Geodesy and Geodynamics xxx (2017) 1–6

Contents lists available at ScienceDirect

Geodesy and Geodynamics

journal homepages: www.keaipublishing.com/en/journals/geog; http://www.jgg09.com/jweb_ddcl_en/EN/volumn/home.shtml

Improvement of the prediction accuracy of polar motion using empirical mode decomposition

Yu Lei [a, b, c, *], Hongbing Cai [a, b], Danning Zhao [a, c]

[a] National Time Service Center, Chinese Academy of Sciences, Xi'an 710600, China
[b] Key Laboratory of Primary Time and Frequency Standards, Chinese Academy of Sciences, Xi'an 710600, China
[c] University of Chinese Academy of Sciences, Beijing 100049, China

A R T I C L E I N F O

Article history:
Received 8 June 2016
Received in revised form
21 September 2016
Accepted 22 September 2016
Available online xxx

Keywords:
Polar motion
Prediction model
Empirical mode decomposition (EMD)
Neural networks (NN)
Extreme learning machine (ELM)

A B S T R A C T

Previous studies revealed that the error of pole coordinate prediction will significantly increase for a prediction period longer than 100 days, and this is mainly caused by short period oscillations. Empirical mode decomposition (EMD), which is increasingly popular and has advantages over classical wavelet decomposition, can be used to remove short period variations from observed time series of pole coordinates. A hybrid model combing EMD and extreme learning machine (ELM), where high frequency signals are removed and processed time series is then modeled and predicted, is summarized in this paper. The prediction performance of the hybrid model is compared with that of the ELM-only method created from original time series. The results show that the proposed hybrid model outperforms the pure ELM method for both short-term and long-term prediction of pole coordinates. The improvement of prediction accuracy up to 360 days in the future is found to be 24.91% and 26.79% on average in terms of mean absolute error (MAE) for the x_p and y_p components of pole coordinates, respectively.

1. Introduction

The Earth orientation parameters (EOP): universal time (UT1-UTC), x_p, y_p pole coordinates and nutation-precession corrections dX, dY, are required for various fields linked to reference frames such as positional astronomy, space navigation, geodesy and precise orbit determinations of artificial Earth satellites [1]. EOP enable the time-varying transform from the International Terrestrial Reference Frame (ITRF) to the International Celestial Reference Frame (ICRF). EOP are obtained predominantly through Global Navigation Satellite System (GNSS), Very Long Baseline Interferometry (VLBI), Satellite Laser Ranging (SLR) and Doppler Orbitography and Radio positioning Integrated by Satellite (DORIS). Owing to complexity of the data processing and measurement model, it is impossible to obtain real-time estimates of EOP. Therefore, regularly generated EOP predictions are published by several international and national services, e.g., the International Earth Rotation and Reference Systems Service (IERS), or the EOP Service of the Institute for Applied Astronomy (IAA) in Saint Petersburg, Russia [2].

The demand to provide better EOP forecasts has promoted more research in this domain. In October 2005, the Earth orientation parameters prediction comparison campaign (EOP PCC) was organized in an effort to assess the various existing prediction methods under same conditions and rules. A main conclusion of the EOP PCC is that the best prediction method is different for different categories (EOP component to be forecasted and prediction interval). In other words, no single prediction approach is superior to the others for all prediction lengths and all EOP components. Generally speaking, the best individual prediction methods participating in the EOP PCC for x_p, y_p pole coordinates are the combination of the least squares (LS) extrapolation of the harmonic model and auto regressive (AR) prediction, combination of the LS extrapolation and spectral analysis and neural networks (NN), while the most accurate prediction approaches for UT1-UTC are the combination of the wavelet decomposition and auto covariance (AC) prediction,

* Corresponding author.
E-mail address: leiyu@ntsc.ac.cn (Y. Lei).
Peer review under responsibility of Institute of Seismology, China Earthquake Administration.

Production and Hosting by Elsevier on behalf of KeAi

http://dx.doi.org/10.1016/j.geog.2016.09.007

adaptive transformation from atmospheric angular momentum (AAM) to LODR and Kalman filter combing AAM forecasts [3].

EOP prediction is very complex to handle due to the non-linear and random characteristics of EOP time-series, and thus it is theoretically much more rational to predict EOP by using NN. It is proved that three-layer feed-forward neural networks can approximate any continuous function to any degree of accuracy [4]. NN need no *a priori* models in advance, and hence this can avoid the model error and make modeling of complicated time-series quite feasible. NN has already been successfully applied for EOP prediction with accuracy equal to or even better than that of other prediction methods [5–11]. This paper focuses on predicting x_p, y_p pole coordinates using the NN technology.

It has been found that the increase of x_p, y_p pole coordinates prediction error up to about 100 days in the future is mostly caused by irregular short period oscillations, which is predominately excited by the equatorial components of joint atmospheric and oceanic excitation functions (AEF and OEF), and it is expected that those shorter period variations would not have as significant influences on prediction error of polar motion [11,12]. The aim of this study is to provide rapid and accurate predictions of polar motion that are of significance for real-time and precision applications such as space navigation and precise orbit determinations of artificial Earth satellites. Since NN are the high potential for predicting quasi-periodic and irregular processes and empirical mode decomposition (EMD) is a powerful tool for separating different frequency bands in time series, the integration of EMD into NN has been realized to improve the prediction accuracy of polar motion. Herein, EMD is used as a pre-processing tool to remove high frequency signals from original time series. After removal of high frequency signals, the pre-processed data are predicted using extreme learning machine (ELM) that is a new learning algorithm for single hidden-layer feed-forward neural networks (SLFNs). The paper is organized as follows. In Section 2, the ELM and EMD algorithms are reviewed and the EMD-NN prediction method is described. The results of the predictions are analyzed and compared with those obtained by other methods in Section 3, followed by a discussion in Section 4.

2. Methodology

2.1. Empirical mode decomposition (EMD)

EMD can cope with non-linear and non-stationary time series. It has some advantages over wavelet decomposition [13,14]. Firstly, EMD is simple to understand and use. Secondly, EMD dose not require to determine a mother function in advance unlike wavelet decomposition. Most important of all, EMD is a fully data-driven technique for decomposing a signal into its natural scale components. In other words, EMD can adaptively decompose a signal into several independent fluctuation components in terms of local characteristics of a signal. According to the above-mentioned merits, EMD is an effective tool for non-linear and non-stationary time series analysis.

The basic principle of EMD is to decompose original time series into a finite and small number of oscillatory functions, namely intrinsic mode functions (IMFs). These IMFs have to satisfy the following two prerequisites [13,14].

(1) In whole data set, the total number of local maxima and minima and the number of zero crossings must either be equal or differ at most by one, and
(2) The mean value of the envelopes defined by local maxima and minima must be zero at all points.

The essence of EMD is a sifting procedure that extracts IMFs from an original signal. The EMD algorithm can be described as bellows [13,14].

Step 1: Identify all local maxima and minima in time series $s(t)(t = 1, 2, \cdots, l)$.

Step 2: Connect all local maxima and minima using a cubic spline line to generate the upper and lower envelopes $e_{\text{up}}(t)$ and $e_{\text{low}}(t)$.

Step 3: Compute the mean value of the point-by-point envelopes from the upper and lower envelopes and then generate the mean envelopes $m(t)$ as

$$m(t) = \left[e_{\text{up}}(t) + e_{\text{low}}(t)\right]/2 \tag{1}$$

Step 4: Calculate the difference between the mean envelopes $m(t)$ and the original time series $s(t)(t = 1, 2, \cdots, l)$, i.e.,

$$h(t) = s(t) - m(t) \tag{2}$$

Step 5: Check whether the proto-intrinsic mode function $h(t)$ meets the characteristics of an IMF. If so, it is treated as the ith IMF and then the original time series $s(t)(t = 1, 2, \cdots, l)$ is replaced by the residuals $r(t) = s(t) - h(t)$. If not, $s(t)$ is replaced by $h(t)$.

Step 6: Steps 1–5 are repeated until the size of the standard deviation calculated from the two consecutive sifting results is smaller than the pre-determined threshold.

By using the above-described sifting process, a set of IMFs can be derived from high frequency to low frequency and thus the original time series $s(t)(t = 1, 2, \cdots, l)$ can be decomposed into n IMFs and one residual as

$$s(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \tag{3}$$

where n is the number of IMFs, $r_n(t)$ is the final residuals representing a trend in $s(t)(t = 1, 2, \cdots, l)$, and $c_i(t)(i = 1, 2, \cdots, n)$ denotes IMFs, which are nearly orthogonal to each other and periodic. Each IMF is independent and specific for expressing local properties of an original signal.

2.2. Extreme learning machine (ELM)

ELM is a novel leaning algorithm for SLFNs, which randomly selects input weight matrix and hidden-layer biases [15]. After input weights and hidden-layer biases are selected randomly, a SLFN can be simply taken as linear systems and output weights (connecting hidden layer and output layer) of a SLFN can be analytically estimated through simple generalized inverse operation of hidden-layer output matrix. In contrast to traditional learning algorithms, ELM can produce good generalization performance at quite fast learning speed.

For a given N arbitrary distinct sample $\boldsymbol{D} = \{(\boldsymbol{x}_i, \boldsymbol{y}_i)_{i=1}^{N}\}$, where $\boldsymbol{x}_i = [x_{i1}, x_{i2}, \cdots, x_{in}]^{\text{T}} \in \boldsymbol{R}^n$ and $\boldsymbol{y}_i = [y_{i1}, y_{i2}, \cdots, y_{im}]^{\text{T}} \in \boldsymbol{R}^m$, If SLFN with L hidden-layer neurons and activation function $g(x)$ can approximate these N samples with zero error, there exist β_i, $\boldsymbol{w}_i$ and b_i such that

$$f_L(\boldsymbol{x}_j) = \sum_{i=1}^{L} \beta_i G(\boldsymbol{w}_i, b_i, \boldsymbol{x}_j) = \boldsymbol{y}_j, \quad j = 1, 2, \cdots, N \tag{4}$$

Eq. (1) can be written compactly as

$$\boldsymbol{H}\beta = \boldsymbol{Y} \tag{5}$$