sciendo
ARTIFICIAL SATELLITES, Vol. 58, No. 2 – 2023
DOI: 10.2478/arsa-2023-0004

MEDIUM- AND LONG-TERM PREDICTION OF POLAR MOTION USING WEIGHTED LEAST SQUARES EXTRAPOLATION AND VECTOR AUTOREGRESSIVE MODELING

Yu LEI[1], Danning ZHAO[2], Min GUO[3]

[1] School of Computer Science and Technology, Xi'an University of Posts and Telecommunications, Xi'an, China
[2] School of Electronic and Electrical Engineering, Baoji University of Arts and Sciences, Baoji, China
[3] Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China

e-mails: leiyu@xupt.edu.cn, zhaodanning31@163.com, gmnwpu@opt.ac.cn

ABSTRACT. This article presents the application of weighted least squares (WLS) extrapolation and vector autoregressive (VAR) modeling in polar motion prediction. A piecewise weighting function is developed for the least squares (LS) adjustment in consideration of the effect of intervals between observation and prediction epochs on WLS extrapolation. Furthermore, the VAR technique is used to simultaneously model and predict the residuals of x_p, y_p pole coordinates for WLS misfit. The simultaneous predictions of x_p, y_p pole coordinates are subsequently computed by the combination of WLS extrapolation of harmonic models for the linear trend, Chandler and annual wobbles, and VAR stochastic prediction of the residuals (WLS+VAR). The 365-day-ahead x_p, y_p predictions are compared with those generated by LS extrapolation+univariate AR prediction and LS extrapolation+VAR modeling. It is shown that the x_p, y_p predictions based on WLS+VAR taking into consideration both the interval effect and correlation between x_p and y_p outperform those generated by two others. The accuracies of the x_p predictions are 13.97 mas, 18.47 mas, and 20.52 mas, respectively for the 150-, 270-, and 365-day horizon in terms of the mean absolute error statistics, 36%, 24.8%, and 33.5% higher than LS+AR, respectively. For the y_p predictions, the 150-, 270-, and 365-day accuracies are 15.41 mas, 21.17 mas, and 21.82 mas respectively, 27.4%, 11.9%, and 21.8% higher than LS+AR respectively. Moreover, the absolute differences of the WLS+VAR predictions and observations are smaller than the differences from LS+VAR and LS+AR, which is practically important to practical and scientific users, although the improvement in accuracies is no more than 10% relative to LS+VAR. The further comparison with the predictions submitted to the 1st Earth Orientation Parameters Prediction Comparison Campaign (1st EOP PCC) shows that while the accuracy of the predictions within 30 days is comparable with that by the most accurate prediction techniques including neural networks and LS+AR participating in the campaign for x_p, y_p pole coordinates, the accuracy of the predictions up to 365 days into the future are better than accuracies by the other techniques except best LS+AR used in the EOP PCC. It is therefore concluded that the medium- and long-term prediction accuracy of polar motion can be improved by modeling x_p, y_p pole coordinates together.

Keywords: polar motion, prediction model, weighted least squares, vector autoregressive.

1. INTRODUCTION

Movement of the Earth's rotation axis with respect to the Earth's surface is denoted as polar motion (PM), which form together with variations of the Earth's rotation rate as well as precession and nutation the whole set of Earth orientation parameters (EOP). The study of EOP provides some information about irregular motion of fluids on a global scale, therefore serves to the requirement for research on geophysical, hydrology, oceanography, meterology, and other disciplines. Daily EOP values are routinely released by the International Earth Rotation and Reference Systems Service (IERS). However, near real-time EOP estimates are currently unavailable for users owing to the delay caused by complex data preprocessing and heavy computation procedures (Gambis and Luzum, 2011). Accurate short-term predictions up to 30 days of the EOP are therefore needed for several real-time applications such as navigation and tracking of interplanetary spacecrafts and precise orbit determination of Earth's satellites (Chin et al., 2004). Medium- and long-term predicted EOP for at least 1 year into the future are also required for climate forecasting and for long-term satellite orbit prediction for the purpose of autonomous satellite navigation, which is of practical importance (Su et al., 2014). The IERS Rapid Service/Prediction Center (RS/PC) provides rapid EOP estimates with degraded accuracies for many real-time applications as well as regularly publishes EOP predictions up 1 year into the future once a week. While errors of EOP estimates have been degraded rapidly over the last decades owing to incorporation of some advanced space geodetic techniques and reach 0.01 ms in length of day and 0.05 mas in PM for finalized combined daily solutions at present (Bizouard et al., 2019), errors of PM predictions are usually two magnitudes greater than determination errors even for prediction horizons up to a few days into the future only (Dill and Dobslaw, 2010). The root mean squares (RMS) of the differences between the PM predictions produced by the EOP RS/PC and the 14 C04 combination solutions are approximately 3.5 mas in the case of the x_p components and 2.5 mas in the case of the y_p component up to 10 days into the future, while the RMS for the x_p and y_p components are larger than 17 mas and 9 mas up to 90 days, respectively (IERS Annual Report 2018). At the European Geosciences Union (EGU) General Assembly 2006, the IERS Working Group on Prediction (IERS WGP) was established to survey the which of the EOP prediction products are desired by the user community (Wooden et al., 2006). The survey confirmed that there are many users who need daily, tabular and one-day spacing predictions with PM accuracies of 1 mas and UT1-UTC accuracies of 0.1 ms up to 30 days. Specially, some academic users have the most stringent accuracy requirements for EOP predictions and would like to improve long-term predictions (years) of EOP. It is therefore necessary to modify or improve the current EOP prediction products to meet more stringent requirements especially for long-term predictions.

Many prediction methods and techniques have been used during the past years to decrease errors of PM predictions. Some of these methods and techniques utilize only information within PM time-series, among which least-squares (LS) extrapolation of the harmonic model in combination with a stochastic method (Akyilmaz and Kutterer, 2004; Kosek et al., 1998; Schuh et al., 2002; Zhao and Lei, 2019) is the most commonly used approach for PM prediction, and some of them use the effective angular momentum from atmosphere, ocean, and hydrology as the explanatory variable to forecast PM (Dill et al., 2019; Dobslaw and Dill, 2017). In October 2005, the 1st Earth Orientation Parameters Prediction Comparison Campaign (1st EOP PCC) was launched to examine fundamental properties of different approaches for EOP prediction (Kalarus et al., 2010). The results of the campaign indicated that each prediction method and technique has certain advantages and weakness, i.e., no individual prediction method or technique performs the best for all prediction ranges and all components of EOP. LS extrapolation+autoregressive (AR) prediction is one of the most accurate prediction techniques participating in the campaign to predict x_{p} and y_{p} pole coordinates, while AR filtering and neural networks are most accurate for short-term prediction of x_{p} and y_{p} pole coordinates besides LS+AR. The requirement for accurate EOP predictions prompted the IERS for more activity on this subject. Starting in 2021, the 2nd EOP PCC is being performed under the auspice of the IERS to evaluate the state-of-the-art of the existing prediction methods and techniques such as accuracies and reliabilities (http://eoppcc.cbk.waw.pl/).

At present, PM predictions are regularly generated in the IERS Bulletin A at daily intervals based on a statistical extrapolation of the most recent PM observations by the combination of LS extrapolation of deterministic models consisting of the linear, Chandler annual and semiannual terms, and AR stochastic predictions of LS residuals, referred to as LS+AR (Dick and Thaller, 2020). This algorithm should be enhanced continually since the accuracies of PM predictions are yet unsatisfactory with the current implementation of the algorithm. Some work has been performed over the past years in an effort to yield improved LS+AR prediction algorithms. Xu and Zhou (2015) studied the effect of length of EOP data on prediction accuracies and then obtained an enhancement for 1 to 90-day predictions of PM using a strategy called as "predictions over optimal data length". Similarly, Wu et al. (2019) found that selecting optimized data for AR modeling can improve PM predictions over spans of 1 to 500 days. To overcome the so-called edge effect on classical LS extrapolation, Guo et al. (2013) presented a technique to enhance LS extrapolation algorithm by fixing the last EOP observation on a polynomial curve for the deterministic components. Zhao and Lei (2020) also proposed a boundary-extension technique to reduce the edge effect. The prediction accuracy of EOP are improved by the two techniques for edge-effect reduction. Xu et al. (2012) employed a combination of a Kalman filter and LS+AR to modify AR modeling of LS residuals, improving 1 to 50-day predictions of PM. For choice of periodic terms in LS extrapolation, Sun et al. (2019) introduced the retrograde annual and semiannual wobbles into a harmonic model for LS extrapolation, effectively enhancing PM predictions. Taking into account of variable annual and Chandler wobbles (AW and CW), Sun et al. (2012) and Wu et al. (2018) utilized the weighted least squares (WLS) for extrapolating the deterministic terms including the CW and AW. In addition, Some studies recently combined singular spectrum analysis (SSA) with LS extrapolation to model the time-variant periodic oscillations (Modiri et al., 2018; Shen et al., 2017; Yang et al., 2022). These studies showed that enhanced medium- and long-term predictions of PM can be obtained by the WLS or SSA in consideration of the time-varying CW and AW.

The Earth's rotation is not constant but changes over time owing to acting internal dynamical process and external gravitational force. The predominant periodicities of the x_{p} and y_{p} components of PM are totally same since the external gravitational force and internal dynamical process act on the Earth's rotational axis simultaneously. However, the abovementioned methods and techniques analyzed and modeled the x_{p} and y_{p} components separately and hence did not take into account the correlation between the two components for medium- and long-term prediction, which need to be further improved. An early study by Schuh et al. (2001) showed that the correlation approach for LS fit can make periodicity extraction more accurately by taking into account a priori correlation between simultaneous x_{p} and y_{p}. Recently, Jin et al. (2021) used multichannel singular spectrum analysis (MSSA) to analyze and model the x_{p} and

y_p components together. Wu et al. (2021) developed a PM prediction method considering the polar coordinates. Wang et al. (2022) proposed a method to predict PM with the difference between the x_p and y_p components, considering the correlation of changes in x_p and y_p. From these studies, enhanced accuracies could be expected in modeling and prediction of PM by combining x_p and y_p. In this study, we therefore attempt to predict x_p together with y_p rather than predict them individually. This study presents a new method to simultaneously model two components by employing LS extrapolation and vector autoregressive (VAR), abbreviated as WLS+VAR. Observed time-series of x_p and y_p is combined with WLS+VAR modeling to take into consideration the correlation between the components of PM. The results of predictions have demonstrated that PM time-series can be predicted 365 days in advance with an enhanced accuracy using the developed WLS+AR approach.

This article is organized as follows: Section 2 and 3 present the methodology of WLS extrapolation and VAR modeling to predict the x_p and y_p components together, respectively. In Section 4, we analyze and discuss the predicted results and compare the prediction accuracy with other methods and technologies. The conclusions are given in Section 5.

2. WEIGHTED LEAST SQUARES EXTRAPOLATION

2.1. Model

The x_p and y_p components contain a few well-known periodicities such as CW and AW. A priori model hence consists of a purely linear part (offset and drift parameters a, b, c, d) and, two periodical elliptical motions representing CW (parameters f_1, $A_{1,x}$, $B_{1,x}$, $A_{1,y}$, $B_{1,y}$) and AW (parameters f_2, $A_{2,x}$, $B_{2,x}$, $A_{2,y}$, $B_{2,y}$), as written in the following.

$$x_\mathrm{p}(t) = a + bt + \sum_{i=1}^{2} A_{i,x}\cos(2\pi f_i t) + B_{i,x}\sin(2\pi f_i t) \tag{1}$$

$$y_\mathrm{p}(t) = c + dt + \sum_{i=1}^{2} A_{i,y}\cos(2\pi f_i t) + B_{i,y}\sin(2\pi f_i t) \tag{2}$$

The semimajor and semiminor axes of an elliptical motion correspond to the amplitude of circular motion. For completeness, bias and drift of the linear part and parameters of the CW and AW in this a priori model are estimated simultaneously from x_p and y_p observations by a WLS fit as

$$\boldsymbol{\beta} = (\boldsymbol{C}^\mathrm{T}\boldsymbol{W}\boldsymbol{C})^{-1}\boldsymbol{C}^\mathrm{T}\boldsymbol{W}\boldsymbol{L} \tag{3}$$

where $\boldsymbol{\beta}$ represents the vector of estimated parameter (Eq. (4)), $\boldsymbol{L}$ represents the matrix of x_p and y_p obervations (Eq. (5)), in which n is the length of observed time-series, $\boldsymbol{C}$ represents the coefficient matrix (Eq. (6)), and $\boldsymbol{W}$ represents a diagonal matrix of the weights, $\boldsymbol{W} = \mathrm{diag}[w_1, w_2, \cdots, w_n]$, in which the weights w_i are greater than zero, as will be described in the next paragraph.

$$\boldsymbol{\beta} = \begin{bmatrix} a & b & A_{1,x} & B_{1,x} & A_{2,x} & B_{2,x} \\ c & d & A_{1,y} & B_{1,y} & A_{2,y} & B_{2,y} \end{bmatrix}^\mathrm{T} \tag{4}$$

$$\boldsymbol{L} = \begin{bmatrix} x_\mathrm{p}(t_1) & x_\mathrm{p}(t_2) & \cdots & x_\mathrm{p}(t_n) \\ y_\mathrm{p}(t_1) & y_\mathrm{p}(t_2) & \cdots & y_\mathrm{p}(t_n) \end{bmatrix}^\mathrm{T} \tag{5}$$

$$
\boldsymbol{C}=\begin{bmatrix} 1 & t_1 & \cos(2\pi f_1 t_1) & \sin(2\pi f_1 t_1) & \cos(2\pi f_2 t_1) & \sin(2\pi f_2 t_1) \\ 1 & t_2 & \cos(2\pi f_1 t_2) & \sin(2\pi f_1 t_2) & \cos(2\pi f_2 t_2) & \sin(2\pi f_2 t_2) \\ & & & \vdots & & \\ 1 & t_n & \cos(2\pi f_1 t_n) & \sin(2\pi f_1 t_n) & \cos(2\pi f_2 t_n) & \sin(2\pi f_2 t_n) \end{bmatrix} \tag{6}
$$

The deterministic models as shown in Eq. (1) and (2) aim to extrapolate the present trend, CW and AW within x_{p} and y_{p} time-series. This can be carried out by the WLS extrapolation. This deterministic models are subsequently utilized for two purposes: (1) to extrapolate the deterministic components in time-series and (2) to get stochastic residuals for WLS misfit (the difference between the deterministic models and data themselves).

2.2. Choice of weighting function

A key issue in WLS fit is the appropriate choice of a weighting matrix $\boldsymbol{W}$, the determination of a proper weight w_i assigned to the observation at time t_i. Schuh et al. (2001) calculated the weights by estimating empirical variance components. Sun and Xu (2012) developed a power weighting function with the optimal power. Wu et al. (2018) determined the weights with the a priori variance of observables. When observations are closer to the epochs of predictions, the greater the effect on predictions are. Considering the effect of intervals between the observation points and prediction epochs on results of LS extrapolation, the weights w_i are selected according to the following two weighting functions in current work.

(1) Equal weighting function: $w_i = 1, i = 1, 2, \cdots, n$. This simple weighting function is widely used for LS analysis, but does not take into consideration the impact of the geometry.

(2) Piecewise weighting function:

$$
w_i = \begin{cases} \frac{1}{3} & 1 \leq i \leq \frac{n}{3} \\ \frac{1}{2} & \frac{n}{3} < i \leq \frac{2n}{3} \\ 1 & \frac{2n}{3} < i \leq n \end{cases} \tag{7}
$$

The function is a very simple method to determine the weights for WLS fit, but takes the geometry into account. It generates the constant weights for relatively short given time-series, which means that a small number of measurements at every interval influence on LS fit in exactly the similar way. The effect of two weighting strategies on PM predictions will be described in the following section.

2.3. Vector autoregressive modeling

The VAR model developed by Love and Zicchino (2006) allows one to account for unobserved individual heterogeneity for the entire time-series via the introduction of fixed effects that enhance the consistency and coherence of measurements. Moreover, this model has some practical advantages that make it a very suitable method to model x_{p} and y_{p} together. First, unlike multivariate autoregressive (MAR) process, the VAR model makes no distinction between exogenous and endogenous variables, in accordance with the interdependence reality. Instead, all variables are treated as endogenous mutually. Each variable that entered a VAR model not only depends on its historical realization but also correlates with other variables, conforming with real simultaneities and a priori correlation between x_{p} and y_{p}. Second, VAR is very uncomplicated for efficient and coherent estimations for a small number of variables, especially for two variables in case of x_{p} and y_{p}. Let us denote the bivariate x_{p} and y_{p} residual data as a vector $\boldsymbol{R}(t) = (r_x(t), r_y(t))'$, where $r_x(t)$ and $r_y(t)$ are x_p and y_p residual time-series, respectively. Modeling the residual bivariate time-series $\boldsymbol{R}(t)$ may be performed utilizing the VAR technique.

Given a vector of residuals $\boldsymbol{R}(t) = (r_x(t), r_y(t))'$ at time t, the basic VAR model of a p-order without deterministic components has the form

$$\boldsymbol{R}(t) = \boldsymbol{\alpha}_1\boldsymbol{R}(t-1) + \boldsymbol{\alpha}_2\boldsymbol{R}(t-2) + \cdots + \boldsymbol{\alpha}_p\boldsymbol{R}(t-p) + \boldsymbol{U}(t) \tag{8}$$

where $\boldsymbol{\alpha}_i$ is the 2×2 matrix of VAR coefficients at lags i ($i = 1, 2, \cdots, p$), and $\boldsymbol{U}(t) = (u_x(t), u_y(t))'$ is an unobservable zero mean white noise vector with time-invariant positive definite covariance matrix, in which $u_x(t)$ and $u_y(t)$ are independent or serially uncorrelated. The model is referenced to briefly as a VAR(p) process as the number of lags is p.

There are a lot of model selection criteria to select an order p. They proceed by fitting VAR(m) models with different orders $m = 1, 2, \cdots, p_\mathrm{max}$ and then select an estimator of the p-order that minimizes some criterion. In this work, we apply the popular Akaike information criterion (AIC) to choose an appropriate order. This criterion for order selection is based on the following statistics (Akaike, 1971).

$$\mathrm{AIC}(p) = \mathrm{In\ det}\left(\widetilde{\sum\nolimits_\mathrm{u}}(\mathrm{m})\right) + \frac{2mk^2}{n} \tag{9}$$

where $\det(\cdot)$ represents the determinant, $\sum_u(m)$ denotes the estimator of the residual covariance matrix for a VAR(m) model, n is the sample data size, and k denotes the dimension of the vector for modeling, specially $k = 2$ in the work. Following Johansen (1995), this work estimates values of the coefficient matrix α_i resulting from fitting the VAR(p) model to the two-dimensional residuals using maximum likelihood.

The fitted VAR model is applied in the process of predicting x_p and y_p together. The one-step prediction is computed as

$$\hat{\boldsymbol{R}}(t+1) = \hat{\boldsymbol{\alpha}}_1\boldsymbol{R}(t) + \hat{\boldsymbol{\alpha}}_2\boldsymbol{R}(t-1) + \cdots + \hat{\boldsymbol{\alpha}}_p\boldsymbol{R}(t-p+1) \tag{10}$$

where $\hat{\boldsymbol{R}}(t+1)$ is the predicted vector at time $t+1$, $\hat{\boldsymbol{\alpha}}_i, i = 1, 2, \cdots, p$, are the estimated coefficient matrices. Multistep predictions of residual vector for $t+2, t+3, t+T$ are calculated recursively using Eq. (10) by taking place at time t by $t+1$.

3. PREDICTION OF POLAR MOTION

3.1. Data description

The IERS regularly releases daily values of the EOP C04 series with a 30-day delay after the combination of solutions of space geodetic technologies including Very Long Baseline Interferometry (VLBI), Satellite Laser Ranging (SLR), Doppler Orbitography and Radiopositioning Integrated by Satellite (DORIS), and Global Navigation Satellite System (GNSS). The combined EOP C04 series is routinely published at one-day intervals at the IRES official website (https://www.iers.org/IERS/EN/DataProducts/data.html). The series consists of observations of the x_p and y_p pole coordinates from 1962 until 30 days from now with a high accuracy. The mean and standard deviation for pole coordinates of the differences between combined solutions derived by the RS/PC and C04 over 2018 are less than 0.06 ms (Dick and Thaller, 2020), and therefore have a negligible impact on PM predictions. The IERS directing board adopted the EOP 14 C04 as the IERS reference series in February 1, 2017 (Bizouard et al., 2019). In this work, the EOP 14 C04 series is taken as data source of the x_{p} and y_{p} pole coordinates.

3.2. Error statistics

The predicted values of x_{p} and y_{p} are calculated as the sum of the WLS extrapolations and univariate or vector predictions of x_{p} and y_{p}. The mean absolute error (MAE), used as the official 1st EOP PCC statistics, is selected for evaluating the prediction accuracy. The MAE is expressed for the ith day into the future as follows (Kalarus et al., 2010).

$$\mathrm{MAE}_i = \frac{1}{N}\sum_{j=1}^{N}|O_{i,j} - P_{i,j}| \tag{11}$$

where N is the number of predictions, $O_{i,j}$ and $P_{i,j}$ are the observed pole coordinates and their ith point of jth prediction, respectively.

3.3. Comparison with the conventional LS+AR

Predictions of x_{p} and y_{p} are computed by means of two prediction techniques: LS+AR and WLS+VAR. The range of all predictions is equal to 365 days. Totally, 500 predictions are generated every 7 days at different starting prediction days from January 1, 2012 to July 25, 2021 to calculate the representative statistics, the number N of prediction is 500. The parameters of the linear part, CW and AW in LS and WLS fit are solved by a sliding window analysis for every 365-day prediction. The window size for LS and WLS fit is set to 12 years, approximately twice the beating period of the CW and AW. The above two weighting functions have been applied in WLS fit to extrapolate the linear trend, CW and AW and obtain the residuals of misfit. The comparison of the predictions is carried out by means of the analysis of: (1) the absolute differences between the x_{p} and y_{p} data and their predictions (Figure 1 and Table 1), and (2) MAE statistics (Figure 2 and Table 2).

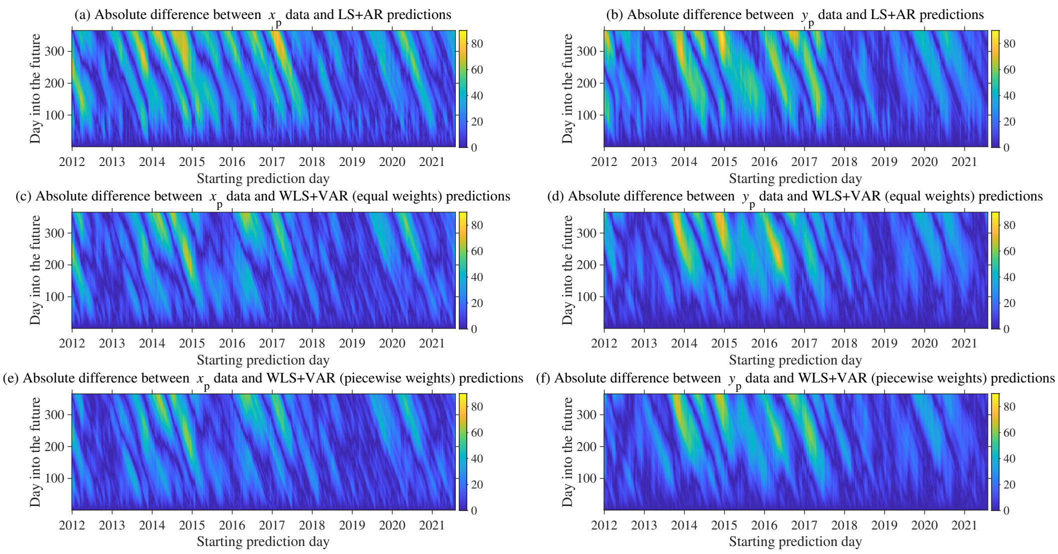

Figure 1. Absolute differences (in mas) between the x_{p} and y_{p} observations and their predictions calculated at different starting prediction days

The absolute values of the difference between the observed x_{p} and y_{p} pole coordinates and their one-year predictions starting at different days do not exceed 85 mas (Figure 1). The larger absolute differences are for the chosen predictions up to prediction horizons greater than 8 months. Those inaccurate predictions with the maximum absolute difference as given in Table 1 are generated for starting prediction epochs in the early days of 2014 (Figure 1), which might correspond to the El Niño event 2014/2015. An El Niño phenomena could cause extreme stochastic perturbations in the seasonal harmonic oscillations and would hence decrease the PM prediction accuracy in time to the El Niño event. In addition, the predictions beginning in the middle of 2016 are inaccurate. During these prediction starting days, the maximum absolute differences are 78.07 mas, 77.84, mas and 66.13 mas respectively calculated by LS+AR, WLS+VAR with equal weights and WLS+VAR with peicewise weights. The interpretation can be correlated with the La Niña 2016/2017. There are also less accurate predictions beginning in the last half of 2020. This can be linked to the La Niña 2020. For the span of tested data, the "uncorrelated" LS+AR solution generates the predictions of x_{p} and y_{p} with the maximum value of difference between the data themselves and predictions of more than 84 mas, while the absolute difference for the "correlated" WLS+VAR techniques is less than 77 mas. These differences and extreme perturbations are substantially smaller than those calculated for the conventional solution, although the worst predictions obtained by the "correlated" techniques also correspond to EI Niño or La Niña events. The cause of the poor accuracy of PM predictions can be the period or phase variation of the annual oscillation in the occurrence of such events. (Kosek et al., 2001). The prediction accuracy may be enhanced in such a way that the period or phase of the annual oscillation is treated as a variable rather than a constant and modeled from short-term PM time-series in the LS adjustment (Guo et al., 2013; Zhang et al., 2012). In the case of WLS+VAR, one can assign larger weights to recent observations to fit the variable annual oscillation.

Table 1 gives the maximum absolute differences for all considered predictions of x_{p} and y_{p} for 10-day, 30-day, 150-day, 270-day, and 365-day into the future. It can be seen that the maximum errors of the medium- and long-term predictions can be noticeably decreased by the proposed WLS+VAR approach, indicting that the error increasing of both x_{p} and y_{p} predictions will be alleviated if two components are modeled and predicted together.

Table 1. Maximum absolute differences (in mas) between the pole coordinate observations and their predictions up to different prediction ranges

Prediction technology	x_{p}					y_{p}				
	10 d	30 d	150 d	270 d	365 d	10 d	30 d	150 d	270 d	365 d
LS+AR	11.28	29.31	67.22	74.84	77.03	9.76	19.66	64.78	73.86	84.96
WLS+VAR (equal weights)	10.63	21.22	51.89	65.71	69.16	8.65	14.16	47.69	71.76	76.66
WLS+VAR (piecewise weights)	11.07	23.57	46.11	61.05	64.70	9.17	13.24	47.29	63.55	70.80

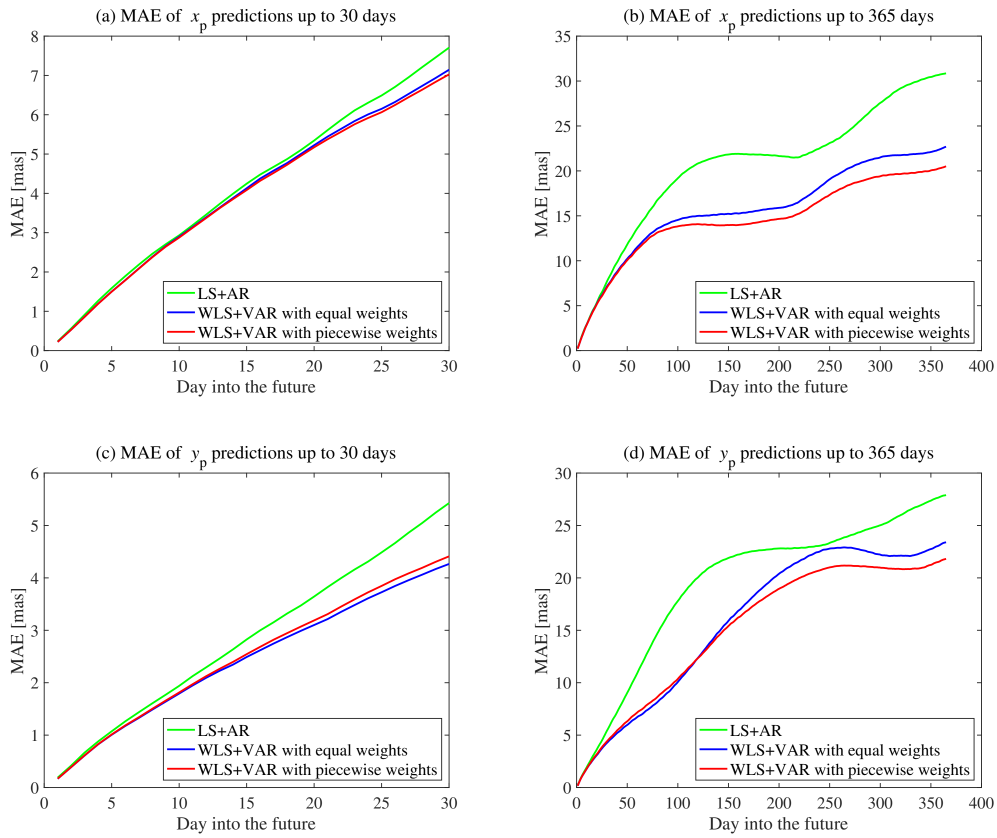

Figure 2. Mean absolute errors (in mas) of the x_p and y_p predictions up to 365 days into the future

The essential results regarding the accuracy of different techniques come from an analysis of the MAE for the x_p and y_p predictions (Figure 2 and Table 2). The results show that the accuracy of three techniques is comparable for the predictions up to 30 days into the future, while the accuracy of the predictions out 30 days computed by WLS+VAR is better than that by the conventional LS+AR approach. In particular, the enhancement in accuracies is mostly seen in the case of the medium- and long-term predictions for horizons of 60 to 365 days based on WLS+VAR. One key reason is that VAR works better than AR in modeling and prediction of the residuals for WLS misfit. Furthermore, it can be seen that the choice of appropriate weights is crucial. The equal weights $w_i = 1$ ignore the effect of intervals between the observation points and prediction epochs on the predicted values. The maximum values of absolute difference between the observations and predictions and MAE values obtained by the equal weights are 76.66 mas and 23.40 mas for horizons of 1 to 365 days respectively, larger than the maximum absolute difference of less than 71 mas and MAE of no more than 22 mas by the piecewise weights as shown in Tables 1 and 2. This is owing to the assignment of different weights to observations that can improve the LS fit for the time-variant CW and AW. That is, if appropriate weights are assigned to PM measurement, WLS fit will work better. The figure and statistics of the MAE not only demonstrates that PM predictions would be biased if the correlation between x_p and y_p was ignored, but also confirms that WLS+VAR with piecewise weights is a more accurate method for PM prediction than two others. The accuracies of the x_p predictions are 13.97 mas, 18.47 mas, and 20.52 mas respectively for 150, 270, and 365 days in terms of the MAE statistics, which are 36%, 24.8%, and 33.5% higher than LS+AR respectively and

8.2%, 9.2% and 9.6% higher than WLS+VAR with equal weights respectively for 150, 270, and 365 days. For the y_p predictions, the 150-, 270- and 365-day accuracies are 15.41 mas, 21.17 mas, and 21.82 mas respectively, which are 27.4%, 11.9%, and 21.8% higher than LS+AR respectively and 3.1%, 7.5%, and 6.8% higher than WLS+VAR with equal weights respectively for the corresponding prediction days.

Table 2. Mean absolute errors (in mas) for different prediction ranges by the LS+AR and WLS+VAR

Prediction technology	x_p					y_p				
	10 d	30 d	150 d	270 d	365 d	10 d	30 d	150 d	270 d	365 d
LS+AR	2.93	7.71	21.83	24.57	30.86	1.94	5.43	21.90	24.04	27.90
WLS+VAR (equal weights)	2.90	7.15	15.22	20.35	22.69	1.79	4.27	15.91	22.88	23.40
WLS+VAR (piecewise weights)	2.87	7.03	13.97	18.47	20.52	1.81	4.41	15.41	21.17	21.82

3.4. Comparison with the 1st EOP PCC

In October 2005, the 1st EOP PCC was launched for making an examination of the property of existing prediction algorithms. There were 13 participants who contributed to the campaign and submitted EOP predictions. At the end of February 2008, this campaign collected about 6,500 submissions of EOP predictions calculated by 20 prediction methods and techniques denoted by ID number. The 1st EOP PCC is a variety of contest where the validation strategy and prediction period were specified clearly in advance. The campaign is refereed to Kalarus et al. (2010) for detail. This work uses WLS+VAR with the piecewise weights to generate the x_p and y_p predictions up to 365 days into the future between January 1, 2005 and February 28, 2008, a same prediction period as that in the 1st EOP PCC, so as to compare the predictions with those submitted to this campaign in a common and objective way. Using the EOP 05 C04 series as a reference, Figure 3 and Table 2 compare the accuracy of the short-term prediction for 30 days ahead and medium- and long-term predictions of x_p and y_p pole coordinates up 365 days into the future based on WLS+AR (piecewise weights) with the results submitted to the campaign by different methods and techniques. What can be said with the information from Figure 3 is that LS extrapolation+AR prediction by ID Kalarus 061 and ID Kosek 051 is the most accurate technique for medium- and long-term x_p, y_p predictions as demonstrated in the 1st EOP PCC, while pure AR prediction (Zotov 091) and neural networks (Zotov 093) are more accurate methods to predict short-term x_p and y_p in addition to LS extrapolation+AR prediction.

Figure 3 and Table 2 indicate that the accuracy of the short-term x_p and y_p predictions up to 30 days obtained by WLS+VAR is comparable with that by LS extrapolation+AR prediction (Kalarus 061 and ID Kosek 051) as well as by pure AR prediction (Zotov 091), and better than the prediction accuracy by neural networks (Zotov 093) participating in the campaign. Moreover, the accuracy of the medium- and long-term predictions up to 365 days is inferior to the best presently available methods in the 1st EOP PCC, LS extrapolation+AR prediction by Kalarus 061 and ID Kosek 051. However, the predictions of both x_p and y_p are more accurate than those submitted to the 1st EOP PCC computed by the other methods and technologies. The MAE values of the x_p and y_p predictions up to 365 days by WLS+VAR are less than 29 mas and 33 mas respectively, while the MAE values of the 365-day horizon are more than 37 mas and 40 mas respectively for the x_p and y_p predictions by the participants except Kosek 051 and Karalus 061.

Table 3. Mean absolute errors (in mas) for different prediction ranges by the WLS+VAR and participants in the 1st EOP PCC

Prediction technology	x_p					y_p				
	10 d	30 d	150 d	270 d	365 d	10 d	30 d	150 d	270 d	365 d
LS+AR (Kosek 051)	3.56	9.52	21.16	23.85	19.78	2.01	6.08	25.65	21.86	28.91
LS+AR (Kalarus 061)	3.33	8.78	19.36	17.46	14.50	1.83	5.01	21.28	25.10	21.32
AR (Zotov 091)	4.17	9.31	—	—	—	2.30	5.64	—	—	—
Neural networks (Zotov 093)	4.21	11.51	—	—	—	2.14	6.08	—	—	—
WLS+VAR (piecewise weights)	3.49	9.20	21.13	29.67	28.04	1.89	5.84	23.18	27.58	32.55

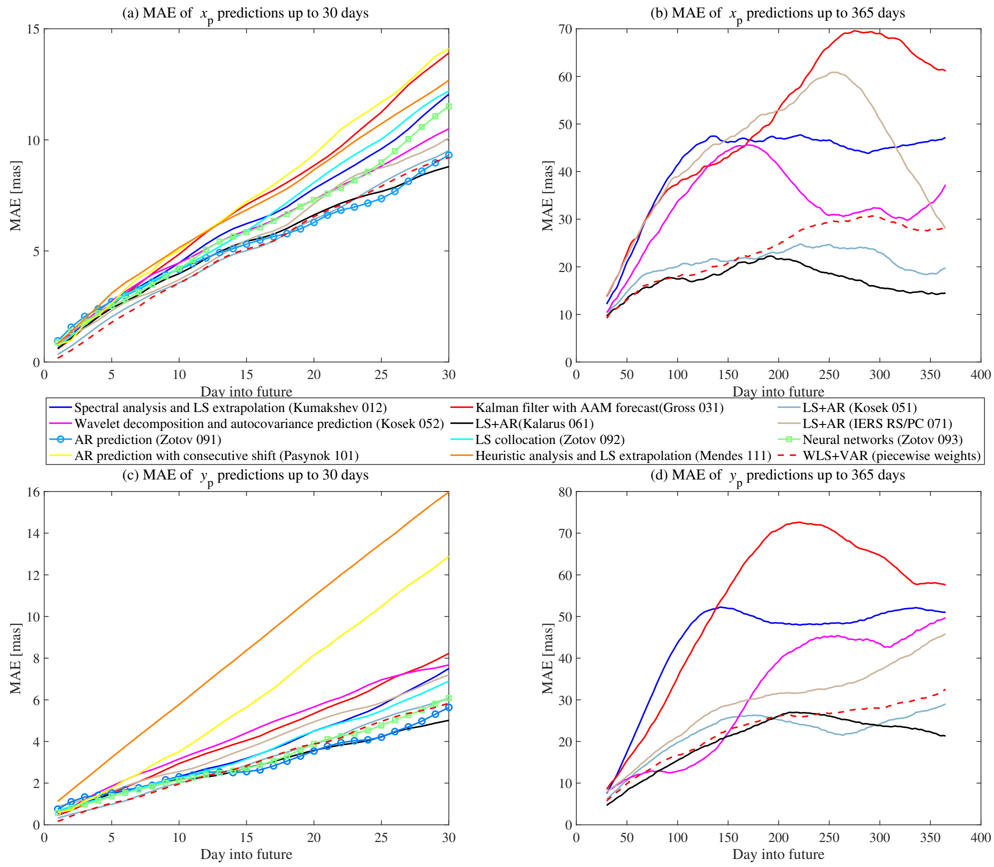

Figure 3. Mean absolute errors (in mas) of the $x_{\rm p}$ and $y_{\rm p}$ predictions up to 365 days into the future

4. CONCLUSIONS

PM time-series can be predicted 365 days in advance into the future with enhanced accuracies by means of the presented WLS+VAR approach. This approach solves simultaneously all parameters of the trend, CW and AW in $x_{\rm p}$ and $y_{\rm p}$ observations by a WLS algorithm for fitting the variabilities of the CW and AW more accurately. For the WLS fit the piecewise weighting function is proposed to take into the effect of intervals between observation and prediction epochs on the results of the WLS fit by assigning larger weights to recent observations. Moreover, considering the correlation between x_p and y_p pole coordinates, the VAR technology is used to model and predict the WLS residuals of x_p and y_p together. A set of predictions indicate that the extrapolation of the harmonic model for the trend, CW and AW supported by the WLS algorithm can improve the PM predictions. The results also show that the VAR technology can be applied successfully to together model and predict the residuals for WLS misfit of x_p and y_p. In the cases analyzed, the VAR predictions are more accurate than the univariate ones. It is found that that the maximum absolute difference of the observations and predictions up to 365 days by the WLS+VAR is no more than 77 mas, smaller than the maximum difference of approximately 85 mas by commonly used LS+AR. Moreover, the MAE of the predictions beyond 30 days are smaller than that by LS+AR especially for medium- and long-term ranges. The WLS+VAR with piecewise weights performs the accuracy of 20.52 mas and 21.82, respectively, for the 365-day predictions of x_p and y_p in terms of the MAE statistics, 33.5$\%$ and 21.8$\%$ higher than LS+AR for the x_p and y_p, respectively. It is also shown that the accuracy of the predictions with piecewise weights slightly outperforms that with equal weights, 9.6$\%$ and 6.8$\%$ higher than equal weights for the 365-day predictions of x_p and y_p, respectively. However, the absolute differences computed by the former are smaller than those by the latter, which is practically important and can be beneficial to practical and scientific users. The comparison with the predictions submitted to the 1st EOP PCC further demonstrates that presented WLS+VAR generates the more accurate predictions than the other techniques except LS+AR, one of the most accurate prediction techniques participating in the campaign for x_p, y_p pole coordinates. The presented approach is very simple to use and can therefore take place of the commonly used LS+AR method to generate PM predictions.

This study supports the claim that PM predictions will be improved if the correlation between x_p and y_p is taken into consideration, even if pole coordinates x_p and y_p are being together modeled and predicted by a linear technology. On the other hand, the strategy for weight assignment needs to be investigated for further improvement in PM predictions.

Acknowledgments. The authors are grateful to the IERS and M. Kalarus for providing the EOP C04 series and the polar motion predictions from the 1st EOP PCC, respectively. This work is supported by the Shaanxi Natural Science Foundation of China (Grant Nos. 2023-JC-YB-057 and 2022-JM031).

REFERENCES

Akaike H. (1971) Autoregressive model fitting for control, *Annals of the Institute of Statistical Mathematics*, Vol. 23, No. 1, 163-180.

Akyilmaz O, Kutterer H. (2004) Prediction of Earth rotation parameters by fuzzy inference systems, *Journal of Geodesy*, Vol. 78, No. 1, 82-93.

Bizouard C, Lambert S, Gattano C, et al. (2019) The IERS EOP 14C04 solution for Earth orientation parameters consistent with ITRF 2014, *Journal of Geodesy*, Vol. 93, No. 5, 621-633.

Chin T, Gross R, Dickey J. (2004) Modeling and forecast of the polar motion excitation functions for short-term polar motion prediction, *Journal of Geodesy*, Vol. 78, No. 6, 343-353.

Dick W, Thaller D. (2020) IERS annual report 2018, *International Earth Rotation and Reference Systems Service*, Central Bureau, Frankfurt am Main.

Dill R, Dobslaw H. (2010) Short-term polar motion forecasts from Earth system modeling data, *Journal of Geodesy*, Vol. 84, No. 9, 529-536.

Dill R, Dobslaw H, Thomas M. (2019) Improved 90-day Earth orientation predictions from angular momentum forecasts of atmosphere, ocean, and terrestrial hydrosphere, *Journal of Geodesy*, Vol. 93, No. 3, 287-295.

Dobslaw H, Dill R. (2017) Predicting Earth rotation variations from global forecasts of atmosphere-hydrosphere dynamics, *Advances in Space Research*, Vol. 61, No. 4, 1047-1054.

Gambis D, Luzum B. (2011) Earth rotation monitoring, UT1determination and prediction, *Metrologia*, Vol. 48, No. 4, S165.

Guo J, Li Y, Dai C, et al. (2013) A technique to improve the accuracy of Earth orientation prediction algorithms based on least squares extrapolation, *Journal of Geodynamics*, Vol. 70, No. 10, 36-48.

IERS Annual Report 2018. Edited by Wolfgang R. Dick and Daniela Thaller. International Earth Rotation and Reference Systems Service, Central Bureau. Frankfurt am Main: Verlag des Bundesamts für Kartographie und Geodäsie, 2020. 207 p., ISBN 978-3-86482-136-3 (print version)

Jin X, Liu X, Guo J. et al. (2021) Analysis and prediction of polar motion using MSSA method, *Earth, Planets and Space*, Vol. 73, No. 1, 1-13.

Kalarus M, Schuh H, Kosek W, et al. (2010) Achievements of the Earth orientation parameters prediction comparison campaign, *Journal of Geodesy*, Vol. 84, No. 10, 587-596.

Kosek W, McCarthy D, Luzum B. (1998) Possible improvement of Earth orientation forecast using autocovariance prediction procedures, *Journal of Geodesy*, Vol. 72, No. 4, 189-199.

Kosek W, McCarthy D, Luzum B. (2001). El Niño impact on polar motion prediction errors, *Studia Geophysica et Geodaetica*, Vol. 45, No. 4, 347-361.

Love I, Zicchino L. (2006) Financial development and dynamic investment behavior: evidence from panel VAR, *The Quarterly Review of Economics and Finance*, Vol. 46, No. 2, 190-210.

Modiri S, Belda S, Heinkelmann R, et al. (2018) Polar motion prediction using the combination of SSA and copula-based analysis, *Earth, Planets and Space*, Vol. 70, No. 1, 1-18.

Schuh H, Nagel S, Seitz T. (2001) Linear drift and periodic variations observed in long time series of polar motion, *Journal of Geodesy*, Vol. 74, No. 10, 701-710.

Schuh H, Ulrich M, Egger D, et al. (2002) Prediction of Earth orientation parameters by artificial neural networks, *Journal of Geodesy*, Vol. 76, No. 5, 247-258.

Shen Y, Guo J, Liu X, et al (2017) One hybrid model combining singular spectrum analysis and LS + ARMA for polar motion prediction, *Advances in Space Research*, Vol. 59, No. 2, 513-523.

Su X, Liu L, Houtse H, et al. (2014) Long-term polar motion prediction using normal time–frequency transform, *Journal of Geodesy*, Vol. 88, No. 2, 145-155.

Sun Z, Xu T (2012) Prediction of Earth rotation parameters based on improved weighted least squares and autoregressive model *Geodesy and Geodynamics*, Vol. 3, No. 3, 57-64.

Sun Z, Xu T, Jiang C, et al. (2019) An improved prediction algorithm for Earth' s polar motion with considering the retrograde annual and semi-annual wobbles based on least squares and autoregressive model, *Acta Geodaetica et Geophysica*, Vol. 54, No. 4, 499-511.

Wang L, Miao W, Wu F. (2022) A new polar motion prediction method combined with the difference between polar motion series *Geodesy and Geodynamics*, Vol. 13, No. 6, 564-572.

Wooden W, van Dam T, Kosek W. (2006) IERS Working Group on prediction plans and activities, EOS Trans. AGU, 87(52), In AGU Fall Meeting Abstracts , pp. G43A-0988

Wu F, Chang G, Deng K, et al. (2019) Selecting data for autoregressive modeling in polar motion prediction, *Acta Geodaetica et Geophysica*, Vol. 54, No. 4, 557-566.

Wu F, Deng K, Chang G, et al. (2018) The application of a combination of weighted least-squares and autoregressive methods in predictions of polar motion parameters, *Acta Geodaetica et Geophysica*, Vol. 53, No. 2, 247-257.

Wu F, Liu Z, Deng K, et al. (2021) A polar motion prediction method considering the polar coordinates, *Advances in Space Research*, Vol. 68, No. 3, 1318-1328.

Xu X, Zhou Y. (2015) EOP prediction using least square fitting and autoregressive filter over optimized data intervals, *Advances in Space Research*, Vol. 56, No. 10, 2248-2253.

Xu X, Zhou Y, Liao X. (2012) Short-term Earth orientation parameters predictions by combination of the least-squares, AR model and Kalman filter, *Journal of Geodynamics*, Vol. 62, No. 12, 83-86.

Yang Y, Nie W, Xu T, et al. (2022) Earth orientation parameters prediction based on the hybrid SSA+LS+SVM model, *Measurement Science and Technology*, Vol. 33, No. 12, 125011.

Zhang H, Wang Q, Zhu J, Zhang X. Application of CLS+AR model polar motion to prediction based on time-varying parameters correction of Chandler wobble, *Geomatics and Information Science of Wuhan University*, Vol. 37, No. 3, 286-289.

Zhao D, Lei Y. (2019) Possible enhancement of Earth' s polar motion predictions using a wavelet-based preprocessing procedure, *Studia Geophysica et Geodaetica*, Vol. 63, No. 1, 83-94.

Zhao D, Lei Y. (2020) A technique to reduce the edge effect in least squares extrapolation for enhanced Earth orientation prediction., *Studia Geophysica et Geodaetica*, Vol. 64, No. 3, 293-305.

Received: *2023-01-05*

Reviewed: *2023-01-20 (undisclosed name); 2023-04-19 (undisclosed name)*

Accepted: *2023-05-17*