the cloud-NARX model more robust than the conventional NARX model.

In summary, the main contribution of this work lies in the cloud-NARX model for (a) describing model structure and parameter uncertainty using a new uncertainty concept “cloud” model; (b) generating a new predicted band, which provides the confidence interval of predicted AE index; and (c) providing a new way to evaluate the model reliability based on uncertainty analysis.

The remainder of this paper is organized as follows. In section 2, the NARX model and cloud model are briefly reviewed. In section 3, the proposed cloud-NARX model is introduced. The AE and solar wind data are described in section 4. Section 5 presents the identified cloud-NARX model with model evaluation and uncertainty analysis. The study is summarized in section 6.

2. Review of NARX Model and Cloud Model

2.1. Construction of the NARX Model

The nonlinear autoregressive moving average with exogenous input (NARMAX) model structure can be described as (Chen & Billings, 1989)

$$y(t) = F[y(t-1), \dots, y(t-n_y), u(t-1), \dots, u(t-n_u), e(t-1), \dots, e(t-n_e)] + e(t), \quad (1)$$

where $y(t)$ and $u(t)$ are systems output and input signals; $e(t)$ is a noise component that can be assumed to be white Gaussian (but such an assumption is not always necessary) in many applications. n_y , n_u , and n_e are the maximum lags for the system output, input, and noise, respectively. $F[\]$ is some nonlinear function. A polynomial NARX model can be written as the following linear-in-the-parameters form:

$$y(t) = \sum_{m=1}^M \theta_m \varphi_m(t) + e(t), \quad (2)$$

where $\varphi_m(t) = \varphi_m(\vartheta(t))$ are the model terms generated from the regressor vector $\vartheta(t) = [y(t-1), \dots, y(t-n_y), u(t-1), \dots, u(t-n_u)]^T$ (T indicates the transpose of the vector), θ_m are the unknown parameters, and M is the number of candidate model terms.

The OFR algorithm is briefly introduced as follows (Chen et al., 1989). First, the regression model and prediction error can be written in a compact matrix form:

$$y = \sum_{m=1}^M \theta_m \varphi_m + e, \quad (3)$$

where $y = [y(1), \dots, y(N)]^T$, $\theta = [\theta(1), \dots, \theta(N)]^T$, $\varphi_m = [\varphi_m(1), \dots, \varphi_m(N)]^T$, and $e = [e(1), \dots, e(N)]^T$. Let $D = \{\varphi_i : 1 \leq i \leq M\}$ be the initial dictionary of all the candidate model terms, the objective of OFR algorithm is to select a number of significant model terms to form a subset, which can be described as $D_n = \{\varphi_{i_1}, \dots, \varphi_{i_n}\}$. The output can then be described with the selected terms as follows:

$$y = \sum_{i=1}^n \theta_{i_i} \varphi_{i_i} + e \quad (4)$$

At first step of the term selection, the ERR index of each candidate model term of the initial dictionary can be calculated by

$$ERR^{(1)}[i] = \frac{(r_0^T \varphi_i)^2}{(r_0^T r_0)(\varphi_i^T \varphi_i)}, \quad (5)$$

where $i = 1, 2, \dots, M$. The first selected model term is the candidate model term with highest ERR value, as

$$l_1 = \arg \max_{1 \leq i \leq M} \{ERR^{(1)}[i]\}. \quad (6)$$

The first selected model term is φ_{l_1} , and its associated orthogonal variable can be defined as $q_1 = \varphi_{l_1}$. The selected term φ_{l_1} is then removed from the initial dictionary, and the dictionary D is then reduced to a

subdictionary D_{M-1} , which consists of $M-1$ model candidates. The residual sum of squares can be calculated as

$$\|r_1\|^2 = \|y\|^2 - \frac{(r_0^T q_1)^2}{q_1^T q_1}. \quad (7)$$

At step s ($s \geq 2$), the $M-s+1$ bases are first transformed into new group of orthogonalized base $[q_1^{(s)}, q_2^{(s)}, \dots, q_{M-s+1}^{(s)}]$ with an orthogonal transformation as below:

$$q_j^{(s)} = \delta_j - \sum_{r=1}^{s-1} \frac{\phi_j^T q_r}{q_r^T q_r} q_r, \quad (8)$$

where q_r ($r = 1, 2, \dots, s-1$) are orthogonal vectors, ϕ_j ($j = 1, 2, \dots, M-s+1$) are the basis of unselected model terms of subset D_{M-s+1} , and $q_j^{(s)}$ ($j = 1, 2, \dots, M-s+1$) are the new orthogonalized bases. The rest of the model terms can then be identified step by step using the ERR index of orthogonalized subsets D_{M-s+1} :

$$ERR^{(s)}[j] = \frac{(y^T q_j^{(s)})^2}{(y^T y)(q_j^{(s)T} q_j^{(s)})}, \quad (9)$$

$$l_s = \arg \max_{1 \leq j \leq M-s+1} \{ERR^{(1)}[j]\}. \quad (10)$$

The s th significant model term of the subset is ϕ_{l_s} , and its associated orthogonal variable can be defined as $q_s = q^{(s)}_{l_s}$. The residual sum of squares can be updated by (Wei et al., 2006; Wei & Billings, 2006):

$$\|r_s\|^2 = \|r_{s-1}\|^2 - \frac{(r_{s-1}^T q_s)^2}{q_s^T q_s}. \quad (11)$$

Recursively, the model terms of the subset $\{\phi_{l_1}, \dots, \phi_{l_n}\}$ can be identified step by step, each at one step. By summing (11) for s from 1 to n yields

$$\|r_n\|^2 = \|y\|^2 - \sum_{s=1}^n \frac{(r_{s-1}^T q_s)^2}{q_s^T q_s}. \quad (12)$$

The $\|r_n\|^2$ is called residual sum of squares, or sum squared error of the final model. The mean square error of the model can be calculated as $\|r_n\|^2/n$. The number of the model terms that should be included in the NARX model can be decided by a modified generalized cross-validation criterion, also known as adjusted predicted sum of squares (APRESS). The APRESS is calculated as (Billings & Wei, 2008; Wei & Billings, 2008)

$$APRESS(k) = p(k)MSE(k) = \left(\frac{1}{1 - \frac{C(k, \alpha)}{N}} \right)^2 MSE(k), \quad (13)$$

where $p(k)$ is a penalty function defined in terms of the cost function $C(k, \alpha) = k \times \alpha$ with α being an tuning parameter.

2.2. Cloud Model and Cloud Transformation

Cloud model is a cognitive model that provides a way of bidirectional transformation between a qualitative concept “cloud” and the quantitative data “cloud drops” (Wang et al., 2014). The concept cloud is described by three numerical characteristics, namely, *ex* (expectation), *en* (entropy), and *he* (hyper entropy). Similar to normal distribution, *ex* is the expectation of all the elements in the set, and *en* is the variance of the distribution. *he* depicts the degree of departure from normal distribution of cloud model (Wang et al., 2014). Based on the theorem that any distribution can be represented by the sum of several normal distributions, the cloud model can be seen as an extension of normal distribution: When *he* equals to 0, the cloud model

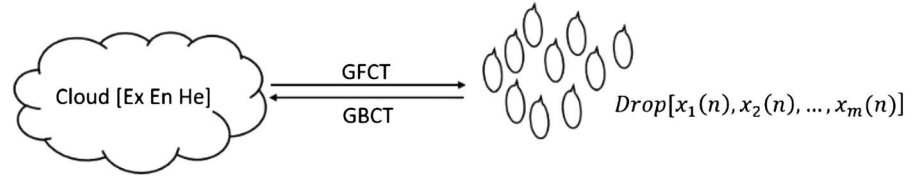


Figure 1. Cloud model and generic forward cloud transformation (GFCT)/generic backward cloud transformation (GBCT).

becomes actually a normal distribution. *he* is often regarded as an extra variable in practical situation, such as psychological quality of an athlete.

The bridges between cloud model and cloud drops is cloud transformation. The commonly used cloud transformation is the generic forward cloud transformation (GFCT) and generic backward cloud transformation (GBCT). The forward transformation is used to generate cloud drops from a known cloud model. The backward transformation is used to identify the cloud model from a sequence of cloud drops. In previous research, an ideal cloud backward transformation is also studied (Zhang et al., 2016). However, it is not feasible in real life as the groups of cloud drops could hardly be obtained in advance. The representation of GFCT and GBCT can be illustrated in Figure 1 and as follows:

$$[x_1, x_2, \dots, x_r] \xrightleftharpoons{\text{generic cloud transformation}} \text{cloud}(ex, en, he), \quad (14)$$

where the notation $\text{cloud}(ex, en, he)$ represents a cloud concept of characteristics modeled from r samples numerical data cloud drops $[x_1, x_2, \dots, x_r]$. The parameter ex , en , and he of the associated cloud model can be calculated from these samples. The GBCT can be described as follows (Wang et al., 2014). First, ex can be calculated by

$$ex = \frac{1}{z} \sum_{i=1}^z x_i. \quad (15)$$

Next, sampling v groups of data $[x_{i1}, x_{i2}, \dots, x_{iu}]$ ($i = 1, 2, \dots, v$) with replacement randomly from $[x_1, x_2, \dots, x_r]$, where u and v are the number of resampling folds parameters for GBCT $u \times v = r$. Then calculate the sample variance of each group ($i = 1, 2, \dots, v$):

$$\gamma_i = \frac{1}{u-1} \sum_{j=1}^u (x_{ij} - ex)^2. \quad (16)$$

Finally, en and he can be calculated as

$$en^2 = \frac{1}{2} \sqrt{4EY^2 - 2DY}, \quad he^2 = EY - en^2, \quad (17)$$

where $EY = \frac{1}{v} \sum_{i=1}^v \gamma_i$ and $DY = \frac{1}{v-1} \sum_{i=1}^v (\gamma_i - EY)^2$ are the sample mean and variance of γ_i .

With an identified cloud model, a series of cloud drops can be generated with GFCT. Let ex , en , and he be the numerical characteristics and n be the number of cloud drops; let a and b be the parameters of number of folds of GFCT ($a \times b = z$). The generated cloud drops x_{ij} with certainty degrees $\mu(x_{ij})$ ($i = 1, 2, \dots, a$, $j = 1, 2, \dots, b$) can be generated by several steps: First, generate a series of normally distributed random numbers γ_i with expectation en and variance he^2 ; next, for each γ_i , generate b normally distributed random numbers x_{ij} with expectation ex and variance γ_i^2 and calculate the certainty degree as

$$\mu(x_{ij}) = \exp \left\{ -\frac{(x_{ij} - ex)^2}{2\gamma_i^2} \right\}. \quad (18)$$

The generic cloud transformation achieves the transformation between intension and extension of the cloud concept. The advantage of cloud model is that it provides a way to describe a distribution with only three parameters that cannot be characterized by traditional normal distribution. The cloud transformation is better and more powerful than normal distribution in that (i) it includes normal distribution as a special case and (ii) many data in real life do not follow a normal distribution.

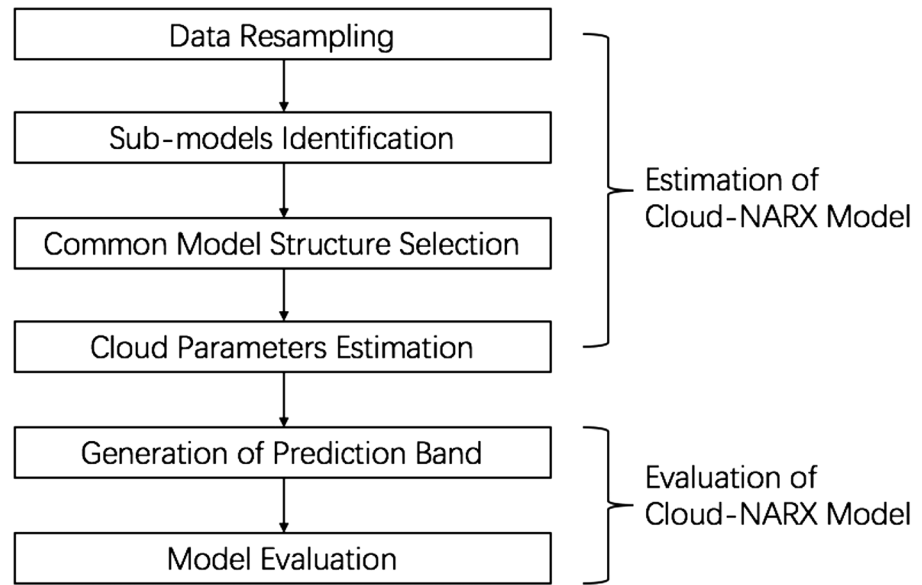


Figure 2. The process of estimation and evaluation of the cloud-nonlinear autoregressive with exogenous input (NARX) model.

3. The Cloud-NARX Model

3.1. The Cloud-NARX Model Structure

Based on cloud model and cloud transformation, a cloud-NARX model is proposed. The idea behind the metrics is to use a new uncertainty “concept” (cloud model) to replace the traditional model parameters. A series of predicted points can be calculated by performing a transformation (generic cloud transformation) to generate a series of model parameters (cloud drops) from the concept. These predicted points form a predicted distribution (surface/band) with confidence intervals, describing the uncertainty and risk brought by model uncertainty. The cloud-NARX model can be described as:

$$y = \sum_{i=1}^n \text{Cloud}_i(ex, en, he) \varphi_i, \quad (19)$$

where $\text{Cloud}_i(ex, en, he)$ ($i = 1, 2, \dots, n$) are the cloud models, which represent the estimated parameters and the uncertainty of these parameters. The parameters ex , en , and he are the characteristics of each cloud model.

3.2. Estimation of the Cloud-NARX Model

The estimation of cloud-NARX model consists of three steps, which are data resampling, submodel identification, and cloud parameter estimation. The general process of estimating the cloud-NARX model is shown in Figure 2.

First, the original data set can be regrouped to form K subdata sets through some resampling methods, for example, random sampling or bootstrap (Wei & Billings, 2009). Assume that the input and output sequence for the k th data set is $\{u^{(k)}(t)\}_{t=1}^{N_k}$ and $\{y^{(k)}(t)\}_{t=1}^{N_k}$, respectively, for $k = 1, 2, \dots, K$. The model terms $[\varphi_1^{(k)}(t), \dots, \varphi_M^{(k)}(t)]$ of the k th data set can be generated from the associated regressor vector relating to the k th data set $[y^{(k)}(t-1), \dots, y^{(k)}(t-n_y), u^{(k)}(t-1), \dots, u^{(k)}(t-n_u)]^T$. The submodel for the k th subdata set can be written in the compact matrix form

$$y^{(k)} = \sum_{i=1}^n \theta_i^{(k)} \varphi_i^{(k)}. \quad (20)$$

Second, for each subdata set, the OFR algorithm can be applied to select a number of significant model terms to establish an individual NARX model. A common model structure can be formed by model terms selected in and important for all the subdata sets. In addition, a RMSS method is developed as an alternative method,

for small-size data modeling problem (Gu & Wei, 2018). With the OFR algorithm and the RMSS method, a common model structure $\{\varphi_{l_1}, \dots, \varphi_{l_n}\}$ can be identified, and the associated model parameters for each subdata set can be estimated as $[\theta_{l_1}^{(1)}, \theta_{l_1}^{(2)}, \dots, \theta_{l_1}^{(K)}]$.

Finally, the cloud models for each model term can be identified from the K groups of model parameters using GBCT.

$$[\theta_{l_1}^{(1)}, \theta_{l_1}^{(2)}, \dots, \theta_{l_1}^{(K)}] \xrightarrow{GBCT} \text{Cloud}_{l_1}(ex, en, he), \quad (21)$$

where $i = 1, 2, \dots, n$; in this way, the cloud-NARX model can be identified.

It is known that when the model structure is perfect, and the data are not corrupted with noises, any of the subdata sets will lead to exact the same model. However, the model structures of the submodels might be different when there is model uncertainty brought by the noises/disturbances/insufficient information. In these situations, any single model might be unreliable. By removing or adding some data points in the K subdata sets, the uncertainty can be quantified by the submodels with different structures and parameters.

3.3. Model Predicted Band and Averaged Prediction

With the identified cloud model of each parameter, K' groups of cloud drops are generated using cloud forward transformation, as follows:

$$\text{Cloud}_{l_1}(ex, en, he) \xrightarrow{GFCT} [\hat{a}_{l_1}^{(1)}, \hat{a}_{l_1}^{(2)}, \dots, \hat{a}_{l_1}^{(K)}], \quad (22)$$

where $\hat{a}_{l_1}^{(k')}$ is the generated parameters for the model term θ_{l_1} with $k' = 1, 2, \dots, K'$. A number of K' predicted time series of output y can then be calculated as

$$\hat{y}^{(k')} = \hat{a}_{l_1}^{(k')} \varphi_{l_1} + \hat{a}_{l_2}^{(k')} \varphi_{l_2} + \dots + \hat{a}_{l_n}^{(k')} \varphi_{l_n}, \quad (23)$$

where k' is the index of predicted time series ($k' = 1, 2, \dots, K'$). The K' model prediction can then form a predicted band with density. The upper and lower boundaries of the predicted band can be determined as

$$\hat{y}_{lower} = \min(\hat{y}^{(1)}, \hat{y}^{(2)}, \dots, \hat{y}^{(K')}), \quad (24)$$

$$\hat{y}_{upper} = \max(\hat{y}^{(1)}, \hat{y}^{(2)}, \dots, \hat{y}^{(K')}). \quad (25)$$

The averaged model prediction can also be calculated as

$$\hat{y}_{averaged} = \frac{1}{K'} \sum_{i=1}^{K'} \hat{y}^{(i)}. \quad (26)$$

3.4. Model Performance Evaluation

To evaluate the averaged prediction of the model, the correlation coefficient (ρ), prediction efficiency (PE), and normalized root-mean square error (NRMSE) are calculated. The PE is defined as

$$PE = 1 - \frac{\sigma_{error}^2}{\sigma_{observed}^2}, \quad (27)$$

where $\sigma_{observed}^2$ is the variance of the observed AE values and σ_{error}^2 is the variance of the error between the predicted AE values and observed AE values. The accuracy of the predicted band can be defined as

$$\gamma = \frac{N_t'}{N_t}, \quad (28)$$

where N_t is the total number of observed data points in test data set and N_t' is number of the observed data points within the predicted band.

Table 1
Descriptions of the Solar Wind Variables and AE Index

Variable	Description
y	AE index
V	solar wind speed/velocity (flow speed; (km/s)
Bst	Interplanetary magnetic field factor (nT)
n	solar wind density (proton density; (cm ⁻³)
p	solar wind pressure (flow pressure; nPa)

Note. $Bst = B_T \sin^6(\theta/2)$ (nT; Boynton et al., 2011).

4. Data

A full description of the solar wind variables and the magnetic indices is given in Table 1. The AE index is one of the most widely used indices for researchers in geomagnetism, aeronomy, and solar-terrestrial physics, to understand the geomagnetic activity. The AE index is the maximum deviation of the horizontal components of geomagnetic field variations from a set of globally distributed ground-based magnetometers located in and near the auroral zone in the Northern Hemisphere (Guarnieri et al., 2018). It increases when a substorm event is happening and repre-

sents the overall disturbance in both eastward and westward ionospheric electrojets located at around 100-km altitude (Davis & Sugiura, 1966).

The AE index and solar wind variables used in this study were all sampled hourly. The AE index and solar wind variables are used as the output and input of the systems modeling, respectively. The amplitude of the solar wind velocity is around 250–1,000 km/s, which is much larger than those of the other input signals. To avoid producing extreme parameter estimations, the solar wind speed/velocity variable is first normalized by $V \rightarrow \dot{V}/1,000$, where \dot{V} is the original signal and V is the normalized signal. Two derived variables, \sqrt{p} and $VBst = V \times B_T \sin^6(\theta/2)$ (Boynton et al., 2011), which are effective in describing the magnetic indices, are also used as input variables for the system model.

5. The Cloud-NARX Model for One-Hour Ahead AE Index Forecasting

5.1. The Cloud-NARX Model

The AE and solar wind data from 2011 to 2013 (around 26,000 sampled data points) were used for training the model, and the data of 2015 (around 9,000 sampled data points) were used for model validation. In the test data set, two time periods of strong magnetic storm, 17–21 March 2015 and 22–26 June 2015 (120 sampled data points for each period) were used as special cases to evaluate the model. The time series of the AE index and solar wind variables of the two interested periods are shown in Figure 3. The figure shows that there were two significant storms on 17 March 2015 and 22 June 2015. Both periods match Interplanetary Coronal Mass Ejections (ICMEs). The first period 17–21 March corresponds to St Patrick storm caused by the CME on 15 March [https://www.swpc.noaa.gov/sites/default/files/images/u33/StPatrick%27sDay_Geomagnetic_Storm.pdf] whereas the second period 22–26 June 2015 corresponds to the ICME registered by Wind [<https://wind.nasa.gov/cycle24.php>].

In order to determine the maximum time lags for both the input and output variables, we have carried out premodeling experiments and simulations; the results suggest that the maximum time lags of the input and output were chosen to be $nu = 2$ and $ny = 2$. The initial full model was chosen to be a polynomial form with nonlinear degree of 2. The input-output data points of training data set were first resampled 100 times with replacement, to form 100 subdata sets. For each subdata set, a NARX model with six model terms is identified. For convenience of description, these single NARX models are referred to as “individual NARX models.” Thus, there are a total number of 100 different individual NARX models, and each has its own parameters. A total of 12 different model terms are selected during the 100 runs, and these terms are used for cloud-NARX model construction. The cloud parameters of each of these selected model terms are shown in Table 2.

It is noteworthy that the cloud-NARX model consists of 12 model terms, rather than six terms; this is because that each individual NARX model has its own structure. There are some common terms that are included in nearly all the individual NARX models, for example, $VBst(t - 02)$ and $y(t - 01)$. Also, some terms, for example, $VBst(t - 02) \times y(t - 01)$, are selected and included in a relatively small number of times out of the 100 runs. These rarely selected model terms are usually ignored in conventional NARX model because they make small contributions to the whole data set. However, in some of the subdata sets, they might be effective in some rare situations, for example, the peak times of the AE index.

Figure 4 shows the normal cloud membership functions of the 12 selected model terms. The estimated parameters of the some model terms are normally distributed, for example, $Bst(t - 02)$, $V(t - 01) * Bst(t - 01)$, $V(t - 01) \times Bst(t - 02)$, $V(t - 01) \times \sqrt{p}(t - 01)$, and $V(t - 02) \times p(t - 01)$. The distributions of the parameters

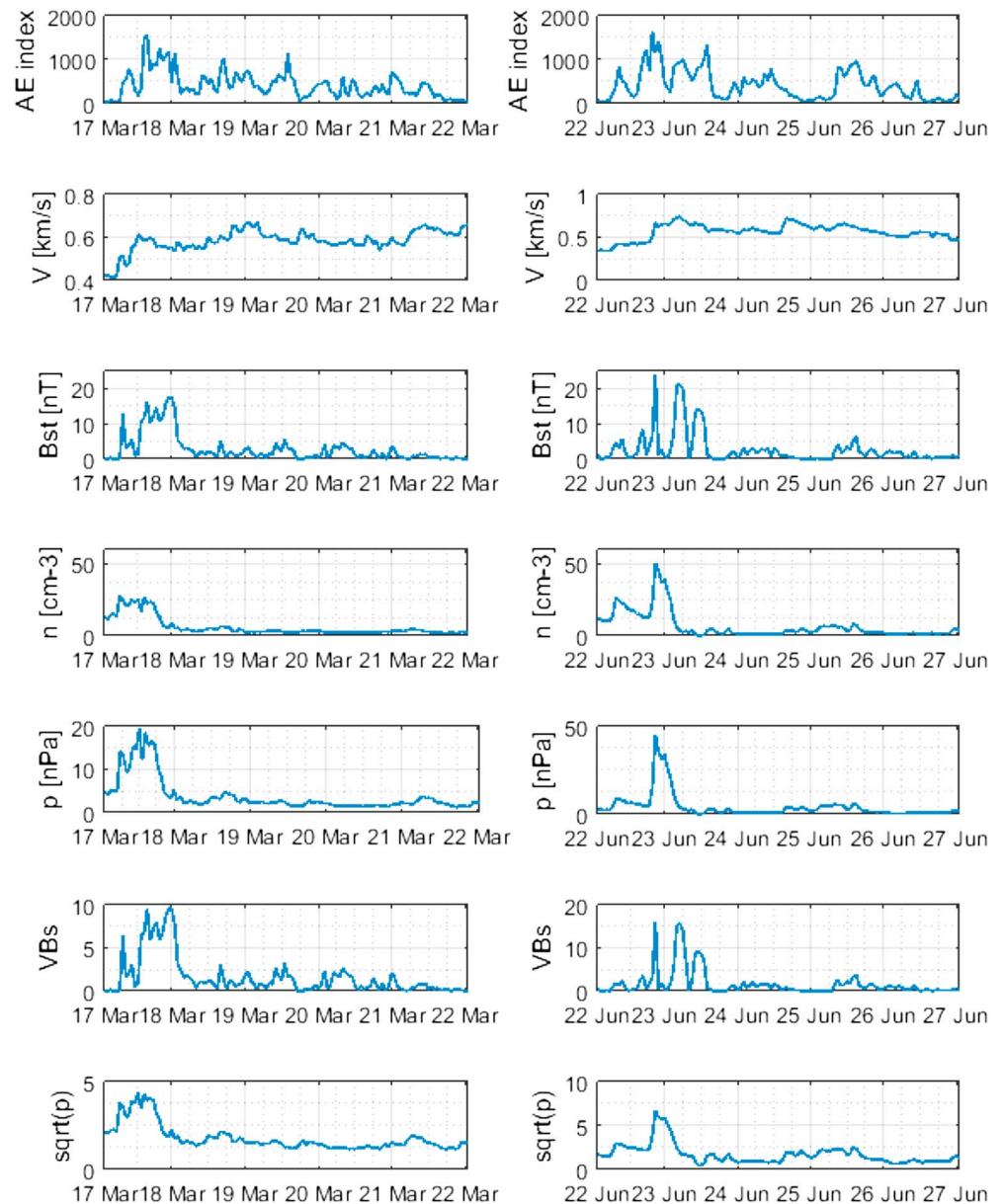


Figure 3. Observation of hourly sampled AE index and solar wind variables of two interested periods of 2015. Note that in the last two panel, the variable $\text{sqrt}(p)$ $\text{sqrt}(p) = \sqrt{p}$.

of some other model terms (e.g., $VBs(t-1)$, $y(t-1)$, $VBs(t-1) \times y(t-1)$) are beyond normal distributions. Note that the normal distributions are not always sufficient to describe the distribution of the estimated parameters of these model terms due to the existence of uncertainty, which do not necessarily follow a normal distribution law. The three characteristics ex , en , and he are used to analyze the uncertainty of each model term. As discussed earlier, ex is the mean of estimated parameters of each model term, which is consistent with the conventional model parameter; en is the variance of the parameter estimation; and he is the hype-parameter to describe the degree of departure of the distribution to normal distribution. The values of en of some model terms (e.g., $y(t-1)$) are quite small, which indicates that the parameters of these model terms in the individual models are very close. In other words, the contributions of these model terms are consistent in each individual model. On the contrary, the values of en of some model terms (e.g., $VBs(t-1)$) are quite large, which means that uncertainty of the estimated parameters of these model terms are strong. In

Table 2
Cloud-NARX Model With Cloud Parameters

Model terms	Cloud parameter		
	<i>ex</i>	<i>en</i>	<i>he</i>
$Bst(t-02)$	-9.7009	8.9120	0.0615
$VBst(t-1)$	6.0214	24.8546	16.3577
$y(t-01)$	0.6252	0.0108	0.0037
$V(t-01) \times Bst(t-01)$	143.6189	33.5311	0.0614
$V(t-01) \times Bst(t-02)$	-19.7937	22.3581	0.6368
$V(t-01) \times \sqrt{p}(t-01)$	14.3895	15.9883	0.1581
$V(t-02) \times p(t-01)$	7.8305	9.0816	0.1808
$V(t-02) \times \sqrt{p}(t-01)$	2.7969	7.8950	2.9935
$Bst(t-2) \times \sqrt{p}(t-02)$	-0.0807	0.5708	0.9887
$p(t-2) \times VBst(t-1)$	-0.0795	0.2808	0.4866
$VBst(t-01) \times VBst(t-01)$	-5.6495	0.4665	0.7133
$VBst(t-02) \times y(t-01)$	-0.0195	0.0029	0.0008

Note. NARX = nonlinear autoregressive with exogenous input; *ex* = expectation; *en* = entropy; *he* = hyper entropy. $Bst = B_7 \sin^6(\theta/2)$ (nT), $VBst = V \times Bst$ (Boynton et al., 2011).

the disturbed periods, the contributions of these model terms are different in each individual model, and the width of the predicted band increases due to the prediction uncertainty. The cloud parameter *he* describes how much the distribution is beyond normal distribution. If the value of *he* is much smaller than that of *en*, it means the estimated parameters of the model term are normally distributed. With the hyper cloud parameter *he*, the cloud model can better describe the estimated model parameters that are not normally distributed.

5.2. One-Hour-Ahead Prediction of AE Cloud-NARX Model

As mentioned earlier, the cloud-NARX model is built on hourly sampled data, so the model can be directly used to generate 1-hr-ahead (i.e., one-step-ahead) predictions of AE index. With the cloud parameters, the generic cloud forward transformation was applied to generate 100 sets of model parameters (that is called cloud drops in the transformation) for all the selected terms. A total number of 100 time series of the AE index prediction were calculated. The average prediction and predicted band are presented in Figure 5. The predicted band

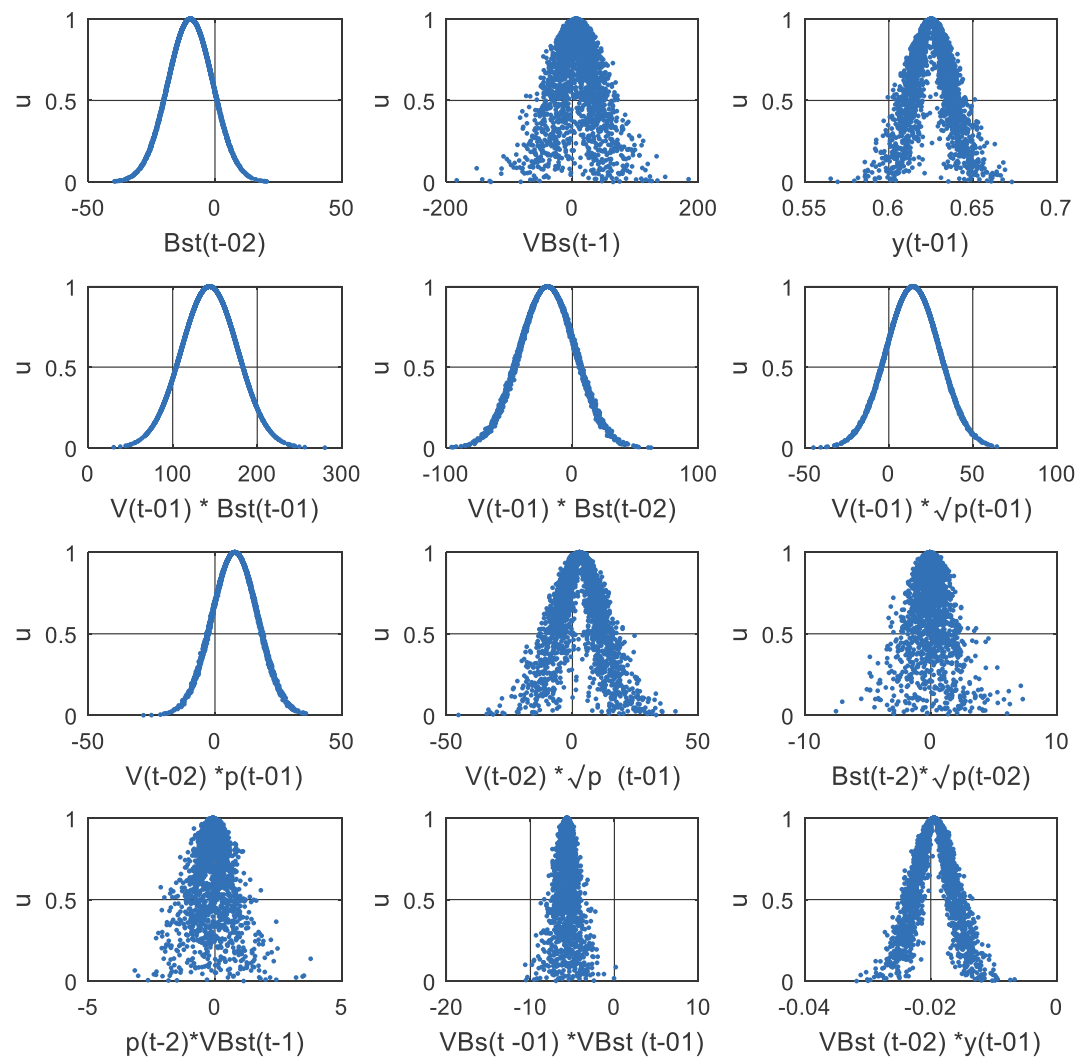


Figure 4. The normal cloud membership functions of the 12 selected model terms.

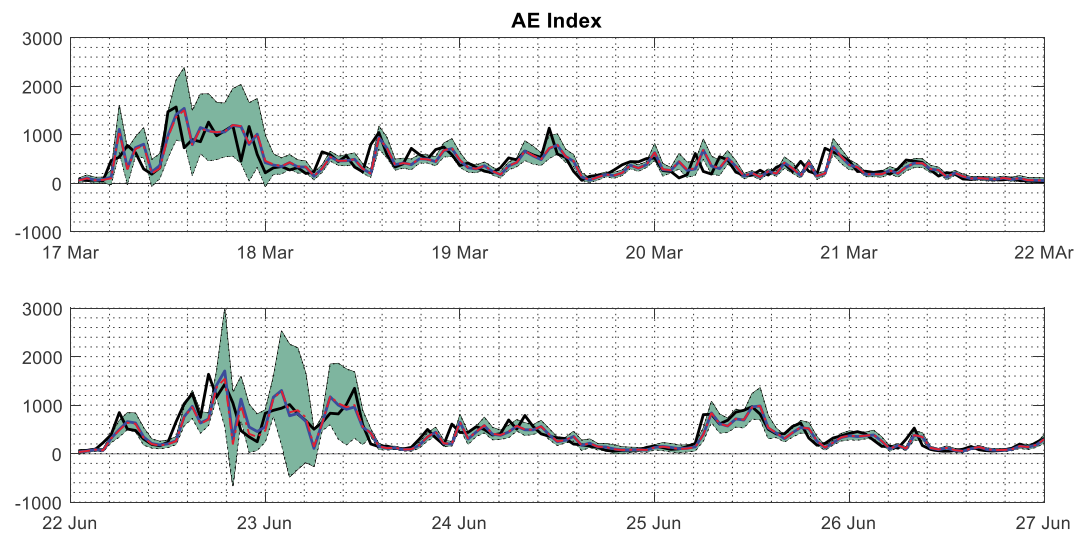


Figure 5. One-hour-ahead predicted band (consists of 80% of generated model predictions) and averaged prediction of AE index over 17–21 March and 22–26 June of 2015 (black line = observation; blue line = averaged prediction; green shadow = predicted band; red line = prediction of conventional nonlinear autoregressive with exogenous input model).

is the quantification of uncertainty throughout the structure detection, parameter estimation, and model prediction. If the model structure is perfect and the parameters are estimated unbiased, the predicted band will be narrow. Otherwise, if there are strong uncertainty in the data itself or the model structure and parameter, the uncertainty will be propagated to model prediction, and the width of predicted band will be increased.

From Figure 5, the predicted band is very wide over the period from 17 Mar 2015 to 22 Jun 2015. This can be explained or understood as follows. First, from the input signals shown in Figure 3, we know that there were interplanetary disturbances over the two periods. It is known that in general most storms last quite a short period in the long-term evolution of the process. As a consequence, most of the training data were sampled at “quiet” times, and only a very small fraction of the training data is for the storm period. This results in that the training data are severely “imbalanced” (Ayala Solares et al., 2016). Therefore, while a single model may well capture the features and dynamics of the system at quiet times, it may not sufficiently capture the system dynamics at the severe situation times. That is why the prediction band is so wide for these stormy periods. Second, the wide prediction band over the period of 17 March 2015 and 22 June 2015 implies that no single model would produce reliable prediction of AE over stormy periods, no matter what/which method is used to build the model. That is why we propose to carry out uncertainty analysis in this study.

Note that the predicted band in Figure 5 provides only rough quantification of the uncertainty. In many situations, the detailed information of the predicted AE index at a specific time point is often needed. Figures 6 and 7 are the predicted bands with density over an 8-hr period on 17 March 2015 and 23 June 2015, respectively. These figures show the probability of the predicted AE index being in each interval. As shown in the two figures, the interval of the predicted band for each time point is divided into 100 bins. The histogram shows the probability (frequency) of a single predicted AE value occurs in each bin. The boundaries of the predicted band are also displayed with the histogram, to visualize the prediction uncertainty and make it easier to understand. In addition, it is straightforward to compare the observation (green line) and averaged prediction (blue line) of AE index in the figure. The overall accuracy of the predicted band on the test data set is 65%. The accuracy of high AE period ($AE > 1000$) is 70%.

The only way to reduce the width of the predicted band is to find a model structure that can better describe the true system. However, it is very hard, if not impossible, to obtain a perfect model structure for real-world system identification data modeling problem in the presence of strong uncertainty. Nevertheless, it should be noted that the performance of the model given by Table 2 outperforms previous models, for example,

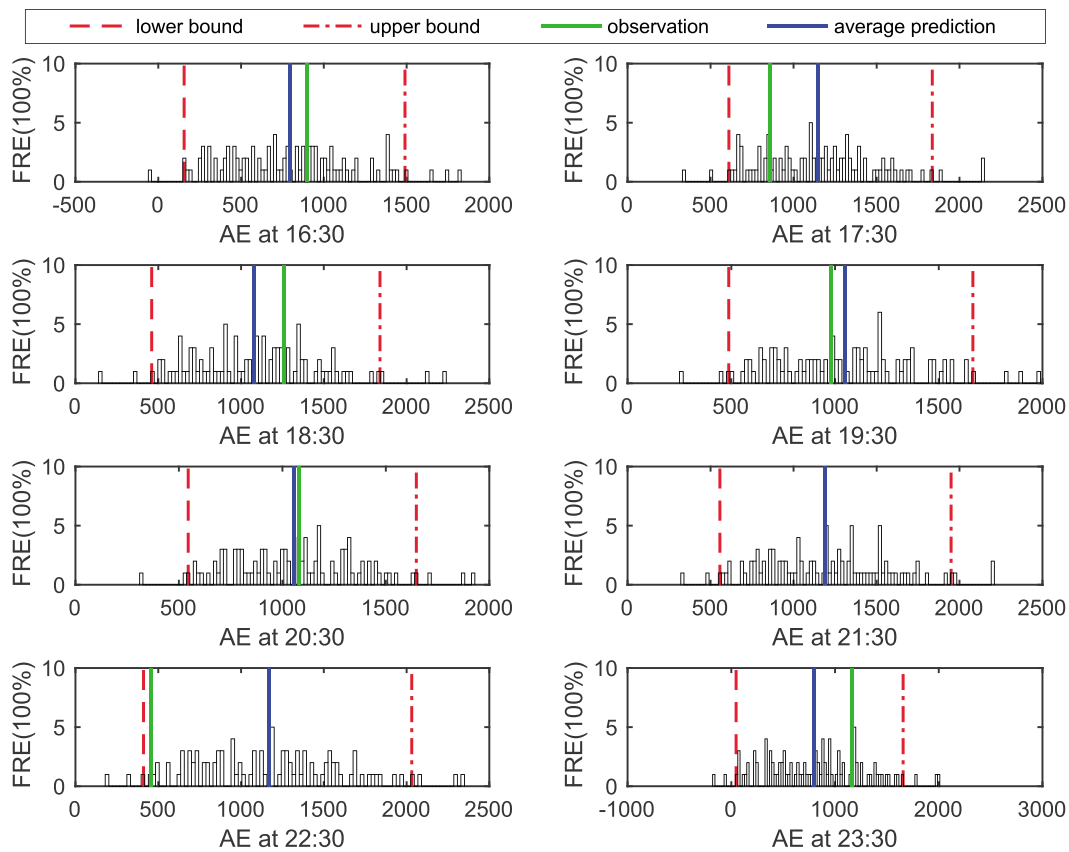


Figure 6. Predicted band with density over an 8-hr period on 17 March 2015 (FRE = the frequency of predicted AE occurrences in each divided bin of the predicted band).

the NN model (Bala & Reiff, 2012; as shown in Table 3). Therefore, a wide predicted band might indicate that a severe situation (interplanetary disturbances) is likely to happen. The property of the predicted band could potentially be used to forecast the arrival of the interplanetary disturbances.

5.3. Performance and Advantage of the Cloud-NARX Model

The performance of the averaged prediction of cloud-NARX model is comparable to the best NARX model with very similar structure but fixed model parameters, as shown in Table 3. Figure 8 presents the scatter plot of the averaged prediction and observation. The correlation coefficient, PE, and NRMSE of the averaged prediction are 0.872, 75.97%, and 0.0589 (for data of year 2015), respectively, which are consistent with the best NARX model. The NARX model outperform the NN model for 1-hr-ahead prediction, as the previous NN model achieves the correlation of 0.83 (Bala & Reiff, 2012). More importantly, the cloud-NARX provides a transparent and parsimonious representation. As shown in Table 2, the NARX model reveals which of the variables/model terms are significant and which are not, for example, the model terms $V(t - 02) \times Bst(t - 01)$ indicates that the dayside reconnection 20–40 min prior (Balikhin et al., 2010) is an important component of the system, and the model terms $y(t - 1)$ suggests that the autoregressive term has a significant effect on the AE index. On the contrary, the NN models are usually very complicated, and the training process involves a huge number of model terms and takes a lot of time.

The cloud-NARX model holds all the good properties of conventional NARX model and possesses an extra advantage. It provides a tool for understanding and analyzing uncertainty in the model structure and forecasting. For example, the uncertainty band in Figure 5 indicates that the model performs well for the period of 18–21 March 2015 and 24–26 June 2015, but the model is insufficient to characterize the dynamics of the process for the period 17 March 2015 and 22–23 June 2015 (i.e., when a sharp change occurs in solar wind

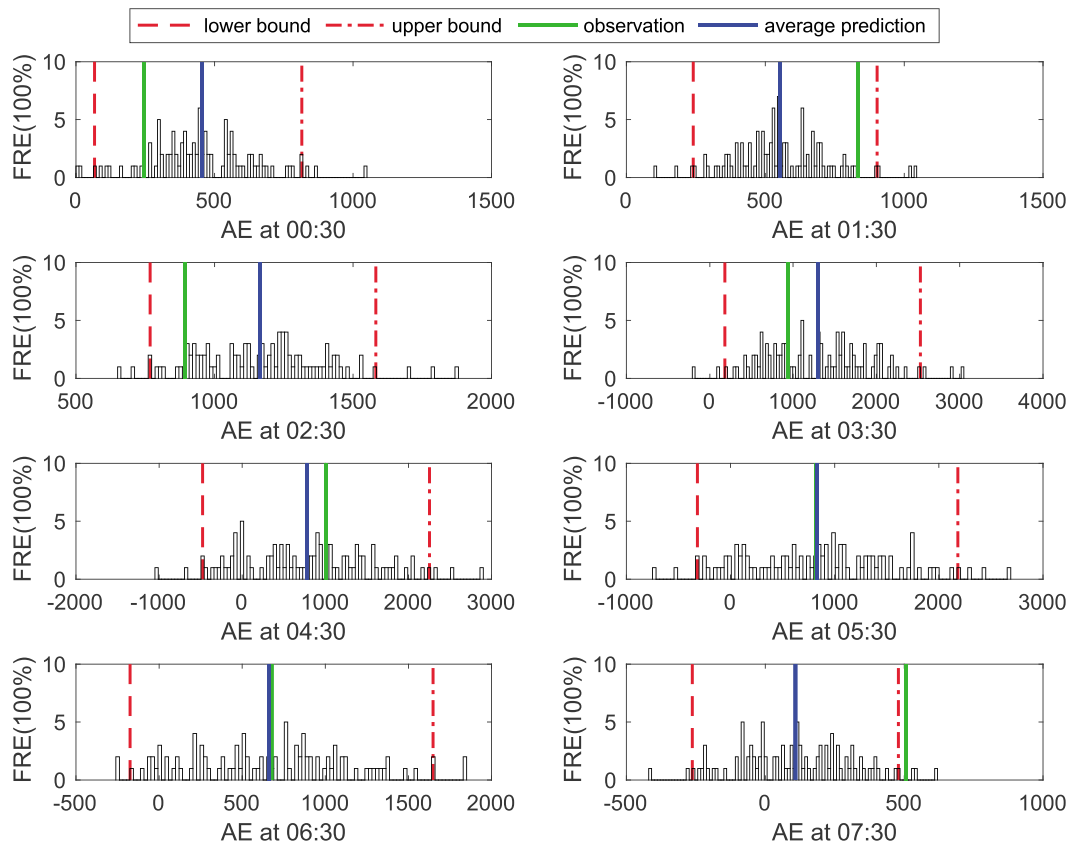


Figure 7. Predicted band with density over an 8-hr period on 23 June 2015 (FRE = the frequency of predicted AE occurrences in each divided bin of the predicted band).

variables, e.g., $Bst/VBst$). As discussed earlier, this property could potentially be used to forecast the arrival of a solar wind storm.

Note that the model also works well and even better in nondisturbed periods. This is because that the model was trained on the data set where most of the data were sampled at nondisturbed period. Therefore, the system behaviors in nondisturbed periods are well captured by the identified model. A comparison between the observed and predicted AE index in two selected nondisturbed periods (23 April to 5 May and 19 October to 1 November) is given in Figure 9. According to the figure, the predicted band is narrow, which means that the uncertainty of the model is not strong. From these results, the cloud-NARX model also generates good prediction results in the nondisturbed times.

The model prediction of the cloud-NARX model and the conventional NARX model are consistent in nondisturbed periods. In some disturbed periods, the prediction performance of the cloud-NARX model is better

than that of the NARX model. The correlation coefficient and NRMSE of cloud-NARX model in disturbed periods ($AE > 1,000$) are 0.3422 and 0.4454, while the conventional NARX model achieves correlation coefficient and NRMSE of 0.3226 and 0.4518 in the same periods. As discussed earlier, the inclusion of some extra selected model terms in the cloud-NARX model can help improve the model robustness in some severe situations. Therefore, compared to the conventional NARX model, the cloud-NARX model can better describe the nonlinear effect in the disturbed periods.

In addition, it is easy to generate long-term prediction using the cloud-NARX model. For example, the 3-hr-ahead AE index forecast can be

Table 3

Comparison of the Performances of the Best NARX Model and Cloud-NARX Model on Test Data of Year 2015

Model	Correlation	PE	NRMSE
Best NARX model	0.8728	0.7611	0.0588
cloud-NARX model	0.8723	0.7597	0.0589
NN model	0.83	/	/

Note. Slash (/) means not available. NARX = nonlinear autoregressive with exogenous input; NN = neural network; PE = prediction efficiency; NRMSE = normalized root-mean square error.

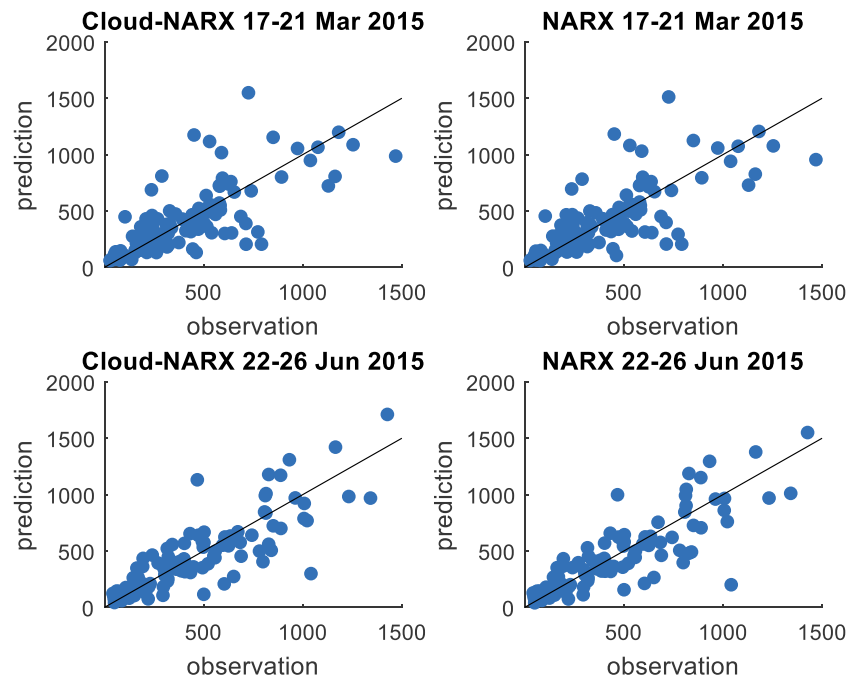


Figure 8. Scatter plot of the averaged prediction and observation of the cloud-nonlinear autoregressive with exogenous input (NARX) model and the best NARX model on two test data sets.

achieved by generating three-step-ahead model predicted output with the cloud-NARX model. The correlation coefficient, prediction efficiency, and NRMSE of the three-step-ahead model predicted output of the cloud-NARX model are 0.8167, 0.6667, and 0.0694, respectively. It is reasonable that the performance of 3-hr-ahead prediction is lower than that of the 1-hr-ahead prediction. It is because that at each step of the multiple-step-ahead prediction, the predicted AE index at previous step is used as the model input (as autoregressive variable). Thus, the prediction uncertainty of long-term prediction is increased due to the propagation of the error.

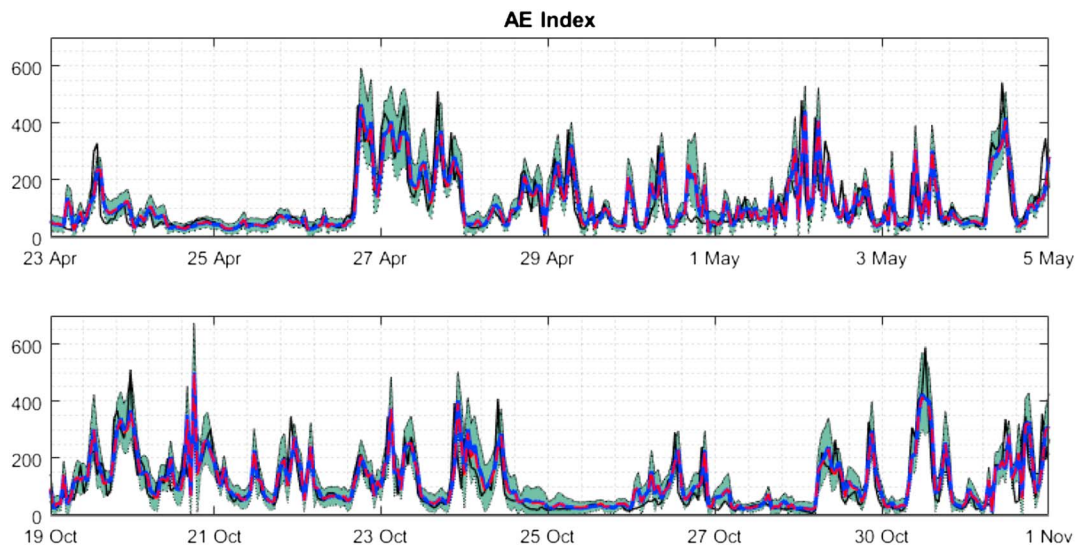


Figure 9. One-hour-ahead predicted band and averaged prediction of AE index over 23 April to 5 May and 19 October to 1 of 2015 (black line = observation; blue line = averaged prediction; green shadow = predicted band; red line = prediction of convention NARX model).

6. Conclusion

In this paper, a new cloud-NARX model was applied to the modeling and forecasting of AE index. Good forecasting results were obtained for 1-hr-ahead AE index prediction. The correlation coefficient between averaged prediction and observation is 0.87 and prediction efficiency of 0.81 when benchmarked for the period of 17–21 March 2015 and 22–26 June 2015, which is nearly identical to that produced by the best NARX model. The cloud-NARX model outperforms the previous models, for example, NNs. More importantly, the cloud-NARX model is capable to quantify the uncertainty of model structure, model parameter, and model prediction. The advantages of this new model can be summarized as follows. First, the model structure of cloud-NARX model is more robust than that of the conventional NARX model, as the model terms of cloud-NARX model are selected from resampled subdata sets. Second, the estimated parameters (*ex*, *en*, and *he*) of cloud-NARX model can provide more comprehensive information on the model parameter uncertainty. Third, based on cloud forward transformation, the cloud-NARX model can generate the predicted band, which clearly indicates the confidence interval of each predicted AE index. It is extremely important for space weather forecast, because when model becomes unreliable under some severe situations, the biased prediction could cause damages and losses. With the predicted band, the bias of model prediction can be identified, and the reliability of model can be evaluated.

One of the limitations of the paper is that there is still room for improvement for the accuracy of predicted band. Because for magnetic storm period and quiet times, the uncertainty is at different levels. Thus, it is highly desirable to further improve the cloud-NARX model to make it more adaptive for the high and low AE data.

Acknowledgments

This work was supported in part by the EU Horizon 2020 Research and Innovation Programme Action Framework under grant agreement 637302 and the Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/I011056/1 and Platform Grant EP/H00453X/1. Solar wind and AE index data were provided by OMNIweb (<https://omniweb.gsfc.nasa.gov/index.html>).

References

- Ala-Lahti, M. M., Kilpua, E. K. J., DImmock, A. P., Osmane, A., Pulkkinen, T., & Souček, J. (2018). Statistical analysis of mirror mode waves in sheath regions driven by interplanetary coronal mass ejection. *Annales Geophysicae*, 36(3), 793–808. <https://doi.org/10.5194/angeo-36-793-2018>
- Ayala Solares, J. R., Wei, H. L., Boynton, R. J., Walker, S. N., & Billings, S. A. (2016). Modeling and prediction of global magnetic disturbance in near-Earth space: A case study for Kp index using NARX models. *Space Weather*, 14, 899–916. <https://doi.org/10.1002/2016SW001463>
- Bala, R., & Reiff, P. (2012). Improvements in short-term forecasting of geomagnetic activity. *Space Weather*, 10, S06001. <https://doi.org/10.1029/2012SW000779>
- Balikhin, M. A., Boynton, R. J., Billings, S. A., Gedalin, M., Ganushkina, N., Coca, D., & Wei, H. (2010). Data based quest for solar wind-magnetosphere coupling function. *Geophysical Research Letters*, 37, L24107. <https://doi.org/10.1029/2010GL045733>
- Balikhin, M. A., Boynton, R. J., Walker, S. N., Borovsky, J. E., Billings, S. A., & Wei, H. L. (2011). Using the NARMAX approach to model the evolution of energetic electrons fluxes at geostationary orbit. *Geophysical Research Letters*, 38, L18105. <https://doi.org/10.1029/2011GL048980>
- Billings, S. A. (2013). *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. New York: Wiley. <https://doi.org/10.1002/9781118535561>
- Billings, S. A., & Wei, H. L. (2005). The wavelet-NARMAX representation: A hybrid model structure combining polynomial models with multiresolution wavelet decompositions. *International Journal of Systems Science*, 36(3), 137–152. <https://doi.org/10.1080/00207720512331338120>
- Billings, S. A., & Wei, H. L. (2008). An adaptive orthogonal search algorithm for model subset selection and non-linear system identification. *International Journal of Control*, 81(5), 714–724. <https://doi.org/10.1080/00207170701216311>
- Boynton, R. J., Balikhin, M. A., Billings, S. A., Wei, H. L., & Ganushkina, N. (2011). Using the NARMAX OLS-ERR algorithm to obtain the most influential coupling functions that affect the evolution of the magnetosphere. *Journal of Geophysical Research*, 116, A05218. <https://doi.org/10.1029/2010JA015505>
- Camporeale, E., Wing, S., Johnson, J., Jackman, C. M., & McGranaghan, R. (2018). Space weather in the machine learning era: A multi-disciplinary approach. *Space Weather*, 16, 2–4. <https://doi.org/10.1002/2017SW001775>
- Chandorkar, M., Camporeale, E., & Wing, S. (2017). Probabilistic forecasting of the disturbance storm time index: An autoregressive Gaussian process approach. *Space Weather*, 15, 1004–1019. <https://doi.org/10.1002/2017SW001627>
- Chen, S., & Billings, S. A. (1989). Representations of non-linear systems: The NARMAX model. *International Journal of Control*, 49(3), 1013–1032. <https://doi.org/10.1080/00207178908559683>
- Chen, S., Billings, S. A., & Luo, W. (1989). Orthogonal least squares methods and their application to non-linear system identification. *International Journal of Control*, 50(5), 1873–1896. <https://doi.org/10.1080/00207178908953472>
- Davis, T. N., & Sugiura, M. (1966). Auroral electrojet activity index AE and its universal time variations. *Journal of Geophysical Research*, 71, 785–801. <https://doi.org/10.1029/JZ071i003p00785>
- Gonzalez, W. D., Joselyn, J. A., Kamide, Y., Kroehl, H. W., Rostoker, G., Tsurutani, B. T., & Vasyliunas, V. M. (1994). What is a geomagnetic storm? *Journal of Geophysical Research*, 99, 5771–5792. <https://doi.org/10.1029/93ja02867>
- Gu, Y., & Wei, H. L. (2018). A robust model structure selection method for small sample size and multiple datasets problems. *Information Sciences*, 451–452, 195–209. <https://doi.org/10.1016/j.ins.2018.04.007>
- Guarnieri, F. L., Tsurutani, B. T., Vieira, L. E. A., Hajra, R., Echer, E., Mannucci, A. J., & Gonzalez, W. D. (2018). A correlation study regarding the AE index and ACE solar wind data for Alfvénic intervals using wavelet decomposition and reconstruction. *Nonlinear Processes in Geophysics*, 25(1), 67–76. <https://doi.org/10.5194/npg-25-67-2018>

- Kamide, Y., Baumjohann, W., Daglis, I., Gonzalez, W. D., Grande, M., Joselyn, J., & Vasyliunas, V. M. (1998). Current understanding of magnetic storms: Storm-substorm relationships. *Journal of Geophysical Research*, 103, 17,705–17,728. <https://doi.org/10.1029/98JA01426>
- Knipp, D. J. (2012). Review of “Future global shocks: Geomagnetic storms”. *Space Weather*, 10, S01001. <https://doi.org/10.1029/2011SW000747>
- Li, Y., Wei, H. L., Billings, S. A., & Sarrianiannis, P. G. (2016). Identification of nonlinear time-varying systems using an online sliding-window and common model structure selection (CMSS) approach with applications to EEG. *International Journal of Systems Science*, 47(11), 2671–2681. <https://doi.org/10.1080/00207721.2015.1014448>
- MacAlester, M. H., & Murtagh, W. (2014). Extreme space weather impact: An emergency management perspective. *Space Weather*, 12, 530–537. <https://doi.org/10.1002/2014SW001095>
- Mayaud, P. N. (1980). Derivation, meaning, and use of geomagnetic indices. *Geophysical Monograph Series*. <https://doi.org/10.1029/GM022>
- Pallochia, G., Amata, E., Consolini, G., Marcucci, M. F., & Bertello, I. (2008). AE index forecast at different time scales through an ANN algorithm based on L1 IMF and plasma measurements. *Journal of Atmospheric and Solar-Terrestrial Physics*, 70(2–4), 663–668. <https://doi.org/10.1016/j.jastp.2007.08.038>
- Schrijver, C. J. (2015). Socio-economic hazards and impacts of space weather: The important range between mild and extreme. *Space Weather*, 13, 524–528. <https://doi.org/10.1002/2015SW001252>
- Takalo, J., & Timonen, J. (1997). Neural network prediction of AE data. *Geophysical Research Letters*, 24, 2403–2406. <https://doi.org/10.1029/97GL02457>
- Tsurutani, B. T., & Gonzalez, W. D. (1987). The cause of high-intensity long-duration continuous AE activity (HILDCAAs): Interplanetary Alfvén wave trains. *Planetary and Space Science* [https://doi.org/10.1016/0032-0633\(87\)90097-3](https://doi.org/10.1016/0032-0633(87)90097-3), 35(4), 405–412.
- Wang, G., Xu, C., & Li, D. (2014). Generic normal cloud model. *Information Sciences*, 280, 1–15. <https://doi.org/10.1016/j.ins.2014.04.051>
- Wei, H. L., & Billings, S. A. (2004). A unified wavelet-based modelling framework for non-linear system identification: The WANARX model structure. *International Journal of Control*, 77(4), 351–366. <https://doi.org/10.1080/0020717042000197622>
- Wei, H. L., & Billings, S. A. (2006). An efficient nonlinear cardinal B-spline model for high tide forecasts at the Venice Lagoon. *Nonlinear Processes in Geophysics*, 13(5), 577–584. <https://doi.org/10.5194/npg-13-577-2006>
- Wei, H. L., & Billings, S. A. (2008). Model structure selection using an integrated forward orthogonal search algorithm assisted by square correlation and mutual information. *International Journal of Modelling, Identification and Control*, 3(4), 341–356. <https://doi.org/10.1504/IJMIC.2008.020543>
- Wei, H. L., & Billings, S. A. (2009). Improved parameter estimates for non-linear dynamical models using a bootstrap method. *International Journal of Control*, 82(11), 1991–2001. <https://doi.org/10.1080/00207170902854118>
- Wei, H. L., Billings, S. A., & Balikhin, M. (2004). Prediction of the Dst index using multiresolution wavelet models. *Journal of Geophysical Research*, 109, A07212. <https://doi.org/10.1029/2003JA010332>
- Wei, H. L., Billings, S. A., & Balikhin, M. A. (2006). Wavelet based non-parametric NARX models for nonlinear input-output system identification. *International Journal of Systems Science*, 37(15), 1089–1096. <https://doi.org/10.1080/00207720600903011>
- Wintoft, P., & Cander, L. R. (2000). Ionospheric storm forecasting using neural networks. *Physics and Chemistry of the Earth, Part C: Solar, Terrestrial and Planetary Science*, 25(4), 267–273. [https://doi.org/10.1016/S1464-1917\(00\)00015-5](https://doi.org/10.1016/S1464-1917(00)00015-5)
- Zhang, H. Y., Ji, P., Wang, J. Q., & Chen, X. H. (2016). A neutrosophic normal cloud and its application in decision-making. *Cognitive Computation*, 8(4), 649–669. <https://doi.org/10.1007/s12559-016-9394-8>