FULL PAPER

Short-term prediction of celestial pole offsets with interpretable machine learning

Mostafa Kiani Shahvandi[1]*, Santiago Belda[2], Siddhartha Mishra[3] and Benedikt Soja[1]

Abstract

The difference between observed and modelled precession/nutation reveals unmodelled signals commonly referred to as Celestial Pole Offsets (CPO), denoted by dX and dY. CPO are currently observed only by Very Long Baseline Interferometry (VLBI), but there is nearly 4 weeks of latency by which the data centers provide the most accurate, final CPO series. This latency problem necessitates predicting CPO for high-accuracy, real-time applications that require information regarding Earth rotation, such as spacecraft navigation. Even though the International Earth Rotation and Reference Systems Service (IERS) provides so-called rapid CPO, they are usually less accurate and therefore, may not satisfy the requirements of the mentioned applications. To enhance the quality of CPO predictions, we present a new methodology based on Neural Additive Models (NAMs), a class of interpretable machine learning algorithms. We formulate the problem based on long short-term memory neural networks and derive simple analytical relations for the quantification of prediction uncertainty and feature importance, thereby enhancing the intelligibility of predictions made by machine learning. We then focus on the short-term prediction of CPO with a forecasting horizon of 30 days. We develop an operational framework that consistently provides CPO predictions. Using the CPO series of Jet Propulsion Laboratory as the input to the algorithm, we show that NAMs predictions improve the IERS rapid products on average by 57% for dX and 25% for dY under fully operational conditions. Our predictions are both accurate and overcome the latency issue of final CPO series and thus, can be used in real-time applications.

Keywords Celestial pole offsets, Interpretable machine learning, Neural additive models, Prediction accuracy
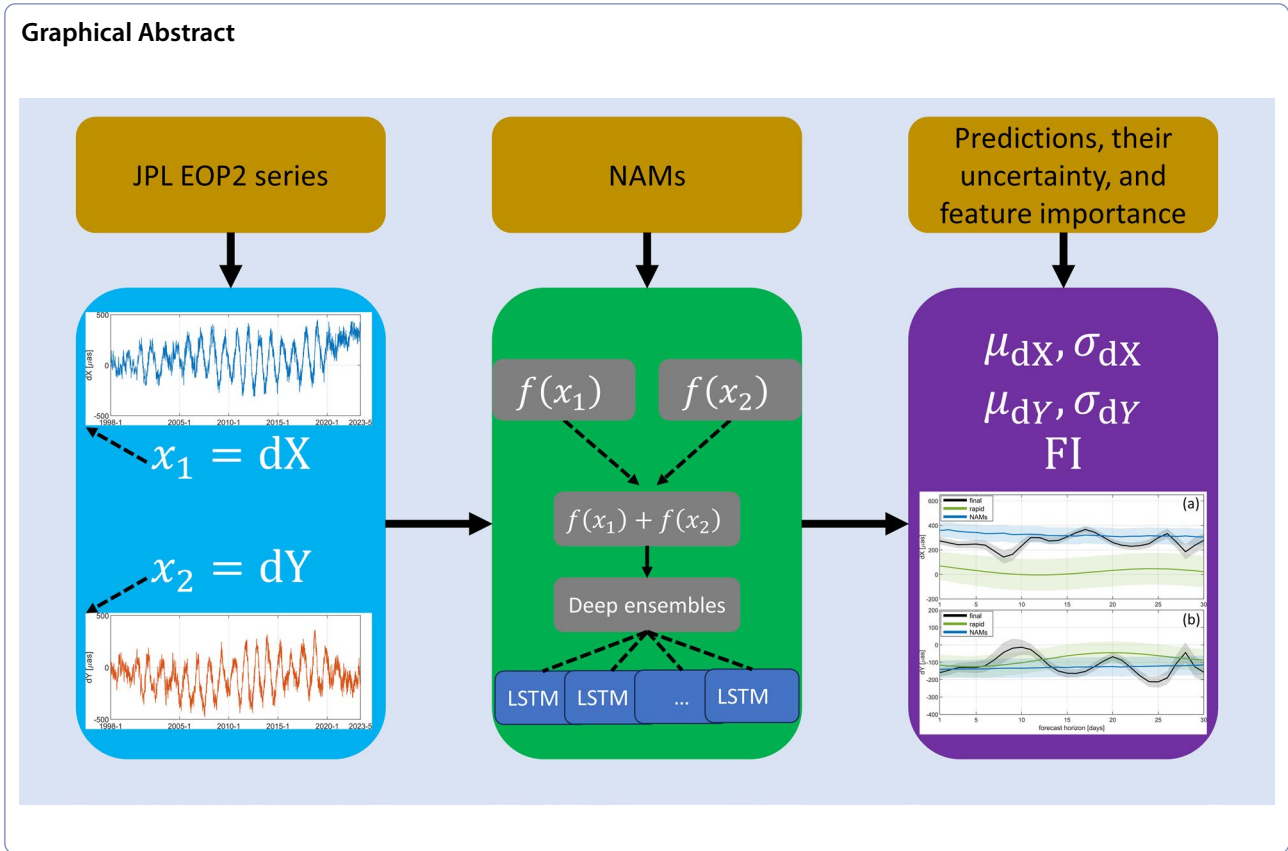
*Correspondence:
Mostafa Kiani Shahvandi
mkiani@ethz.ch
Full list of author information is available at the end of the article

Graphical Abstract

Introduction

Gravitational effects of the Sun and Moon induce variations in the Earth rotation axis, commonly known as precession and nutation (Gross 2015). In addition, mass redistribution within the Earth system also perturbs the rotation axis of the Earth. The latter can be modulated onto the nutation, generating small oscillatory motions. These motions are irregular and depend on various parameters, including geodynamics of the Earth. Therefore, in contrast to the major components of precession and nutation that are modelled with rigorous theories (e.g. Wahr 1981; Matthews et al. 1991a, b, 2002), the mentioned irregular motions are left unmodelled. With the advent of Very Long Baseline Interferometry (VLBI; Sovers et al. 1998) researchers were enabled to derive these unmodelled signals (Herring et al. 1991), which are commonly referred to as Celestial Pole Offsets (CPO). CPO represent by definition a two-dimensional motion, the components of which are denoted by dX and dY. These components vary in time and are dominated by a retrograde oscillatory motion referred to as Free Core Nutation (FCN; Wahr 1988). The primary cause of FCN is the misalignment between the rotation axis of the fluid core and mantle and it is suggested to be excited by a combination of atmospheric and oceanic processes (Sasao and

Wahr 1981). The period of FCN in the celestial frame is around 431 days, but by complex demodulation at diurnal frequency (Brzeziński 1994) it is translated to the terrestrial frame as a retrograde diurnal motion; thus, it is sometimes referred to as the "nearly diurnal free wobble" (Sasao and Wahr 1981). As such, a simple harmonic function (based on sine and cosine) can be fitted to the CPO time series and subsequently used to predict CPO. This has been the basis of various methodologies (e.g. Belda et al. 2016, 2017, 2018) that attempt to model and predict CPO time series.

In addition to FCN, there are some other, more irregular parts in CPO which are difficult to model and predict, including perhaps the effect of geomagnetic jerks (Shirai et al. 2005). Furthermore, VLBI observations of CPO are contaminated by high degrees of noise. Therefore, robust modelling and prediction of CPO time series has been challenging (Belda et al. 2017). The mentioned methodologies based on extrapolation of harmonic functions may not present the highest accuracy. Other methodologies such as Kalman filtering that are based on stochastic modelling may be more accurate (Nastula et al. 2020). However, even the mentioned methods may not take into account the full dynamics of CPO time series including the potential nonlinear evolution (e.g. nonstationarity

due to possible impacts from the aforementioned geomagnetic jerks). To overcome this problem, the use of machine learning algorithms is suggested to be beneficial (Kiani Shahvandi and Soja 2021, 2022a; Kiani Shahvandi et al. 2022b). In fact, to compare the prediction capability of different algorithms for the prediction of Earth Orientation Parameters (EOPs), the second Earth Orientation Parameters Prediction Comparison Campaign (EOPPCC) was organized (Śliwińska et al. 2023). One of the primary conclusions of this campaign was that the machine learning algorithms present the most accurate CPO predictions (Wińska et al. 2023) in operational settings. Motivated by this success, we intend to present a new methodology based on machine learning that can accurately predict CPO time series under operational conditions. These predictions can then be used in applications that require real-time information regarding Earth rotation, including satellite navigation (Kiani Shahvandi et al. 2022a).

Even though machine learning algorithms are a powerful approach for time series prediction (Lim and Zohren 2021), they lack the interpretability of traditional, simpler models. One of the most common characteristics of the conventional methods is that they are mostly linear in nature (such as fitting a line to observational data). In contrast, machine learning algorithms are nonlinear, which makes their interpretation quite challenging. To address this problem, so-called interpretable machine learning algorithms have been developed (e.g. Molnar 2023). We particularly focus on one such algorithm that has been proposed recently, called Neural Additive Models (NAMs; Agarwal et al. 2021). NAMs are suggested to be highly accurate and easy to interpret. The key idea behind NAMs is to use a neural network for each input feature and then sum up the output of all these neural networks (i.e., neural network has one input feature to also predict only one feature). This implies that NAMs are based on the so-called generalized additive models (Hastie and Tibshirani 1986), where for each input feature the function that relates this feature to the output can be nonlinear, but the final output of the algorithm is linearly related to the output of each individual model. However, the prediction uncertainty (Kiani Shahvandi et al. 2023) and feature importance analysis methodologies (Kiani Shahvandi et al. 2022a) are not available yet for NAMs. The high level of noise in CPO data necessitates the quantification of prediction uncertainty to assign a measure of reliability to the predictions (Kiani Shahvandi and Soja 2022b). On the other hand, it is important to analyze which input features (for example dX vs dY) contribute the most to the prediction accuracy. We therefore focus on developing the theoretical frameworks for prediction uncertainty and feature importance

and subsequently use them to predict CPO time series in short-term forecasting horizon, i.e., up to 30 days to the future. We focus on this forecasting horizon because we intend to fully cover the 4 week latency of final EOP series provided by the International Earth Rotation and Reference Systems Service (IERS). In short, the goals of the present study are as follows:

- Presenting a new methodology based on NAMs for the prediction of CPO, as well as analytically formulating the quantification of prediction uncertainty and analysis of feature importance
- Analyzing the methodology under operational settings to realistically examine the prediction performance

The rest of this paper is organized as follows. In "Methods" the methodology is developed. In "Data description" the data used in the study are described. "Results and discussions" is devoted to presenting the results and their interpretation, whilst "Conclusions" to stating the conclusions.

Methods

Let us assume here that two features x_1 and x_2 describe a time series y. If the relationship between x_1, x_2 and y is through a function denoted by f, then NAMs present y as equivalent to the linear summation of $f(x_1)$ and $f(x_1)$ as follows

$$y = f(x_1) + f(x_2). \tag{1}$$

This simple idea enables interpretation of the role of x_1 and x_2 on the prediction of y through the analysis of feature importance (Kiani Shahvandi et al. 2022a). Since the output of the model is linear, we define the importance of features x_1 and x_2 based on the absolute value of linear correlation coefficients of x_1 and x_2 with y. Considering that this correlation coefficient is based on the variance of $f(x_1)$ and $f(x_2)$, we need to estimate the prediction uncertainty (Kiani Shahvandi and Soja 2022b; Kiani Shahvandi et al. 2023). To quantify the prediction uncertainty, we follow the deep ensemble approach (Lakshminarayanan et al. 2016) described in Equation (2), where several models with different initial parameters are fitted simultaneously to the same data. The size of this ensemble is denoted here by M and the individual ensemble member by the index j, where $j = 1, ..., M$. We use $M = 10$ as suggested by Kiani Shahvandi et al. (2023), because it proved to be quite robust in presenting the most accurate predictions whilst keeping the computational time reasonable (i.e., low computational complexity).

It is important to note that here we use Long Short-Term Memory (LSTM; Hochreiter and Schmidhuber 1997) neural networks for the function *f*. This choice is motivated by the efficiency of LSTM in accurately capturing the temporal dependency in the time series and presenting highly accurate predictions as shown by Lara-Benitez et al. (2021). For each of the ensemble members, we define two LSTM neural networks $\mu_j(x_1) = \text{LSTM}_{\mu_{x_1}}$ and $\mu_j(x_2) = \text{LSTM}_{\mu_{x_2}}$ that represent the function *f* in relation (1). These are shown in Equations (2a)–(2b). The uncertainties in the prediction are also defined based on two LSTM neural networks, denoted by $\sigma_j^2(x_1) = \text{LSTM}_{\sigma_{x_1}}$ and $\sigma_j^2(x_2) = \text{LSTM}_{\sigma_{x_2}}$ in Equation (2c)–(2d). The difference between these neural networks and $\mu_j(x_1)$, $\mu_j(x_2)$ is twofold: (1) each of these neural networks has its own specific, learnable parameters, denoted by $W_{\mu_{x_1}}$, $W_{\mu_{x_2}}$, $W_{\sigma_{x_1}}$, and $W_{\sigma_{x_2}}$; (2) a softplus activation function (Zheng et al. 2015) is applied to $\sigma_j^2(x_1)$ and $\sigma_j^2(x_2)$, ensuring that the value of uncertainty is positive. A numerical stabilizer $\epsilon = 10^{-8}$ is added to the output of the mentioned softplus function (Kiani Shahvandi et al. 2023) to avoid division by zero in the second term of the loss function in Equation (2g). According to Equation (1), the output of NAMs is defined as the linear summation given in Equation (2e). The uncertainty assigned to NAMs is defined according to the linear propagation of uncertainty as in Equation (2f) based on covariance (denoted by cov). The loss function is defined according to Equation (2g) and subsequently minimized by the Adam optimizer (Kingma and Ba 2015) with respect to the learnable parameters $W_{\mu_{x_1}}$, $W_{\mu_{x_2}}$, $W_{\sigma_{x_1}}$, and $W_{\sigma_{x_2}}$. The learning rate and the number of training epochs for optimization are 5×10^{-4} and 500, respectively, proven to be effective in other studies as well (Gou et al. 2023). Averaging over the individual ensemble members, we derive the NAMs output and its uncertainty, given in Equation (2h)–(2i).

$$\sigma_j^2(x_1) = \log\left(1 + \exp\left(\text{LSTM}_{\sigma_{x_1}}(W_{\sigma_{x_1}}, j, x_1)\right)\right) + \varepsilon \tag{2c}$$

$$\sigma_j^2(x_2) = \log\left(1 + \exp\left(\text{LSTM}_{\sigma_{x_2}}(W_{\sigma_{x_2}}, j, x_2)\right)\right) + \varepsilon \tag{2d}$$

$$\mu_j = \mu_j(x_1) + \mu_j(x_2) \tag{2e}$$

$$\sigma_j^2 = \sigma_j^2(x_1) + \sigma_j^2(x_2) + 2\text{cov}\left(\mu_j(x_1), \mu_j(x_2)\right) \tag{2f}$$

$$\ell_j = \frac{1}{2}\log\sigma_j^2 + \frac{1}{2}\frac{(F - \mu_j)^2}{\sigma_j^2} \quad \rightarrow \text{minimize} \tag{2g}$$

$$\mu = \frac{1}{M}\sum_{j=1}^{M}\mu_j \tag{2h}$$

$$\sigma^2 = -\mu^2 + \frac{1}{M}\sum_{j=1}^{M}\left[\sigma_j^2 + \mu_j^2\right] \tag{2i}$$

After having the estimates of uncertainty in the predictions, i.e., $\sigma_j^2(x_1)$ and $\sigma_j^2(x_2)$ in Equation (2c)–(2d), it can be simply shown (see Appendix A) that the importance of features x_1 and x_2 (denoted by FI_{x_1} and FI_{x_2} respectively) is derived using Equation (3). Note that for each ensemble member a distinct feature importance can be derived, Equation (3a). Computing the mean and standard deviation of these individual values, we derive the ensemble FI_{x_1} and FI_{x_2} together with their uncertainties, Equation (3b). It is important to note that FI_{x_1} and FI_{x_2} vary in time and also for different forecasting horizons. The sum over forecasting horizons and its average over all the predictions gives a representative value of feature importance. We therefore present the results for both cases, i.e., the overall feature importance and its temporal variability.

$$\text{FI}_{j,x_k} = \left|\frac{\sigma_j^2(x_k) + \text{cov}\left(\mu_j(x_1), \mu_j(x_2)\right)}{\sigma_j(x_k)\sqrt{\sigma_j^2(x_1) + \sigma_j^2(x_2) + 2\text{cov}\left(\mu_j(x_1), \mu_j(x_2)\right)}}\right|, \quad k = 1, 2 \tag{3a}$$

$$\mu_j(x_1) = \text{LSTM}_{\mu_{x_1}}(W_{\mu_{x_1}}j, x_1) \tag{2a}$$

$$\mu_j(x_2) = \text{LSTM}_{\mu_{x_2}}(W_{\mu_{x_2}}j, x_2) \tag{2b}$$

$$\text{FI}_{x_k} = \frac{1}{M}\sum_{j=1}^{M}\text{FI}_{j,x_k} \pm \sqrt{\frac{1}{M-1}\sum_{j=1}^{M}(\text{FI}_{j,x_k} - \frac{1}{M}\sum_{j=1}^{M}\text{FI}_{j,x_k})^2} \tag{3b}$$

It should be mentioned that NAMs require dX and dY to be predicted separately to retain the idea of

interpretability. This implies that even though the input features to the algorithm are both dX and dY, the output is either dX or dY. Thus, two separate NAMs are defined for dX and dY. Note also that since we defined NAMs in terms of LSTM, we need to define the length of input and output sequence (Gou et al. 2023). These are the number of previous values of the time series used as the input and the forecasting horizon, respectively. An advantage of NAMs is that the input and output sequence lengths are equal, implying that in our case they are both equivalent to 30. The number of hidden neurons used in LSTM is 10, taken from Gou et al. (2023). This proved to be quite efficient in our setting, as it provided the highest prediction accuracy compared to models with fewer or more neurons. The method is implemented using the Tensor-Flow library (Abadi et al. 2015).

To evaluate the prediction performance, we use the Mean Absolute Error (MAE) criterion, employed extensively in EOP studies (Kiani Shahvandi and Soja 2022a; Kiani Shahvandi et al. 2022a, b, 2023). MAE is defined as in Equation (4)

$$\text{MAE}_k = \frac{1}{N} \sum_{i=1}^{N} \left| \mu^{(k,i)} - F^{(k,i)} \right|, \quad k = 1, ..., 30, \quad (4)$$

where k is the forecasting horizon, $N = 731$ is the number of predictions made, $\mu^{(k,i)}$ is the NAMs prediction at i^{th} prediction day for k^{th} forecasting horizon, and $F^{(k,i)}$ (shown also in Equation (2g)) are the values that we evaluate $\mu^{(k,i)}$ on at i^{th} prediction day for k^{th} forecasting horizon. These evaluation values are typically chosen to be the IERS EOP series, which include the CPO time series. Further details about this series are given in "Data description".

Data description

The input data to our algorithm are the CPO time series. However, it should be mentioned various institutions/services such as Jet Propulsion Laboratory (JPL) and IERS provide unique CPO solutions, leaving us with several possible choices. Our previous studies under operational settings (Soja et al. 2023) have shown that for the prediction of CPO it is beneficial to use the JPL solution (Chin et al. 2009; Ratcliff and Gross 2022) as the input since it results in the highest prediction performance. We therefore use JPL CPO data, namely the EOP2 series, to train our algorithm. The JPL EOP2 series is an inter-technique solution, where the observations of various space-geodetic techniques are combined to generate the EOP series using a Kalman filtering approach (Ratcliff and Gross 2022). The CPO in JPL EOP2 series start from January 1, 1998, placing a lower bound on the available training data. The evaluations are commonly

done against the official products of IERS (Kiani Shahvandi et al. 2022a; Gou et al. 2023; Kiani Shahvandi et al. 2023) and therefore, we follow this approach here as well. We use the IERS 20 C04 series, which is the most recent EOP series provided by IERS and consistent with the latest realization of the International Terrestrial Reference Frame (ITRF2020; Altamimi et al. 2023). This is further justified considering that IERS 20 C04 is proven to be a more consistent EOP series (Kiani Shahvandi et al. 2023) than its predecessor IERS 14 C04 (Bizouard et al. 2019), implying that the rapid data provided by IERS agree better with IERS 20 C04 compared to IERS 14 C04. It should be mentioned that JPL and IERS 20 C04 CPO data are based on versions 2 and 3 of the International Celestial Reference Frame (ICRF2 and ICRF3, respectively; Fey et al. 2015; Charlot et al. 2020). However, our proposed algorithm can model any potential systematic effect stemming from the difference in the ICRF versions.

Using the mentioned data, we have created an operational framework where the predictions of CPO are generated daily since May 20, 2021. We analyze the predictions in the range May 20, 2021 to May 20, 2023, i.e., 2 years (731 days) in total. The choice of this prediction interval is restricted by the fact that JPL EOP2 series are updated daily (i.e., overwritten) and only since May 20, 2021 have we archived these files. At each prediction epoch, we retrain our algorithm from January 1, 1998 up to the prediction day to take advantage of the most recently available CPO data. An alternative approach is to train the model only once and then predict in time. Even though this approach is slightly faster, it results in less accurate predictions (as much as 20%), because the recent variations may not be captured by the algorithm.

In Fig. 1 we have shown the JPL EOP2 and IERS 20 C04 series in the range January 1, 1998 to May 20, 2023, in the unit of microarcseconds (μas). The mean and standard deviation of differences between JPL and IERS 20 C04 series are approximately 24 μas and 97 μas for dX, and -12 μas and 100 μas for dY, respectively. However, as mentioned before, our algorithm can cover these differences to a certain extent because in the training the JPL series is mapped to the IERS 20 C04 series. In Fig. 2 we have summarized the steps in applying the methodology designed in this paper.

Results and discussion

In Fig. 3a we have shown the prediction accuracy of NAMs for both dX and dY. The dX predictions are more accurate than those of dY: the average accuracy of dX across all days is 65 μas, whilst 93 μas for dY. The MAE values for dX do not show any particular trend, whereas the values for dY show an increasing trend after day 10 and then reaching a plateau at around day 20 onward.
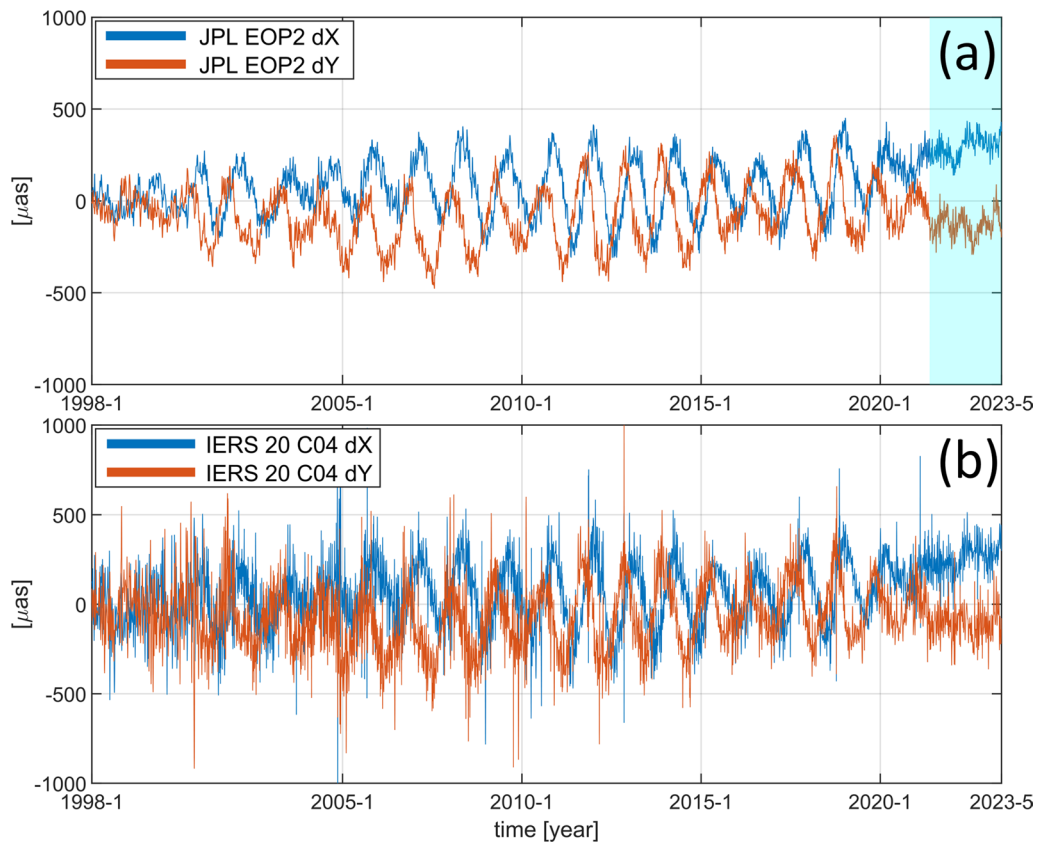
Fig. 1 a JPL EOP2 series from January 1, 1998 to May 20, 2023, used for training NAMs. The shaded blue vertical area shows the evaluation period, where the predictions made in this interval are compared with the corresponding values in IERS 20 C04 time series shown in **b**. The units are μas
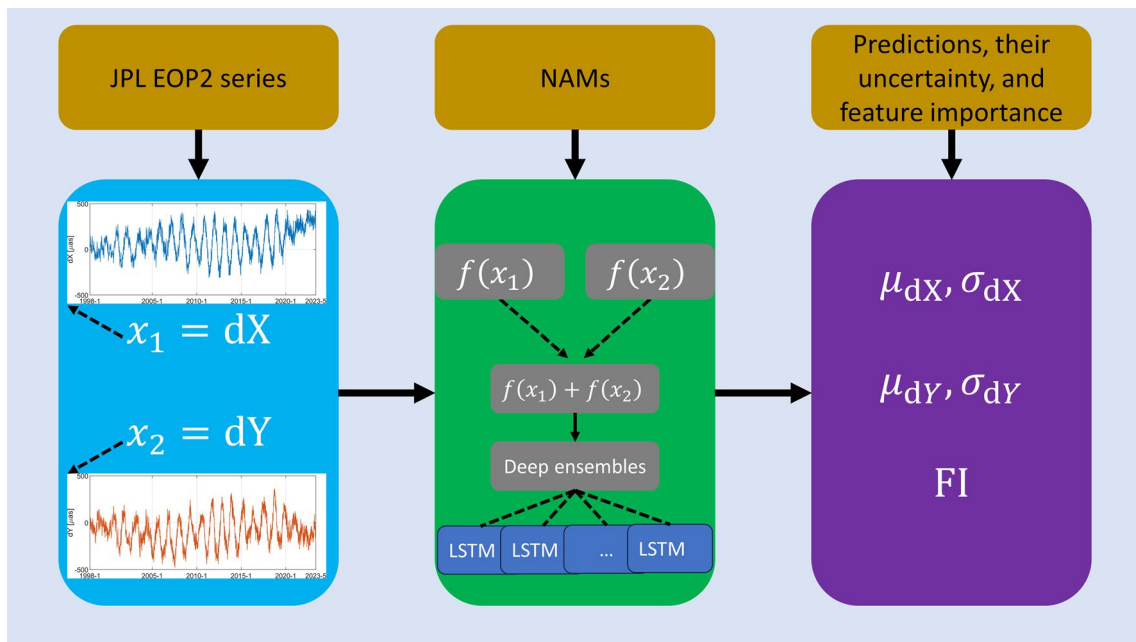

Fig. 2 The steps to apply the designed methodology in this paper. First, the JPL CPO series are inserted to the NAMs algorithm. Then, the methodology predicts the next values of the time series (μ) and assigns a prediction uncertainty (σ) to each value. The importance of features are also analyzed
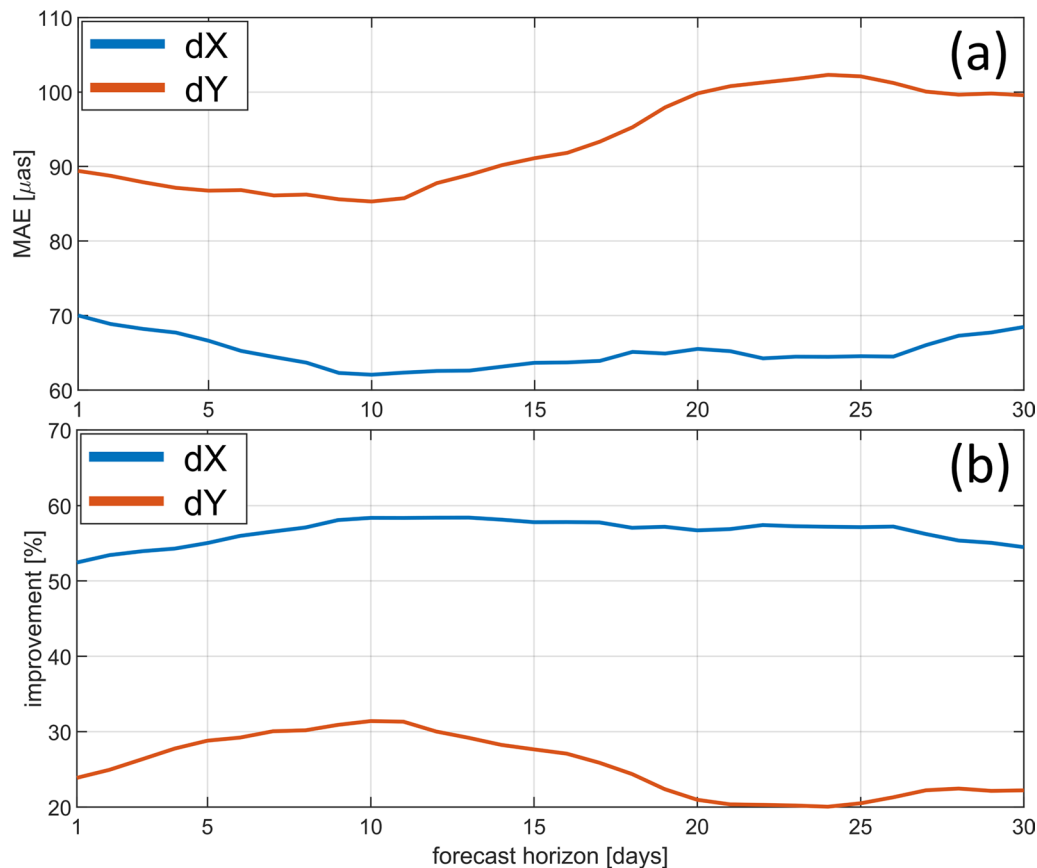
Fig. 3 a Prediction accuracy of NAMs in terms of μas for the dX and dY components separately. **b** Improvements achieved with respect to the rapid data provided by IERS

This has been observed in other studies as well (Belda et al. 2018; Nastula et al. 2020) and implies that the long-term predictions of CPO are not accurate and CPO components are predictable mainly in short-term horizons. This can be due to the nature of the CPO, where both the amplitude and period of FCN are variable in time (Cui et al. 2018), making it exceedingly difficult to predict these irregular variations long ahead.

We also show in Fig. 3b the improvement (e.g. Kiani Shahvandi et al. 2023) of our results with respect to the predictions provided by IERS (i.e., the rapid data). The improvements for dX are much larger than those of dY across all forecasting horizons. The average improvements for dX and dY are 57% and 25%, respectively.

To analyze the NAMs predictions more thoroughly, we present the prediction errors for all the 30 forecasting horizons and the 2-year-long evaluation interval, Fig. 4. These errors are computed by subtracting the NAMs predictions from the corresponding final IERS 20 C04 values. Apart from certain structures especially in dX, the errors seem to be random especially in dY: the strips that are visible are the errors in the input data at a certain epoch that persist almost for a month throughout the predictions. This is due to the fact that the input sequence length is 30 and if there is an error in the input series it usually persists until the anomalous value is outside of the input sequence (which has the length 30). Excluding these values, we can observe that most of these difference values are small. By extension, this implies that we have been able to adequately capture the main features of CPO time series.

In Fig. 5 we visualize the NAMs predictions, together with the final IERS 20 C04 and rapid data for an arbitrary date (November 20, 2022), to gain a better understanding of the NAMs performance. From this figure it can be understood that the predictions made by NAMs follow the final values more accurately compared to rapid data, especially in dX. One important observation is that the rapid dX data seem to contain a bias with respect to the final data and that this bias persists throughout most of the predictions (Kiani Shahvandi et al. 2024). This is in fact a major source of error of rapid data provided by IERS. The average bias for dX component over all the 731 predictions range between 116 μas and 123 μas for
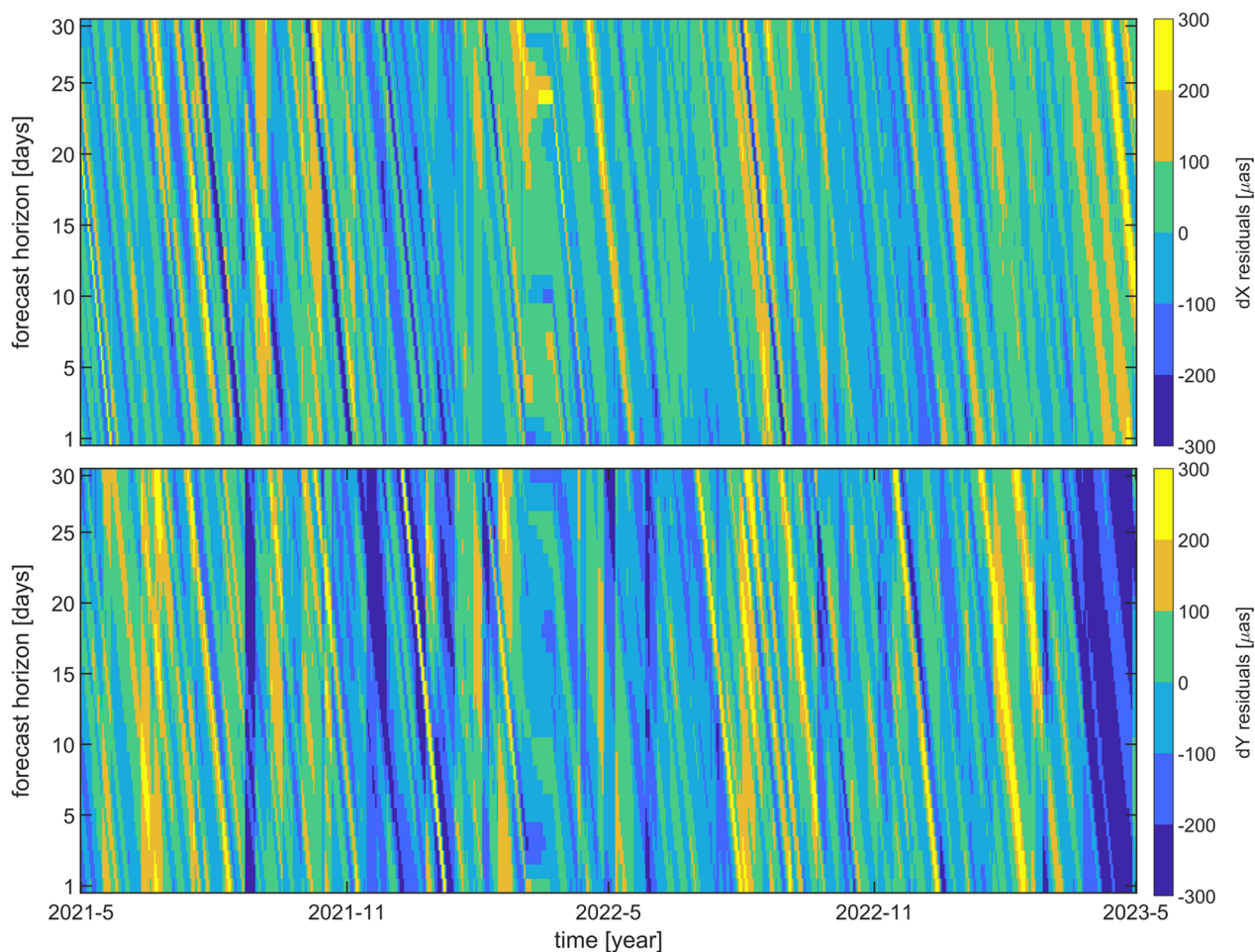
Fig. 4 Prediction errors defined as the difference between NAMs prediction and final CPO values in IERS 20 C04 series. The upper and lower panels are for dX and dY, respectively. The unit is μas

the first and last prediction horizons. The corresponding values for dY component are -36 μas and -51 μas. With NAMs, we have been able to adequately capture this bias and thus, significantly improve the prediction performance.

We show the overall feature importance analysis in Fig. 6. This figure shows the importance of both dX and dY on their own prediction, as well as the mutual influence of dX for the prediction of dY and vice versa. The uncertainties of these values are also shown. From this figure it can be understood that dX predictions are influenced by previous values of dX by 73%±15% and previous values of dY by 27%±9%. Similarly dY predictions are influenced by previous values of dX by 18%±9% and previous values of dY by 82%±17%. As expected, for the prediction of dX (dY) the most important features are dX (dY), whilst a certain contribution comes from dY (dX). This has also been observed in the case of polar motion (Kiani Shahvandi et al. 2022a). As mentioned in

"Methods", it is also possible to track the variation of feature importance in time and for different prediction horizons. In Fig. 7 we have shown these variations. It can be understood that the pattern of importance of dX for the prediction of dY and vice versa are similar to each other (although the importance of dX for the prediction of dY is smaller in amplitude), and that is also reflected in the overall feature importance shown in Fig. 6. Furthermore, it is noteworthy that the importance of dX for the prediction of dX itself ($FI_{dX,dX}$) exhibits rapid variations for different forecasting horizons over the prediction interval. In contrast, the importance of dY for the prediction of dY itself ($FI_{dY,dY}$) is comparatively stable. This can be explained by the fact that the IERS 20 C04 observations of dX in the prediction interval show a more anomalous behaviour compared to those of dY (that is dX shows a strong upward trend in this interval, whereas although seen also in dY, it is less severe).
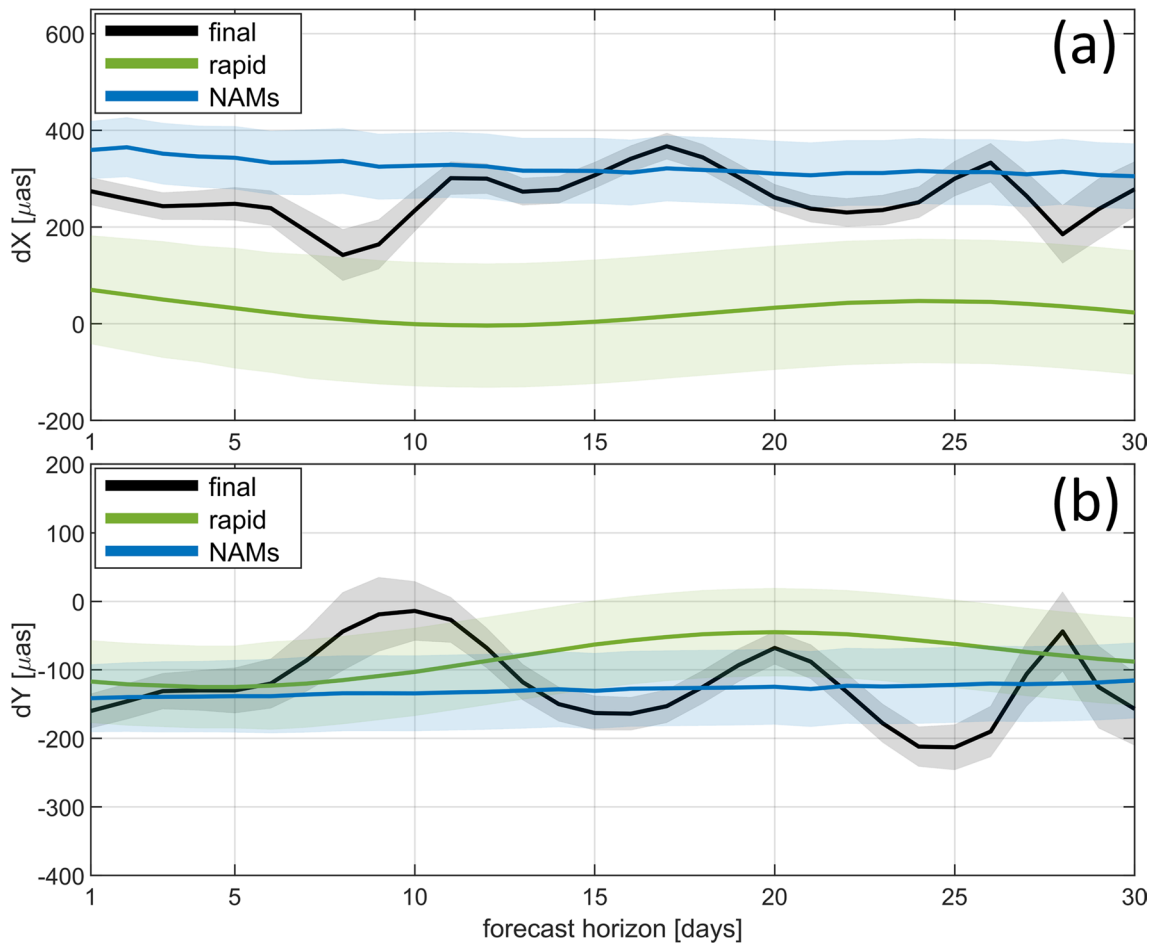
Fig. 5 NAMs predictions, rapid, and final IERS 20 C04 data, for November 20, 2022. dX values are shown in **a**, whilst dY in **b**. The shaded envelopes show the uncertainties
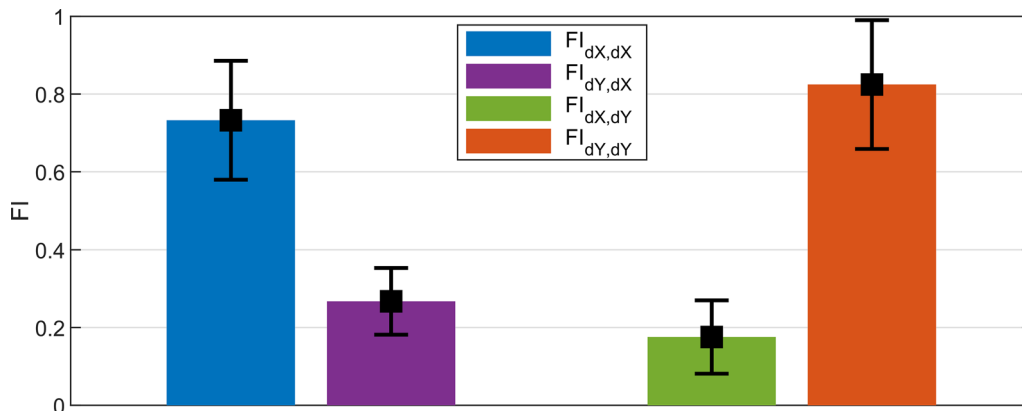

Fig. 6 Overall feature importance for the NAMs predictions, together with the uncertainty assigned to the values (in the form of error bars at the level of one standard deviation) computed using Equation (3). $FI_{p,q}$, p, q = dX, dY represent the importance of feature p for the prediction of feature q. The bars display the uncertainties in the values of feature importance
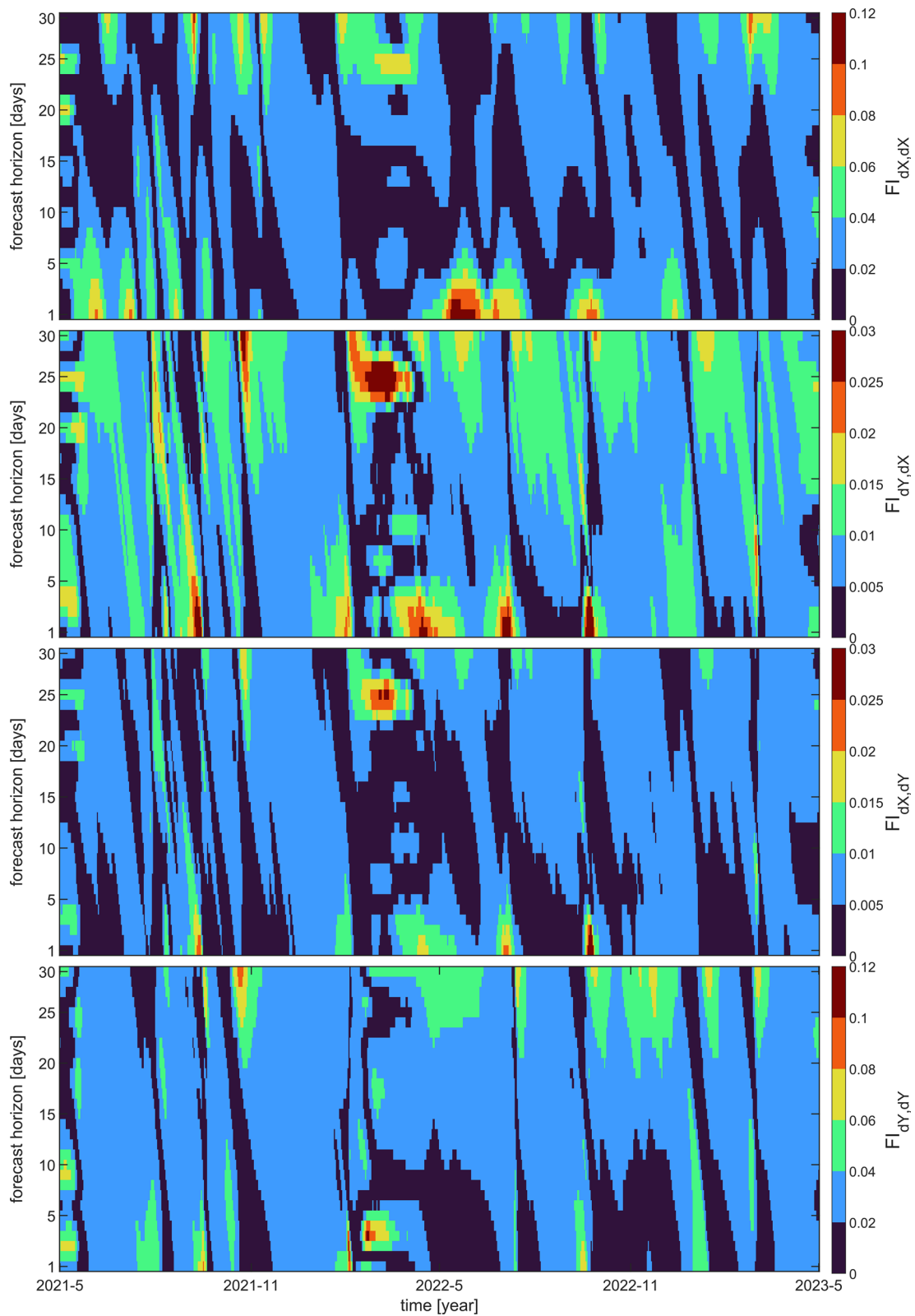
Fig. 7 Temporal variations of feature importance for the NAMs predictions. $FI_{p,q}$, $p, q = dX, dY$ represent the importance of feature p for the prediction of feature q

It should be mentioned that some other institutions also provide predictions of CPO, including SYstèmes de Référence Temps-Espace (SYRTE). To compare the prediction accuracy of NAMs with that of predictions provided by SYRTE, we first compute the MAE of SYRTE predictions with respect to final IERS 20 C04 and subsequently compute the improvement of our predictions with respect to those of SYRTE. In other words, we compare the MAE of NAMs with respect to IERS 20 C04 with the MAE of SYRTE with respect to IERS 20 C04. We show these improvements in Fig. 8. Here, we observe that we can generally improve the dY component more compared to the dX component (in contrast to comparison with respect to IERS 20 C04 in Fig. 3). The average improvements for dX and dY are 15% and 30%, respectively.

Finally, it is also possible to add other features to our prediction algorithm. In specific, it is suggested that the so-called ansatz models might improve the prediction performance (Caro et al. 2022). For this purpose, we need to extend the methodology to more than two features, which is presented in Appendix B. A reasonable anstaz can be the harmonic model suggested by IERS (Petit and Luzum 2010). This is based on the fact that CPO time series is dominated by FCN and fitting a harmonic function with FCN frequency captures the major variations of CPO. Our analyses (Appendix B) show that even though this approach is slightly beneficial for the prediction of dY component, it degrades the prediction performance of dX. Therefore, overall we suggest to rather use only the two features dX and dY.

Conclusions

We have focused on the short-term prediction of CPO, which is an important problem for various applications in space geodesy, such as satellite navigation. For this purpose, we have used NAMs, which belong to the category of interpretable neural networks. We have developed mathematical formulas for the estimation of prediction uncertainty and analysis of feature importance, improving the intelligibility of machine learning predictions. We have applied our algorithm to the JPL EOP2 series and evaluated the predictions against the IERS 20 C04 series. By comparing the NAMs prediction accuracy with the rapid data provided by IERS we significantly improve the CPO prediction accuracy, by 57% and 25% for dX and dY components, respectively. In addition, our predictions are more accurate than those provided by SYRTE by as much as 15% and 30% for dX and dY, respectively.

Since prediction of CPO is an important task, the EOP prediction community organized the second EOP-PCC (Śliwińska et al. 2023), in which the prediction performance of various methods was analyzed. Since we achieved the highest prediction performance amongst all the methods for the prediction of CPO (Wińska et al. 2023), we were motivated to improve our methodology and also provide the CPO prediction on a regular basis in an operational setting. Using NAMs, we have achieved this goal. Therefore, we encourage other researchers in the field to use interpretable machine learning as a method for highly accurate and interpretable prediction of EOPs. A suitable research direction is to include the geophysical data and constraints into the algorithm, as suggested by Kiani Shahvandi et al. (2024).
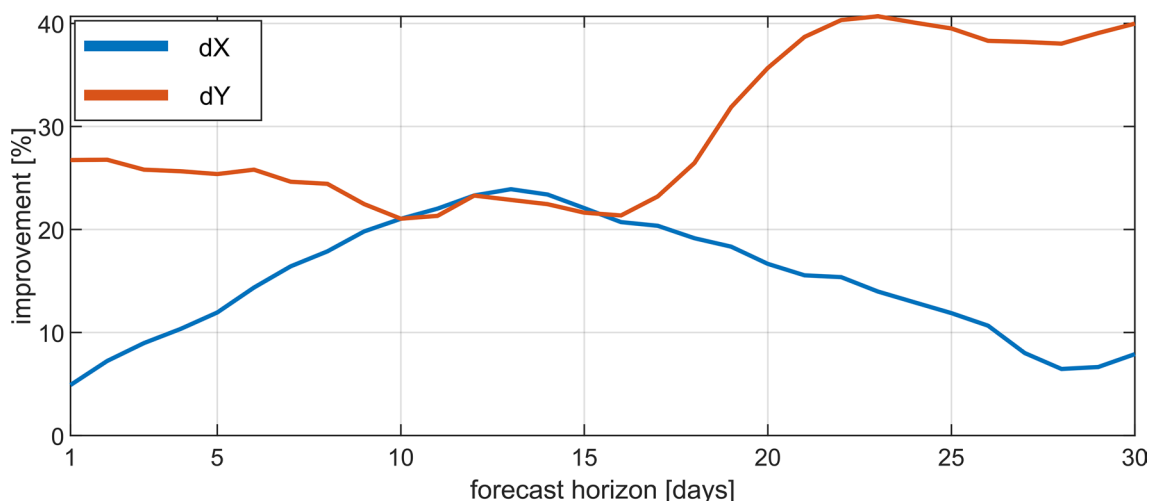

Fig. 8 Improvement of NAMs predictions with respect to the predictions provided by SYRTE. Both NAMs and SYRTE predictions are evaluated against the IERS 20 C04

Appendix A. Derivation of the relation of feature importance

The proof of Eq. 3 is given below. The importance in NAMs is defined as the absolute value of correlation coefficient ρ between y and $f(x_1)$, $f(x_2)$. Based on relation 1, the correlation coefficient is defined as in Equation (A.1a). Computation of $\text{cov}\big(y, f(x_k)\big)$ $k = 1, 2$ and σ_y based on Equation (A.1b)–(A.1c) results in Equation (A.1e), which is equivalent to Equation (3b). Note that the value of $\sigma_{f(x_k)}$ is similar to $\sigma_j(x_k)$ taken from Equation (2c)–(2d). Also, the final result presented in Equation (3b) is based on the ensemble mean and standard deviation of individual FI_{j,x_k}.

$$\rho_{y,f(x_k)} = \frac{\text{cov}\big(y, f(x_k)\big)}{\sigma_y \sigma_{f(x_k)}}, \qquad k = 1, 2 \qquad \text{(A.1a)}$$

$$\begin{aligned} \text{cov}(y, f(x_k)) &= \text{cov}(f(x_1) + f(x_2), f(x_k)) \\ &= \sigma_j^2(x_k) + \text{cov}\big(\mu_j(x_1), \mu_j(x_2)\big) \end{aligned}$$
(A.1b)

$$\begin{aligned} \sigma_y &= \sqrt{\text{cov}(f(x_1) + f(x_2), f(x_1) + f(x_2))} \\ &= \sqrt{\sigma_j^2(x_1) + \sigma_j^2(x_2) + 2\text{cov}(\mu_j(x_1), \mu_j(x_2))} \end{aligned}$$
(A.1c)

$$\sigma_{f(x_k)} = \sigma_j(x_k) \qquad \text{(A.1d)}$$

$$\text{FI}_{j,x_k} = \left| \rho_{y,f(x_k)} \right| = \left| \frac{\sigma_j^2(x_k) + \text{cov}\big(\mu_j(x_1), \mu_j(x_2)\big)}{\sigma_j(x_k)\sqrt{\sigma_j^2(x_1) + \sigma_j^2(x_2) + 2\text{cov}\big(\mu_j(x_1), \mu_j(x_2)\big)}} \right| \qquad \text{(A.1e)}$$

Since we defined the feature importance in terms of the absolute value of correlation coefficient, we have $0 \leq \text{FI}_{j,x_k} \leq 1$.

Appendix B. Extension of methodology to more than two features

Here we extend the methodology presented in the main text to an arbitrary number of features. Assuming s features $x_1, ..., x_s$ (where $x_i \in \mathbb{R}^n$, $i = 1, ..., s$ and n is the length of input and output sequences), then the equivalent definition of NAMs in Equation (1) is given as in Equation (B.1).

$$y = \sum_{k=1}^{s} f(x_k) \qquad \text{(B.1)}$$

The extension of Equation (2) can be simply given as the following relations in Equation (B.2).

$$\mu_j(x_k) = \text{LSTM}_{\mu_{x_k}}(W_{\mu_{x_k}, j}, x_k), \qquad k = 1, ..., s \qquad \text{(B.2a)}$$

$$\sigma_j^2(x_k) = \log(1 + \exp(\text{LSTM}_{\sigma_{x_k}}(W_{\sigma_{x_k}, j}, x_k))) + \varepsilon \qquad \text{(B.2b)}$$

$$\mu_j = \sum_{k=1}^{s} \mu_j(x_k) \qquad \text{(B.2c)}$$

$$\sigma_j^2 = \sum_{k=1}^{s} \sigma_j^2(x_k) + 2 \sum_{p=1}^{s} \sum_{q>p}^{s} \text{cov}\big(\mu_j(x_p), \mu_j(x_q)\big) \qquad \text{(B.2d)}$$

$$\ell_j = \frac{1}{2}\log\sigma_j^2 + \frac{1}{2}\frac{(F - \mu_j)^2}{\sigma_j^2} \quad \rightarrow \text{minimize} \qquad \text{(B.2e)}$$

$$\mu = \frac{1}{M}\sum_{j=1}^{M} \mu_j \qquad \text{(B.2f)}$$

$$\sigma^2 = -\mu^2 + \frac{1}{M}\sum_{j=1}^{M}\left[\sigma_j^2 + \mu_j^2\right] \qquad \text{(B.2g)}$$

To test whether adding other features would help in improving the prediction performance, we perform the same analysis in the main text but with the difference that in addition to dX and dY as the input features, we also add the following two features: $\cos(\sigma_f t)$ and $\sin(\sigma_f t)$ where $\sigma_f = 0.014578$ is the approximate value of frequency of FCN and t is time. This is based on the model recommended in IERS conventions (Petit and Luzum 2010) for the variations of CPO.

Using the mentioned four features, we train our model and compare the predictions with the observations. The MAE of these predictions is shown in Fig. 9. Comparing these results with those in Fig. 3, we observe that whilst the addition of the two features $\cos(\sigma_f t)$ and $\sin(\sigma_f t)$ has slightly improved the prediction performance of dY, that of dX has been significantly reduced. We conclude that the addition of more features does not necessarily improve the prediction performance of NAMs. One fundamental reason is that the higher the number of input features, the larger the number of model parameters and therefore, the more difficult to properly train the model. In fact, this is one
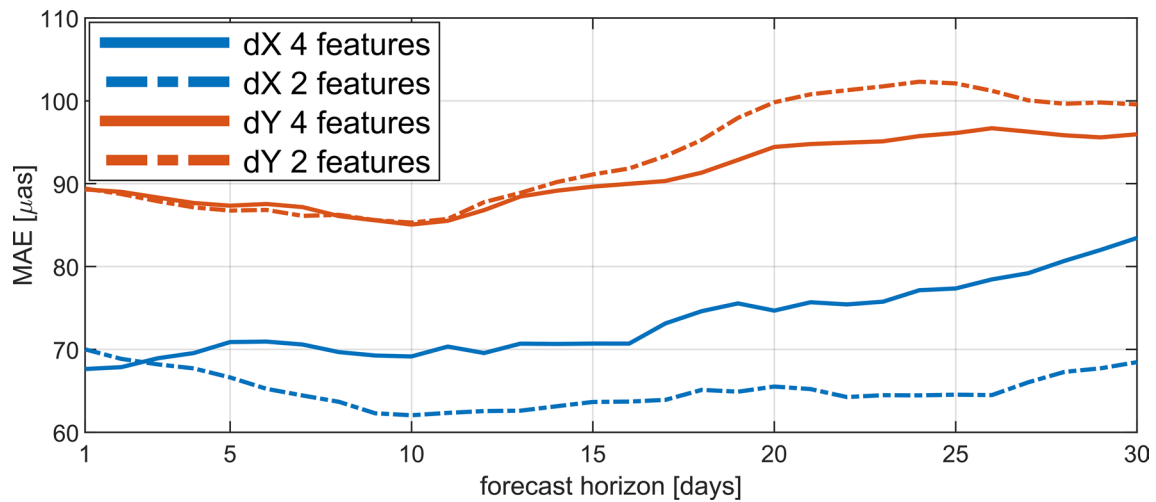
Fig. 9 Prediction accuracy of NAMs in terms of µas for the dX and dY components separately. Two cases are presented: using two input features (dX and dY); four input features (dX, dY, $\cos(\sigma_f t)$, and $\sin(\sigma_f t)$). The former are equivalent to the results presented in Fig. 3

Table 1 The models used in NAMs and the number of parameters of each model based on the number of features (*s*), hidden neurons (*H*), and input and output sequence length *n*

Model	Number of parameters
$\mu_j(x_k)$	$4H^2 + 5nH + 4H + n$
$\sigma_j^2(x_k)$	$4H^2 + 5nH + 4H + n$
μ_j	$s(4H^2 + 5nH + 4H + n)$
σ_j^2	$s(4H^2 + 5nH + 4H + n)$
NAMs	$2s(4H^2 + 5nH + 4H + n)$

of the disadvantages of NAMs. Finally, it should be mentioned that even by trying to adapt the value of σ_f we still could not improve the prediction performance. Another possibility is to use the temporal variation of phase and amplitude of FCN as the input features (c.f. Belda et al. 2016, 2017). Using these as additional features, we can improve the prediction accuracy of both dX and dY, maximum by as much as 8%. However, these improvements are small and the computational cost of adding additional features outweighs the benefit of improving the prediction accuracy.

Finally, we show the architecture of NAMs for *s* number of input features, with input sequence length *n*, and number of hidden neurons *H*. A short description of mathematical formulas of LSTM is given in Gou et al. (2023). Based on these formulas, we can present Table 1, where the models and the number of parameters of them are represented. It can be understood that for two features and the model architecture used in this study ($n = 30$, $H = 10$), NAMs have 7880 parameters. The

similar number for four features is 15760. Even though machine learning algorithms can be highly parameterized, the large number of parameters used in NAMs imply that for a proper training, large number of data samples are required. We therefore expect that over time, the prediction accuracy of NAMs would increase, since the length of CPO time series increases and we would have more data for training.

Abbreviations

CPO	Celestial pole offsets
EOPs	Earth orientation parameters
IERS	International earth rotation and reference systems service
JPL	Jet propulsion laboratory
MAE	Mean absolute error
µas	Microarcseconds
NAMs	Neural additive models
SYRTE	SYstèmes de Référence Temps-Espace
VLBI	Very long baseline interferometry

Author contributions
MKS conceived the research, implemented the method, analyzed the results, and wrote the original draft. SB and SM helped with methodology. All the authors read the manuscript and commented upon it.

Availability of data and materials
The datasets generated and/or analyzed during the current study are available in the GitHub repository https://github.com/mkianis/NAMs-for-CPO-prediction.git, as well as updated daily at ETH Zurich geodetic prediction center (Soja et al. 2022) website https://gpc.ethz.ch/products/EOP/

eop-nams-30d-jpl-eth/ and https://gpc.ethz.ch/EOP/report/eop-nams-30d-jpl-eth/index.html. JPL long EOP2 series can be found on https://eop2-external.jpl.nasa.gov/. IERS rapid and final EOP series can be found on https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html. SYRTE data are available on https://hpiers.obspm.fr/eop-pc/index.php.

Declarations

Competing interests
The authors declare that they do not have competing interests.

Author details
[1]Institute of Geodesy and Photogrammetry, ETH Zürich, Zürich, Switzerland.
[2]Department of Applied Mathematics, University of Alicante, Alicante, Spain.
[3]Seminar for Applied Mathematics, Department of Mathematics, ETH AI Center, ETH Zürich, Zürich, Switzerland.

References

Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow I, Harp A, Irving G, Isard M, Jozefowicz R, Jia Y, Kaiser L, Kudlur M, Levenberg J, Mané D, Schuster M, Monga R, Moore S, Murray D, Olah C, Shlens J, Steiner B, Sutskever I, Talwar K, Tucker P, Vanhoucke V, Vasudevan V, Viégas F, Vinyals O, Warden P, Wattenberg M, Wicke M, Yu Y, Zheng X (2015) TensorFlow large-scale machine learning on heterogeneous systems. ArXiv 1603:00447

Agarwal R, Melnick L, Frosst N, Zhang X, Lengerich B, Caruana R, Hinton G (2021) Neural additive models: interpretable machine learning with neural nets. Neural Inform Process Syst 2021(34):4699

Altamimi Z, Rebischung P, Collilieux X, Métivier L, Chanard K (2023) ITRF2020: an augmented reference frame refining the modeling of nonlinear station motions. J Geodesy 97:47

Belda S, Ferrándiz JM, Heinkelmann R, Nilsson T, Schuh H (2016) Testing a new free core nutation empirical model. J Geodynam 94:59–67

Belda S, Heinkelmann R, Ferrándiz JM, Karbon M, Nilsson T, Schuh H (2017) An improved empirical harmonic model of the celestial intermediate pole offsets from a global VLBI solution. The Astronomical J 154:154–166

Belda S, Ferrándiz JM, Heinkelmann R, Schuh H (2018) A new method to improve the prediction of the celestial pole offsets. Sci Rep 8(1):13861

Bizouard C, Lambert S, Gattano C, Becker O, Richard JY (2019) The IERS EOP 14C04 solution for earth orientation parameters consistent with ITRF 2014. J Geodesy 93:621–633

Brzeziński A (1994) Polar motion excitation by variations of the effective angular momentum function: II. extended model. Manuscripta Geodaetica 19:157–171

Caro MC, Huang HY, Cerezo M, Sharma K, Sornborger A, Cincio L, Coles PJ (2022) Generalization in quantum machine learning from few training data. Nature Commun 13(1):4919

Charlot P, Jacobs CS, Gordon D, Lambert S, de Witt A, Böhm J, Fey AL, Heinkelmann R, Skurikhina E, Titov O, Arias EF, Bolotin S, Bourda G, Ma C, Malkin Z, Nothnagel A, Mayer D, MacMillan DS, Nilsson T, Gaume R (2020) The third realization of the international celestial reference frame by very long baseline interferometry. Astronom Astrophys 644:159

Chin TM, Gross RS, Boggs DH, Ratcliff JT (2009) Dynamical and observation models in the Kalman earth orientation filter. The Interpl Network Progress Rep 42:1–25

Cui X, Sun H, Xu J, Zhu J, Chen X (2018) Detection of free core nutation resonance variation in earth tide from global superconducting gravimeter observations. Earth Planets Space 70:199

Fey, AL, Gordon D, Jacobs CS, Ma C, Gaume RA, Arias EF, Bianco G, Boboltz DA, Böckmann S, Bolotin S, Charlot P, Collioud A, Engelhardt G, Gipson J, Gontier AM, Heinkelmann R, Kurdubov S, Lambert S, Lytvyn S, MacMillan DS, Malkin Z, Nothnagel A, Ojha R, Skurikhina E, Sokolova J, Souchay J, Sovers OJ, Tesmer V, Titov O, Wang G, Zharov V, (2015) The second realization of the international celestial reference frame by very long baseline interferometry. Astronom J 150(2):58

Gou J, Kiani Shahvandi M, Hohensinn R, Soja B (2023) Ultra-short-term prediction of LOD using LSTM neural networks. J Geodesy 97(52):52

Gross RS (2015) Earth rotation variations - long period. In: Schubert G (ed) Treatise on Geophysics. Elsevier, Amsterdam

Hastie T, Tibshirani R (1986) Generalized additive models. Statis Sci 1(3):297–310

Herring TA, Buffet BA, Matthews PM, Shapiro II (1991) Forced nutations of the Earth: influence of inner core dynamics 2. 3. Very long interferometry data analysis. J Geophys Res 96:8259–8273

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780

Kiani Shahvandi M, Soja B (2021). Modified deep transformers for GNSS time series prediction. IGARSS 2021 - 2021 IEEE International Geoscience and Remote Sensing Symposium: 8313-8316

Kiani Shahvandi M, Soja B (2022a) Small geodetic datasets and deep networks: attention-based residual LSTM autoencoder stacking for geodetic time series. Int Conf Machine Learning Optimization Data Sci 2:296–307

Kiani Shahvandi M, Soja B (2022b) Inclusion of data uncertainty in machine learning and its application in geodetic data science, with case studies for the prediction of Earth orientation parameters and GNSS station coordinate time series. Adv Space Res 70(3):563–575

Kiani Shahvandi M, Schartner M, Soja B (2022a) Neural ODE differential learning and its application in polar motion prediction. J Geophys Res Solid Earth 127(11):e2022JB024775

Kiani Shahvandi M, Gou J, Schartner M, Soja B (2022b). Data driven approaches for the prediction of Earth's effective angular momentum functions. IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium: 6550-6553

Kiani Shahvandi M, Dill R, Dobslaw H, Kehm A, Bloßfeld M, Schartner M, Mishra S, Soja B (2023) Geophysically informed machine learning for improving rapid estimation and short-term prediction of Earth orientation parameters. J Geophys Res Solid Earth 128(10):e2023JB026720

Kiani Shahvandi M, Belda S, Karbon M, Mishra S, Soja B (2024) Deep ensemble geophysics-informed neural networks for the prediction of celestial pole offsets. Geophys J Int 236(1):480–493

Kingma DP, Ba J (2015) Adam: a method for stochastic optimization. Int Conf Learn Represent 1412:6980

Lakshminarayanan B, Pritzel A, Blundell C (2016). Simple and scalable predictive uncertainty estimation using deep ensembles. Adv Neural Inform Process Syst (NeurIPS 2016)

Lara-Benitez P, Carranza-García M, Riquelme JC (2021) An experimental review on deep learning architectures for time series forecasting. Int J Neural Syst 31(3):2130001

Lim B, Zohren S (2021) Time-series forecasting with deep learning: a survey. Philosophical Trans Royal Soc A 379(2194):20200209

Matthews PM, Buffet BA, Herring TA, Shapiro II (1991) Forced nutations of the earth: influence of inner core dynamics 1. theory. J Geophys Res Solid Earth 96:8219–8242

Matthews PM, Buffet BA, Herring TA, Shapiro II (1991) Forced nutations of the earth: influence of inner core dynamics 2. numerical results and comparisons. J Geophys Res Solid Earth 96:8243–8257

Matthews PM, Herring TA, Buffet BA (2002) Modeling of nutation and precession: new nutation series for nonrigid earth and insights into the Earth's interior. J Geophys Res Solid Earth 107:ETG-3-1-ETG−3-26

Molnar C (2023) Interpretable machine learning: a guide for making black box models explainable(2nd ed.). christophm.github.io/interpretable-ml-book/

Nastula J, Chin TM, Gross R, Śliwińska J, Wińska M (2020) Smoothing and predicting celestial pole offsets using a Kalman filter and smoother. J Geodesy 94(1):17

Petit G, Luzum B (2010) IERS Conventions. Frankfurt am Main: Verlag des Bundesamts für Kartographie und Geodäsie

Ratcliff J, Gross RS (2022) Combinations of earth orientation measurements: SPACE2021, COMB2021, and POLE2021. JPL Publications, Pasadena

Sasao T, Wahr JM (1981) An excitation mechanism for the free 'core nutation'. Geophys J Int 64(3):729–746

Śliwińska J, Kur T, Wińska M, Nastula J, Dobslaw H, Partyka A (2023) Second earth orientation parameters prediction comparison campaign (2nd EOP PCC): overview. Artif Satellites 57:237–253

Shirai T, Fukushima T, Malkin Z (2005) Detection of phase disturbances of free core nutation of the Earth and their concurrence with geomagnetic jerks. Earth Planets Space 57:151–155

Soja B, Kiani Shahvandi M, Schartner M, Gou J, Kłopotek G, Crocetti L, Awadaljeed M (2022) The new geodetic prediction center at ETH Zurich. EGU General Assembly 2022. https://doi.org/10.5194/egusphere-egu22-9285

Soja B, Kiani Shahvandi M, Schartner M, Gou J (2023) Comparison of machine-learning-based predictions of Earth orientation parameters using different input data. Second Earth Orientation Parameters Prediction Comparison Campaign (2nd EOP PCC). https://doi.org/10.3929/ethz-b-000603468

Sovers OJ, Fanselow JL, Jacobs CS (1998) Astrometry and geodesy with radio interferometry: experiments, models, results. Rev Modern Phys 70:1393

Wińska M, Śliwińska J, Kur T, Nastula J, Dobslaw H, Partyka A (2023) Assessment of precession-nutation predictions based on the results of the Second Earth Orientation Parameters Prediction Comparison Campaign (2nd EOP PCC). https://doi.org/10.13140/RG.2.2.30474.85447

Wahr JM (1981) The forced nutations of an elliptical, rotating, elastic and oceanless Earth. Geophys J Int 64(3):705–727

Wahr JM (1988) The Earth's rotation. Ann Rev Earth Planetary Sci 16:231–249

Zheng H, Yang Z, Liu W, Liang J, Li Y (2015) Improving deep neural networks using softplus units. 2015 International Joint Conference on Neural Networks (IJCNN). https://10.1109/IJCNN.2015.7280459

Publisher's Note