ResLearner: geophysically-informed machine learning for improving the accuracy of rapid Earth orientation parameters

Mostafa Kiani Shahvandi[1], Robert Dill[2], Henryk Dobslaw[3], Alexander Kehm[4], Mathis Bloßfeld[4], Matthias Schartner[1], Siddhartha Mishra[5], and Benedikt Soja[6]

[1]Institute of geodesy and photogrammetry, ETH Zurich
[2]Deutsches GeoForschungsZentrum GFZ
[3]GeoForschungsZentrum Potsdam
[4]Deutsches Geodätisches Forschungsinstitut (DGFI-TUM), Technical University of Munich
[5]ETH Zurich
[6]Institute of Geodesy and Photogrammetry, ETH Zurich

May 4, 2023

Abstract

Rapid provision of Earth Orientation Parameters (EOPs, here polar motion and dUT1) is indispensable in many geodetic applications and also for spacecraft navigation. There are, however, discrepancies between the rapid EOPs and the final EOPs that have a higher latency, but the highest accuracy. To reduce these discrepancies, we focus on a data-driven approach, present a novel method named ResLearner, and use it in the context of deep ensemble learning. Furthermore, we introduce a geophysically-constrained approach for ResLearner. We show that the most important geophysical information to improve the rapid EOPs is the effective angular momentum functions of atmosphere, ocean, land hydrology, and sea level. In addition, semi-diurnal, diurnal, and long-period tides coupled with prograde and retrograde tidal excitations are important features. The influence of some climatic indices on the prediction accuracy of dUT1 is discussed and El Ni\~{n}o Southern Oscillation is found to be influential. We developed an operational framework, providing the improved EOPs on a daily basis with a prediction window of 63 days to fully cover the latency of final EOPs. We show that under the operational conditions and using the rapid EOPs of the International Earth Rotation and Reference Systems Service (IERS) we achieve improvements as high as 60\%, thus significantly reducing the differences between rapid and final EOPs. Furthermore, we discuss how the new final series IERS 20 C04 is preferred over 14 C04. Finally, we compare against EOP hindcast experiments of European Space Agency, on which ResLearner presents comparable improvements.

ResLearner: geophysically-informed machine learning for improving the accuracy of rapid Earth orientation parameters

Mostafa Kiani Shahvandi[1], Robert Dill[2], Henryk Dobslaw[2], Alexander Kehm[3], Mathis Bloßfeld[3], Matthias Schartner[1], Siddhartha Mishra[4], Benedikt Soja[1]

[1]Institute of Geodesy and Photogrammetry, ETH Zurich
[2]Section for Earth System Modelling, GFZ German Research Centre for Geosciences
[3]Deutsches Geodätisches Forschungsinstitut (DGFI-TUM), Technical University of Munich
[4]Seminar for Applied Mathematics, Department of Mathematics, and ETH AI Center, ETH Zurich

Key Points:

- We introduce a novel machine learning algorithm named ResLearner to improve the accuracy of rapid Earth orientation parameters
- We also present geophysically-constrained ResLearner, using Earth's effective angular momentum functions, tides, and climatic indices
- Besides prediction, ResLearner is also able to effectively correct deficits in rapidly processed EOPs with respect to final EOPs

Corresponding author: Mostafa Kiani Shahvandi, `mkiani@ethz.ch`

Abstract

Rapid provision of Earth Orientation Parameters (EOPs, here polar motion and dUT1) is indispensable in many geodetic applications and also for spacecraft navigation. There are, however, discrepancies between the rapid EOPs and the final EOPs that have a higher latency, but the highest accuracy. To reduce these discrepancies, we focus on a data-driven approach, present a novel method named ResLearner, and use it in the context of deep ensemble learning. Furthermore, we introduce a geophysically-constrained approach for ResLearner. We show that the most important geophysical information to improve the rapid EOPs is the effective angular momentum functions of atmosphere, ocean, land hydrology, and sea level. In addition, semi-diurnal, diurnal, and long-period tides coupled with prograde and retrograde tidal excitations are important features. The influence of some climatic indices on the prediction accuracy of dUT1 is discussed and El Niño Southern Oscillation is found to be influential. We developed an operational framework, providing the improved EOPs on a daily basis with a prediction window of 63 days to fully cover the latency of final EOPs. We show that under the operational conditions and using the rapid EOPs of the International Earth Rotation and Reference Systems Service (IERS) we achieve improvements as high as 60%, thus significantly reducing the differences between rapid and final EOPs. Furthermore, we discuss how the new final series IERS 20 C04 is preferred over 14 C04. Finally, we compare against EOP hindcast experiments of European Space Agency, on which ResLearner presents comparable improvements.

Plain Language Summary

The International Earth Rotation and Reference Systems Service (IERS) provides rapid Earth Orientation Parameters (EOPs) using different space geodetic techniques to bridge the latency of the final, most accurate EOPs solution. However, these rapid EOPs are not in full agreement with the final EOPs. In order to reduce the differences between the rapid and final EOPs, we focus on the application of machine learning and present a novel method named ResLearner, which is based on geodetic data and geophysical constraints. We present the method in the context of deep ensemble learning, focusing on a prediction window of 63 days. We also attempt to link informative geophysical effects to these discrepancies. We show that they are linked to a mixture of atmospheric, oceanic, hydrological, and sea level effective angular momentum functions, dominance of the GNSS-derived polar motion, and various short- and long-term tidal excitations. El Niño Southern Oscillation is also relevant for dUT1 prediction. The methodology can provide significant improvements of up to 60% in operational settings with respect to rapid EOPs provided by IERS. Additional validation is done by using the data of Jet Propulsion Laboratory final EOP series and also EOP series provided by the European Space Agency.

1 Introduction

Earth Orientation Parameters (EOPs) represent variations of Earth's rotation axis in time (Lambeck, 1980; Gross, 1997). Among these parameters, polar motion components, (xp, yp), and the difference between universal time and coordinated universal time, dUT1, are of great interest, because of their importance for applications such as satellite and spacecraft navigation and orientation of deep-space telescopes (Dobslaw & Dill, 2019b). These EOPs are routinely provided at different latencies, of which two are considered here: rapid and final (Kehm et al., 2023). Final EOPs require a combination of different data sources (Bizouard et al., 2019; Ratcliff & Gross, 2022) such as Global Navigation Satellite Systems (GNSS), Very Long Baseline Interferometry (VLBI), and Lunar and Satellite Laser Ranging (LLR, SLR). Some of the techniques require longer processing time and therefore, delays of up to several weeks are expected, by which the data are provided to the scientific community. The current uncertainty level in final EOPs provided by International Earth Rotation and Reference Systems Service (IERS) is around 20-30 micro-arcseconds [µas] for polar motion components, and 9-10 micro-seconds [µs] for dUT1 in terms of formal errors.

Rapid EOPs provided by the IERS are determined through a combination of the most recent Global Positioning System (GPS) and VLBI 24-hour and intensive sessions data, augmented with Atmospheric Angular Momentum (AAM). These rapid data contain polar motion components (xp, yp) and dUT1, bridging the latency of final EOPs by providing 90 days of rapid combined EOPs to the past and 90 days of predicted EOPs into the future, with respect to the date the data are provided at. The uncertainty in the estimations is also provided. Currently, the level of these uncertainties varies across different days and also for combined and predicted EOPs. For the rapid combined EOPs, it can be several times bigger than that of final EOPs, but mostly below 1 milli-arcseconds [mas]. Predictions into the future are based on extrapolation of mathematical functions such as harmonic models. For longer prediction horizons, the accuracy is degraded significantly and can be up to several milli-arcseconds.

There are some routines performed on the mentioned datasets before operationally providing the rapid EOPs data. These include systematic corrections and smoothing. Systematic corrections are used to mitigate the impact of different VLBI baseline solutions on polar motion and dUT1. For instance, based on different VLBI solutions of the United States Naval Observatory (USNO), corrections are added to the polar motion and dUT1 of 24-hour sessions, and similar corrections to dUT1 of intensive sessions. Smoothing algorithms are applied to remove the high-frequency noise, usually by a Lagrangian interpolation scheme. It is important to note that ocean tidal effects are dealt with in the rapid EOPs as otherwise, the accuracy would be significantly degraded because of the systematic effect of tides. Furthermore, AAM data that are used for the improved determination of rapid EOPs contain some errors. Errors in the removal of tides and also the addition of AAM with its associated errors would result in inaccuracies in the rapid data, and therefore, inconsistencies w.r.t the final EOPs. These discrepancies can easily exceed the current uncertainty level of final polar motion and dUT1 mentioned above, thus suggesting the need for some type of calibration.

There are several deficiencies in the rapid data that are currently provided by the IERS. First, as mentioned the errors in the removal of tides can propagate to the rapid EOPs. Furthermore, only AAM is used, which is essentially one type of the Effective Angular Momentum (EAM) functions (Barnes et al., 1983). It is shown that Oceanic Angular Momentum (OAM), Hydrological Angular Momentum (HAM), and Sea Level Angular Momentum (SLAM) can have a non-negligible effect on polar motion and dUT1 as well (Dahlen, 1976; Nastula & Ponte, 1999; Brzezinski & Nastula, 2002; Chin et al., 2004; Gross, 2008; Dobslaw et al., 2010; Dill & Dobslaw, 2010; Bizouard & Seoane, 2010; Luo et al., 2022; Kiani-Shahvandi et al., 2022). Furthermore, phenomena such as El Niño Southern Oscillation (ENSO) can have some influence on the rate of dUT1 (Raut et al., 2022; Xu et al., 2022). This can be analyzed using climatic indices (CI) like the multivariate ENSO index (MEI, Wolter & Timlin, 1993), the Madden Julian Oscillation index (MJI, Kiladis et al., 2014), and the North Atlantic Oscillation index (NAI, Visbeck, Hurrell, Polvani, & Cullen, 2001). It is important to mention that the included AAM may not have fully covered the atmospheric effects and a calibration is also needed for this. In addition, the effect of EAM functions is non-tidal, but it can get mixed with the tidal effects during the application of routines. Disentangling the causes of discrepancies between rapid and final EOPs could be challenging and might require specifically-designed algorithms, especially in the absence of physical or analytical models for calibration. As the mixture of tidal and non-tidal effects, systematic corrections, and smoothing can be in a non-linear fashion, one needs to potentially use non-linear models for the purpose of disentanglement. Furthermore, the historical data of rapid EOPs can be utilized to present data-driven approaches that eliminate the need for an analytical calibration approach. These arguments imply that a machine learning algorithm is potentially well suitable for this problem, which is the approach followed in this paper.

There have been successful applications of machine learning for the analysis and prediction of EOPs (Dill et al., 2021; Kiani-Shahvandi & Soja, 2021, 2022; Kiani-Shahvandi et al., 2022). Here, however, we need to consider the specific aspects of the problem and develop a new machine learning algorithm. These specific aspects include 1) the calibration characteristic, 2) the need for non-linear uncertainty estimation, and 3) the importance analysis of different features included in the model.

The first aspect of the problem, namely the calibration characteristic, relates to the fact that the goal of the problem is to reduce the discrepancies between rapid and final EOPs, or in other words, calibration of rapid EOPs w.r.t final EOPs. This implies that the input to the machine learning model should contain the rapid EOPs themselves. These rapid EOPs are already close to the final EOPs in a sense, therefore making the problem similar to an identity mapping by machine learning. This can be difficult for non-linear machine learning algorithms (He et al., 2016), and it has been shown that a better approach would be to consider a residual learning framework (He et al., 2016). Inspired by this approach, we develop our new method in a residual learning manner, in which the overall output (final EOPs) is the summation of rapid EOPs and the output a neural network (having rapid EOPs and other geophysical information either as inputs or constraints). The mentioned neural network can then learn the calibration, enabling us also to use further geophysical information and constraints in the model. Note that self-calibration algorithms can also be considered (Minderer et al., 2021), in which the errors in different variables in the model are potentially reduced by trying to simultaneously learn the calibration effects.

The second aspect of the problem, i.e., uncertainty estimation, is an important task in the field of geodetic science (Kiani-Shahvandi & Soja, 2022), as these uncertainties provide a measure of the reliability of predictions. However, this can be challenging because of the potential non-linearity in neural networks. In this paper, deep ensembles (Lakshminarayanan et al., 2016; Ganaie et al., 2022) are used, which can reduce the epistemic uncertainty in the models. In deep ensembles, a series of neural networks are simultaneously trained to find the mean and standard deviation in the predictions. Since the output is the average of the predictions of all models, the epistemic uncertainty is reduced and mainly the aleatoric uncertainty remains (due to the uncertainty of input data).

Finally, it is important to use algorithms that support the importance analysis of different variables included in the model. Using this approach, we are able to analyze the potential sources of errors in the rapid EOPs.

The following points summarize the goals of the current paper:

- Developing a new machine learning algorithm specifically designed for the problem of improving rapid EOPs accuracy, which can also provide information on uncertainties in the predictions,
- Using geophysically-constrained neural networks as an additional approach in the context of the method,
- Analyzing the geophysical causes of discrepancies between rapid and final EOPs.

The rest of this paper is organized as follows. In Section 2, the ResLearner methodology is introduced. In Section 3, the data used for the numerical results presented in the paper are described. Section 4 is devoted to results and discussions. Conclusions are given in Section 5.

2 ResLearner methodology

This section describes the ResLearner method, including the general approach and its architecture.

2.1 Introducing ResLearner

As mentioned in Section 1, the idea of ResLearner is to calibrate the rapid EOPs (henceforward denoted by R) with respect to the final EOPs (denoted by F) in a residual manner using neural networks (NN). This implies that the conceptual representation of ResLearner can be described by Equation (1)

$$F = R + \mathrm{NN}(\theta, R, X) \tag{1}$$

in which NN is a neural network with parameters θ, and X a set of geophysical data. In the present study, X includes EAM functions (AAM, OAM, HAM, and SLAM), tides, tidal excitations, and MEI, MJI, and NAI. For the architecture of the neural network NN, we have observed that a nonlinear Multi-Layer Perceptron (MLP, Bishop, 2006) with two layers is sufficient to produce the best results. The first and second layers have 1 and 63 hidden neurons (for predicting 63 days), respectively. The activation function of the first layer is tangent hyperbolic, whereas for the second layer, it is linear. An important point regarding the architecture is that linear models can also present competitive results (Kiani-Shahvandi et al., 2022). For the purpose of comparison of the architectures, we use three different linear models: Ridge regression with cross-validation, (RidgeCV, Marquardt & Snee, 1975; S. Liu & Dobriban, 2020), Random Sample Consensus (RANSAC, Fischler & Bolles, 1981), and Ordinary Least Squares (OLS, Teunissen, 2003). The reason for this choice is that RidgeCV and RANSAC are robust against outliers and less sensitive to the possible high variability of rapid data across different days. Out of these, OLS is the simplest method that can present competitive results. Note that we analyzed several other algorithms including Huber (Huber, 1964, 1973; Sun et al., 2020), but they turned out to be computationally expensive and less accurate.

2.2 ResLearner in deep ensembles

We use ResLearner in the context of deep ensembles (Lakshminarayanan et al., 2016). Therefore, a series of neural networks are trained simultaneously based on the same data, and the final prediction would be the average of the prediction of all the individual models. This reduces the epistemic uncertainty (Sullivan, 2015), which is due to errors in the utilized model. The mathematical formulation of deep ensembles (Lakshminarayanan et al., 2016) is based on the assumption that the data can be represented by a heteroscedastic Gaussian distribution. The variance and mean of the distribution are then solved for, following the minimization of the logarithm of the likelihood function $\ell(F, R, X)$ as the loss function. The formulation of the deep ensembles for the calibration of rapid EOPs is given in Equations (2a)-(2f).

$$\mu_j(R,X) = \mathrm{NN}_\mu(\theta_{\mu,j}, R, X) \tag{2a}$$

$$\sigma_j^2(R,X) = \log(1+\exp(\mathrm{NN}_\sigma(\theta_{\sigma,j}, R, X))) + \epsilon \tag{2b}$$

$$\ell_j(F,R,X) = \frac{1}{2}\log\sigma_j^2(R,X) + \frac{1}{2}\frac{(F-R-\mu_j(R,X))^2}{\sigma_j^2(R,X)} \tag{2c}$$

$$\ell_j(F,R,X) \longrightarrow \text{minimize} \tag{2d}$$

$$\mu(R,X) = \frac{1}{M}\sum_{j=1}^{M}\mu_j(R,X) \tag{2e}$$

$$\sigma^2(R,X) = -\mu^2(R,X) + \frac{1}{M}\sum_{j=1}^{M}\sigma_j^2(R,X) + \mu_j^2(R,X) \tag{2f}$$

where $\mu(R,X)$ and $\sigma^2(R,X)$ are the ensemble mean and variance, being the average of M individual members of the ensembles with mean and variance $\mu_j(R,X)$ in Equation (2a) and $\sigma_j^2(R,X)$ in Equation (2b), respectively. In our case, we observed that $M = 10$ is sufficient and results in the highest accuracy. Using significantly more than 10 models seems to be unnecessary, while being drastically more computationally expensive, and at the same time, resulting in no significant gains in accuracy (below the current uncertainty level in EOPs). $\mu_j(R,X)$ and $\sigma_j^2(R,X)$ are modelled by two different neural networks $\mathrm{NN}_\mu(\theta_{\mu,j}, R, X)$ and $\mathrm{NN}_\sigma(\theta_{\sigma,j}, R, X)$ with different learnable parameters $\theta_{\mu,j}$ and $\theta_{\sigma,j}$, respectively, as in Equations (2a) and (2b). Since the variance has to be positive, the softplus function (Szandała, 2021) is applied to the neural network $\mathrm{NN}_\sigma(\theta_{\sigma,j}, R, X)$, i.e., Equation (2b). The term ϵ is a constant for numerical stability. In our problem, we observed that a value of $\epsilon = 10^{-8}$ performs sufficiently well. The loss function $\ell_j(F,R,X)$ is minimized for each individual model separately using Adam optimizer (Kingma & Ba, 2015) with 200 epochs. Finally, it is worthwhile to mention that we implement the method using the TensorFlow library in Python (Abadi et al., 2016).

2.3 Unmixing and self-calibration approaches: geophysical information and constraints

In order to investigate the causes of discrepancies between rapid and final EOPs, one can explicitly model some of the known effects. Here, we model the effect of errors in EAM functions, ocean tides, and tidal excitations. The discrepancies between rapid and final polar motion, denoted by δxp and δyp, and rapid and final dUT1, denoted by δdUT1, are the sum of individual discrepancies due to EAM functions δEAM, ocean tides δT, tidal excitations δTE (for polar motion), and additional effects δU, which include smoothing, systematic correction, and unknown effects. δEAM, δT, and δTE are related to the variable X in the neural network in Equation (1). It is also important to note that the component-wise summation of individual EAM functions is used (Kiani-Shahvandi et al., 2022).

Both the polar motion components and dUT1 are affected by ocean tides and libration in terms of diurnal and subdiurnal variations (Sections 5.5 and 8.2 of Petit & Luzum, 2010). Moreover, polar motion is affected by long-period ocean (both prograde and retrograde) tides which are conventionally modelled with periods from 9 days to 18.6 years (Section 8.3 of Petit & Luzum, 2010). However, dUT1 is affected by zonal tides (i.e., the effect of tidal deformation), which are modelled with periods from 5 days to 18.6 years (Section 8.1 of Petit & Luzum, 2010).

The general approach to include the tidal effects in our model is to consider the harmonic functions with fixed frequencies through Delaunay parameters (Petit & Luzum, 2010), but with variable, estimable amplitudes. This is due to the fact that in rapid EOPs tides are already taken care of, and we need to compensate for the potential erroneous effect of tides included in the model. Therefore, δT and δTE can be modelled as in Equation (3)

$$\delta\mathrm{T},\ \delta\mathrm{TE} = \sum_{i=1}^{K} A_i \cos\Theta(t) + B_i \sin\Theta(t) \tag{3}$$

in which K is the number of tidal constituents considered, A and B the coefficients that should be determined by the neural networks, and $\Theta(t)$ the time-dependent argument of the harmonic functions based on the Delaunay parameters (Petit & Luzum, 2010). In the case of subdiurnal polar motion and dUT1, $K = 30$ constituents are added as features for each of xp, yp, and dUT1. For the diurnal tides, this number is $K = 41$ for each EOP. For the long period ocean tides and tidal excitations specific to polar motion the number is $K = 10$ for both xp and yp, and for the prograde and retrograde motions. The zonal tides specific to dUT1 have $K = 62$ constituents (Petit & Luzum, 2010).

δEAM is decomposed into two parts: equatorial components $\delta\chi_1, \chi_2$ and the axial part $\delta\chi_3$ of the excitations. These two parts can be modelled with two groups of neural networks ($\mathrm{NN}_{\chi_1}, \mathrm{NN}_{\chi_2}$) and NN_{χ_3}. Additional constraints can be applied to NN_{χ_1}, NN_{χ_2} and NN_{χ_3}. For instance, we apply the Liouville equation (Chin et al., 2004) for δP (in the imaginary domain, $\delta P = \delta\mathrm{xp} - i\delta\mathrm{yp}$) to investigate if there are additional parts that are not available in EAM data or the tidal effects that result in errors δxp, δyp in the polar motion components. Similarly, for the rate of dUT1 a linear combination of mass (pressure: p) and motion (wind: w) terms of the χ_3 component of the EAM functions would be considered, bearing physical meaning for example concerning mantle anelasticity (Dickman, 2003; Dobslaw & Dill, 2019b). In addition, a neural network denoted by $\mathrm{NN}_s(\theta_s, R, \chi_3)$ should learn the remaining signals in the rate of dUT1 (i.e., periods larger than annual), including its interannual trend. Furthermore, since EAM data used in the study are both observations and forecasts, NN_{χ_1}, NN_{χ_2}, and NN_{χ_3} can be used to minimize the difference between forecasts and their corresponding observations simultaneously with the minimization of the difference between rapid and final EOPs.

Depending on the effects included, we have to consider two aspects, namely the unmixing problem and the self-calibration. The unmixing problem occurs when the tidal effects and EAM functions are included in the model and investigated for their impact on the reduction of differences between rapid and final EOPs. If, in addition, we try to calibrate the EAM forecasts simultaneously with the calibration of rapid EOPs, we have to introduce a self-calibration approach. In mathematical terms, this concept is described in Equations (4a)-(4f):

$$\delta\text{xp},\ \delta\text{yp} = \delta\chi_1,\ \delta\chi_2 + \delta\text{T} + \delta\text{TE} + \delta\text{U} \tag{4a}$$

$$\begin{aligned}
&\delta P + \frac{i}{\sigma_{cw}}\frac{d}{dt}\delta P = \delta\chi_1 + i\delta\chi_2\\
&\delta P = \delta\text{xp} - i\delta\text{yp}\\
&\sigma_{cw} = \frac{2\pi}{T}(1+\frac{i}{2Q})\\
&T = 434.2\\
&Q = 100\\
&i = \sqrt{-1}
\end{aligned} \tag{4b}$$

$$\delta\chi_{1,o},\ \delta\chi_{2,o} = \delta\chi_{1,f},\ \delta\chi_{2,f} + \text{NN}_{\chi_1,\chi_2}(\theta_{\chi_{1,2}}, R, \chi_{1,f}, \chi_{2,f}) \tag{4c}$$

$$\delta\text{dUT1} = \delta\chi_3 + \delta\text{T}' + \delta\text{U}' \tag{4d}$$

$$\frac{d}{dt}\delta\text{dUT1} = \alpha\delta_{\chi_3^p} + \beta\delta_{\chi_3^w} + \text{NN}_s(\theta_s, R, \chi_3) \tag{4e}$$

$$\delta\chi_{3,o} = \delta\chi_{3,f} + \text{NN}_{\chi_3}(\theta_{\chi_3}, R, \chi_3) \tag{4f}$$

In Equation (4a), the error terms in polar motion δxp and δyp result from the errors in the equatorial components of the excitation functions $\delta\chi_1, \chi_2$, ocean tides, long period ocean tides and tidal excitations, and the remaining errors (smoothing, systematic correction, or unknown). NN_{χ_1}, NN_{χ_2} are used to calibrate the EAM forecasts used in the model with respect to the corresponding observations as in Equation (4c). These calibrated values can then be used in Equation (4b) to improve the prediction accuracy. A similar condition can be considered for dUT1 based on the differentiation of dUT1 and the mass and motion terms of the axial component of EAM $\delta\chi_3^p$, $\delta\chi_3^w$, through the linear equation (4e), with learnable parameters α and β. Crucial to mention is the presence of the neural network NN_s that learns the remaining signals in the rate of dUT1, including the interannual trend. Note that the errors in dUT1 (c.f. Equation (4d)) come from the errors in the axial component of the excitation functions $\delta\chi_3$, subdiurnal and diurnal tides $\delta T''$, long-period (zonal) tides $\delta Z'$ and the remaining errors $\delta\text{U}'$ ($\delta T' = \delta\text{T}'' + \delta\text{Z}'$). Similar to the case of polar motion, here also the difference between forecasts and their corresponding observations is simultaneously minimized with the calibration of rapid EOPs–Equation (4f). Finally, it is worthwhile mentioning that the methods used for polar motion use both xp and yp as the feature in the model, since this is shown to result in better prediction accuracy (Kiani-Shahvandi et al., 2022).

2.4 Feature importance methodology

For the analysis of feature importance, the goal of which is to investigate the importance of different input features in making accurate predictions, we use the method of deep feature ranking (Maksymilian & Chen, 2020). This method eliminates the need for combinatorial optimization (Bengio et al., 2021) for feature importance. This is advantageous since the importance of different features can be simultaneously analyzed, instead of analyzing individual or combinations of different features. Therefore, a large number of features can be investigated. The choice is furthermore justified since the ResLearner approach is mainly non-linear.

We define the feature importance (FI) as the relative contribution to the results. This means that FI in the first approximation is the ratio of the standard deviation of the method with or without the k-th feature $\sigma^{(k)}$ relative to the standard deviation of the output σ^F, as in Equation (5)

$$\text{FI}_k = \frac{\sigma^{(k)}}{\sigma^F} \tag{5}$$

Note that $\sigma^{(k)},\ k = 1, ...$ are the output of the deep feature ranking method (Maksymilian & Chen, 2020).

2.5 Geophysically-constrained neural networks: introducing ResLearner PhycoRNN

In addition to the unmixing and self-calibration problems, the concept of Physically Constrained Neural Networks (PCNN, Geneva & Zabaras, 2020) can be used for directly applying the physical constraints to the problem using Recurrent Neural Networks (RNN, Rumelhart, Hinton, & Williams, 1986). It has been shown that PCNN methods like PhyLSTM (Zhang et al., 2020), which is based on long short-term memory (LSTM, Hochreiter & Schmidhuber, 1997) and the physical conditions of the problem, could present state-of-the-art prediction performance. As LSTM is the base of PhyLSTM, one can think of replacing it with more modern architectures. We investigated several state-of-the-art architectures for the problem, including PhyLSTM itself, coupled oscillatory RNN (coRNN, Rusch & Mishra, 2021) and Long Expressive Memory (LEM, Rusch, Mishra, Erichson, & Mahoney, 2022). The coRNN architecture achieved the best performance and therefore we chose it to replace the LSTM cell in PhyLSTM. Using this approach, we devise a new architecture called PhycoRNN. The architecture is shown in Figure 1. In this architecture, there are two coRNN cells. The input $I = (R, \mathrm{EAM})$, containing rapid EOPs and EAM, passes through the first coRNN cell and generates two outputs V_1, V_2 which are subsequently passed through a Dense layer (Bishop, 2006) to generate the output G. The squared difference between G and the output F containing final EOPs data should be minimized, which can be called the mathematical loss, denoted by Loss_m. V_1 and V_2 are additionally passed through the second coRNN cell to generate the two outputs Z_1 and Z_2, which by applying another Dense layer to them would generate the output H. The geophysical constraints are then applied to H.

The geophysical constraint in the case of polar motion is the Liouville equation presented in Equation (4b), while for dUT1 rate is the linear combination presented in Equation (4e). In this case, α and β can be written as the following Equation (6) (Dobslaw & Dill, 2019b).

$$
\begin{aligned}
\alpha &= 2\pi\Omega\frac{k_r}{C_{\mathrm{eff}}}(1 + k'_{2,\mathrm{eff}} + \Delta k'_{\mathrm{an,eff}}) \\
\beta &= 2\pi\frac{k_r}{C_{\mathrm{eff}}}
\end{aligned}
\tag{6}
$$

in which $\Omega = 7.292115 \times 10^{-5}$ $[\frac{1}{s}]$ is the rotation rate of the Earth, $k_r = 0.9976$ the effect of rotational deformation, $C_{\mathrm{eff}} = 7.118246 \times 10^{37}$ [kgm^2] the effective axial moment of inertia, and $k'_{2,\mathrm{eff}} = -0.2415,\ \Delta k'_{\mathrm{an,eff}} = -0.0087$ the effective load Love number and the mantle anelasticity, respectively.

The mentioned geophysical constraints constitute the so-called physical loss, denoted by Loss_p. The total loss is the summation of the mathematical loss and the physical loss. To optimize the parameters of the neural networks we use the so-called LBFGS algorithm (D. Liu & Nocedal, 1989) since it has been shown to be quite efficient in PCNN problems. Finally, it should be noted that we investigated the number of time steps (input sequence length) used in the coRNN cell and a value of 3 was chosen since it resulted in the best prediction accuracy. Here, 200 epochs of training were used. The method was implemented using the PyTorch library (Paszke et al., 2019).

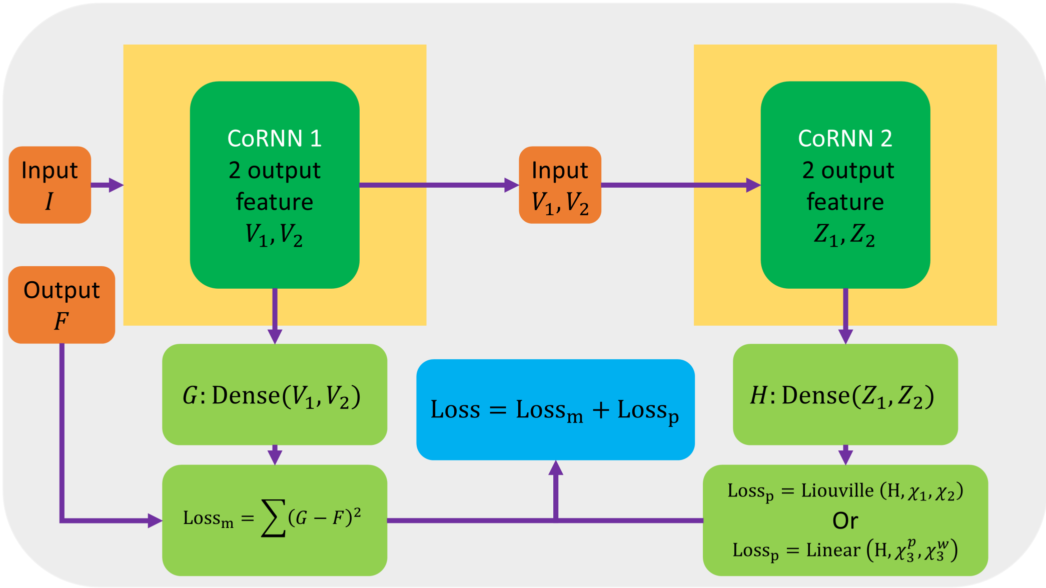


Figure 1: PhycoRNN architecture as a geophysically-constrained neural network, devised and used in the study.

2.6 Prediction accuracy metric

In order to evaluate the prediction accuracy, we use the mean absolute error (MAE) metric, which is commonly used in EOP prediction studies (Kalarus et al., 2010; Modiri et al., 2018; Kiani-Shahvandi et al., 2022). This is done for each day individually.

The quantification of improvement is based on the change in MAE for different days. If the MAE of one method is smaller than the baseline of rapid data themselves, we achieve an improvement. The MAE and improvement are defined in Equations (7a) and (7b):

$$\text{MAE}_k = \frac{1}{N}\sum_{i=1}^{N} |R_{i,k}^C - F_i|, \qquad k = -31, ..., 31 \tag{7a}$$

$$\text{improvement}_k = 100\% \frac{\text{MAE}_k^B - \text{MAE}_k}{\text{MAE}_k^B} \tag{7b}$$

In these equations, the index k is used for the day number, which is from -31 to 31. The number of predictions made is denoted by N. The predictions are denoted by $R_{i,k}^C$ (superscript C referring to calibration) for the i-th prediction and k-th day ahead. F_i denotes the corresponding final EOPs. The improvement is calculated by the percentage change in the MAE across different days, relative to the baseline (superscript B).

2.7 Summary of the concepts and optimal characteristics for ResLearner

A summary of the optimal characteristics of the ResLearner method is presented in Table 1, as determined in extended tests.

Table 1: Optimal characteristics for the ResLearner machine learning algorithm used for the calibration of rapid EOPs with respect to final EOPs

characteristic	choice/description
primary type of neural networks	non-linear MLP with two layers. 1 and 63 hidden neurons in layers, with tangent hyperbolic and linear activation functions for first and second layers, respectively
alternative type of neural networks	linear models: RANSAC, RidgdeCV, OLS
grouping of EOPs	equatorial and axial, i.e., for the prediction of xp or yp: both xp and yp used as feature; for the prediction of dUT1: only dUT1
non-linear uncertainty estimation	deep ensembles with M=10 simultaneous neural networks
feature importance analysis	deep feature ranking
evaluation metric	MAE
EAM functions considered	atmosphere, ocean, hydrology, and sea level
tidal effects	subdiurnal, diurnal, long period and tidal excitations, and long-period (zonal, for dUT1 only) with K= 30, 41, 10, 62 constituents, respectively
CI	MEI, NAI, MJI
PhycoRNN number of time steps	3
geophysical conditions for PhycoRNN	Liouville equation for rotational dynamics and polar motion; Earth rotation rate for first derivative of dUT1
unmixing	importance analysis of different features included in the model for their impact on the discrepancies between rapid and final EOPs
self-calibration	simultaneous calibration of EAM forecasts and the rapid EOPs

3 Data description

Here we describe the data used for the numerical results presented in the paper. Essentially, there are seven groups of data used in the study

- IERS rapid and final EOP 14 C04 series
- IERS final EOP 20 C04 series
- Jet Propulsion Laboratory (JPL) final EOP series (EOP2)
- European Space Agency (ESA) rapid and final EOP series
- ETH Zurich 14-day EAM forecasts
- GFZ German Research Center for Geosciences EAM analysis products
- National Oceanic and Atmospheric Administration (NOAA) MEI, NAI, MJI

IERS final 14 C04 EOP series (Bizouard et al., 2019) is the result of the combination of different space geodetic techniques including GNSS and VLBI and acts as the baseline to evaluate the various predictions against. This EOP time series is available from 1962 onward. Similar final EOPs data that are consistent with the latest International Terrestrial Reference Frame (ITRF2020) are provided by SYstèmes de Référence Temps-Espace (SYRTE). As mentioned in Section 1, IERS rapid EOPs (Dick & Thaller, 2018) are provided by using the most recent GPS and VLBI (24-hour and intensive session) data. The data are updated daily, but not archived publicly (daily finals). We have saved the rapid files since January 2015. Therefore, approximately 8 years of data is available for training and evaluation of the ResLearner algorithm. JPL series 2 of final EOPs are provided daily and contain the EOPs from 1976 onward, with less latency compared to the final IERS data. The JPL final series can act as the target in the training phase, i.e., IERS rapid EOPs are mapped to the final JPL EOPs. This creates another solution in addition to the one with final IERS data as the target.

For the purpose of additional validation, we use final, rapid and predicted EOPs provided by ESA and derived within the framework of the ESA project on "Independent Generation of Earth Orientation Parameters" (ESA-EOP, Dill et al., 2020; Kehm et al., 2023). The data result from series of hindcast experiments, in which the final EOPs are combined from GNSS, SLR, VLBI and DORIS and the rapid EOPs are combined from GNSS and VLBI only. Predictions are based on deterministic signals derived from the final and rapid EOPs time series in combination with EAM analysis and prediction data (as available on the assumed start date of prediction). Two series of hindcast scenarios from the study were provided, namely a realistic scenario and an ideal scenario. While the realistic scenario (scenario H1 in Kehm et al., 2023) assumes that the VLBI contribution to rapid (combined) EOPs solely relies on intensive data, the ideal scenario (scenario H2 in Kehm et al., 2023) assumes both 24-hour and intensive data to be available for the rapid combination. Each hindcast scenario is provided in the form of a data set containing 656 daily files for a time span from January 2018 up to January 2020. Thereby, each daily file contains final EOPs from around January 2009 up to a prediction horizon of about -28 days, rapid (combined) EOPs up to the day before the prediction start, and predicted EOPs up to a prediction horizon of +90 days. Here, we will use both scenarios for validation.

Regarding the EAM data, both the observations and forecasts are used, since forecasts can help significantly to improve the EOP prediction performance (Modiri et al., 2020; Kiani-Shahvandi et al., 2022). Since the horizon of the forecasts is also a determining factor (Kur et al., 2022), we use 14-day forecasts of ETH Zurich (Kiani Shahvandi et al., 2022) since they are both accurate and cover a reasonable forecasting horizon for short-term EOP prediction (i.e., suitable for accurate real-time purposes). Note that EAM predictions from all 14 days are used, since based on our analysis it results in the best performance (for instance, using 10-day forecasts results in less improvement). The EAM analysis files are taken from GFZ German Research Center for Geosciences (Dobslaw & Dill, 2018; Dill et al., 2019a). All four types of EAM functions, i.e., AAM, OAM, HAM, and SLAM, are used as geophysical features in the ResLearner algorithm.

We use CI provided by NOAA. Climatic index MEI is provided bimonthly by an empirical orthogonal function that combines different variables including sea surface pressure and temperature (Wolter & Timlin, 1993; Timmermann et al., 2018; Di Lorenzo et al., 2023). Since the data are bimonthly, they should be interpolated to generate daily values to be used as an additional feature for the prediction of dUT1. We also use NAI and MJI suspected for their influence on the rate of dUT1 (Hendon, 1995; Mazzarella, 2007).

Several investigations are presented in Section 4. In Figure 2, we show the rapid xp, yp, and dUT1 time series as well as the training and evaluation intervals for five different studies presented in this paper. The first study (S1) is similar to the subsequent three, but it is done operationally, with retraining at each prediction epoch. The starting date of evaluation is 20 May 2021 to be consistent with operational EAM forecasts (Kiani Shahvandi et al., 2022). The next three (S2, S3, S4) are hindcast studies that use IERS rapid EOPs as the input and IERS final 14 C04 or JPL EOP2 as the output. The purpose of these studies is to analyze the performance of the algorithm in the past. The final study (S5) is based on the ESA and IERS rapid and final EOPs. This is also only possible in a hindcast study. Crucial to mention is that hindcast studies observe the rules of real-time prediction (i.e., no future information being available), but with the prediction time in the past.

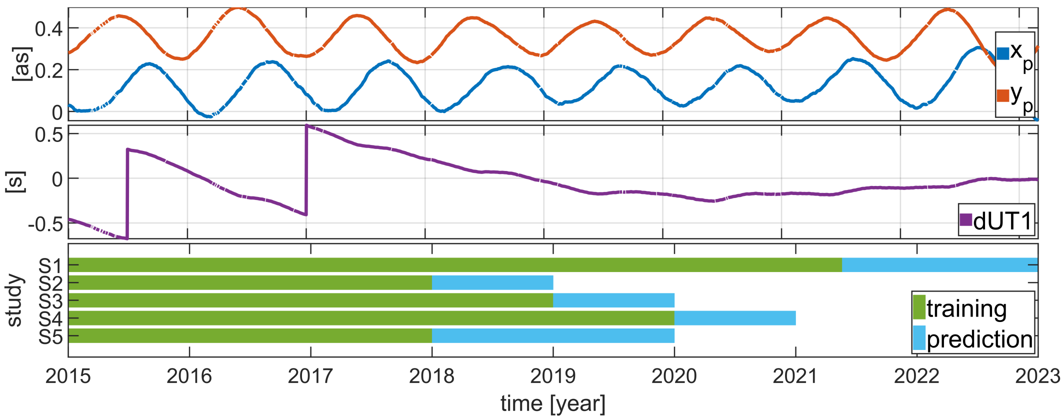


Figure 2: Top and middle panels show the polar motion and dUT1 series used in the study. The bottom panel shows the training and prediction intervals for each of the five studies (S1)-(S5) presented in Section 4.

4 Results and discussions

4.1 Analysis for the operational results in 2021-2022

Here, we present the performance analysis of the methods discussed in Section 2 based on the data described in Section 3. Note that the analysis refers to the study number 1 (S1) in Figure 2. The following points summarize the study configuration:

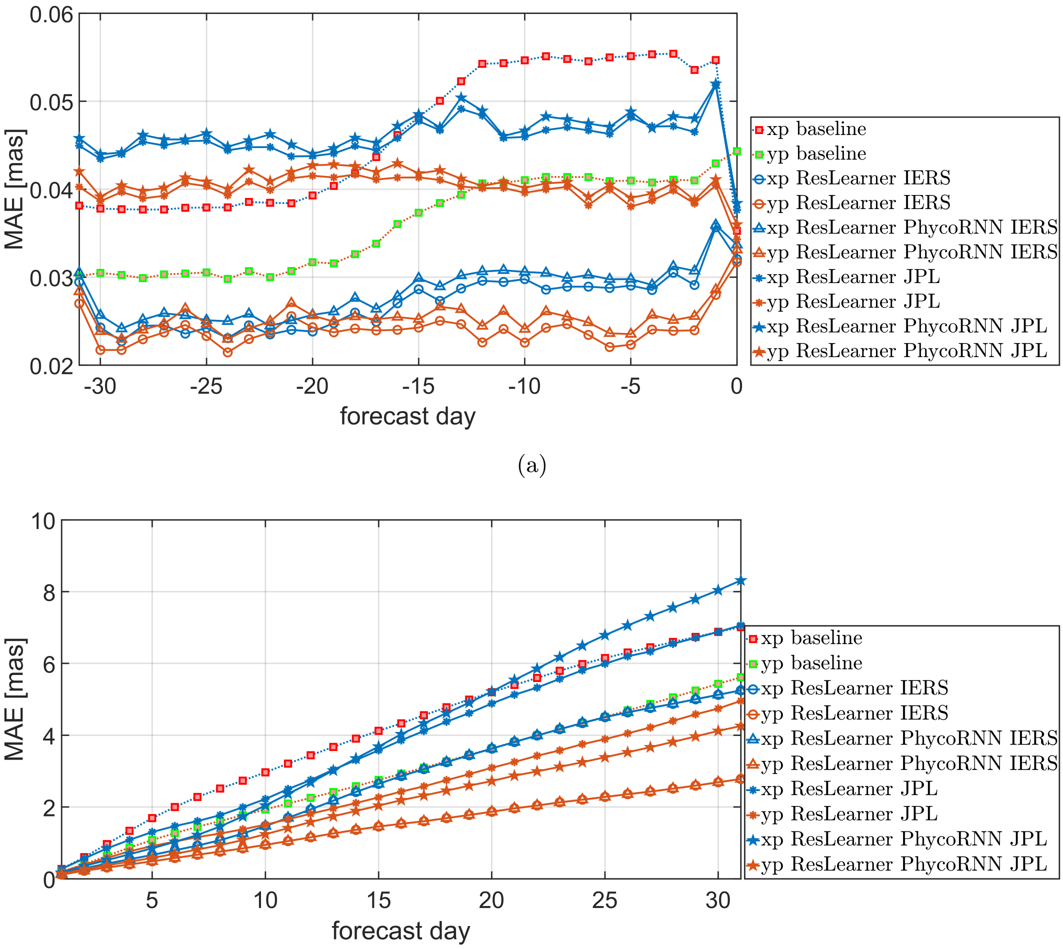
- The baseline solution is rapid EOPs as provided by IERS,
- Methods are trained on both IERS and JPL final EOPs,
- The final IERS 14 C04 EOP series is used for evaluation.

4.1.1 Prediction accuracy and improvement

Figures 3 and 4 present the results of applying both the ResLearner and ResLearner PhycoRNN algorithms to the study interval shown in Figure 2. For better visualization of the performances, the prediction interval is divided into two parts: days -31 to 0 and days 1 to 31. The improvements with respect to the IERS baseline are presented in Figure 5 for polar motion and Figure 6 for dUT1. Based on Figures 3-6, several important points become evident.

First, the results of ResLearner PhycoRNN from days 1 onward seem to be identical to those of ResLearner when IERS 14 C04 is used for training. They are also very similar on days -31 to day 0, but not identical. This proves that for methods trained on IERS 14 C04, both PhycoRNN and ResLearner can be used. However, when JPL EOPs is used in the training, the results of ResLearner PhycoRNN and ResLearner are different. In this case, ResLearner PhycoRNN works better in yp, but worse in xp, approximately after day 13. This can be explained by the fact that ResLearner PhycoRNN has focused more on the yp component because of its larger amplitude and thus is performing worse on xp. Note, however, this is the best architecture for ResLearner PhycoRNN, implying that it cannot outperform ResLearner in xp, but only in yp. We tried to weight the loss functions so that the amplitudes of the errors of xp and yp be in the same range, but this did not improve the results. Regarding the difference between the results using JPL and IERS data as target, it becomes clear that the PhycoRNN has been able to capture the physics, but there is not as meaningful geophysical information in the mapping from rapid to JPL as from rapid to IERS. This is because the PhycoRNN is effectively transforming between EAM and GAM (Geodetic Angular Momentum), which as Dill et al. (2020) also point out, are not in full agreement with the JPL combined EOP series, especially for the equatorial components. This implies that having the Liouville equation as a hard constraint would not be beneficial if the EAM and EOPs series do not correspond to each other. In this case, a more mathematical-based approach would present better results, which is the case with ResLearner. We conclude that if the EOP and EAM series correspond to each other, the results of ResLearner and ResLearner PhycoRNN are almost identical, thereby suggesting physical and mathematical information have been adequately captured. Otherwise, ResLearner PhycoRNN does not perform well, since the geophysical constraints are less informative. This happens mostly for polar motion, but not for dUT1, which is due to the better agreement on the axial components of the GAM derived from different EOPs series (Dobslaw & Dill, 2019b).

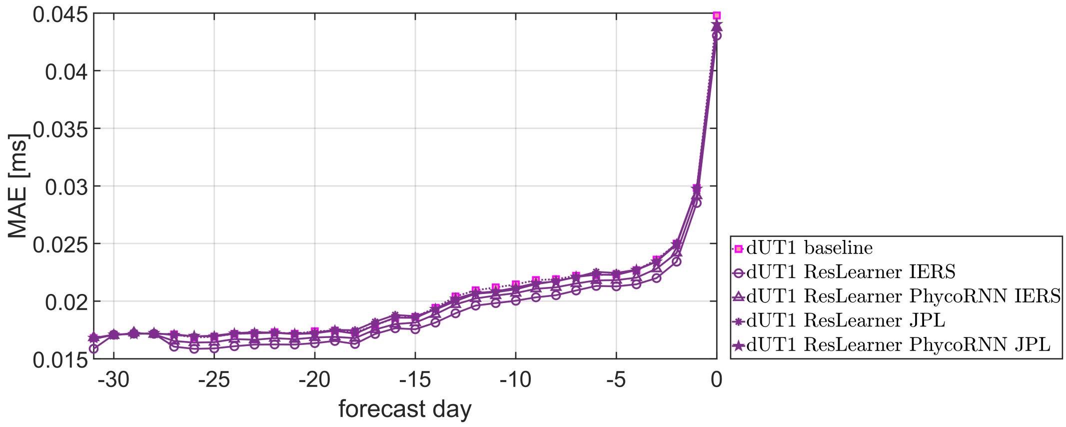
Second, the improvement for polar motion components reaches 60% and generally remains above 40% for days -15 to 13. This is achieved by training the data on IERS 14 C04 final series, but not on JPL. Reasons for this discrepancy may include the longer interval that JPL provides the data for, which results in less informative data as a result of the degraded accuracy. More importantly, as mentioned GAM derived from IERS and JPL using EAM data do not fully correspond and can have large discrepancies, resulting in a reduction in accuracy of PhycoRNN predictions with JPL data as target. The improvements for dUT1 are generally smaller than those for polar motion. But they tend to increase for longer prediction horizons. The accuracy of both ResLearner PhycoRNN and ResLearner in days -31 to 0 for polar motion is almost below or at the uncertainty level of the polar motion data. This confirms that the methods can deliver results with an uncertainty level similar to that of the polar motion data. Finally, it is important to note that the accuracy of the IERS baseline and most of the methods is better at day 0 than at day -1. This behavior is more pronounced in polar motion compared to dUT1, meaning that the improvement for polar motion drops significantly at this day. We suspect that the reason for this anomalous behavior lies within the data and not in the applied models, as it is also visible in the IERS baseline, and might be related to a dominance of GNSS-derived polar motion information in the final IERS product and on the final day of the rapid combination (Kehm et al., 2023). The ResLearner unmixer algorithm (Section 2.3) can be used to further investigate this anomalous behavior.



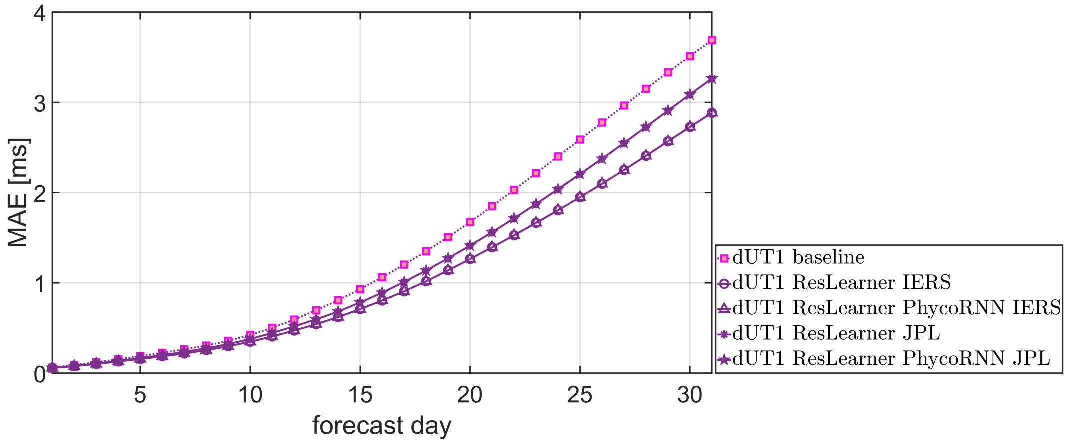
(a)

(b)

Figure 3: Prediction accuracy of polar motion components xp, yp for the first study (S1), in terms of MAE [mas]. ResLearner and ResLearner PhycoRNN are trained on both JPL and IERS final EOPs. (a) shows the MAE across days -31 to 0, while (b) focuses on days 1 to 31.

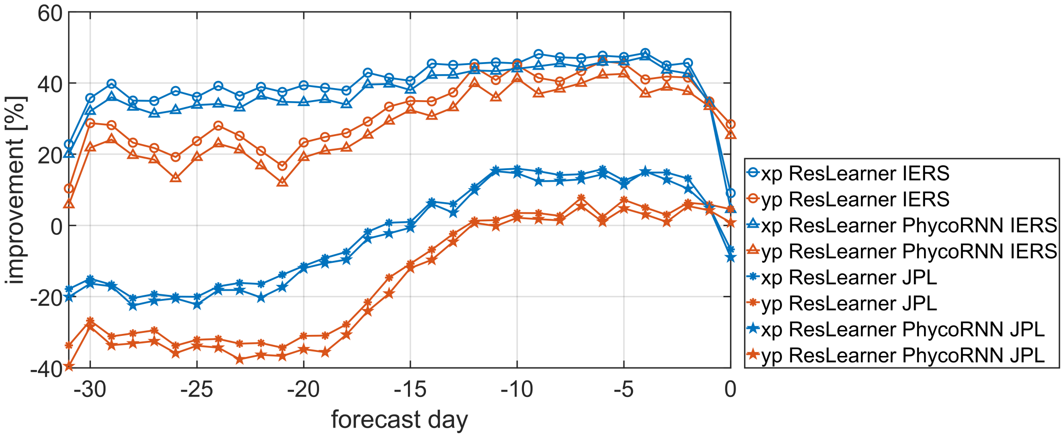


(a)

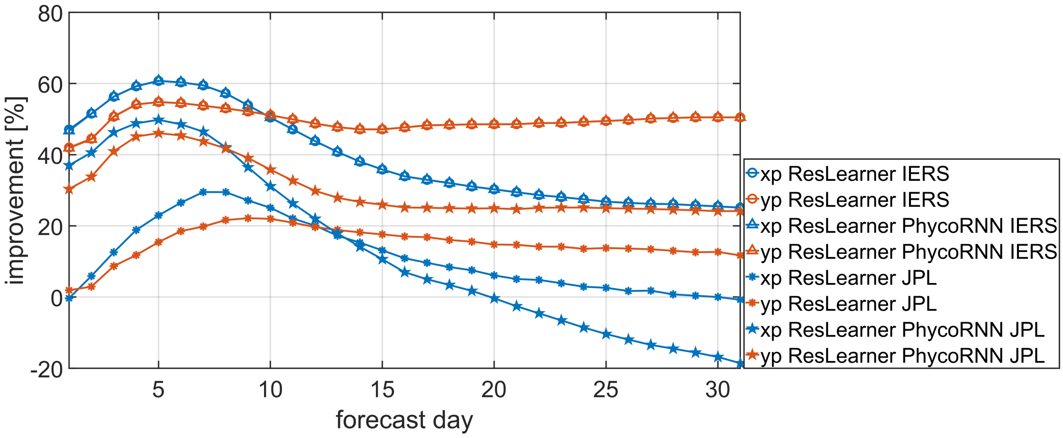


(b)

Figure 4: Prediction accuracy of dUT1 for the first study (S1), in terms of MAE [ms]. ResLearner and ResLearner PhycoRNN are trained on both JPL and IERS final EOPs. (a) shows the MAE across days -31 to 0, while (b) focuses on days 1 to 31.

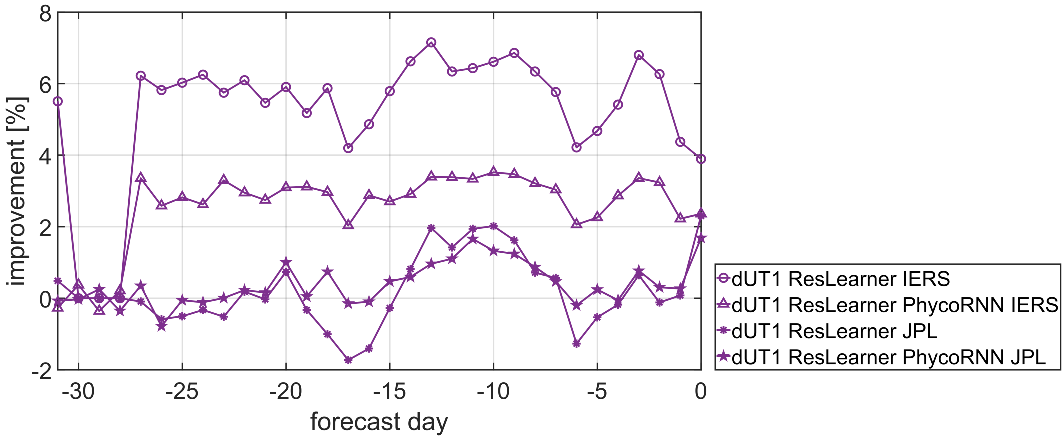


(a)

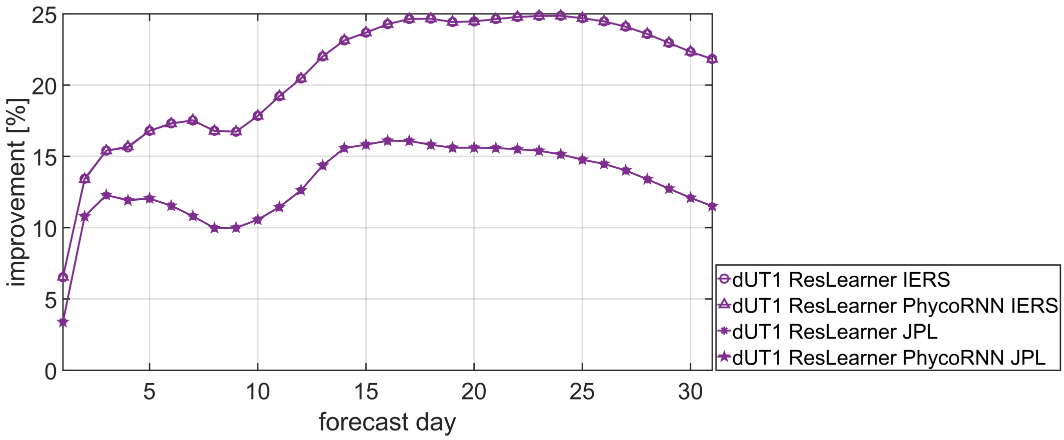


(b)

Figure 5: Improvement of prediction accuracy of polar motion components xp, yp for the first study (S1), in terms of percentage [%], computed according to Equation (7) based on the MAE of the baseline and that of ResLearner and ResLearner PhycoRNN. (a) shows the improvement across days -31 to 0, while (b) focuses on days 1 to 31.



(a)



(b)

Figure 6: Improvement of prediction accuracy of dUT1 for the first study (S1) presented in Figure 2, in terms of percentage [%], computed according to Equation (7) based on the MAE of baseline and that of ResLearner and ResLearner PhycoRNN. (a) shows the improvement across days -31 to 0, while (b) focuses on days 1 to 31. Note that the improvements are with respect to the IERS rapid data.

4.1.2 Importance of geophysical information

We find that EAM functions are one of the most important features that contribute to the discrepancies between rapid and final EOPs. As an example, in Figure 7, the Kendall correlations between the differences between rapid and final EOP IERS 14 C04, and the equatorial components of the individual EAM functions are shown. AAM and OAM (particularly the motion terms) present the highest correlation with these differences, thereby suggesting the importance of EAM for the ResLearner unmixer. Furthermore, even though in the rapid data AAM is included, the presence of the correlation suggests errors in accounting for AAM in the processes. In Figure 8 the importance of different features (FI) used in the model is presented, based on the methodology presented in Section 2 and according to Equation (5). For polar motion, Figure 8 gives the importance of the features xp, yp, EAM, and tides (semi-diurnal, diurnal, long-period tidal excitations combined), while for dUT1, it gives the importance of the features dUT1, EAM, tides (semi-diurnal,

diurnal, and long-period (zonal) combined), and CI. The individual CI components, i.e., MEI, NAI, and MJI are also displayed. Besides xp, yp, and dUT1 themselves, the EAM and tides are the most important features, confirmed also by other studies (Kiani-Shahvandi et al., 2022). Figure 7 also shows that AAM and OAM are the most important EAM functions for this problem (both mass and motion terms). Among CI, MEI seems to be the most relevant and can have effects several times bigger than the uncertainty level of dUT1. However, NAI and MJI have only a minor importance for the short-term prediction of dUT1. We therefore recommend only using MEI among the various climatic indices. We consider this to be in alignment with the observation that ENSO has a significant impact on the rate of dUT1, especially on interannual time scales (Chao, 1984).

We furthermore analyze the relationship between MEI and the physical condition on the rate of dUT1. In Figure 9, we show the negative of the rate of dUT1, i.e., $-\frac{d}{dt}dUT1$ (IERS rapid data) and the reproduced trend (which is in fact, rather an interannual signal in view of the limited time-period considered), the χ_3^p and χ_3^w components of the EAM functions, and MEI. Most of the signal in the rate can be explained by χ_3^w which is due to the zonal winds (Volland, 1996). However, the reproduced MEI also seems to be able to explain parts of the signal, especially around mid-2022. This can potentially be attributed to a La Niña event, which occurred in mid-2022. La Niña events have been shown to influence the rotation rates of the Earth (Xu et al., 2022). We can therefore state that ResLearner has been able to link the geophysical information to the input data. Note, however, that in short-term prediction the importance of MEI is smaller than that of other features, including χ_3^p and χ_3^w. But in the long-term, using MEI results in better training and prediction by ResLearner.

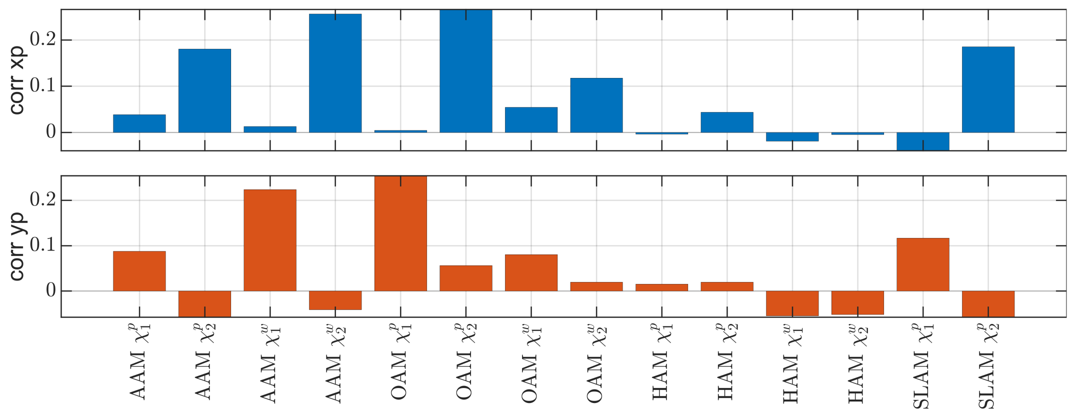
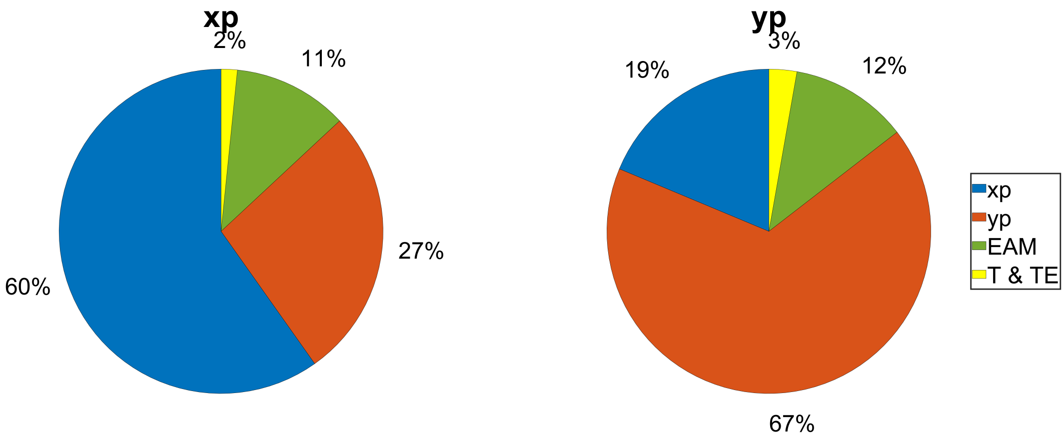
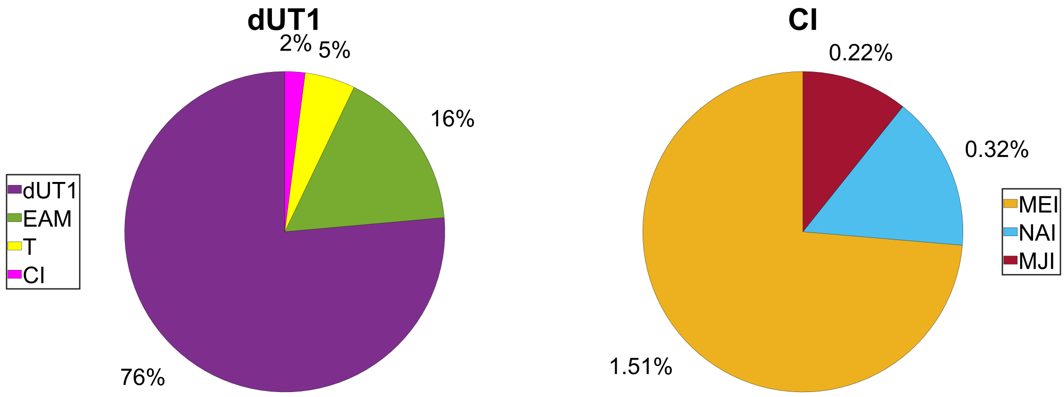


Figure 7: Kendall correlation (shown as corr in the figure) between the differences between rapid and final IERS EOPs, and the equatorial components of the individual EAM functions. Note that mass and motion terms (χ_i^p, χ_i^w $i = 1, 2$) are analyzed separately.



(a)



(b)

Figure 8: Feature importance analysis based on the algorithm presented in Section 2.4 and according to Equation (5). For polar motion components (a), features include xp, yp, equatorial components of EAM, T and TE (i.e., semi-diurnal, diurnal, and long-period tides and tidal excitations). For dUT1 (b), the features are dUT1, axial component of EAM, tides (semi-diurnal, diurnal, and zonal), and CI (climatic indices). CI is further decomposed into its components, i.e., MEI, NAI, MJI.

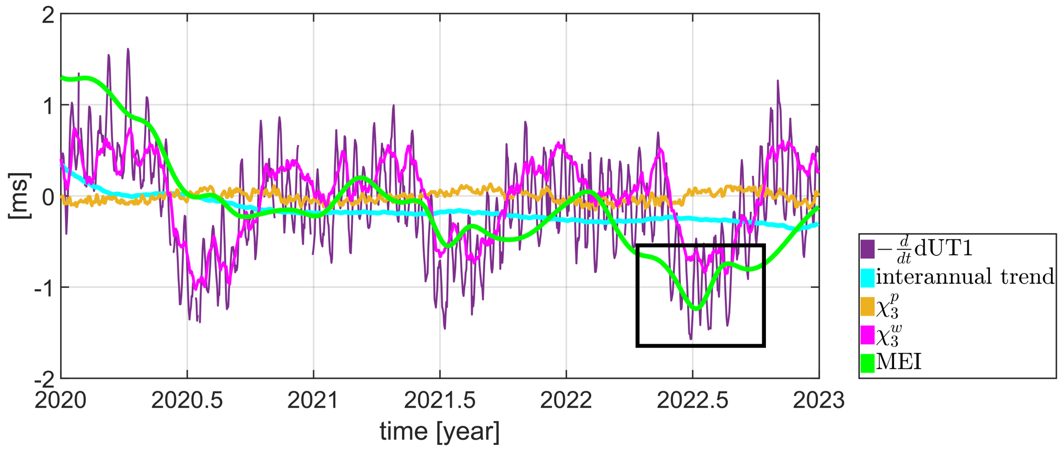


Figure 9: Negative rate of dUT1 (IERS rapid), $-\frac{d}{dt}$dUT1, together with the regressed interannual trend, χ_3^p, χ_3^w components of the EAM functions, and MEI, as obtained from the ResLearner algorithm. The interannual trend is solved during the training process and predicted accordingly at the prediction epoch. MEI here refers to what the ResLearner sees during training, i.e., the input feature MEI. Similarly, χ_3^p and χ_3^w are reproduced, but they are almost identical to their input form, because of their high feature importance. The mid-2022 La Niña event is highlighted by a black box.

4.1.3 Unmixing: on the potential causes of errors in rapid EOP data

Building upon the results of feature importance analysis in Figures 7 and 8, the ResLearner unmixer algorithm can be applied to find the individual components of the EAM and tides that contribute most to the discrepancies between rapid and final EOPs. The corresponding results are presented in Figure 10, based on FI as given in Equation (5). In order to asses their significance, we also show their corresponding 95% confidence intervals. We have grouped the contributions into 1) tides and EAM (δT, δEAM) and 2) remaining errors (δU, systematic correction, smoothing, and unknown). Panel (a) gives the relative contributions of these two groups. The effect of the first group is bigger, thereby suggesting that the potential causes of discrepancy lie within tides and EAM. The five most important features among the first group are further investigated in panel (b).

It is important to clarify that based on Figure 10 one can conclude that the most important features contributing to the anomaly observed at day 0 are (in the order of importance) δEAM at day 0, δU (including the dominance of the GNSS-derived polar motion), and δT. Regarding tides in polar motion, subdiurnal and diurnal tides, retrograde 13.63 and 27.56 days, and prograde 13.66 and 27.56 days long-period tides and tidal excitations are important. For dUT1, however, zonal tides of periods 13.78, 14.77, 23.89 days, and subdiurnal tides are relevant. For δU the approximate FI, together with their 95% confidence intervals are summarized in Table 2. Note that for δEAM and δT, the approximate values of importance are computed by multiplying the FI in panel (a) and (b), based on the fundamental rule of probability.

Table 2: The approximate FI and corresponding 95% confidence intervals for δEAM, δU, δT for the potential causes of discrepancies between the rapid and final EOP IERS 14 C04 series.

EOP	δEAM	δU	δT
xp	37±20%	33±6%	29±18%
yp	47±23%	30±7%	23±15%
dUT1	54±28%	26±8%	21±11%

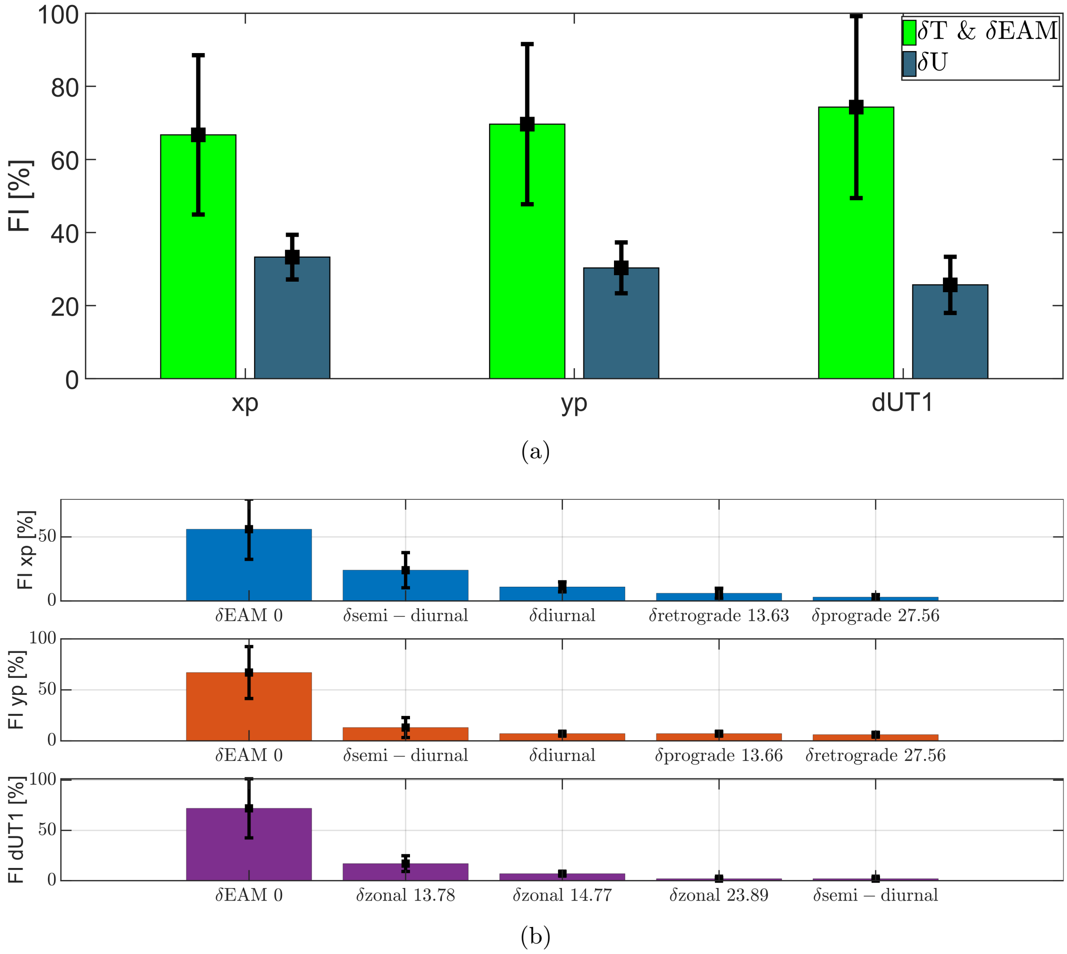
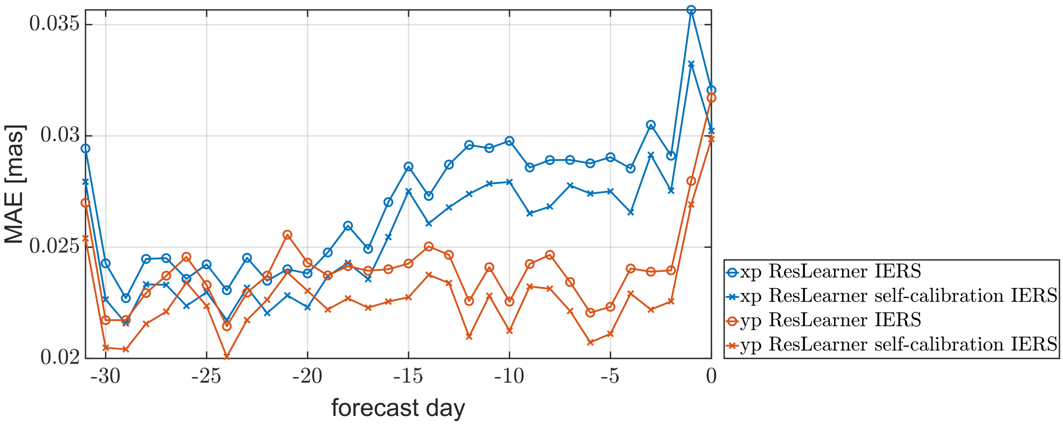


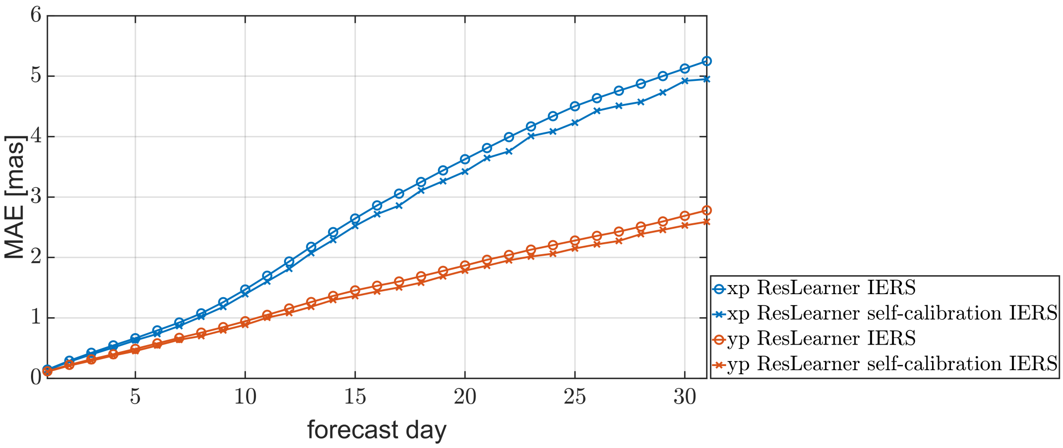
Figure 10: (a) FI computed according to Equation (5) for two groups 1) tides and EAM (δT & δEAM), 2) rest of errors (δU, systematic correction, smoothing, unknown) ; (b) FI computed for EAM and various tidal constituents resulting in the discrepancies between rapid and final EOPs, based on the methodology presented in Section 2. The uncertainties shown in the form of error bars are for 95% confidence interval. The analysis is for day 0 of prediction, containing the anomalous behaviour.

4.1.4 Self-calibration

After identifying the causes of errors in rapid data as from Figure 10, we apply the ResLearner self-calibration algorithm described in Section 2.3 in order to reduce the erroneous effects of the EAM functions. The results are shown in Figures 11 and 12 against the output of ResLearner algorithm without self-calibration. ResLearner self-calibration slightly improves the prediction performance (on average 5.5%). The improvement is achieved on both polar motion and dUT1, thereby suggesting the success of ResLearner self-calibration in reducing the errors.

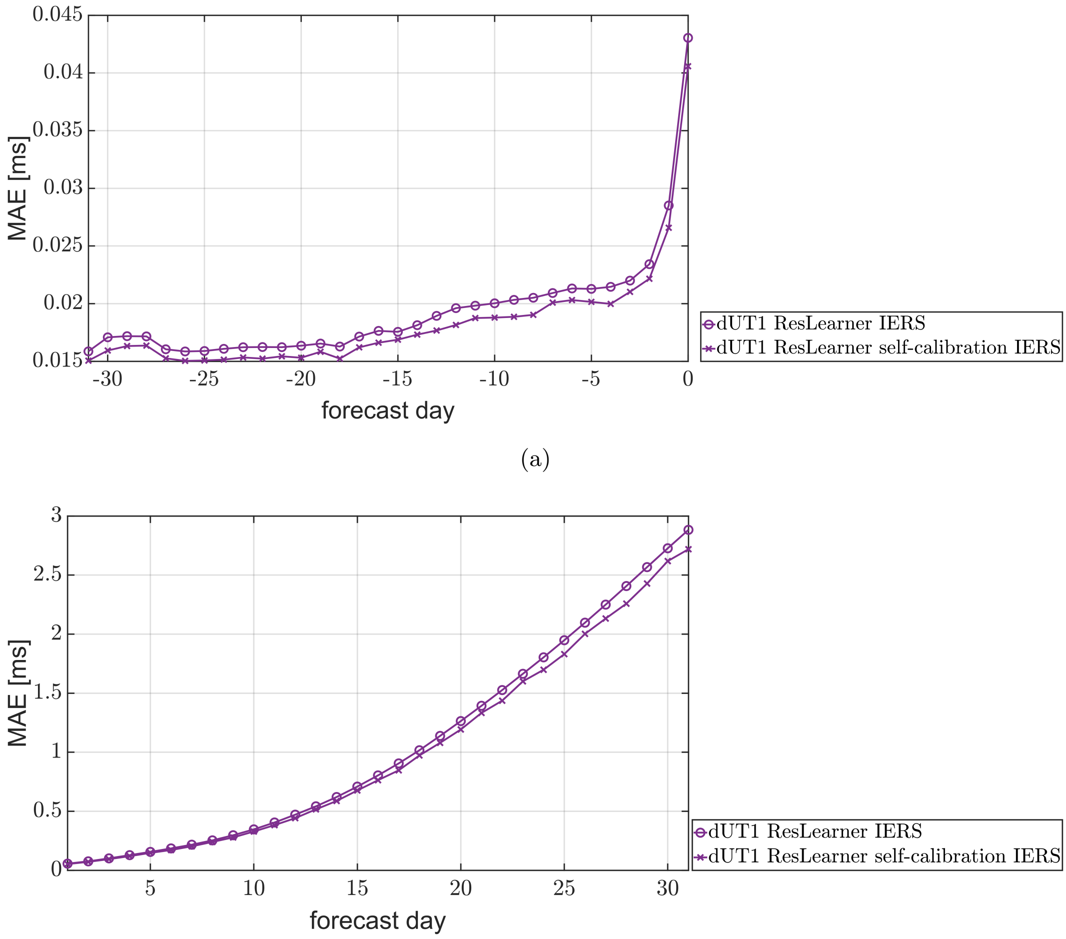


(a)



(b)

Figure 11: ResLearner self-calibration algorithm for the polar motion components against the ResLearner without self-calibration. (a) shows the comparison of days -31 to 0 while (b) displays that of days 1 to 31.



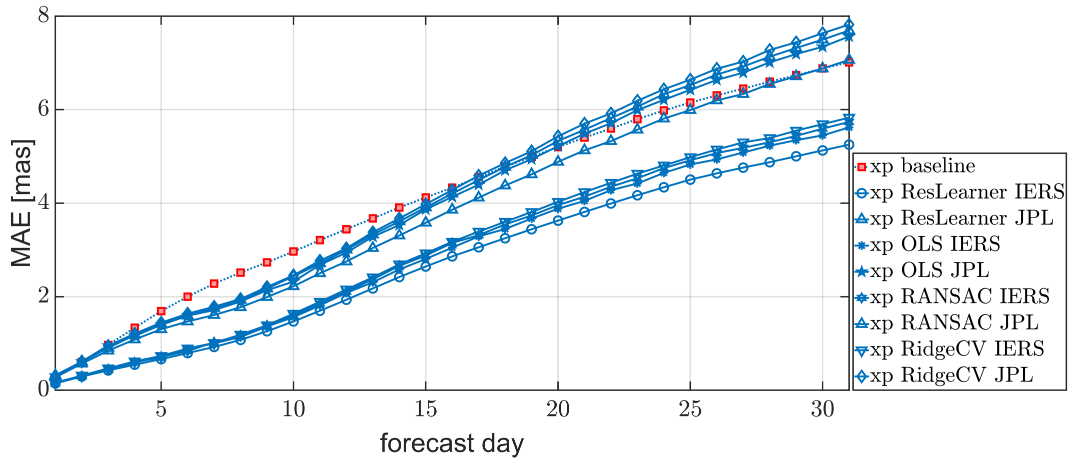
(a)

(b)

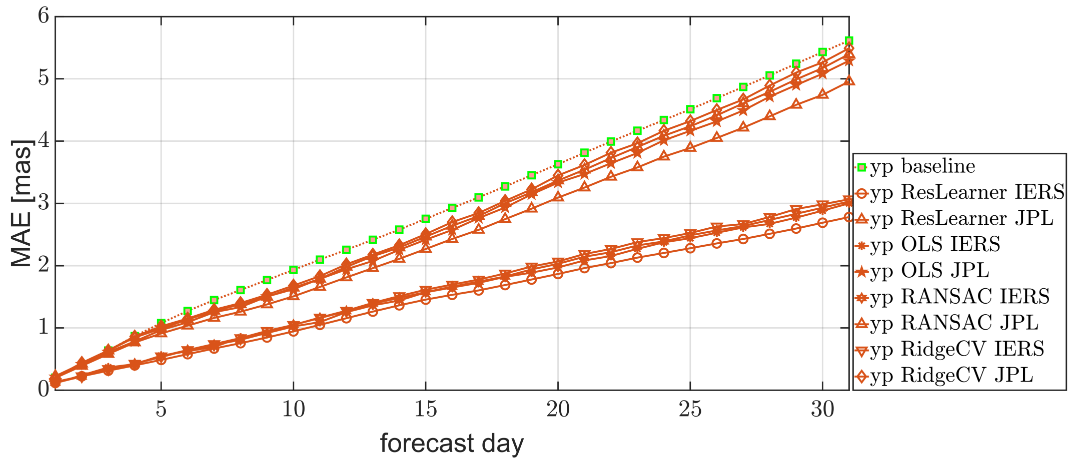
Figure 12: ResLearner self-calibration algorithm for dUT1 against the ResLearner without self-calibration. (a) shows the comparison of days -31 to 0 while (b) displays that of days 1 to 31.

4.1.5 Comparative analysis: linear models

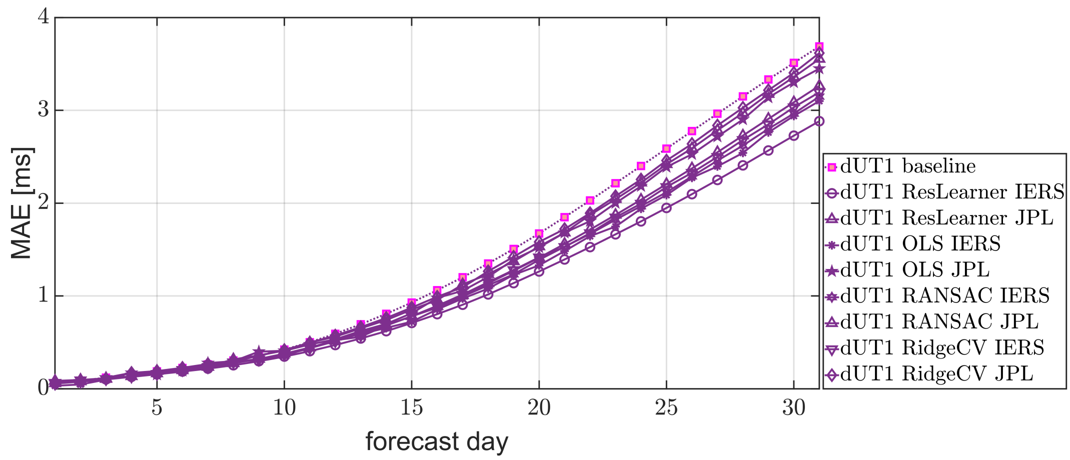
As mentioned in Section 2, linear ResLearner models can also present competitive results, i.e., close to the prediction performance of the state-of-the-art algorithms. The goal of this analysis is to illuminate the role of non-linearity in the model. Three different methods are considered: OLS, RANSAC, and RidgeCV. The results are compared with the non-linear ResLearner. The results of the comparative analysis are summarized in Figure 13. The results are shown only for days 1 to 31 since it is only on these days that we see a clear pattern of superiority of non-linear models. On days -31 to 0, the results are mixed: methods like OLS may outperform non-linear ones on some days, while on the rest of the days, the non-linear models outperform OLS. This analysis confirms that in this study, the non-linearity results in a gain in prediction performance. Furthermore, it is by non-linearity that the unmixing and self-calibration problems can capture almost all the signals in the input data.



(a)



(b)

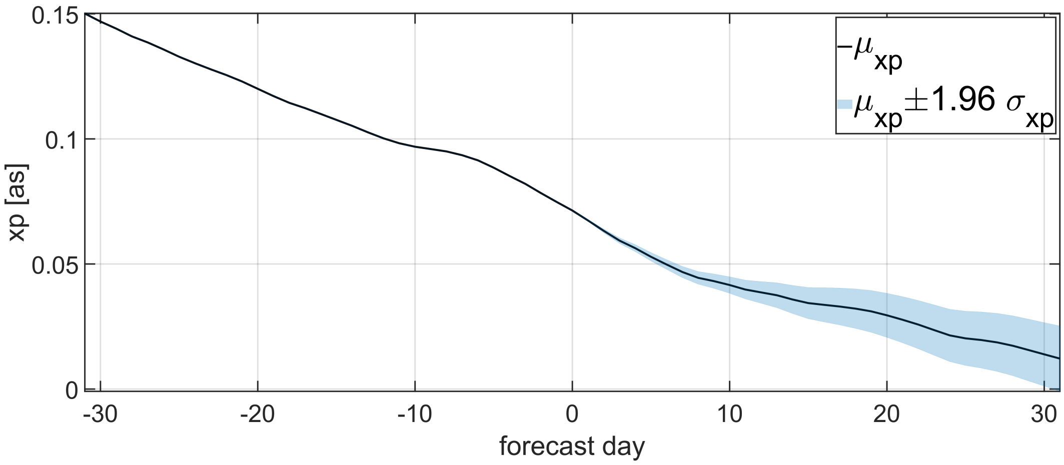


(c)

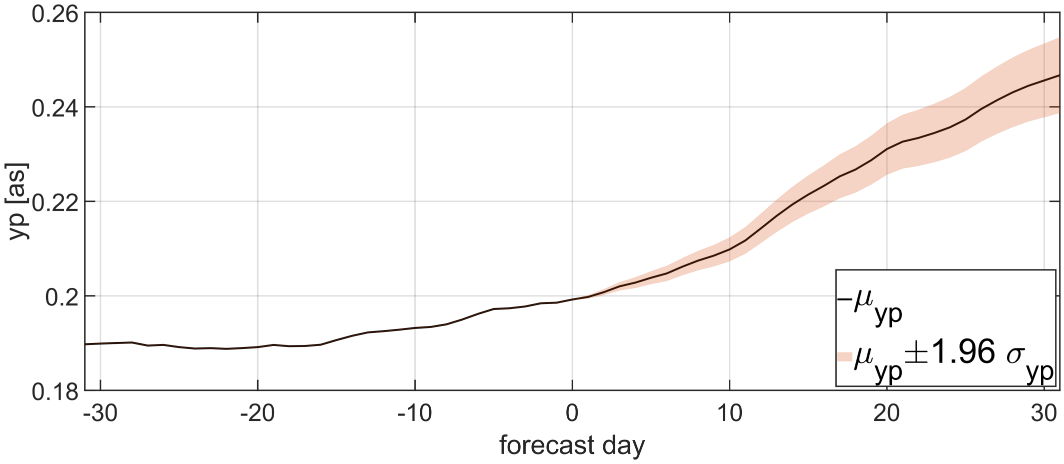
Figure 13: Comparison between the linear and non-linear ResLearner algorithms. Three different linear models are analyzed: OLS, RANSAC, RidgeCV. (a) shows the results for polar motion xp component, (b) for polar motion yp component, and (c) for dUT1.

4.1.6 Prediction uncertainty

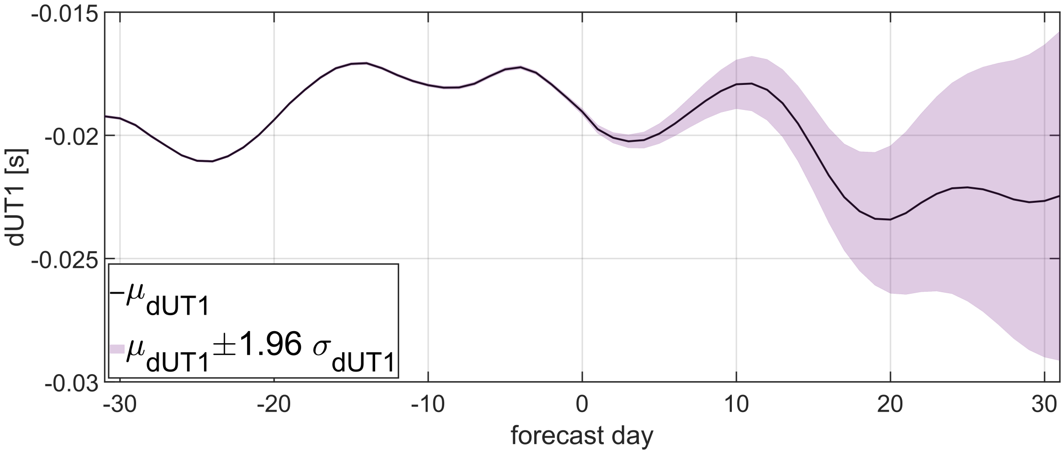
The ResLearner methodology implemented in the context of deep ensembles can provide uncertainties in the predictions. As an example, Figure 14 shows the predictions of polar motion and dUT1 together with their associated uncertainties, plotted for 2022-12-31. The mean values are given by μ, while the standard deviations are given by σ. The prediction uncertainties shown represent a 95% confidence ($\pm 1.96\sigma$) interval. Note that the derived prediction uncertainties depend on the respective day, but are usually close to the uncertainties in the rapid data. This confirms that ResLearner models in deep ensembles have been able to effectively reduce the epistemic uncertainty due to model errors. The reason is, the ResLearner is essentially a parametric model, the parameters of which are derived through optimization schemes. As a result, there is inevitably some uncertainties in the model parameters, which translate to the uncertainty in the predictions. Using the ensemble approach, we can effectively reduce this type of uncertainty and allow the model to predict more accurately and confidently.



(a)



(b)



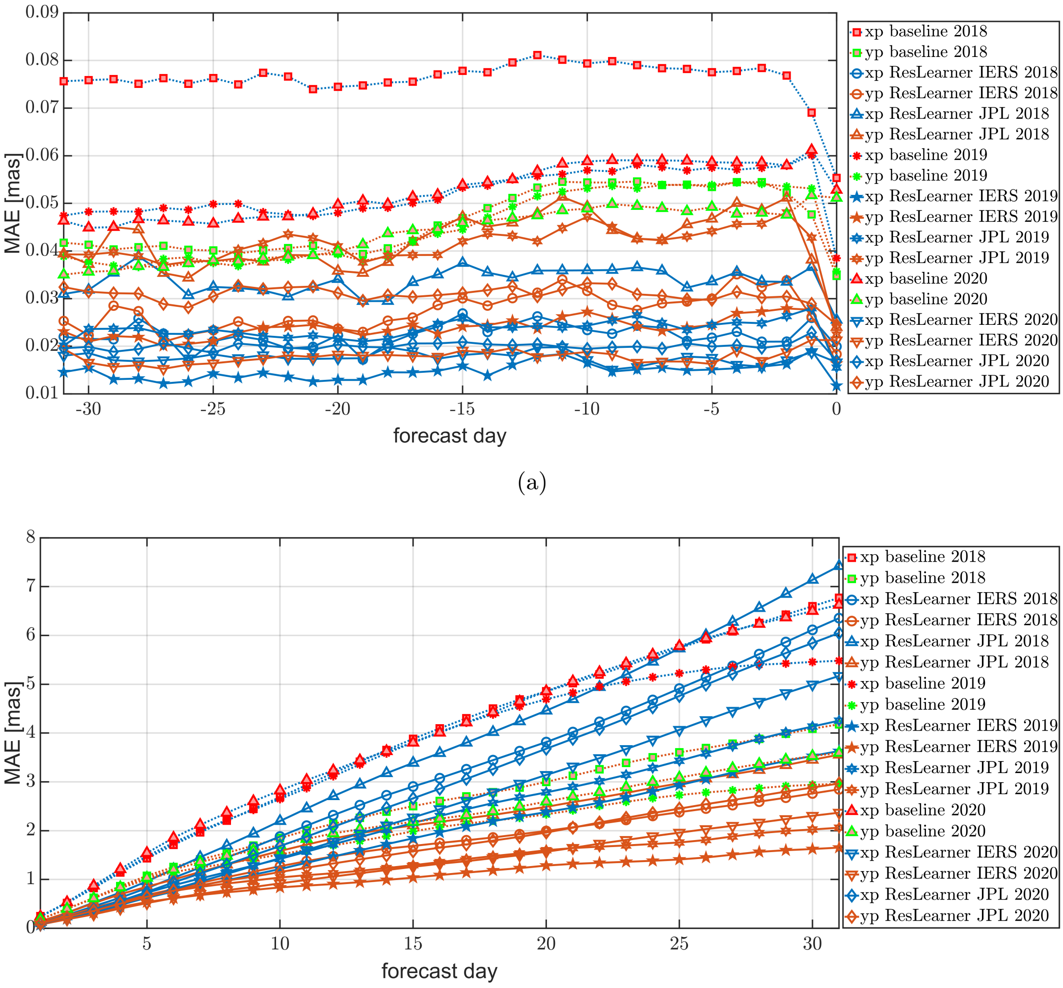
(c)

Figure 14: Prediction uncertainty for (a) polar motion xp, (b) polar motion yp, and (c) dUT1 for the date 2022-12-31, using ResLearner in the context of deep ensembles. μ_{xp}, μ_{yp}, and μ_{dUT1} are the mean values of the prediction, while σ_{xp}, σ_{yp}, and σ_{dUT1} are the associated standard deviations. The confidence interval is 95% ($\pm 1.96\sigma$).

4.2 Hindcast analysis: 2018, 2019, 2020

We analyze the performance of the ResLearner method in hindcast scenarios, corresponding to the second, third, and fourth analyses (S2, S3, and S4) shown in Figure 2. The same conditions as in the first study (S1) are applied here as well, i.e., using the rapid IERS as the baseline, training on both IERS 14 C04 and JPL final EOPs 2 data, and evaluating against the IERS 14 C04 series.

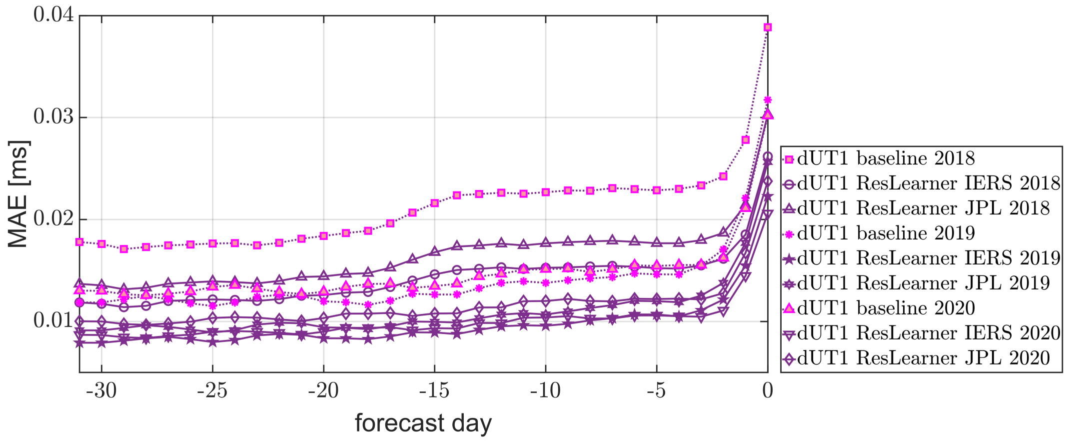
Applying the same ResLearner architecture to these intervals, we get the results displayed in Figures 15 and 16. The results are divided into two parts: days -31 to 0 and days 1 to 31. Two important points can be deduced from these results. First, the accuracies are different from year to year and they do not show a clear reduction with increasing training intervals. This means that ResLearner tends to improve the prediction accuracy even when the training time span is shorter. Thus, the algorithm does not critically depend on the amount of data fed to it (c.f. Kiani-Shahvandi & Soja, 2021). This can be explained by the fact that the architecture is designed in a way that does not include too many learnable parameters, which can therefore be well trained. Second, the anomalous behavior of the polar motion components at day 0 also appears here, suggesting that the problem with rapid data also existed during earlier years.



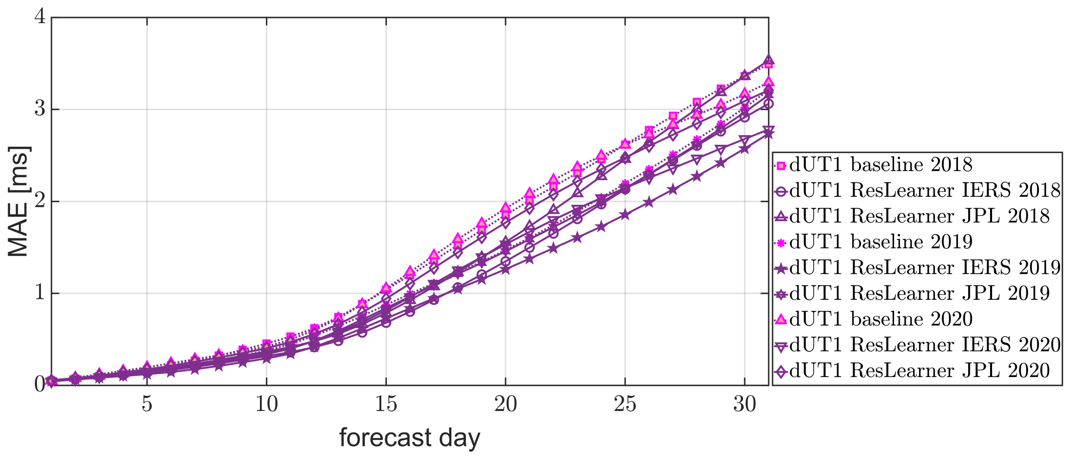
(a)

(b)

Figure 15: Prediction accuracy of hindcast studies S2, S3, and S4 for polar motion components xp and yp, in terms of MAE [mas]. Only the ResLearner is used (but not ResLearner PhycoRNN since they are similar). (a) displays the results for the days -31 to 0 and (b) for the days 1 to 31.



(a)



(b)

Figure 16: Prediction accuracy of hindcast studies S2, S3, and S4 for dUT1, in terms of MAE [ms]. Only the ResLearner is used. (a) displays the results for the days -31 to 0, while (b) for the days 1 to 31.

4.3 Analysis of ESA EOP data: a hindcast study

This analysis corresponds to the last study (S5) in Figure 2, the role of which is to validate our approach against an independent dataset of EOPs. The following points are important regarding this study.

- The prediction horizon is 31 days, i.e., days -15 to 15
- Two baselines are considered: the rapid EOPs as provided by IERS and by ESA EOPs
- The final ESA EOPs are used for evaluation
- Validation is done against both the ideal and realistic ESA hindcast scenarios described in Section 3

We perform three different evaluations, namely:

- evaluation 1: training only on IERS final EOPs up to the end of 2022,
- evaluation 2: training only on IERS final EOPs up to the respective time of ESA EOPs, using retraining at each epoch,
- evaluation 3: training on a combination of IERS and ESA EOPs, similarly with retaining.

The first evaluation is a hindcast study based on the pre-trained models. This means that no retraining is needed and predictions are made all at once. The second evaluation is more of operational nature, although in the past. The training period is thereby assumed to extend from 2015 up to the prediction day. In the third evaluation, IERS data from 2015 up to the end of 2017 are used for the training and first prediction. For each subsequent prediction, the ESA final data are added day-by-day to the training.

We analyze both the ideal and realistic scenarios mentioned in Section 3. First, we discuss the ideal case. The results of these evaluations are shown in Figures 17-18. Considering these results, we would like to highlight the following points: First, ResLearner is able to further improve the prediction accuracy based on ESA data, confirming its flexibility for different datasets. Second, there is not much difference between the results of the three evaluations. Only evaluation 1 presents minor superiority over the other evaluations. This is expected, however, as in this case, the model has seen not only the past but also the future final IERS EOPs. Third, all evaluations, as well as the ideal ESA baseline, show a significant improvement compared to the IERS baseline. Moreover, they show a more realistic behavior of the error of day 0, omitting the anomalous behavior seen in the IERS baseline (the error of day 0 being smaller than that of day -1). Application of ResLearner unmixer here points mostly again towards the EAM as the culprit. Furthermore, it shows that ESA and IERS data are slightly inconsistent at day 0, with the rapid IERS baseline accuracy being better when evaluated against IERS 14 C04. This, however, does not have an impact on the high prediction accuracy of both ESA baseline scenarios, which is close to that achieved with ResLearner.

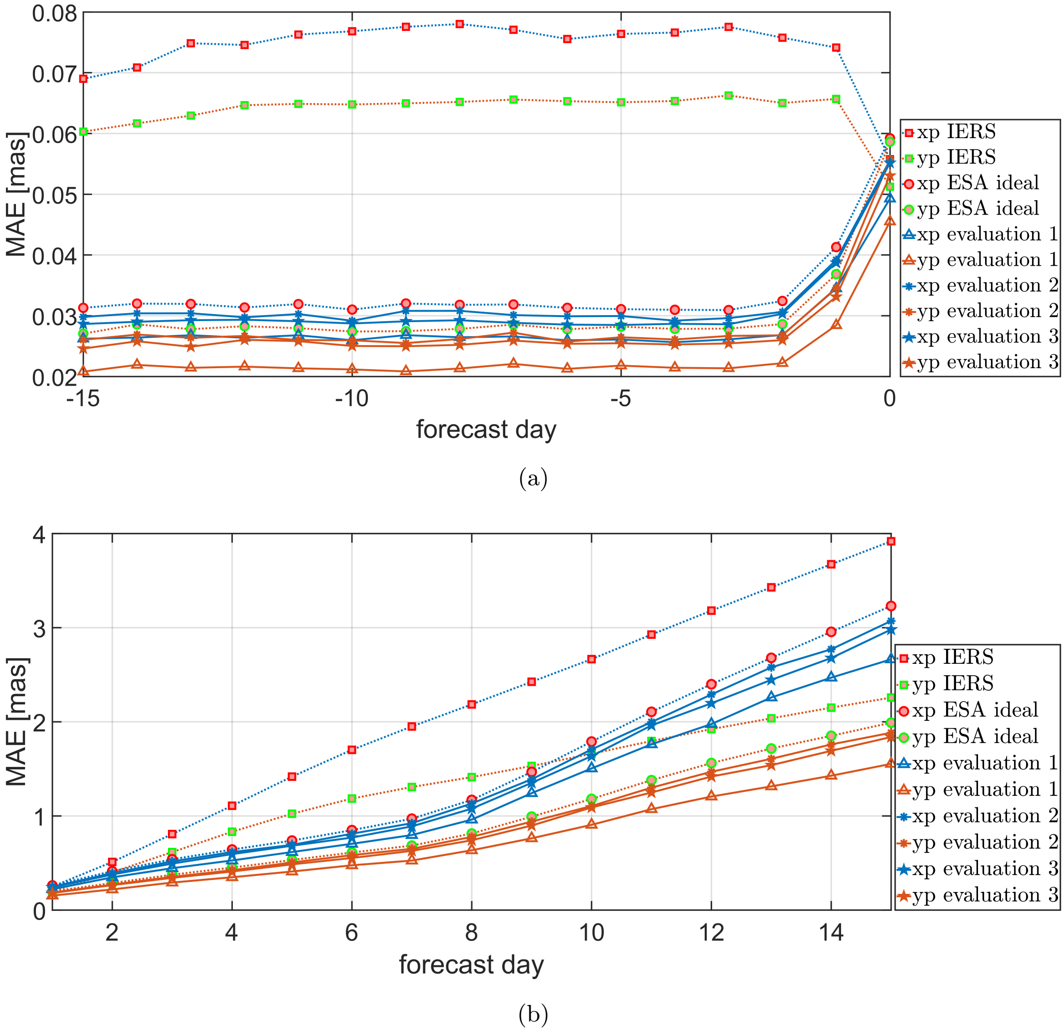
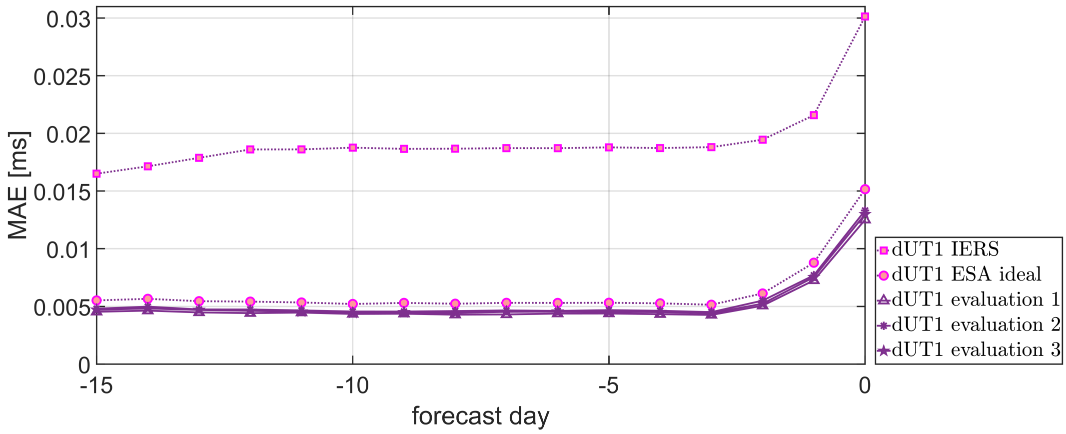
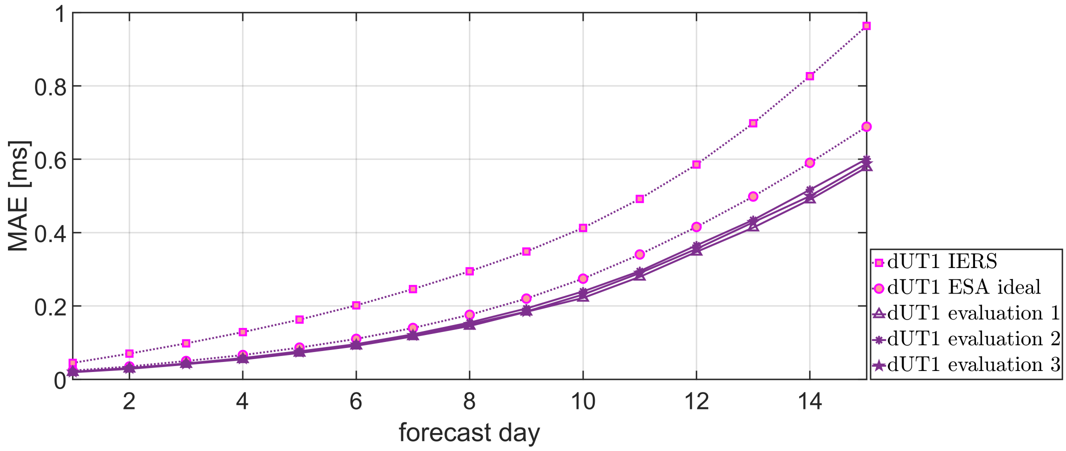


Figure 17: Prediction accuracy of the ResLearner algorithm for polar motion components xp and yp, based on study 5 (S5) and for three different evaluations: 1) training only IERS final EOPs up to the end of 2022, 2) evaluation 2: training only IERS final EOPs up to the respective time of ESA, 3) evaluation 3: training on a combination of IERS and ESA ideal data. Two baselines are presented: rapid IERS and rapid ESA ideal scenario. The data are evaluated against the final ESA ideal data. (a) shows the results for prediction days -31 to 0, while (b) for days 1 to 31.



(a)



(b)

Figure 18: Prediction accuracy of the ResLearner algorithm for dUT1, based on study 5 (S5) and for three different evaluations: 1) training only IERS final EOPs up to the end of 2022, 2) evaluation 2: training only IERS final EOPs up to the respective time of ESA, 3) evaluation 3: training on a combination of IERS and ESA ideal data. Two baselines are presented: rapid IERS and rapid ESA ideal scenario. The data are evaluated against the final ESA ideal data. (a) shows the results for prediction days -31 to 0, while (b) for days 1 to 31.

Figure 19 presents the results of the ESA realistic scenario for dUT1. While there is no significant difference between the ESA ideal and realistic scenarios for polar motion, dUT1 shows a clear reduction in prediction accuracy for days -15 to 0 compared to the ESA ideal scenario. This can be related to the missing of VLBI 24-hour data on these days, as the ESA realistic scenario only considers VLBI intensive sessions and GNSS rapids in the rapid combination. However, the change in prediction accuracy from days 1 to 15 is insignificant.

For ResLearner trained on the ESA realistic data, the prediction horizons between -15 and 0 days show a significant improvement compared to the ESA realistic scenario. This is in contrast to the results achieved by training on the ESA ideal scenario, where the additional improvement achieved by ResLearner is only minor. Thus, the results suggest that ResLearner can contribute to mitigating the effect of the processing latency of 24-hour VLBI sessions, which are crucial for a reliable determination of dUT1.

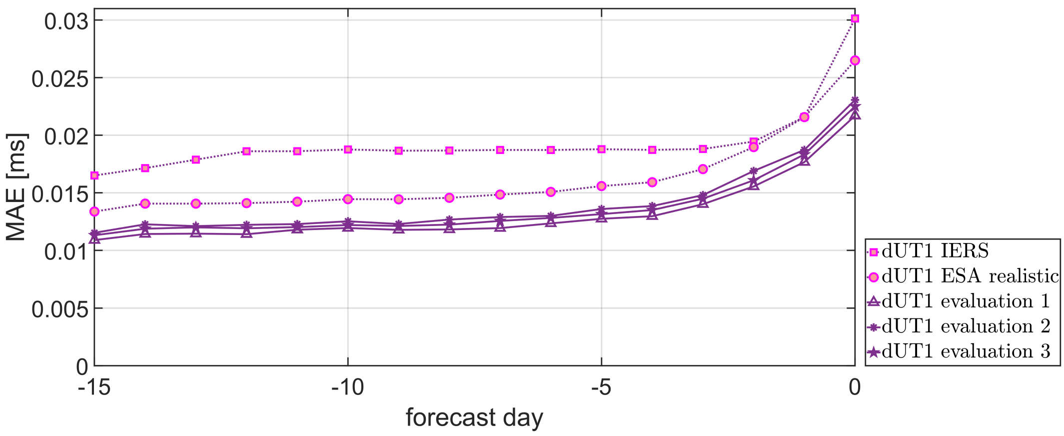


Figure 19: Prediction accuracy of the ResLearner algorithm for dUT1, based on study 5 (S5) and for three different evaluations: 1) training only IERS final EOPs up to the end of 2022, 2) evaluation 2: training only IERS final EOPs up to the respective time of ESA, 3) evaluation 3: training on a combination of IERS and ESA realistic data. Two baselines are presented: rapid IERS and rapid ESA realistic scenario. The data are evaluated against the final ESA realistic data.

4.4 Further discussions and recommendations

Several consequences arise from the results presented above. First, in order to analyze the sensitivity of the anomalous behavior at day 0 between the rapid and final IERS EOP series for evaluation, we evaluate the results of ResLearner and ResLearner PhycoRNN against the IERS 20 C04 series. This is similar to what is presented in Figure 5, but the reference EOP series is different. The results are shown in Figure 20. Comparing Figures 5 and 20, we observe that the anomalous behavior at day 0 is less severe. This further shows the dependence of the results on the version of IERS final and confirms that the choice of reference evaluation series is important when evaluating in general, and in this case especially for day 0. Note that we also trained the algorithms based on the IERS 20 C04 series and observed that the anomalous behavior at day 0 is less severe. This attests to the suitability of IERS 20 C04 to address this problem to a certain extent.

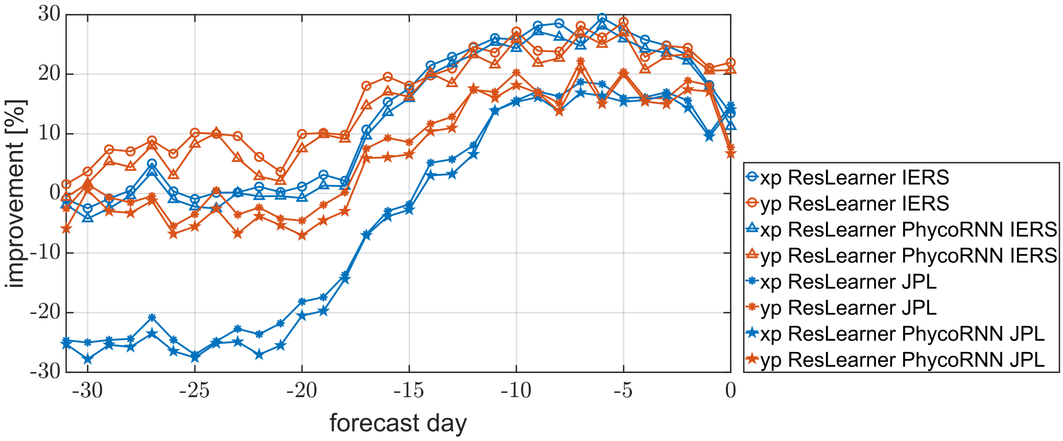


Figure 20: Improvement in prediction accuracy of polar motion components xp, yp for the first study presented in Figure 2, in terms of percentage. This is similar to Figure 5, but evaluated against the IERS 20 C04 instead of IERS 14 C04. Only the days -31 to 0 are shown to check for anomalous behavior at day 0.

In addition, since there are several types of the ResLearner method, we can compute an ensemble of all types based on IERS 20 C04 as target EOPs. A simple weighted ensemble is used, with the weights computed based on the overall prediction performance of individual types of ResLearner. We call this type of ResLearner the full ensemble ResLearner. The results of improvement for the full ensemble ResLearner are shown in Figure 21. The problem at day 0 is almost eliminated and we achieve up to 50% improvement in accuracy compared to the IERS rapid data. Note, however, that the improvements for days 1 to 31 are smaller compared to those presented in Figure 5, thereby suggesting that using the full ensemble approach is only beneficial in days -31 to 0. Crucial to note is that training a similar full ensemble based on IERS 14 C04 is not beneficial as the error would still persist.

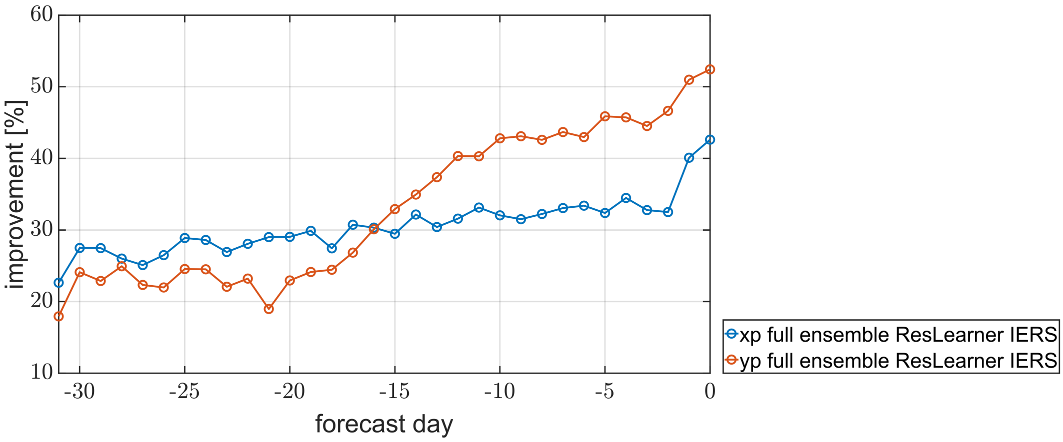


Figure 21: Improvement in prediction accuracy of polar motion components xp, yp for the first study presented in Figure 2, in terms of percentage. This is similar to Figures 5 and 20, but a weighted ensemble of the types of ResLearner algorithm is used. Only the days -31 to 0 are shown to check for anomalous behavior at day 0.

Based on our thorough analyses, we present some recommendations regarding the improvement of rapid EOP data, summarized in Table 3.

Table 3: Recommendations based on the numerical results presented in Section 4.

characteristics	recommendation
type of ResLearner	non-linear ResLearner with self-calibration
most relevant features	EAM, semi-diurnal, diurnal, zonal tides, MEI
EOP series for training and evaluation	IERS 20 C04

5 Conclusions

We devised a new machine learning method called ResLearner for the purpose of reducing errors in rapid EOPs w.r.t. final EOPs. The method is essentially non-linear and has a physically-constrained form called ResLearner PhycoRNN based on coupled oscillatory recurrent neural networks. Additionally, we also investigated the linear form of the method. Unmixing and self-calibration problems are analyzed as well, used for finding the causes of discrepancies between rapid and final EOPs, and calibrating the errors in the input features. Extensive numerical investigations are performed on both IERS and JPL final data, as well as validations against independent series of ESA hindcast experiments. The results show the superiority of non-linear ResLearner compared to the linear methods. Furthermore, ResLearner PhycoRNN can outperform ResLearner in the yp component of polar motion, while ResLearner is better in the xp component. Generally, the improvement in the accuracy of both polar motion components is over 40% across a large portion of the prediction horizon and can reach up to 60%. For dUT1, the improvement in prediction accuracy is smaller, but becomes larger for later prediction days, reaching up to 25%. In this context, validation against the ESA hindcast experiments demonstrates the capability of ResLearner to partially compensate for quality limitations in rapid dUT1 determination that are related to the latency of 24-hour VLBI data. As technical limitations will not allow for a faster availability of these data in the foreseeable future, ResLearner could become a valuable component in enhancing the quality of this parameter crucial for low-latency and real-time applications.

There is an anomalous behavior in the IERS rapid EOP data at day 0, where the consistency with the IERS finals appears to be better than at day -1. The unmixing algorithm suggests that errors in EAM, dominance of GNSS-derived polar motion, and tides are the main causes of this behavior. By applying the ResLearner self-calibration to the data, the errors are reduced and further improvement is achieved. Furthermore, using the IERS 20 C04 series either as the target in the training phase or as reference series for evaluation reduces this anomalous behavior, which suggests the superiority of the IERS 20 C04 over the 14 C04 EOP series. This is further justified when an ensemble of all types of ResLearner methods is used, in which case we no longer observe this anomalous behavior.

We further discussed the importance of geophysical information and found that besides EAM functions, tidal corrections and CI contribute to the prediction performance. Subdiurnal, diurnal, and long-period (zonal) tides in the oceans are all found to be relevant. Furthermore, the multivariate ENSO index is found to be the most relevant CI. Further investigation in this context should focus on each individual component in order to judge whether errors assigned to a certain part of a (conventional) model are actually to be related to it. In this context, feature importance can give hints on where model deficiencies might have an impact on the quality of current EOP determination.

Up to now, the ResLearner-based EOP determination realises a rapid EOP product that does not have a seamless transition from the corresponding final EOPs. This is in contrast to the EOP series realised by the ESA approach, where final and rapid EOPs combined from space-geodetic observations are directly complemented by a prediction that uses the last set of rapid (combined) EOPs as initial values. Further investigation might put focus on incorporating ML-based features already as conditions into the combination of the space-geodetic techniques, thereby realising a seamless EOP time series from the past into the future.

Since the method developed in this paper is based on the concept of physically-constrained neural networks, by modifying the geophysical constraints it can be used for other adjustment and prediction problems as well. One such problem in the field of Earth rotation is the long-term prediction of changes in the length-of-day. We hope that the results presented in this paper stimulate further research in this direction to combine the mathematical rigor of neural networks and the strength of geophysical information.

Acknowledgments

The authors acknowledge the European Space Agency (ESA) for providing series of hindcast experiments derived within the ESA project on “Independent Generation of Earth Orientation Parameters” (ESA-EOP; ESA Contract 4000120430/17/D/SR).

Declarations

Conflict of interest: None

Data availability

The improved rapid EOPs based on the methodology presented in this paper are operationally available on the ETH Zurich Geodetic Prediction Center (GPC) website at `https://gpc.ethz.ch/EOP/Rapid/`. The 14-day forecasts of EAM functions can be accessed at the ETH Zurich GPC website at `https://gpc.ethz.ch/EAM/`. EAM analysis products of GFZ German Research Center for Geosciences are available for download at `http://rz-vm115.gfz-potsdam.de:8080/repository`. IERS rapid and final EOPs (series 14 C04) are available at `https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html`. EOP series 20 C04, consistent with ITRF 2020, can be accessed via `https://hpiers.obspm.fr/iers/eop/eopc04_20/eopc04.1962-now`. The JPL final EOP series can be obtained via `https://eop2-external.jpl.nasa.gov/`. ESA data used in the study has been provided on request for this study (cf. Kehm et al., 2023). The developed software is available at `https://doi.org/10.5281/zenodo.7712379`. Information regarding the rapid files processing strategy can be accessed at `https://maia.usno.navy.mil/ser7/archive.notes` and `https://maia.usno.navy.mil/information/iers-gaz13.txt`. The multivariate ENSO index can be accessed via `https://psl.noaa.gov/enso/mei/` and the MJI data via `https://www.psl.noaa.gov/mjo/mjoindex/`. Data regarding NAI are available at `https://www.ncei.noaa.gov/access/monitoring/nao/`.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., … Zheng, X. (2016). Tensorflow: A system for large-scale machine learning. In *12th usenix symposium on operating systems design and implementation (OSDI*

16), usenix association (p. 265-283). doi: https://dl.acm.org/doi/10.5555/3026877.3026899

Barnes, R., Hide, R., White, A., & Wilson, C. (1983). Atmospheric angular momentum fluctuations, length-of-day changes and polar motion. *Proceedings of the Royal Society A*, *387*. doi: https://doi.org/10.1098/rspa.1983.0050

Bengio, Y., Lodi, A., & Prouvost, A. (2021). Machine learning for combinatorial optimization: A methodological tour d'horizon. *European Journal of Operational Research*, *290*(2), 405-421. doi: https://doi.org/10.1016/j.ejor.2020.07.063

Bishop, C. (2006). *Pattern recognition and machine learning.* Springer Information Science and Statistics.

Bizouard, C., Lambert, S., Gattano, C., Becker, O., & Richard, J. (2019). The IERS EOP 14C04 solution for Earth orientation parameters consistent with itrf 2014. *Journal of Geodesy*, *93*, 621–633. doi: https://doi.org/10.1007/s00190-018-1186-3

Bizouard, C., & Seoane, L. (2010). Atmospheric and oceanic forcing of the rapid polar motion. *Journal of Geodesy*, *84*, 19–30. doi: https://doi.org/10.1007/s00190-009-0341-2

Brzezinski, A., & Nastula, J. (2002). Oceanic excitation of the Chandler wobble. *Advances in Space Research*, *30*, 195–200. doi: https://doi.org/10.1016/S0273-1177(02)00284-3

Chao, B. (1984). Interannual length-of-day variation with relation to the southern oscillation/El Nino. *Geophysical Research Letters*, *11*, 541–544. doi: https://doi.org/10.1029/GL011i005p00541

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*, *78*, 343–353. doi: https://doi.org/10.1007/s00190-004-0411-4

Dahlen, F. (1976). The passive influence of the oceans upon the rotation of the Earth. *Geophysical Journal International*, *46*, 363–406. doi: https://doi.org/j.1365-246X.1976.tb04163.x

Dick, R., & Thaller, D. (2018). *IERS Annual Report 2018.* IERS.

Dickman, S. (2003). Evaluation of "effective angular momentum function" formulations with respect to core-mantle coupling. *Journal of Geophysical Research: Solid Earth*, *108*. doi: https://doi.org/10.1029/2001JB001603

Dill, R., & Dobslaw, H. (2010). Short-term polar motion forecasts from Earth system modeling data. *Journal of Geodesy*, *84*, 529–536. doi: https://doi.org/10.1007/s00190-010-0391-5

Dill, R., Dobslaw, H., Hellmers, H., Kehm, A., Bloßfeld, M., Thomas, M., … Schönemann, E. (2020). Evaluating processing choices for the geodetic estimation of Earth orientation parameters with numerical models of global geophysical fluids. *Journal of Geophysical Research: Solid Earth*, *125*. doi: https://doi.org/10.1029/2020JB020025

Dill, R., Dobslaw, H., & Thomas, M. (2019a). Improved 90-day Earth orientation predictions from angular momentum forecasts of atmosphere, ocean, and terrestrial hydrosphere. *Journal of Geodesy*, *93*, 287–295. doi: https://doi.org/10.1007/s00190-018-1158-7

Dill, R., Saynisch-Wagner, J., Irrgang, C., & Thomas, M. (2021). Improving atmospheric angular momentum forecasts by machine learning. *Earth and Space Science*, *8*(12). doi: https://doi.org/10.1029/2021EA002070

Di Lorenzo, E., Xu, T., Zhao, Y., Newman, M., Capotondi, A., Stevenson, S., … Zhang, H. (2023). Modes and mechanisms of pacific decadal-scale variability. *Annual Review of Marine Science*, *15*(1), null. doi: https://doi.org/10.1146/annurev-marine-040422-084555

Dobslaw, H., & Dill. (2018). Predicting Earth orientation changes from global forecasts of atmosphere-hydrosphere dynamics. *Advances in Space Research*, *61*, 1047–1054. doi: https://doi.org/10.1016/j.asr.2017.11.044

Dobslaw, H., & Dill, R. (2019b). Effective angular momentum functions from Earth system modelling at GeoForschungsZentrum in Potsdam. In *Product description document: GFZ German Research Centre for Geosciences Department 1: Geodesy Section 1.3: Earth System Modelling.* Retrieved from `http://rz-vm115.gfz-potsdam.de:8080/repository`

Dobslaw, H., Dill, R., Grötzsch, A., Brzezinski, A., & Thomas, M. (2010). Seasonal polar motion excitation from numerical models of atmosphere, ocean, and continental hydrosphere. *Journal of Geophysical Research: Solid Earth*, *115*. doi: https://doi.org/10.1029/2009JB007127

Fischler, M., & Bolles, R. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, *24*, 381-395. doi: https://doi.org/10.1145/358669.358692

Ganaie, M., Hu, M., Malik, A., Tanveer, M., & Suganthan, P. (2022). Ensemble deep learning: A review. *Engineering Applications of Artificial Intelligence*, *115*. doi: https://doi.org/10.1016/j.engappai.2022.105151

Geneva, N., & Zabaras, N. (2020). Modeling the dynamics of PDE systems with physics-constrained deep auto-regressive networks. *Journal of Computational Physics*, *403*, 109056. doi: https://doi.org/10.1016/j.jcp.2019.109056

Gross, R. (1997). Earth rotation variations-long period. *Treatise on Geophysics*, *3*, 239–294. doi: https://doi.org/10.1016/B978-044452748-6.00057-2

Gross, R. (2008). An improved empirical model for the effect of long-period ocean tides on polar motion. *Journal of Geodesy*, *83*, 635–644. doi: https://doi.org/10.1007/s00190-008-0277-y

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE conference on computer vision and pattern recognition (CVPR).*

Hendon, H. (1995). Length of day changes associated with the Madden-Julian Oscillation. *Journal of the Atmospheric Sciences*, *52*, 2373–2383. doi: https://doi.org/10.1175/1520-0469(1995)052⟨2373:LODCAW⟩2.0.CO;2

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*, 1735–1780. doi: https://doi.org/10.1162/neco.1997.9.8.1735

Huber, P. (1964). Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, *35*, 73-101. doi: https://doi.org/10.1214/aoms/1177703732

Huber, P. (1973). Robust regression: Asymptotics, conjectures and Monte Carlo. *The Annals of Statistics*, *1*, 799-821. doi: https://doi.org/10.1214/aos/1176342503

Kalarus, M., Schuh, H., Kosek, W., Akyilmaz, O., Bizouard, C., Gambis, D., ... Zotov, L. (2010). Achievements of the Earth orientation parameters prediction comparison campaign. *Journal of Geodesy*, *84*, 587–596. doi: https://doi.org/10.1007/s00190-010-0387-1

Kehm, A., Hellmers, H., Bloßfeld, M., Dill, R., Angermann, D., Seitz, F., ... Enderle, W. (2023). Combination strategy for consistent final, rapid and predicted Earth rotation parameters. *Journal of Geodesy*, *97*. doi: https://doi.org/10.1007/s00190-022-01695-w

Kiani Shahvandi, M., Gou, J., Schartner, M., & Soja, B. (2022). Data driven approaches for the prediction of Earth's effective angular momentum functions. *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, *127*, 6550–6553. doi: https://doi.org/10.1109/IGARSS46834.2022.9883545

Kiani-Shahvandi, M., Schartner, M., & Soja, B. (2022). Neural ODE differential learning and its application in polar motion prediction. *Journal of Geophysical Research: Solid Earth*, *127*. doi: https://doi.org/10.1029/2022JB024775

Kiani-Shahvandi, M., & Soja, B. (2021). Small geodetic datasets and deep networks: Attention-based residual LSTM autoencoder stacking for geodetic time series.

In *7th international conference on machine learning, optimization, and data science.* doi: https://doi.org/10.1007/978-3-030-95467-3_22

Kiani-Shahvandi, M., & Soja, B. (2022). 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. *Advances in Space Research*, *70*. doi: https://doi.org/10.1016/j.asr.2022.05.042

Kiladis, G., Dias, J., Straub, K., Wheeler, M., Tulich, S., Kikuchi, K., ... Ventrice, M. (2014). A comparison of OLR and circulation-based indices for tracking the MJO. *Monthly Weather Review*, *142*(5), 1697–1715. doi: https://doi.org/10.1175/MWR-D-13-00301.1

Kingma, D., & Ba, J. (2015). Adam: A method for stochastic optimization. In *International conference on learning representations (iclr).*

Kur, T., Dobslaw, H., Śliwińska, J., Nastula, J., Wińska, M., & Partyka, A. (2022). Evaluation of selected short-term predictions of UT1-UTC and LOD collected in the second Earth orientation parameters prediction comparison campaign. *Earth, Planets and Space*, *74*. doi: https://doi.org/10.1186/s40623-022-01753-9

Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2016). Simple and scalable predictive uncertainty estimation using deep ensembles. In *Advances in neural information processing systems.* Curran Associates, Inc. Retrieved from `https://doi.org/10.48550/arXiv.1612.01474`

Lambeck, K. (1980). The Earth's variable rotation: Geophysical causes and consequences. *Cambridge University Press*. doi: https://doi.org/10.1017/CBO9780511569579

Liu, D., & Nocedal, J. (1989). On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, *45*, 503–528. doi: https://doi.org/10.1007/BF01589116

Liu, S., & Dobriban, E. (2020). Ridge regression: Structure, cross-validation, and sketching. In *International conference on learning representations.* Retrieved from `https://openreview.net/forum?id=HklRwaEKwB`

Luo, J., Chen, W., Ray, J., & Li, J. (2022). Short-term polar motion forecast based on the Holt-Winters algorithm and angular momenta of global surficial geophysical fluids. *Surveys in Geophysics*. doi: https://doi.org/10.1007/s10712-022-09733-0

Maksymilian, W., & Chen, K. (2020). Feature importance ranking for deep learning. In *Advances in neural information processing systems 33.* Morgan Kaufmann Publishers.

Marquardt, D., & Snee, R. (1975). Ridge regression in practice. *The American Statistician*, *29*, 3-20.

Mazzarella, A. (2007). The 60-year solar modulation of global air temperature: the Earth's rotation and atmospheric circulation connection. *Theoretical and Applied Climatology*, 193–199. doi: https://doi.org/10.1007/s00704-005-0219-z

Minderer, M., Djolonga, J., Romijnders, R., Hubis, F., Zhai, X., Houlsby, N., ... Lucic, M. (2021). Revisiting the calibration of modern neural networks. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, & J. W. Vaughan (Eds.), *Advances in neural information processing systems* (Vol. 34, pp. 15682–15694). Curran Associates, Inc. Retrieved from `https://proceedings.neurips.cc/paper/2021/file/8420d359404024567b5aefda1231af24-Paper.pdf`

Modiri, S., Belda, S., Heinkelmann, R., Hoseini, M., Ferrandiz, J., & Schuh, H. (2018). Polar motion prediction using the combination of SSA and Copula-based analysis. *Earth, Planets and Space*, *70*. doi: https://doi.org/10.1186/s40623-018-0888-3

Modiri, S., Belda, S., Hoseini, M., Heinkelmann, R., Ferrandiz, J., & Schuh, H. (2020). A new hybrid method to improve the ultra-short-term prediction of LOD. *Journal of Geodesy*, *94*. doi: https://doi.org/10.1007/s00190-020-01354-y

Nastula, J., & Ponte, R. (1999). Further evidence for oceanic excitation of polar motion. *Geophysical Journal International*, *139*, 123–130. doi: https://doi.org/10.1046/j.1365-246X.1999.00930.x

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., … Chintala, S. (2019). Pytorch: An imperative style, high-performance deep learning library. In *Advances in neural information processing systems 32* (pp. 8024–8035). Curran Associates, Inc. Retrieved from `http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf`

Petit, G., & Luzum, B. (2010). *IERS Technical Note 36*. IERS.

Ratcliff, J., & Gross, R. (2022). Combinations of Earth orientation measurements: SPACE2021, COMB2021, and POLE2021. *JPL Publications*.

Raut, S., Modiri, S., Heinkelmann, R., Balidakis, K., Belda, S., Kitpracha, C., & Schuh, H. (2022). Investigating the relationship between length of day and El-Niño using wavelet coherence method. In (pp. 1–6). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: https://doi.org/10.1007/1345_2022_167

Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning internal representations by error propagation. In *Parallel distributed processing* (pp. 318–362). MIT Press. doi: https://dl.acm.org/doi/10.5555/104279.104293

Rusch, T., & Mishra, S. (2021). Coupled oscillatory recurrent neural network (coRNN): An accurate and (gradient) stable architecture for learning long time dependencies. In *International conference on learning representations.*

Rusch, T., Mishra, S., Erichson, N., & Mahoney, M. (2022). Long expressive memory for sequence modeling. In *International conference on learning representations.*

Sullivan, T. (2015). *Introduction to uncertainty quantification* (Vol. 63). Springer, Texts in Applied Mathematics.

Sun, Q., Zhou, W., & Fan, J. (2020). Adaptive Huber regression. *Journal of the American Statistical Association*, *115*. doi: https://doi.org/10.1080/01621459.2018.1543124

Szandała, T. (2021). Review and comparison of commonly used activation functions for deep neural networks. *Bio-inspired Neurocomputing*, 203–224. doi: https://doi.org/10.1007/978-981-15-5495-7_11

Teunissen, P. (2003). *Adjustment theory: an introduction.* VSSD Series on Mathematical Geodesy and Positioning.

Timmermann, A., An, S., Kug, J., Jin, F., Cai, W., Capotondi, A., … Zhang, X. (2018). El Niño–Southern Oscillation complexity. *Nature*, *559*(7715), 535–545. doi: https://doi.org/10.1038/s41586-018-0252-6

Visbeck, M., Hurrell, J., Polvani, L., & Cullen, H. (2001). The North Atlantic Oscillation: Past, present, and future. *Proceedings of the Natural Academy of Sciences*, *98*, 12876–12877. doi: https://doi.org/10.1073/pnas.231391598

Volland, H. (1996). Atmosphere and Earth's rotation. *Surveys in Geophysics*, *17*, 101–144. doi: https://doi.org/10.1007/BF01904476

Wolter, K., & Timlin, M. (1993). Monitoring ENSO in COADS with a seasonally adjusted principal component index. Norman: Proceedings of the 17th Climate Diagnostics Workshop.

Xu, X., Zhou, Y., Duan, P., Fang, M., Kong, Z., Xu, C., & An, X. (2022). Contributions of oceanic and continental AAM to interannual variation in ΔLOD with the detection of 2020–2021 La Nina event. *Journal of Geodesy*, *96*. doi: https://doi.org/10.1007/s00190-022-01632-x

Zhang, R., Liu, Y., & Sun, H. (2020). Physics-informed multi-LSTM networks for metamodeling of nonlinear structures. *Computer Methods in Applied Mechanics and Engineering*, *369*, 113226. doi: https://doi.org/10.1016/

j.cma.2020.113226

ResLearner: geophysically-informed machine learning for improving the accuracy of rapid Earth orientation parameters

Mostafa Kiani Shahvandi[1], Robert Dill[2], Henryk Dobslaw[2], Alexander Kehm[3], Mathis Bloßfeld[3], Matthias Schartner[1], Siddhartha Mishra[4], Benedikt Soja[1]

[1]Institute of Geodesy and Photogrammetry, ETH Zurich
[2]Section for Earth System Modelling, GFZ German Research Centre for Geosciences
[3]Deutsches Geodätisches Forschungsinstitut (DGFI-TUM), Technical University of Munich
[4]Seminar for Applied Mathematics, Department of Mathematics, and ETH AI Center, ETH Zurich

Key Points:

- We introduce a novel machine learning algorithm named ResLearner to improve the accuracy of rapid Earth orientation parameters
- We also present geophysically-constrained ResLearner, using Earth's effective angular momentum functions, tides, and climatic indices
- Besides prediction, ResLearner is also able to effectively correct deficits in rapidly processed EOPs with respect to final EOPs

Corresponding author: Mostafa Kiani Shahvandi, mkiani@ethz.ch

Abstract

Rapid provision of Earth Orientation Parameters (EOPs, here polar motion and dUT1) is indispensable in many geodetic applications and also for spacecraft navigation. There are, however, discrepancies between the rapid EOPs and the final EOPs that have a higher latency, but the highest accuracy. To reduce these discrepancies, we focus on a data-driven approach, present a novel method named ResLearner, and use it in the context of deep ensemble learning. Furthermore, we introduce a geophysically-constrained approach for ResLearner. We show that the most important geophysical information to improve the rapid EOPs is the effective angular momentum functions of atmosphere, ocean, land hydrology, and sea level. In addition, semi-diurnal, diurnal, and long-period tides coupled with prograde and retrograde tidal excitations are important features. The influence of some climatic indices on the prediction accuracy of dUT1 is discussed and El Niño Southern Oscillation is found to be influential. We developed an operational framework, providing the improved EOPs on a daily basis with a prediction window of 63 days to fully cover the latency of final EOPs. We show that under the operational conditions and using the rapid EOPs of the International Earth Rotation and Reference Systems Service (IERS) we achieve improvements as high as 60%, thus significantly reducing the differences between rapid and final EOPs. Furthermore, we discuss how the new final series IERS 20 C04 is preferred over 14 C04. Finally, we compare against EOP hindcast experiments of European Space Agency, on which ResLearner presents comparable improvements.

Plain Language Summary

The International Earth Rotation and Reference Systems Service (IERS) provides rapid Earth Orientation Parameters (EOPs) using different space geodetic techniques to bridge the latency of the final, most accurate EOPs solution. However, these rapid EOPs are not in full agreement with the final EOPs. In order to reduce the differences between the rapid and final EOPs, we focus on the application of machine learning and present a novel method named ResLearner, which is based on geodetic data and geophysical constraints. We present the method in the context of deep ensemble learning, focusing on a prediction window of 63 days. We also attempt to link informative geophysical effects to these discrepancies. We show that they are linked to a mixture of atmospheric, oceanic, hydrological, and sea level effective angular momentum functions, dominance of the GNSS-derived polar motion, and various short- and long-term tidal excitations. El Niño Southern Oscillation is also relevant for dUT1 prediction. The methodology can provide significant improvements of up to 60% in operational settings with respect to rapid EOPs provided by IERS. Additional validation is done by using the data of Jet Propulsion Laboratory final EOP series and also EOP series provided by the European Space Agency.

1 Introduction

Earth Orientation Parameters (EOPs) represent variations of Earth's rotation axis in time (Lambeck, 1980; Gross, 1997). Among these parameters, polar motion components, (xp, yp), and the difference between universal time and coordinated universal time, dUT1, are of great interest, because of their importance for applications such as satellite and spacecraft navigation and orientation of deep-space telescopes (Dobslaw & Dill, 2019b). These EOPs are routinely provided at different latencies, of which two are considered here: rapid and final (Kehm et al., 2023). Final EOPs require a combination of different data sources (Bizouard et al., 2019; Ratcliff & Gross, 2022) such as Global Navigation Satellite Systems (GNSS), Very Long Baseline Interferometry (VLBI), and Lunar and Satellite Laser Ranging (LLR, SLR). Some of the techniques require longer processing time and therefore, delays of up to several weeks are expected, by which the data are provided to the scientific community. The current uncertainty level in final EOPs provided by International Earth Rotation and Reference Systems Service (IERS) is around 20-30 micro-arcseconds [µas] for polar motion components, and 9-10 micro-seconds [µs] for dUT1 in terms of formal errors.

Rapid EOPs provided by the IERS are determined through a combination of the most recent Global Positioning System (GPS) and VLBI 24-hour and intensive sessions data, augmented with Atmospheric Angular Momentum (AAM). These rapid data contain polar motion components (xp, yp) and dUT1, bridging the latency of final EOPs by providing 90 days of rapid combined EOPs to the past and 90 days of predicted EOPs into the future, with respect to the date the data are provided at. The uncertainty in the estimations is also provided. Currently, the level of these uncertainties varies across different days and also for combined and predicted EOPs. For the rapid combined EOPs, it can be several times bigger than that of final EOPs, but mostly below 1 milli-arcseconds [mas]. Predictions into the future are based on extrapolation of mathematical functions such as harmonic models. For longer prediction horizons, the accuracy is degraded significantly and can be up to several milli-arcseconds.

There are some routines performed on the mentioned datasets before operationally providing the rapid EOPs data. These include systematic corrections and smoothing. Systematic corrections are used to mitigate the impact of different VLBI baseline solutions on polar motion and dUT1. For instance, based on different VLBI solutions of the United States Naval Observatory (USNO), corrections are added to the polar motion and dUT1 of 24-hour sessions, and similar corrections to dUT1 of intensive sessions. Smoothing algorithms are applied to remove the high-frequency noise, usually by a Lagrangian interpolation scheme. It is important to note that ocean tidal effects are dealt with in the rapid EOPs as otherwise, the accuracy would be significantly degraded because of the systematic effect of tides. Furthermore, AAM data that are used for the improved determination of rapid EOPs contain some errors. Errors in the removal of tides and also the addition of AAM with its associated errors would result in inaccuracies in the rapid data, and therefore, inconsistencies w.r.t the final EOPs. These discrepancies can easily exceed the current uncertainty level of final polar motion and dUT1 mentioned above, thus suggesting the need for some type of calibration.

There are several deficiencies in the rapid data that are currently provided by the IERS. First, as mentioned the errors in the removal of tides can propagate to the rapid EOPs. Furthermore, only AAM is used, which is essentially one type of the Effective Angular Momentum (EAM) functions (Barnes et al., 1983). It is shown that Oceanic Angular Momentum (OAM), Hydrological Angular Momentum (HAM), and Sea Level Angular Momentum (SLAM) can have a non-negligible effect on polar motion and dUT1 as well (Dahlen, 1976; Nastula & Ponte, 1999; Brzezinski & Nastula, 2002; Chin et al., 2004; Gross, 2008; Dobslaw et al., 2010; Dill & Dobslaw, 2010; Bizouard & Seoane, 2010; Luo et al., 2022; Kiani-Shahvandi et al., 2022). Furthermore, phenomena such as El Niño Southern Oscillation (ENSO) can have some influence on the rate of dUT1 (Raut et al., 2022; Xu et al., 2022). This can be analyzed using climatic indices (CI) like the multivariate ENSO index (MEI, Wolter & Timlin, 1993), the Madden Julian Oscillation index (MJI, Kiladis et al., 2014), and the North Atlantic Oscillation index (NAI, Visbeck, Hurrell, Polvani, & Cullen, 2001). It is important to mention that the included AAM may not have fully covered the atmospheric effects and a calibration is also needed for this. In addition, the effect of EAM functions is non-tidal, but it can get mixed with the tidal effects during the application of routines. Disentangling the causes of discrepancies between rapid and final EOPs could be challenging and might require specifically-designed algorithms, especially in the absence of physical or analytical models for calibration. As the mixture of tidal and non-tidal effects, systematic corrections, and smoothing can be in a non-linear fashion, one needs to potentially use non-linear models for the purpose of disentanglement. Furthermore, the historical data of rapid EOPs can be utilized to present data-driven approaches that eliminate the need for an analytical calibration approach. These arguments imply that a machine learning algorithm is potentially well suitable for this problem, which is the approach followed in this paper.

There have been successful applications of machine learning for the analysis and prediction of EOPs (Dill et al., 2021; Kiani-Shahvandi & Soja, 2021, 2022; Kiani-Shahvandi et al., 2022). Here, however, we need to consider the specific aspects of the problem and develop a new machine learning algorithm. These specific aspects include 1) the calibration characteristic, 2) the need for non-linear uncertainty estimation, and 3) the importance analysis of different features included in the model.

The first aspect of the problem, namely the calibration characteristic, relates to the fact that the goal of the problem is to reduce the discrepancies between rapid and final EOPs, or in other words, calibration of rapid EOPs w.r.t final EOPs. This implies that the input to the machine learning model should contain the rapid EOPs themselves. These rapid EOPs are already close to the final EOPs in a sense, therefore making the problem similar to an identity mapping by machine learning. This can be difficult for non-linear machine learning algorithms (He et al., 2016), and it has been shown that a better approach would be to consider a residual learning framework (He et al., 2016). Inspired by this approach, we develop our new method in a residual learning manner, in which the overall output (final EOPs) is the summation of rapid EOPs and the output a neural network (having rapid EOPs and other geophysical information either as inputs or constraints). The mentioned neural network can then learn the calibration, enabling us also to use further geophysical information and constraints in the model. Note that self-calibration algorithms can also be considered (Minderer et al., 2021), in which the errors in different variables in the model are potentially reduced by trying to simultaneously learn the calibration effects.

The second aspect of the problem, i.e., uncertainty estimation, is an important task in the field of geodetic science (Kiani-Shahvandi & Soja, 2022), as these uncertainties provide a measure of the reliability of predictions. However, this can be challenging because of the potential non-linearity in neural networks. In this paper, deep ensembles (Lakshminarayanan et al., 2016; Ganaie et al., 2022) are used, which can reduce the epistemic uncertainty in the models. In deep ensembles, a series of neural networks are simultaneously trained to find the mean and standard deviation in the predictions. Since the output is the average of the predictions of all models, the epistemic uncertainty is reduced and mainly the aleatoric uncertainty remains (due to the uncertainty of input data).

Finally, it is important to use algorithms that support the importance analysis of different variables included in the model. Using this approach, we are able to analyze the potential sources of errors in the rapid EOPs.

The following points summarize the goals of the current paper:

- Developing a new machine learning algorithm specifically designed for the problem of improving rapid EOPs accuracy, which can also provide information on uncertainties in the predictions,
- Using geophysically-constrained neural networks as an additional approach in the context of the method,
- Analyzing the geophysical causes of discrepancies between rapid and final EOPs.

The rest of this paper is organized as follows. In Section 2, the ResLearner methodology is introduced. In Section 3, the data used for the numerical results presented in the paper are described. Section 4 is devoted to results and discussions. Conclusions are given in Section 5.

2 ResLearner methodology

This section describes the ResLearner method, including the general approach and its architecture.

2.1 Introducing ResLearner

As mentioned in Section 1, the idea of ResLearner is to calibrate the rapid EOPs (henceforward denoted by R) with respect to the final EOPs (denoted by F) in a residual manner using neural networks (NN). This implies that the conceptual representation of ResLearner can be described by Equation (1)

$$F = R + \mathrm{NN}(\theta, R, X) \tag{1}$$

in which NN is a neural network with parameters θ, and X a set of geophysical data. In the present study, X includes EAM functions (AAM, OAM, HAM, and SLAM), tides, tidal excitations, and MEI, MJI, and NAI. For the architecture of the neural network NN, we have observed that a nonlinear Multi-Layer Perceptron (MLP, Bishop, 2006) with two layers is sufficient to produce the best results. The first and second layers have 1 and 63 hidden neurons (for predicting 63 days), respectively. The activation function of the first layer is tangent hyperbolic, whereas for the second layer, it is linear. An important point regarding the architecture is that linear models can also present competitive results (Kiani-Shahvandi et al., 2022). For the purpose of comparison of the architectures, we use three different linear models: Ridge regression with cross-validation, (RidgeCV, Marquardt & Snee, 1975; S. Liu & Dobriban, 2020), Random Sample Consensus (RANSAC, Fischler & Bolles, 1981), and Ordinary Least Squares (OLS, Teunissen, 2003). The reason for this choice is that RidgeCV and RANSAC are robust against outliers and less sensitive to the possible high variability of rapid data across different days. Out of these, OLS is the simplest method that can present competitive results. Note that we analyzed several other algorithms including Huber (Huber, 1964, 1973; Sun et al., 2020), but they turned out to be computationally expensive and less accurate.

2.2 ResLearner in deep ensembles

We use ResLearner in the context of deep ensembles (Lakshminarayanan et al., 2016). Therefore, a series of neural networks are trained simultaneously based on the same data, and the final prediction would be the average of the prediction of all the individual models. This reduces the epistemic uncertainty (Sullivan, 2015), which is due to errors in the utilized model. The mathematical formulation of deep ensembles (Lakshminarayanan et al., 2016) is based on the assumption that the data can be represented by a heteroscedastic Gaussian distribution. The variance and mean of the distribution are then solved for, following the minimization of the logarithm of the likelihood function $\ell(F, R, X)$ as the loss function. The formulation of the deep ensembles for the calibration of rapid EOPs is given in Equations (2a)-(2f).

$$\mu_j(R,X) = \mathrm{NN}_\mu(\theta_{\mu,j}, R, X) \tag{2a}$$
$$\sigma_j^2(R,X) = \log(1+\exp(\mathrm{NN}_\sigma(\theta_{\sigma,j}, R, X))) + \epsilon \tag{2b}$$
$$\ell_j(F,R,X) = \frac{1}{2}\log\sigma_j^2(R,X) + \frac{1}{2}\frac{(F-R-\mu_j(R,X))^2}{\sigma_j^2(R,X)} \tag{2c}$$
$$\ell_j(F,R,X) \longrightarrow \text{minimize} \tag{2d}$$
$$\mu(R,X) = \frac{1}{M}\sum_{j=1}^{M}\mu_j(R,X) \tag{2e}$$
$$\sigma^2(R,X) = -\mu^2(R,X) + \frac{1}{M}\sum_{j=1}^{M}\sigma_j^2(R,X) + \mu_j^2(R,X) \tag{2f}$$

where $\mu(R,X)$ and $\sigma^2(R,X)$ are the ensemble mean and variance, being the average of M individual members of the ensembles with mean and variance $\mu_j(R,X)$ in Equation (2a) and $\sigma_j^2(R,X)$ in Equation (2b), respectively. In our case, we observed that $M = 10$ is sufficient and results in the highest accuracy. Using significantly more than 10 models seems to be unnecessary, while being drastically more computationally expensive, and at the same time, resulting in no significant gains in accuracy (below the current uncertainty level in EOPs). $\mu_j(R,X)$ and $\sigma_j^2(R,X)$ are modelled by two different neural networks $\mathrm{NN}_\mu(\theta_{\mu,j}, R, X)$ and $\mathrm{NN}_\sigma(\theta_{\sigma,j}, R, X)$ with different learnable parameters $\theta_{\mu,j}$ and $\theta_{\sigma,j}$, respectively, as in Equations (2a) and (2b). Since the variance has to be positive, the softplus function (Szandała, 2021) is applied to the neural network $\mathrm{NN}_\sigma(\theta_{\sigma,j}, R, X)$, i.e., Equation (2b). The term ϵ is a constant for numerical stability. In our problem, we observed that a value of $\epsilon = 10^{-8}$ performs sufficiently well. The loss function $\ell_j(F,R,X)$ is minimized for each individual model separately using Adam optimizer (Kingma & Ba, 2015) with 200 epochs. Finally, it is worthwhile to mention that we implement the method using the TensorFlow library in Python (Abadi et al., 2016).

2.3 Unmixing and self-calibration approaches: geophysical information and constraints

In order to investigate the causes of discrepancies between rapid and final EOPs, one can explicitly model some of the known effects. Here, we model the effect of errors in EAM functions, ocean tides, and tidal excitations. The discrepancies between rapid and final polar motion, denoted by δxp and δyp, and rapid and final dUT1, denoted by δdUT1, are the sum of individual discrepancies due to EAM functions δEAM, ocean tides δT, tidal excitations δTE (for polar motion), and additional effects δU, which include smoothing, systematic correction, and unknown effects. δEAM, δT, and δTE are related to the variable X in the neural network in Equation (1). It is also important to note that the component-wise summation of individual EAM functions is used (Kiani-Shahvandi et al., 2022).

Both the polar motion components and dUT1 are affected by ocean tides and libration in terms of diurnal and subdiurnal variations (Sections 5.5 and 8.2 of Petit & Luzum, 2010). Moreover, polar motion is affected by long-period ocean (both prograde and retrograde) tides which are conventionally modelled with periods from 9 days to 18.6 years (Section 8.3 of Petit & Luzum, 2010). However, dUT1 is affected by zonal tides (i.e., the effect of tidal deformation), which are modelled with periods from 5 days to 18.6 years (Section 8.1 of Petit & Luzum, 2010).

The general approach to include the tidal effects in our model is to consider the harmonic functions with fixed frequencies through Delaunay parameters (Petit & Luzum, 2010), but with variable, estimable amplitudes. This is due to the fact that in rapid EOPs tides are already taken care of, and we need to compensate for the potential erroneous effect of tides included in the model. Therefore, δT and δTE can be modelled as in Equation (3)

$$\delta\mathrm{T},\ \delta\mathrm{TE} = \sum_{i=1}^{K} A_i \cos\Theta(t) + B_i \sin\Theta(t) \tag{3}$$

in which K is the number of tidal constituents considered, A and B the coefficients that should be determined by the neural networks, and $\Theta(t)$ the time-dependent argument of the harmonic functions based on the Delaunay parameters (Petit & Luzum, 2010). In the case of subdiurnal polar motion and dUT1, $K = 30$ constituents are added as features for each of xp, yp, and dUT1. For the diurnal tides, this number is $K = 41$ for each EOP. For the long period ocean tides and tidal excitations specific to polar motion the number is $K = 10$ for both xp and yp, and for the prograde and retrograde motions. The zonal tides specific to dUT1 have $K = 62$ constituents (Petit & Luzum, 2010).

δEAM is decomposed into two parts: equatorial components $\delta\chi_1, \chi_2$ and the axial part $\delta\chi_3$ of the excitations. These two parts can be modelled with two groups of neural networks ($\mathrm{NN}_{\chi_1}, \mathrm{NN}_{\chi_2}$) and NN_{χ_3}. Additional constraints can be applied to NN_{χ_1}, NN_{χ_2} and NN_{χ_3}. For instance, we apply the Liouville equation (Chin et al., 2004) for δP (in the imaginary domain, $\delta P = \delta\mathrm{xp} - i\delta\mathrm{yp}$) to investigate if there are additional parts that are not available in EAM data or the tidal effects that result in errors δxp, δyp in the polar motion components. Similarly, for the rate of dUT1 a linear combination of mass (pressure: p) and motion (wind: w) terms of the χ_3 component of the EAM functions would be considered, bearing physical meaning for example concerning mantle anelasticity (Dickman, 2003; Dobslaw & Dill, 2019b). In addition, a neural network denoted by $\mathrm{NN}_s(\theta_s, R, \chi_3)$ should learn the remaining signals in the rate of dUT1 (i.e., periods larger than annual), including its interannual trend. Furthermore, since EAM data used in the study are both observations and forecasts, NN_{χ_1}, NN_{χ_2}, and NN_{χ_3} can be used to minimize the difference between forecasts and their corresponding observations simultaneously with the minimization of the difference between rapid and final EOPs.

Depending on the effects included, we have to consider two aspects, namely the unmixing problem and the self-calibration. The unmixing problem occurs when the tidal effects and EAM functions are included in the model and investigated for their impact on the reduction of differences between rapid and final EOPs. If, in addition, we try to calibrate the EAM forecasts simultaneously with the calibration of rapid EOPs, we have to introduce a self-calibration approach. In mathematical terms, this concept is described in Equations (4a)-(4f):

$$\delta \text{xp},\ \delta \text{yp} = \delta\chi_1,\ \delta\chi_2 + \delta \text{T} + \delta \text{TE} + \delta \text{U} \tag{4a}$$

$$\begin{aligned}
&\delta P + \frac{i}{\sigma_{cw}} \frac{d}{dt}\delta P = \delta\chi_1 + i\delta\chi_2 \\
&\delta P = \delta \text{xp} - i\delta \text{yp} \\
&\sigma_{cw} = \frac{2\pi}{T}(1 + \frac{i}{2Q}) \\
&T = 434.2 \\
&Q = 100 \\
&i = \sqrt{-1}
\end{aligned} \tag{4b}$$

$$\delta\chi_{1,o},\ \delta\chi_{2,o} = \delta\chi_{1,f},\ \delta\chi_{2,f} + \text{NN}_{\chi_1,\chi_2}(\theta_{\chi_{1,2}}, R, \chi_{1,f}, \chi_{2,f}) \tag{4c}$$

$$\delta \text{dUT1} = \delta\chi_3 + \delta \text{T}' + \delta \text{U}' \tag{4d}$$

$$\frac{d}{dt}\delta \text{dUT1} = \alpha\delta_{\chi_3^p} + \beta\delta_{\chi_3^w} + \text{NN}_s(\theta_s, R, \chi_3) \tag{4e}$$

$$\delta\chi_{3,o} = \delta\chi_{3,f} + \text{NN}_{\chi_3}(\theta_{\chi_3}, R, \chi_3) \tag{4f}$$

In Equation (4a), the error terms in polar motion δxp and δyp result from the errors in the equatorial components of the excitation functions $\delta\chi_1, \chi_2$, ocean tides, long period ocean tides and tidal excitations, and the remaining errors (smoothing, systematic correction, or unknown). NN_{χ_1}, NN_{χ_2} are used to calibrate the EAM forecasts used in the model with respect to the corresponding observations as in Equation (4c). These calibrated values can then be used in Equation (4b) to improve the prediction accuracy. A similar condition can be considered for dUT1 based on the differentiation of dUT1 and the mass and motion terms of the axial component of EAM $\delta\chi_3^p$, $\delta\chi_3^w$, through the linear equation (4e), with learnable parameters α and β. Crucial to mention is the presence of the neural network NN_s that learns the remaining signals in the rate of dUT1, including the interannual trend. Note that the errors in dUT1 (c.f. Equation (4d)) come from the errors in the axial component of the excitation functions $\delta\chi_3$, subdiurnal and diurnal tides $\delta T''$, long-period (zonal) tides $\delta Z'$ and the remaining errors $\delta\text{U}'$ ($\delta T' = \delta\text{T}'' + \delta\text{Z}'$). Similar to the case of polar motion, here also the difference between forecasts and their corresponding observations is simultaneously minimized with the calibration of rapid EOPs–Equation (4f). Finally, it is worthwhile mentioning that the methods used for polar motion use both xp and yp as the feature in the model, since this is shown to result in better prediction accuracy (Kiani-Shahvandi et al., 2022).

2.4 Feature importance methodology

For the analysis of feature importance, the goal of which is to investigate the importance of different input features in making accurate predictions, we use the method of deep feature ranking (Maksymilian & Chen, 2020). This method eliminates the need for combinatorial optimization (Bengio et al., 2021) for feature importance. This is advantageous since the importance of different features can be simultaneously analyzed, instead of analyzing individual or combinations of different features. Therefore, a large number of features can be investigated. The choice is furthermore justified since the ResLearner approach is mainly non-linear.

We define the feature importance (FI) as the relative contribution to the results. This means that FI in the first approximation is the ratio of the standard deviation of the method with or without the k-th feature $\sigma^{(k)}$ relative to the standard deviation of the output σ^F, as in Equation (5)

$$\text{FI}_k = \frac{\sigma^{(k)}}{\sigma^F} \tag{5}$$

Note that $\sigma^{(k)},\ k = 1, ...$ are the output of the deep feature ranking method (Maksymilian & Chen, 2020).

2.5 Geophysically-constrained neural networks: introducing ResLearner PhycoRNN

In addition to the unmixing and self-calibration problems, the concept of Physically Constrained Neural Networks (PCNN, Geneva & Zabaras, 2020) can be used for directly applying the physical constraints to the problem using Recurrent Neural Networks (RNN, Rumelhart, Hinton, & Williams, 1986). It has been shown that PCNN methods like PhyLSTM (Zhang et al., 2020), which is based on long short-term memory (LSTM, Hochreiter & Schmidhuber, 1997) and the physical conditions of the problem, could present state-of-the-art prediction performance. As LSTM is the base of PhyLSTM, one can think of replacing it with more modern architectures. We investigated several state-of-the-art architectures for the problem, including PhyLSTM itself, coupled oscillatory RNN (coRNN, Rusch & Mishra, 2021) and Long Expressive Memory (LEM, Rusch, Mishra, Erichson, & Mahoney, 2022). The coRNN architecture achieved the best performance and therefore we chose it to replace the LSTM cell in PhyLSTM. Using this approach, we devise a new architecture called PhycoRNN. The architecture is shown in Figure 1. In this architecture, there are two coRNN cells. The input $I = (R, \mathrm{EAM})$, containing rapid EOPs and EAM, passes through the first coRNN cell and generates two outputs V_1, V_2 which are subsequently passed through a Dense layer (Bishop, 2006) to generate the output G. The squared difference between G and the output F containing final EOPs data should be minimized, which can be called the mathematical loss, denoted by Loss_m. V_1 and V_2 are additionally passed through the second coRNN cell to generate the two outputs Z_1 and Z_2, which by applying another Dense layer to them would generate the output H. The geophysical constraints are then applied to H.

The geophysical constraint in the case of polar motion is the Liouville equation presented in Equation (4b), while for dUT1 rate is the linear combination presented in Equation (4e). In this case, α and β can be written as the following Equation (6) (Dobslaw & Dill, 2019b).

$$\begin{aligned} \alpha &= 2\pi\Omega\frac{k_r}{C_{\mathrm{eff}}}(1 + k'_{2,\mathrm{eff}} + \Delta k'_{\mathrm{an,eff}}) \\ \beta &= 2\pi\frac{k_r}{C_{\mathrm{eff}}} \end{aligned} \tag{6}$$

in which $\Omega = 7.292115 \times 10^{-5}\ [\frac{1}{s}]$ is the rotation rate of the Earth, $k_r = 0.9976$ the effect of rotational deformation, $C_{\mathrm{eff}} = 7.118246 \times 10^{37}$ [kgm^2] the effective axial moment of inertia, and $k'_{2,\mathrm{eff}} = -0.2415,\ \Delta k'_{\mathrm{an,eff}} = -0.0087$ the effective load Love number and the mantle anelasticity, respectively.

The mentioned geophysical constraints constitute the so-called physical loss, denoted by Loss_p. The total loss is the summation of the mathematical loss and the physical loss. To optimize the parameters of the neural networks we use the so-called LBFGS algorithm (D. Liu & Nocedal, 1989) since it has been shown to be quite efficient in PCNN problems. Finally, it should be noted that we investigated the number of time steps (input sequence length) used in the coRNN cell and a value of 3 was chosen since it resulted in the best prediction accuracy. Here, 200 epochs of training were used. The method was implemented using the PyTorch library (Paszke et al., 2019).

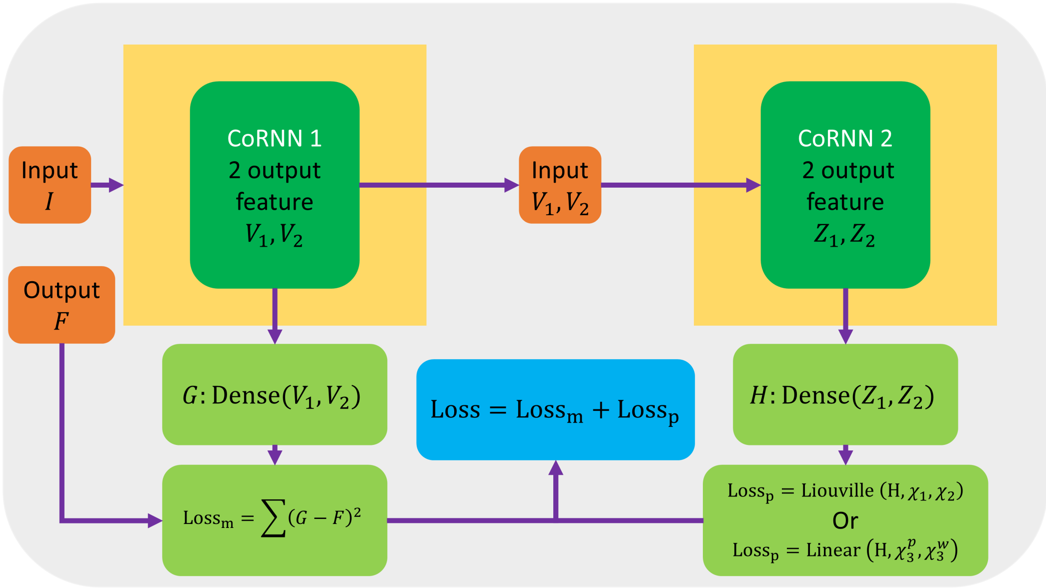


Figure 1: PhycoRNN architecture as a geophysically-constrained neural network, devised and used in the study.

2.6 Prediction accuracy metric

In order to evaluate the prediction accuracy, we use the mean absolute error (MAE) metric, which is commonly used in EOP prediction studies (Kalarus et al., 2010; Modiri et al., 2018; Kiani-Shahvandi et al., 2022). This is done for each day individually.

The quantification of improvement is based on the change in MAE for different days. If the MAE of one method is smaller than the baseline of rapid data themselves, we achieve an improvement. The MAE and improvement are defined in Equations (7a) and (7b):

$$\text{MAE}_k = \frac{1}{N}\sum_{i=1}^{N} |R^C_{i,k} - F_i|, \qquad k = -31, ..., 31 \tag{7a}$$

$$\text{improvement}_k = 100\% \frac{\text{MAE}^B_k - \text{MAE}_k}{\text{MAE}^B_k} \tag{7b}$$

In these equations, the index k is used for the day number, which is from -31 to 31. The number of predictions made is denoted by N. The predictions are denoted by $R^C_{i,k}$ (superscript C referring to calibration) for the i-th prediction and k-th day ahead. F_i denotes the corresponding final EOPs. The improvement is calculated by the percentage change in the MAE across different days, relative to the baseline (superscript B).

2.7 Summary of the concepts and optimal characteristics for ResLearner

A summary of the optimal characteristics of the ResLearner method is presented in Table 1, as determined in extended tests.

Table 1: Optimal characteristics for the ResLearner machine learning algorithm used for the calibration of rapid EOPs with respect to final EOPs

characteristic	choice/description
primary type of neural networks	non-linear MLP with two layers. 1 and 63 hidden neurons in layers, with tangent hyperbolic and linear activation functions for first and second layers, respectively
alternative type of neural networks	linear models: RANSAC, RidgdeCV, OLS
grouping of EOPs	equatorial and axial, i.e., for the prediction of xp or yp: both xp and yp used as feature; for the prediction of dUT1: only dUT1
non-linear uncertainty estimation	deep ensembles with M=10 simultaneous neural networks
feature importance analysis	deep feature ranking
evaluation metric	MAE
EAM functions considered	atmosphere, ocean, hydrology, and sea level
tidal effects	subdiurnal, diurnal, long period and tidal excitations, and long-period (zonal, for dUT1 only) with K= 30, 41, 10, 62 constituents, respectively
CI	MEI, NAI, MJI
PhycoRNN number of time steps	3
geophysical conditions for PhycoRNN	Liouville equation for rotational dynamics and polar motion; Earth rotation rate for first derivative of dUT1
unmixing	importance analysis of different features included in the model for their impact on the discrepancies between rapid and final EOPs
self-calibration	simultaneous calibration of EAM forecasts and the rapid EOPs

3 Data description

Here we describe the data used for the numerical results presented in the paper. Essentially, there are seven groups of data used in the study

- IERS rapid and final EOP 14 C04 series
- IERS final EOP 20 C04 series
- Jet Propulsion Laboratory (JPL) final EOP series (EOP2)
- European Space Agency (ESA) rapid and final EOP series
- ETH Zurich 14-day EAM forecasts
- GFZ German Research Center for Geosciences EAM analysis products
- National Oceanic and Atmospheric Administration (NOAA) MEI, NAI, MJI

IERS final 14 C04 EOP series (Bizouard et al., 2019) is the result of the combination of different space geodetic techniques including GNSS and VLBI and acts as the baseline to evaluate the various predictions against. This EOP time series is available from 1962 onward. Similar final EOPs data that are consistent with the latest International Terrestrial Reference Frame (ITRF2020) are provided by SYstèmes de Référence Temps-Espace (SYRTE). As mentioned in Section 1, IERS rapid EOPs (Dick & Thaller, 2018) are provided by using the most recent GPS and VLBI (24-hour and intensive session) data. The data are updated daily, but not archived publicly (daily finals). We have saved the rapid files since January 2015. Therefore, approximately 8 years of data is available for training and evaluation of the ResLearner algorithm. JPL series 2 of final EOPs are provided daily and contain the EOPs from 1976 onward, with less latency compared to the final IERS data. The JPL final series can act as the target in the training phase, i.e., IERS rapid EOPs are mapped to the final JPL EOPs. This creates another solution in addition to the one with final IERS data as the target.

For the purpose of additional validation, we use final, rapid and predicted EOPs provided by ESA and derived within the framework of the ESA project on "Independent Generation of Earth Orientation Parameters" (ESA-EOP, Dill et al., 2020; Kehm et al., 2023). The data result from series of hindcast experiments, in which the final EOPs are combined from GNSS, SLR, VLBI and DORIS and the rapid EOPs are combined from GNSS and VLBI only. Predictions are based on deterministic signals derived from the final and rapid EOPs time series in combination with EAM analysis and prediction data (as available on the assumed start date of prediction). Two series of hindcast scenarios from the study were provided, namely a realistic scenario and an ideal scenario. While the realistic scenario (scenario H1 in Kehm et al., 2023) assumes that the VLBI contribution to rapid (combined) EOPs solely relies on intensive data, the ideal scenario (scenario H2 in Kehm et al., 2023) assumes both 24-hour and intensive data to be available for the rapid combination. Each hindcast scenario is provided in the form of a data set containing 656 daily files for a time span from January 2018 up to January 2020. Thereby, each daily file contains final EOPs from around January 2009 up to a prediction horizon of about -28 days, rapid (combined) EOPs up to the day before the prediction start, and predicted EOPs up to a prediction horizon of +90 days. Here, we will use both scenarios for validation.

Regarding the EAM data, both the observations and forecasts are used, since forecasts can help significantly to improve the EOP prediction performance (Modiri et al., 2020; Kiani-Shahvandi et al., 2022). Since the horizon of the forecasts is also a determining factor (Kur et al., 2022), we use 14-day forecasts of ETH Zurich (Kiani Shahvandi et al., 2022) since they are both accurate and cover a reasonable forecasting horizon for short-term EOP prediction (i.e., suitable for accurate real-time purposes). Note that EAM predictions from all 14 days are used, since based on our analysis it results in the best performance (for instance, using 10-day forecasts results in less improvement). The EAM analysis files are taken from GFZ German Research Center for Geosciences (Dobslaw & Dill, 2018; Dill et al., 2019a). All four types of EAM functions, i.e., AAM, OAM, HAM, and SLAM, are used as geophysical features in the ResLearner algorithm.

We use CI provided by NOAA. Climatic index MEI is provided bimonthly by an empirical orthogonal function that combines different variables including sea surface pressure and temperature (Wolter & Timlin, 1993; Timmermann et al., 2018; Di Lorenzo et al., 2023). Since the data are bimonthly, they should be interpolated to generate daily values to be used as an additional feature for the prediction of dUT1. We also use NAI and MJI suspected for their influence on the rate of dUT1 (Hendon, 1995; Mazzarella, 2007).

Several investigations are presented in Section 4. In Figure 2, we show the rapid xp, yp, and dUT1 time series as well as the training and evaluation intervals for five different studies presented in this paper. The first study (S1) is similar to the subsequent three, but it is done operationally, with retraining at each prediction epoch. The starting date of evaluation is 20 May 2021 to be consistent with operational EAM forecasts (Kiani Shahvandi et al., 2022). The next three (S2, S3, S4) are hindcast studies that use IERS rapid EOPs as the input and IERS final 14 C04 or JPL EOP2 as the output. The purpose of these studies is to analyze the performance of the algorithm in the past. The final study (S5) is based on the ESA and IERS rapid and final EOPs. This is also only possible in a hindcast study. Crucial to mention is that hindcast studies observe the rules of real-time prediction (i.e., no future information being available), but with the prediction time in the past.

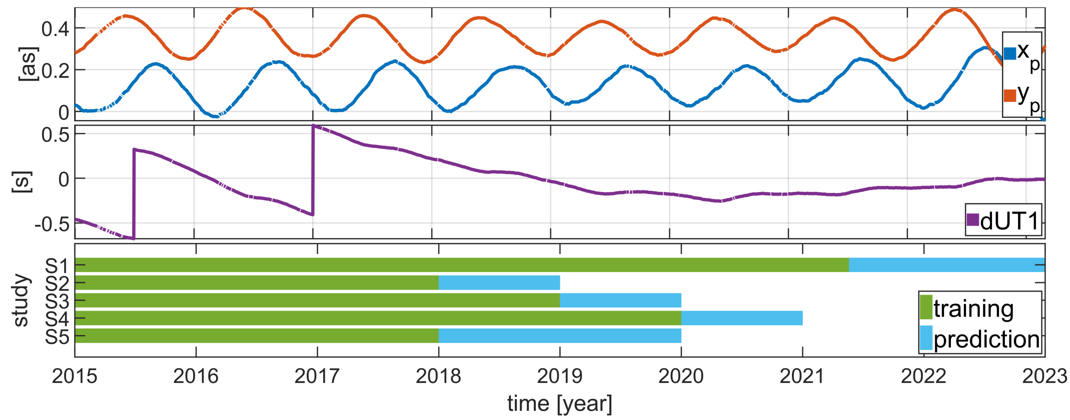


Figure 2: Top and middle panels show the polar motion and dUT1 series used in the study. The bottom panel shows the training and prediction intervals for each of the five studies (S1)-(S5) presented in Section 4.

4 Results and discussions

4.1 Analysis for the operational results in 2021-2022

Here, we present the performance analysis of the methods discussed in Section 2 based on the data described in Section 3. Note that the analysis refers to the study number 1 (S1) in Figure 2. The following points summarize the study configuration:

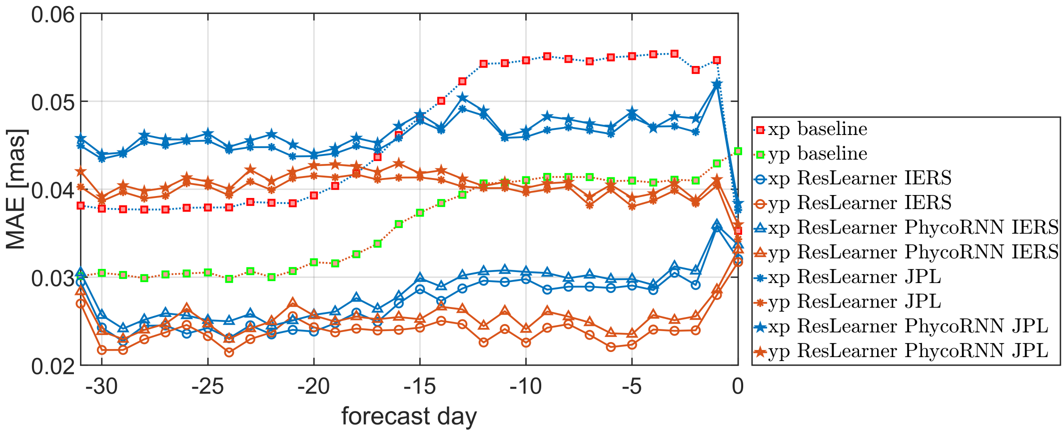
- The baseline solution is rapid EOPs as provided by IERS,
- Methods are trained on both IERS and JPL final EOPs,
- The final IERS 14 C04 EOP series is used for evaluation.

4.1.1 Prediction accuracy and improvement

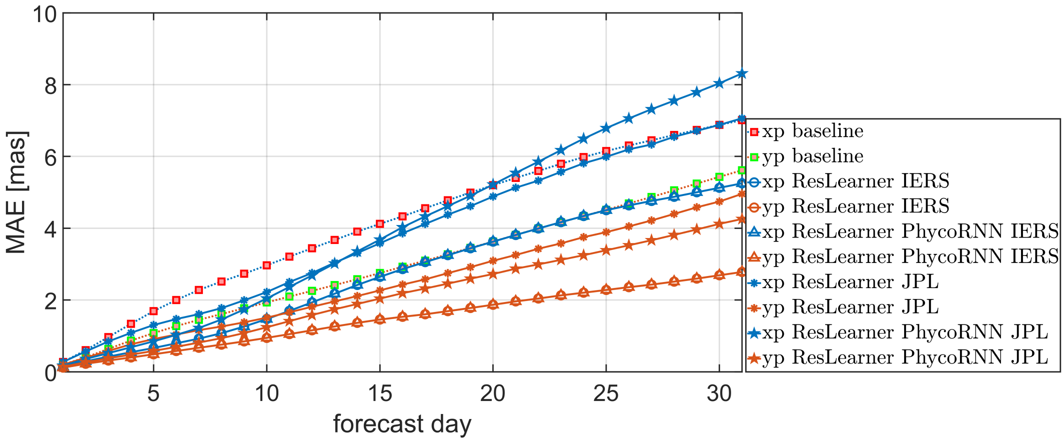
Figures 3 and 4 present the results of applying both the ResLearner and ResLearner PhycoRNN algorithms to the study interval shown in Figure 2. For better visualization of the performances, the prediction interval is divided into two parts: days -31 to 0 and days 1 to 31. The improvements with respect to the IERS baseline are presented in Figure 5 for polar motion and Figure 6 for dUT1. Based on Figures 3-6, several important points become evident.

First, the results of ResLearner PhycoRNN from days 1 onward seem to be identical to those of ResLearner when IERS 14 C04 is used for training. They are also very similar on days -31 to day 0, but not identical. This proves that for methods trained on IERS 14 C04, both PhycoRNN and ResLearner can be used. However, when JPL EOPs is used in the training, the results of ResLearner PhycoRNN and ResLearner are different. In this case, ResLearner PhycoRNN works better in yp, but worse in xp, approximately after day 13. This can be explained by the fact that ResLearner PhycoRNN has focused more on the yp component because of its larger amplitude and thus is performing worse on xp. Note, however, this is the best architecture for ResLearner PhycoRNN, implying that it cannot outperform ResLearner in xp, but only in yp. We tried to weight the loss functions so that the amplitudes of the errors of xp and yp be in the same range, but this did not improve the results. Regarding the difference between the results using JPL and IERS data as target, it becomes clear that the PhycoRNN has been able to capture the physics, but there is not as meaningful geophysical information in the mapping from rapid to JPL as from rapid to IERS. This is because the PhycoRNN is effectively transforming between EAM and GAM (Geodetic Angular Momentum), which as Dill et al. (2020) also point out, are not in full agreement with the JPL combined EOP series, especially for the equatorial components. This implies that having the Liouville equation as a hard constraint would not be beneficial if the EAM and EOPs series do not correspond to each other. In this case, a more mathematical-based approach would present better results, which is the case with ResLearner. We conclude that if the EOP and EAM series correspond to each other, the results of ResLearner and ResLearner PhycoRNN are almost identical, thereby suggesting physical and mathematical information have been adequately captured. Otherwise, ResLearner PhycoRNN does not perform well, since the geophysical constraints are less informative. This happens mostly for polar motion, but not for dUT1, which is due to the better agreement on the axial components of the GAM derived from different EOPs series (Dobslaw & Dill, 2019b).

Second, the improvement for polar motion components reaches 60% and generally remains above 40% for days -15 to 13. This is achieved by training the data on IERS 14 C04 final series, but not on JPL. Reasons for this discrepancy may include the longer interval that JPL provides the data for, which results in less informative data as a result of the degraded accuracy. More importantly, as mentioned GAM derived from IERS and JPL using EAM data do not fully correspond and can have large discrepancies, resulting in a reduction in accuracy of PhycoRNN predictions with JPL data as target. The improvements for dUT1 are generally smaller than those for polar motion. But they tend to increase for longer prediction horizons. The accuracy of both ResLearner PhycoRNN and ResLearner in days -31 to 0 for polar motion is almost below or at the uncertainty level of the polar motion data. This confirms that the methods can deliver results with an uncertainty level similar to that of the polar motion data. Finally, it is important to note that the accuracy of the IERS baseline and most of the methods is better at day 0 than at day -1. This behavior is more pronounced in polar motion compared to dUT1, meaning that the improvement for polar motion drops significantly at this day. We suspect that the reason for this anomalous behavior lies within the data and not in the applied models, as it is also visible in the IERS baseline, and might be related to a dominance of GNSS-derived polar motion information in the final IERS product and on the final day of the rapid combination (Kehm et al., 2023). The ResLearner unmixer algorithm (Section 2.3) can be used to further investigate this anomalous behavior.

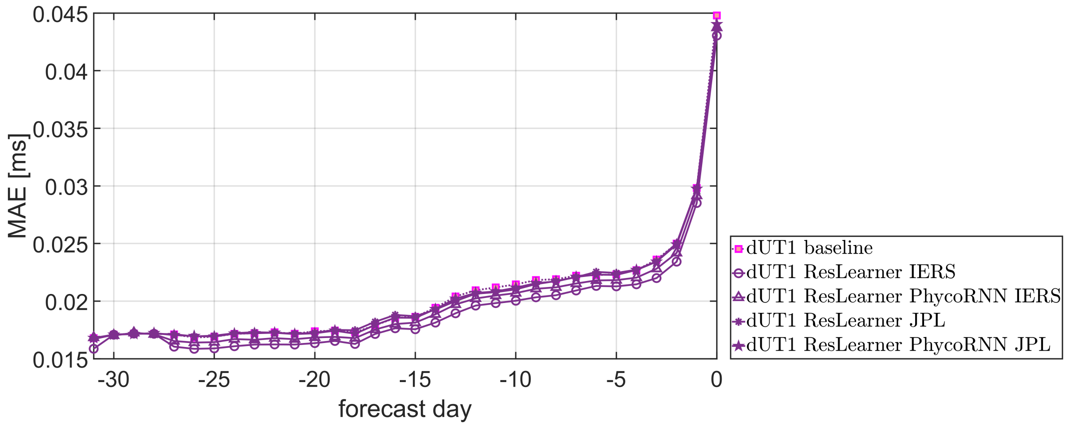


(a)

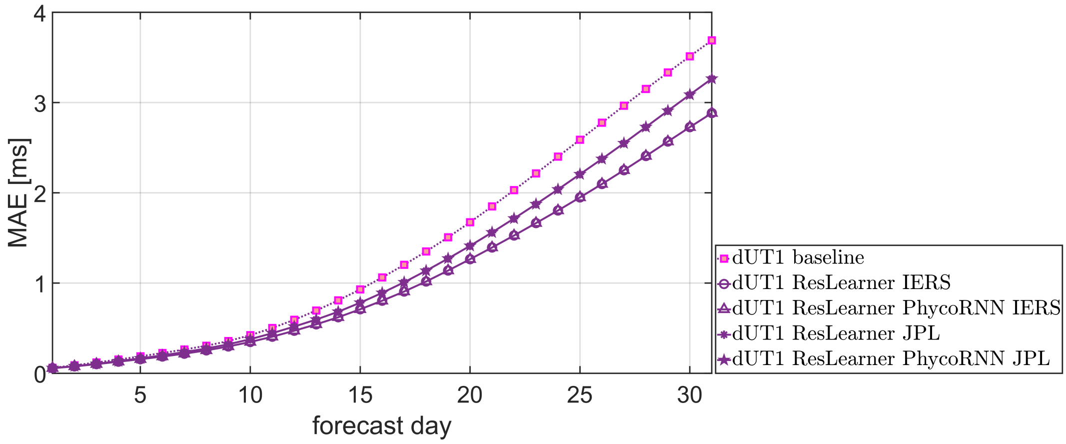


(b)

Figure 3: Prediction accuracy of polar motion components xp, yp for the first study (S1), in terms of MAE [mas]. ResLearner and ResLearner PhycoRNN are trained on both JPL and IERS final EOPs. (a) shows the MAE across days -31 to 0, while (b) focuses on days 1 to 31.

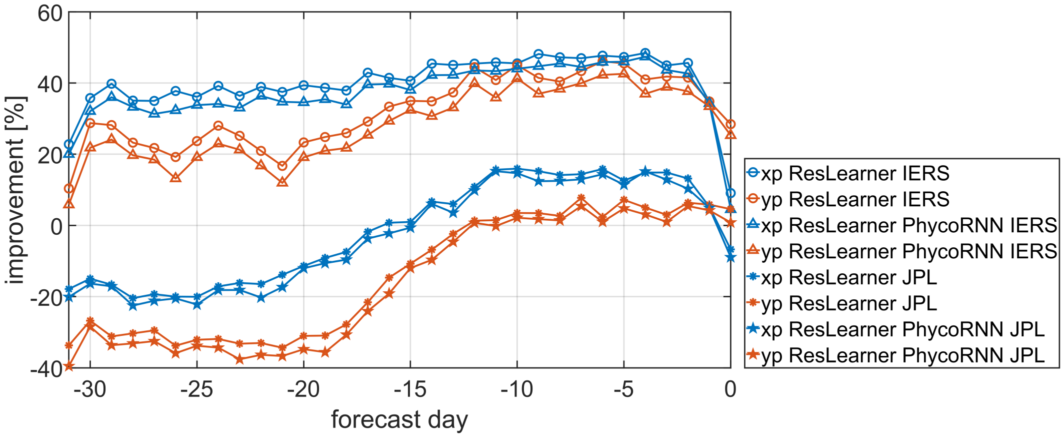


(a)

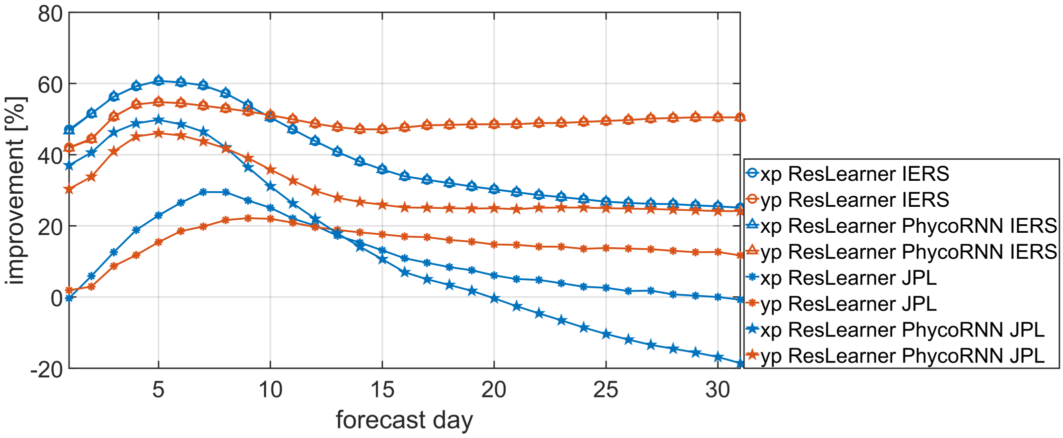


(b)

Figure 4: Prediction accuracy of dUT1 for the first study (S1), in terms of MAE [ms]. ResLearner and ResLearner PhycoRNN are trained on both JPL and IERS final EOPs. (a) shows the MAE across days -31 to 0, while (b) focuses on days 1 to 31.

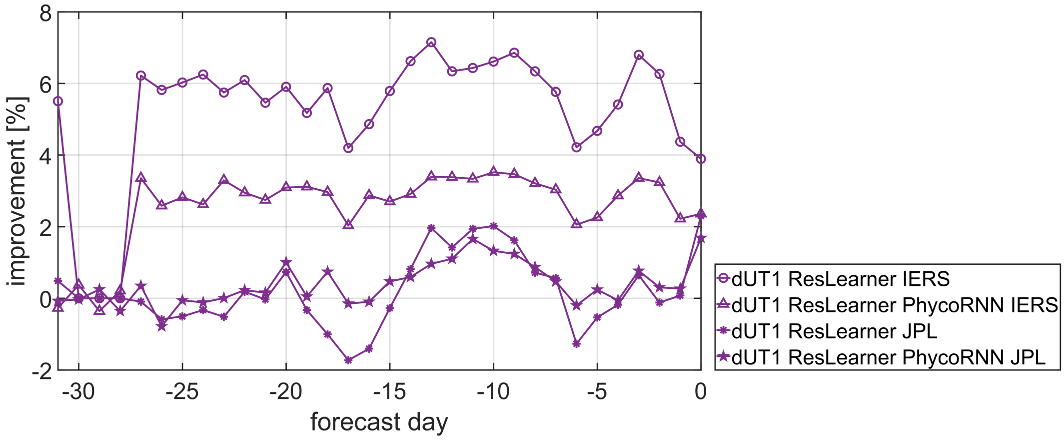


(a)

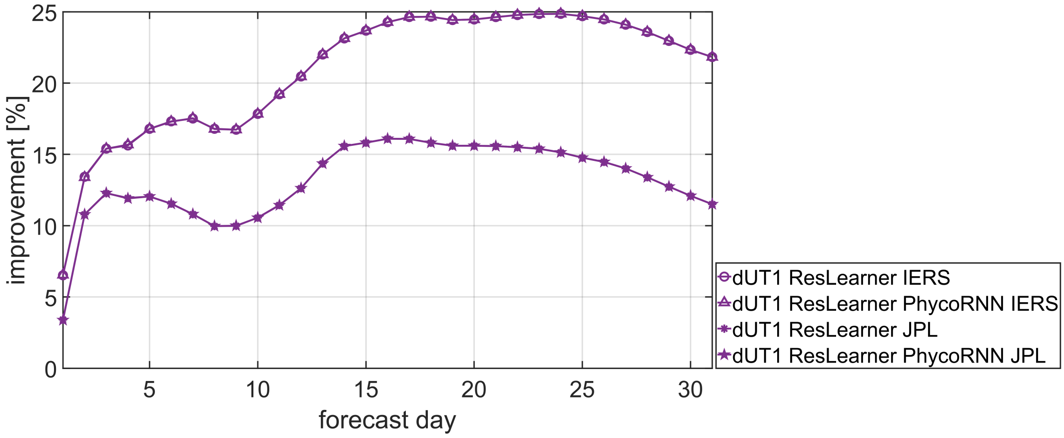


(b)

Figure 5: Improvement of prediction accuracy of polar motion components xp, yp for the first study (S1), in terms of percentage [%], computed according to Equation (7) based on the MAE of the baseline and that of ResLearner and ResLearner PhycoRNN. (a) shows the improvement across days -31 to 0, while (b) focuses on days 1 to 31.



(a)



(b)

Figure 6: Improvement of prediction accuracy of dUT1 for the first study (S1) presented in Figure 2, in terms of percentage [%], computed according to Equation (7) based on the MAE of baseline and that of ResLearner and ResLearner PhycoRNN. (a) shows the improvement across days -31 to 0, while (b) focuses on days 1 to 31. Note that the improvements are with respect to the IERS rapid data.

4.1.2 Importance of geophysical information

We find that EAM functions are one of the most important features that contribute to the discrepancies between rapid and final EOPs. As an example, in Figure 7, the Kendall correlations between the differences between rapid and final EOP IERS 14 C04, and the equatorial components of the individual EAM functions are shown. AAM and OAM (particularly the motion terms) present the highest correlation with these differences, thereby suggesting the importance of EAM for the ResLearner unmixer. Furthermore, even though in the rapid data AAM is included, the presence of the correlation suggests errors in accounting for AAM in the processes. In Figure 8 the importance of different features (FI) used in the model is presented, based on the methodology presented in Section 2 and according to Equation (5). For polar motion, Figure 8 gives the importance of the features xp, yp, EAM, and tides (semi-diurnal, diurnal, long-period tidal excitations combined), while for dUT1, it gives the importance of the features dUT1, EAM, tides (semi-diurnal,

diurnal, and long-period (zonal) combined), and CI. The individual CI components, i.e., MEI, NAI, and MJI are also displayed. Besides xp, yp, and dUT1 themselves, the EAM and tides are the most important features, confirmed also by other studies (Kiani-Shahvandi et al., 2022). Figure 7 also shows that AAM and OAM are the most important EAM functions for this problem (both mass and motion terms). Among CI, MEI seems to be the most relevant and can have effects several times bigger than the uncertainty level of dUT1. However, NAI and MJI have only a minor importance for the short-term prediction of dUT1. We therefore recommend only using MEI among the various climatic indices. We consider this to be in alignment with the observation that ENSO has a significant impact on the rate of dUT1, especially on interannual time scales (Chao, 1984).

We furthermore analyze the relationship between MEI and the physical condition on the rate of dUT1. In Figure 9, we show the negative of the rate of dUT1, i.e., $-\frac{d}{dt}dUT1$ (IERS rapid data) and the reproduced trend (which is in fact, rather an interannual signal in view of the limited time-period considered), the χ_3^p and χ_3^w components of the EAM functions, and MEI. Most of the signal in the rate can be explained by χ_3^w which is due to the zonal winds (Volland, 1996). However, the reproduced MEI also seems to be able to explain parts of the signal, especially around mid-2022. This can potentially be attributed to a La Niña event, which occurred in mid-2022. La Niña events have been shown to influence the rotation rates of the Earth (Xu et al., 2022). We can therefore state that ResLearner has been able to link the geophysical information to the input data. Note, however, that in short-term prediction the importance of MEI is smaller than that of other features, including χ_3^p and χ_3^w. But in the long-term, using MEI results in better training and prediction by ResLearner.

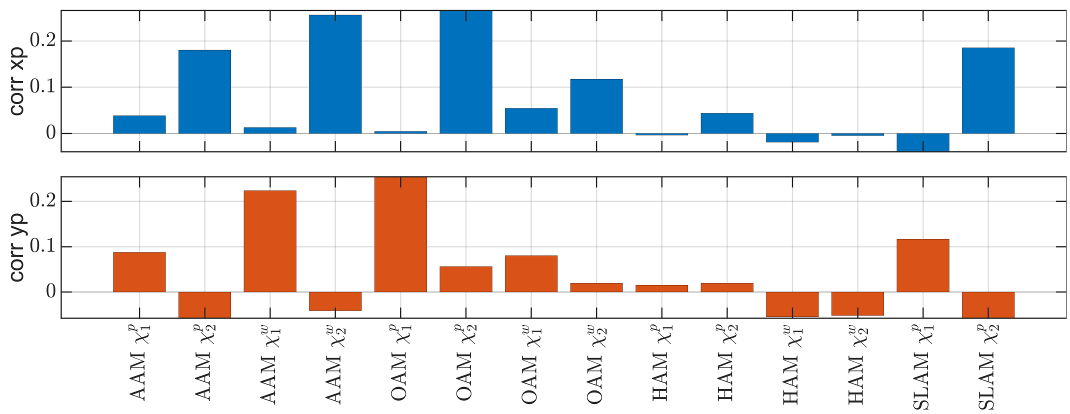


Figure 7: Kendall correlation (shown as corr in the figure) between the differences between rapid and final IERS EOPs, and the equatorial components of the individual EAM functions. Note that mass and motion terms (χ_i^p, χ_i^w $i = 1, 2$) are analyzed separately.

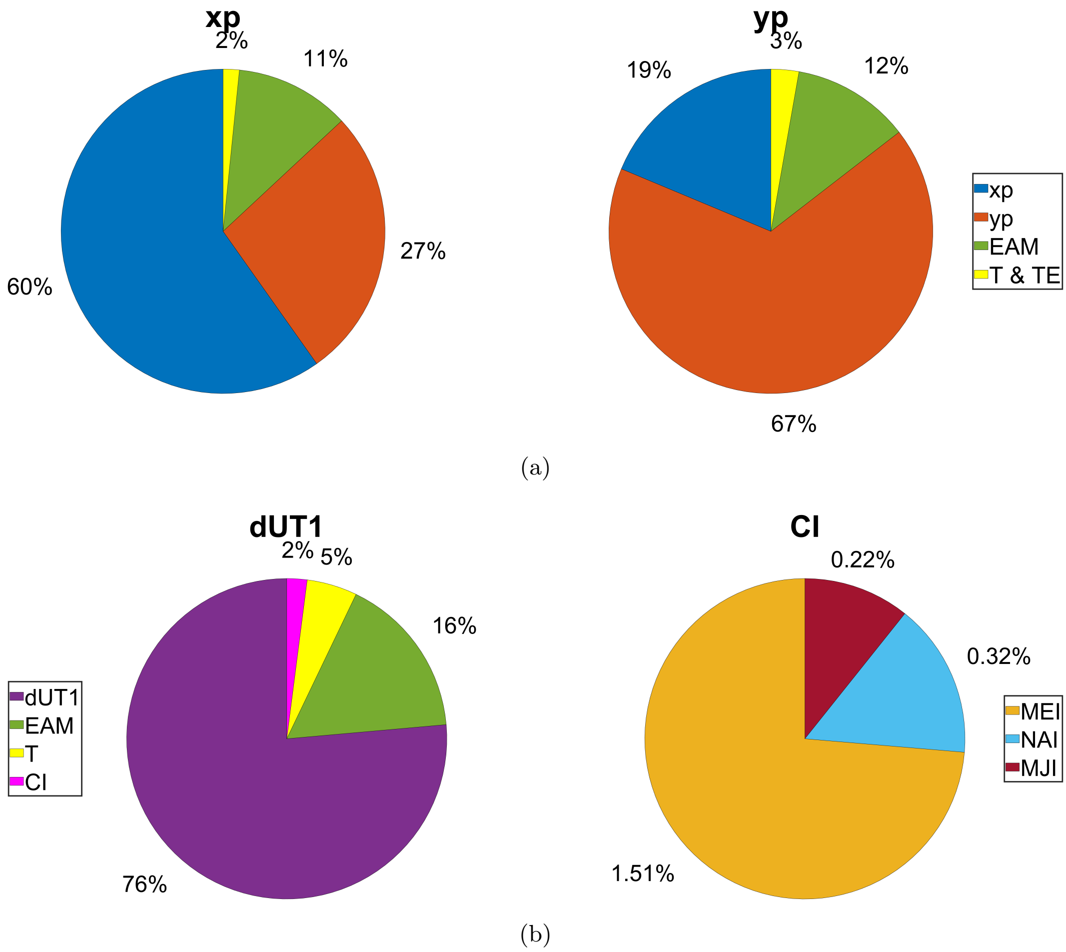


Figure 8: Feature importance analysis based on the algorithm presented in Section 2.4 and according to Equation (5). For polar motion components (a), features include xp, yp, equatorial components of EAM, T and TE (i.e., semi-diurnal, diurnal, and long-period tides and tidal excitations). For dUT1 (b), the features are dUT1, axial component of EAM, tides (semi-diurnal, diurnal, and zonal), and CI (climatic indices). CI is further decomposed into its components, i.e., MEI, NAI, MJI.

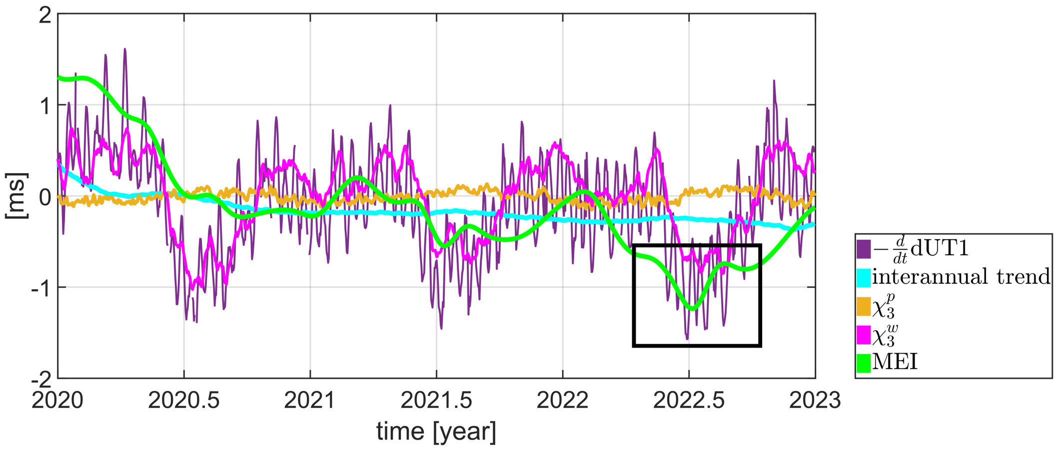


Figure 9: Negative rate of dUT1 (IERS rapid), $-\frac{d}{dt}$dUT1, together with the regressed interannual trend, χ_3^p, χ_3^w components of the EAM functions, and MEI, as obtained from the ResLearner algorithm. The interannual trend is solved during the training process and predicted accordingly at the prediction epoch. MEI here refers to what the ResLearner sees during training, i.e., the input feature MEI. Similarly, χ_3^p and χ_3^w are reproduced, but they are almost identical to their input form, because of their high feature importance. The mid-2022 La Niña event is highlighted by a black box.

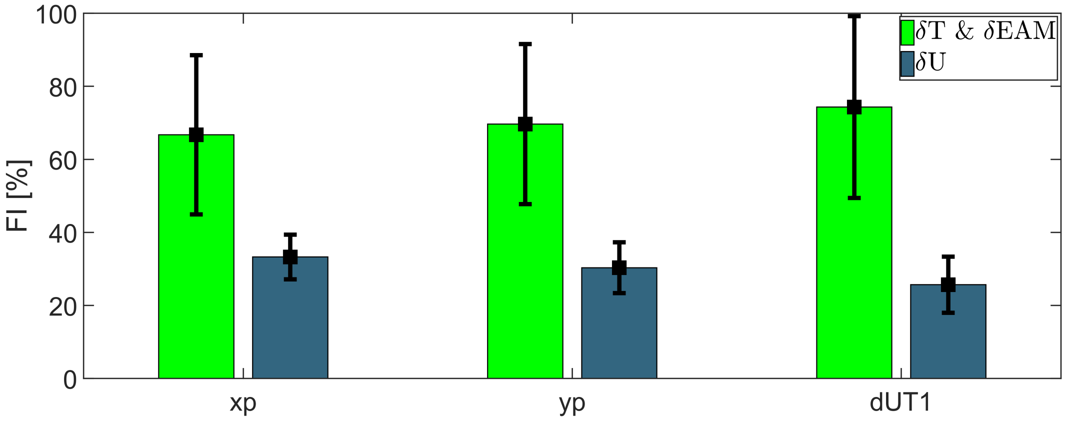
4.1.3 Unmixing: on the potential causes of errors in rapid EOP data

Building upon the results of feature importance analysis in Figures 7 and 8, the ResLearner unmixer algorithm can be applied to find the individual components of the EAM and tides that contribute most to the discrepancies between rapid and final EOPs. The corresponding results are presented in Figure 10, based on FI as given in Equation (5). In order to asses their significance, we also show their corresponding 95% confidence intervals. We have grouped the contributions into 1) tides and EAM (δT, δEAM) and 2) remaining errors (δU, systematic correction, smoothing, and unknown). Panel (a) gives the relative contributions of these two groups. The effect of the first group is bigger, thereby suggesting that the potential causes of discrepancy lie within tides and EAM. The five most important features among the first group are further investigated in panel (b).

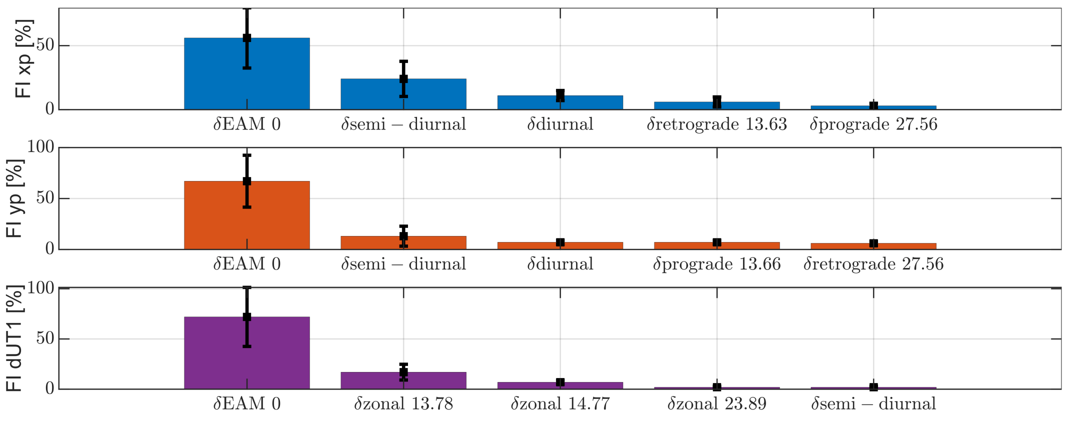
It is important to clarify that based on Figure 10 one can conclude that the most important features contributing to the anomaly observed at day 0 are (in the order of importance) δEAM at day 0, δU (including the dominance of the GNSS-derived polar motion), and δT. Regarding tides in polar motion, subdiurnal and diurnal tides, retrograde 13.63 and 27.56 days, and prograde 13.66 and 27.56 days long-period tides and tidal excitations are important. For dUT1, however, zonal tides of periods 13.78, 14.77, 23.89 days, and subdiurnal tides are relevant. For δU the approximate FI, together with their 95% confidence intervals are summarized in Table 2. Note that for δEAM and δT, the approximate values of importance are computed by multiplying the FI in panel (a) and (b), based on the fundamental rule of probability.

Table 2: The approximate FI and corresponding 95% confidence intervals for δEAM, δU, δT for the potential causes of discrepancies between the rapid and final EOP IERS 14 C04 series.

EOP	δEAM	δU	δT
xp	37±20%	33±6%	29±18%
yp	47±23%	30±7%	23±15%
dUT1	54±28%	26±8%	21±11%



(a)

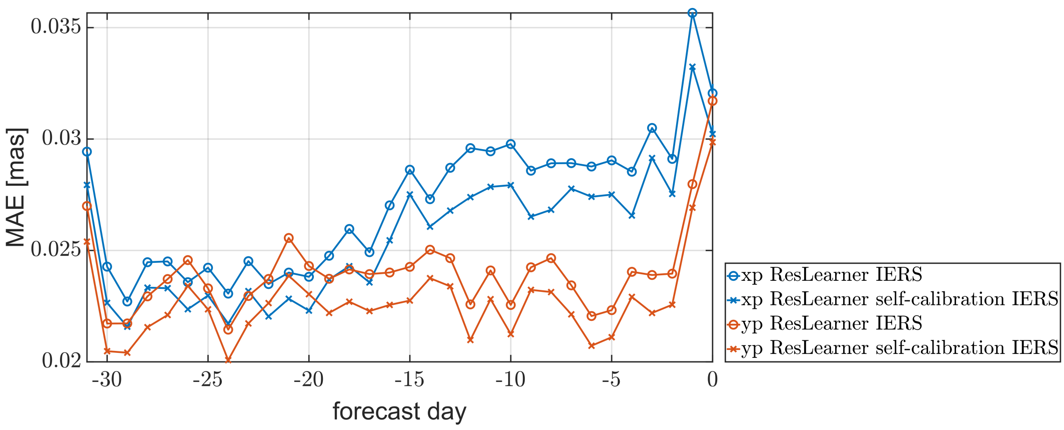


(b)

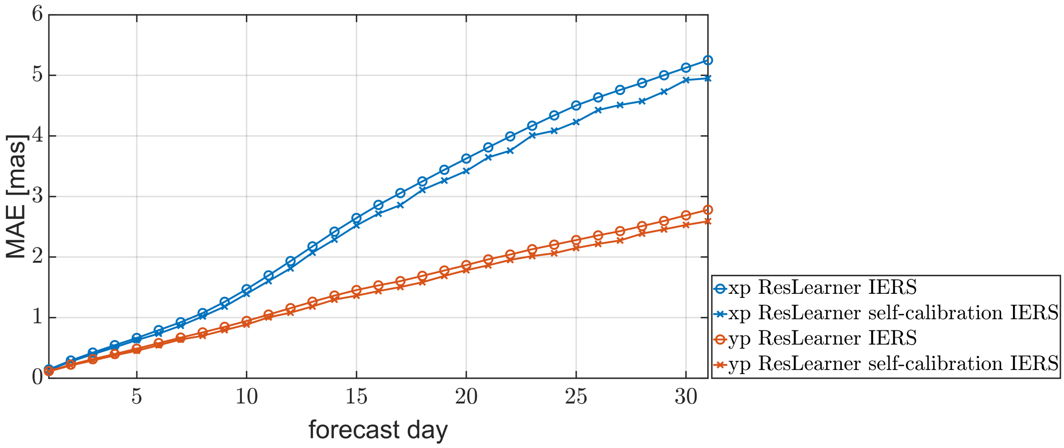
Figure 10: (a) FI computed according to Equation (5) for two groups 1) tides and EAM (δT & δEAM), 2) rest of errors (δU, systematic correction, smoothing, unknown) ; (b) FI computed for EAM and various tidal constituents resulting in the discrepancies between rapid and final EOPs, based on the methodology presented in Section 2. The uncertainties shown in the form of error bars are for 95% confidence interval. The analysis is for day 0 of prediction, containing the anomalous behaviour.

4.1.4 Self-calibration

After identifying the causes of errors in rapid data as from Figure 10, we apply the ResLearner self-calibration algorithm described in Section 2.3 in order to reduce the erroneous effects of the EAM functions. The results are shown in Figures 11 and 12 against the output of ResLearner algorithm without self-calibration. ResLearner self-calibration slightly improves the prediction performance (on average 5.5%). The improvement is achieved on both polar motion and dUT1, thereby suggesting the success of ResLearner self-calibration in reducing the errors.

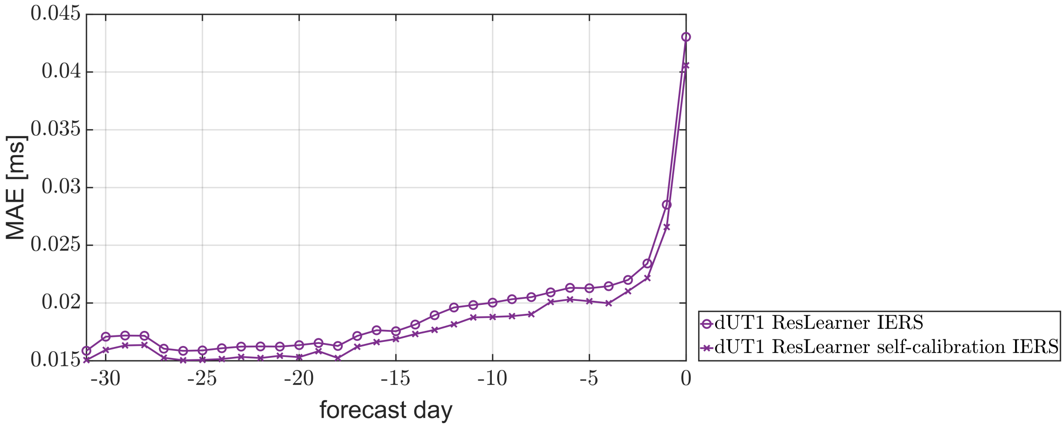


(a)

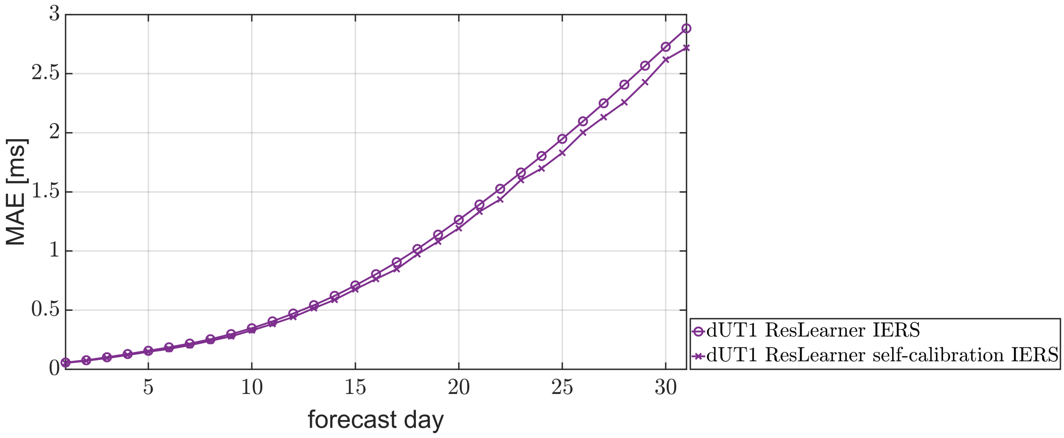


(b)

Figure 11: ResLearner self-calibration algorithm for the polar motion components against the ResLearner without self-calibration. (a) shows the comparison of days -31 to 0 while (b) displays that of days 1 to 31.



(a)

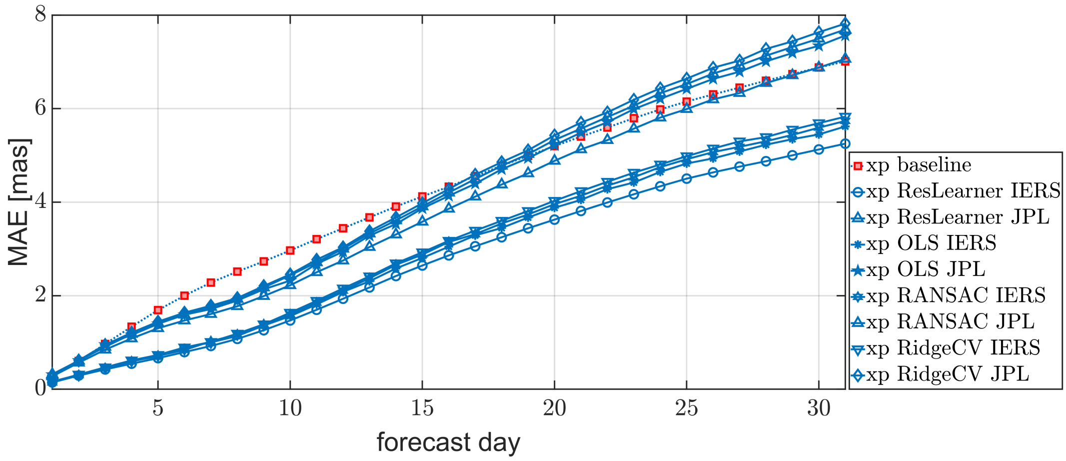


(b)

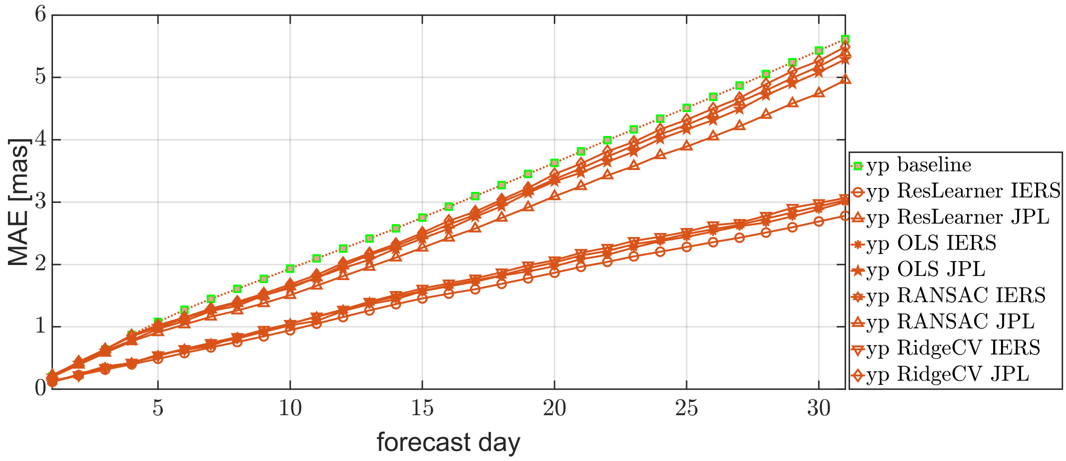
Figure 12: ResLearner self-calibration algorithm for dUT1 against the ResLearner without self-calibration. (a) shows the comparison of days -31 to 0 while (b) displays that of days 1 to 31.

562 #### *4.1.5 Comparative analysis: linear models*

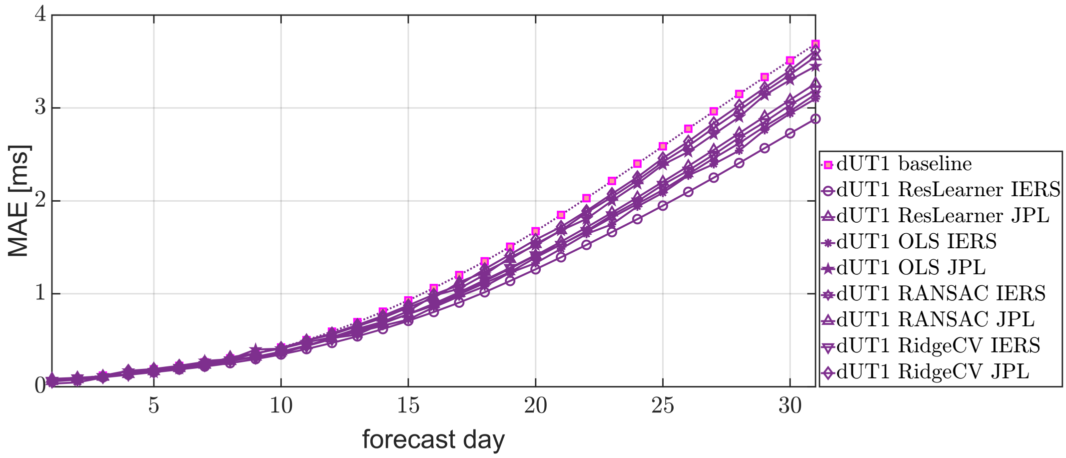
As mentioned in Section 2, linear ResLearner models can also present competitive results, i.e., close to the prediction performance of the state-of-the-art algorithms. The goal of this analysis is to illuminate the role of non-linearity in the model. Three different methods are considered: OLS, RANSAC, and RidgeCV. The results are compared with the non-linear ResLearner. The results of the comparative analysis are summarized in Figure 13. The results are shown only for days 1 to 31 since it is only on these days that we see a clear pattern of superiority of non-linear models. On days -31 to 0, the results are mixed: methods like OLS may outperform non-linear ones on some days, while on the rest of the days, the non-linear models outperform OLS. This analysis confirms that in this study, the non-linearity results in a gain in prediction performance. Furthermore, it is by non-linearity that the unmixing and self-calibration problems can capture almost all the signals in the input data.



(a)



(b)

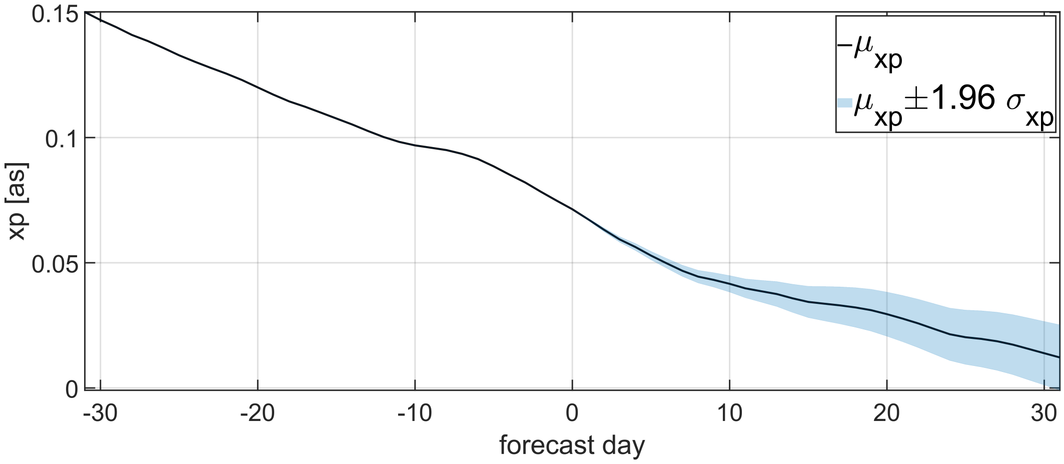


(c)

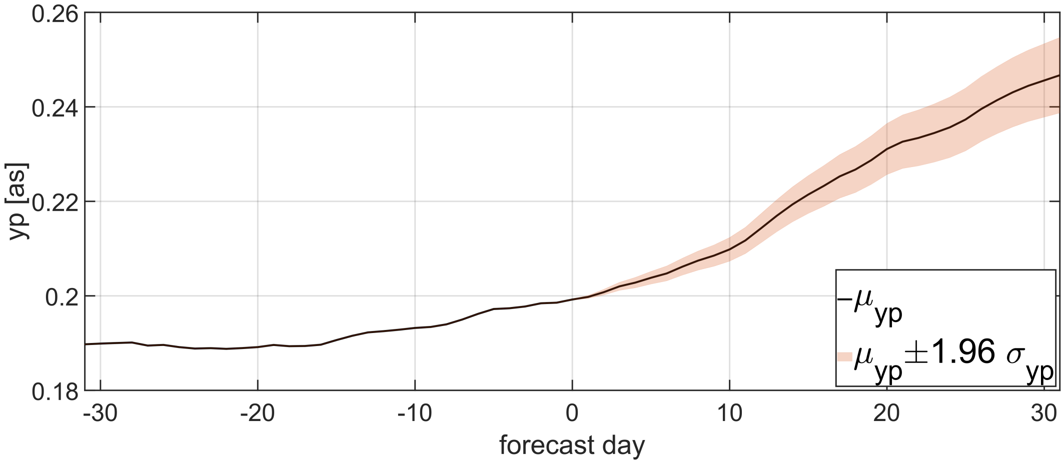
Figure 13: Comparison between the linear and non-linear ResLearner algorithms. Three different linear models are analyzed: OLS, RANSAC, RidgeCV. (a) shows the results for polar motion xp component, (b) for polar motion yp component, and (c) for dUT1.

4.1.6 Prediction uncertainty

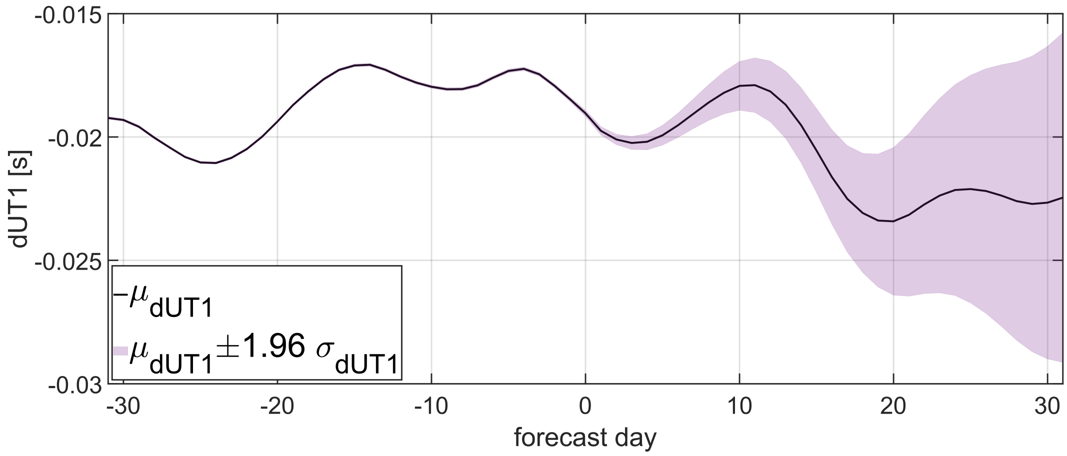
The ResLearner methodology implemented in the context of deep ensembles can provide uncertainties in the predictions. As an example, Figure 14 shows the predictions of polar motion and dUT1 together with their associated uncertainties, plotted for 2022-12-31. The mean values are given by μ, while the standard deviations are given by σ. The prediction uncertainties shown represent a 95% confidence ($\pm 1.96\sigma$) interval. Note that the derived prediction uncertainties depend on the respective day, but are usually close to the uncertainties in the rapid data. This confirms that ResLearner models in deep ensembles have been able to effectively reduce the epistemic uncertainty due to model errors. The reason is, the ResLearner is essentially a parametric model, the parameters of which are derived through optimization schemes. As a result, there is inevitably some uncertainties in the model parameters, which translate to the uncertainty in the predictions. Using the ensemble approach, we can effectively reduce this type of uncertainty and allow the model to predict more accurately and confidently.



(a)



(b)



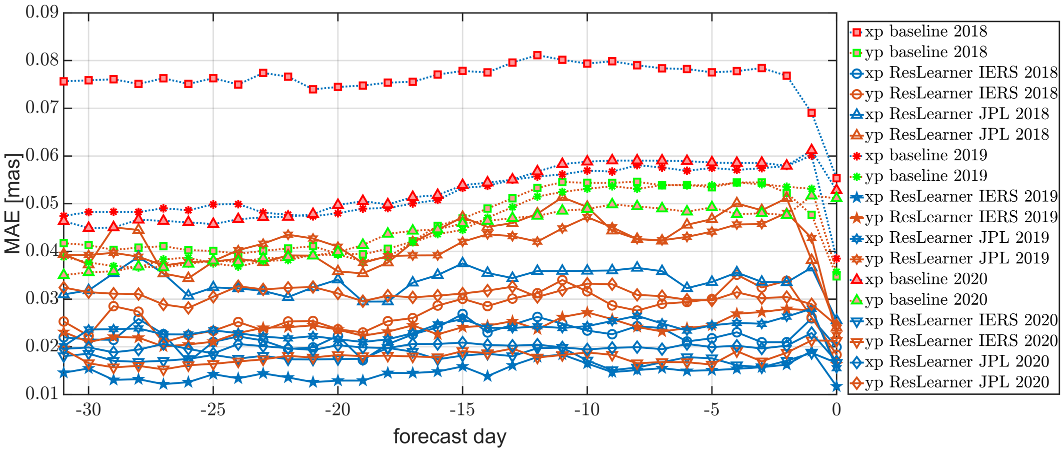
(c)

Figure 14: Prediction uncertainty for (a) polar motion xp, (b) polar motion yp, and (c) dUT1 for the date 2022-12-31, using ResLearner in the context of deep ensembles. μ_{xp}, μ_{yp}, and μ_{dUT1} are the mean values of the prediction, while σ_{xp}, σ_{yp}, and σ_{dUT1} are the associated standard deviations. The confidence interval is 95% ($\pm 1.96\sigma$).

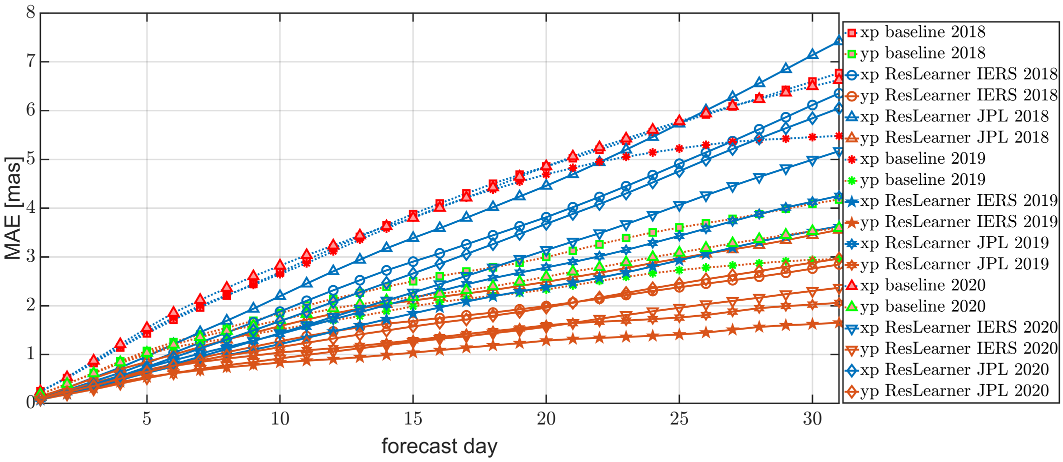
4.2 Hindcast analysis: 2018, 2019, 2020

We analyze the performance of the ResLearner method in hindcast scenarios, corresponding to the second, third, and fourth analyses (S2, S3, and S4) shown in Figure 2. The same conditions as in the first study (S1) are applied here as well, i.e., using the rapid IERS as the baseline, training on both IERS 14 C04 and JPL final EOPs 2 data, and evaluating against the IERS 14 C04 series.

Applying the same ResLearner architecture to these intervals, we get the results displayed in Figures 15 and 16. The results are divided into two parts: days -31 to 0 and days 1 to 31. Two important points can be deduced from these results. First, the accuracies are different from year to year and they do not show a clear reduction with increasing training intervals. This means that ResLearner tends to improve the prediction accuracy even when the training time span is shorter. Thus, the algorithm does not critically depend on the amount of data fed to it (c.f. Kiani-Shahvandi & Soja, 2021). This can be explained by the fact that the architecture is designed in a way that does not include too many learnable parameters, which can therefore be well trained. Second, the anomalous behavior of the polar motion components at day 0 also appears here, suggesting that the problem with rapid data also existed during earlier years.

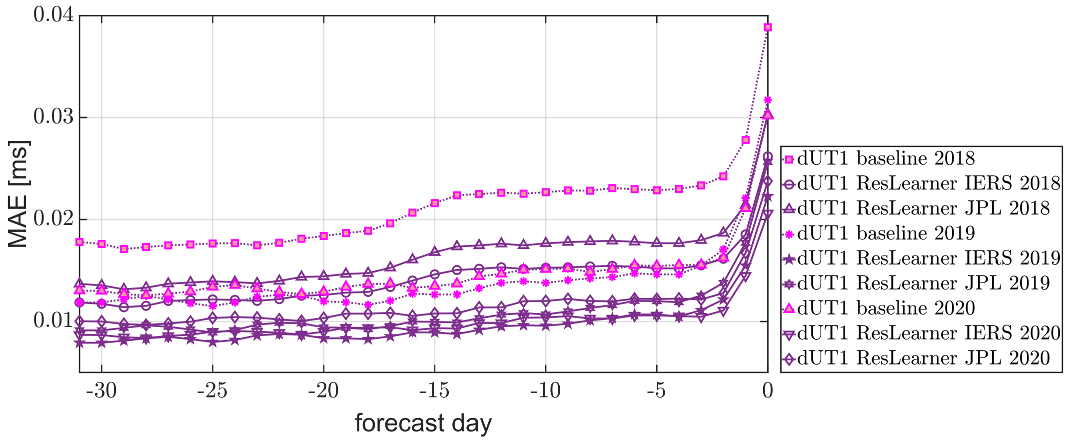


(a)

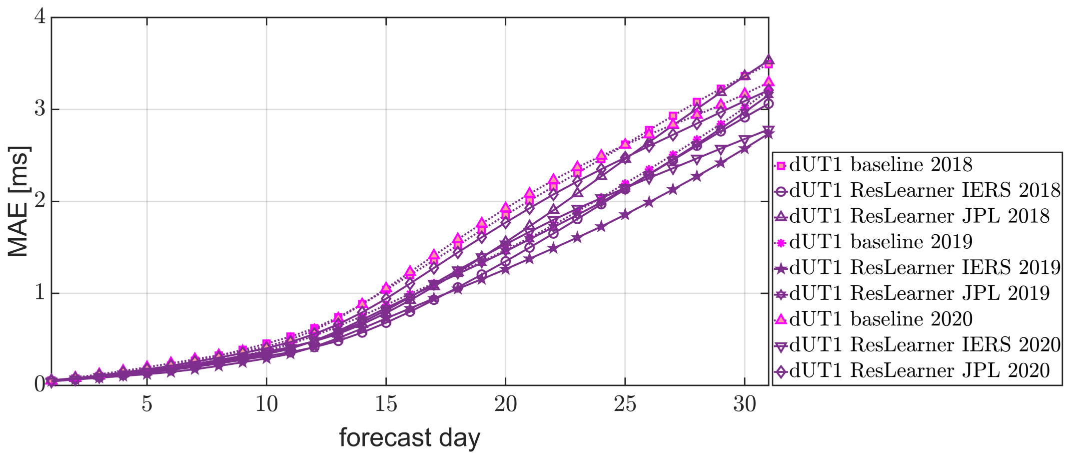


(b)

Figure 15: Prediction accuracy of hindcast studies S2, S3, and S4 for polar motion components xp and yp, in terms of MAE [mas]. Only the ResLearner is used (but not ResLearner PhycoRNN since they are similar). (a) displays the results for the days -31 to 0 and (b) for the days 1 to 31.



(a)



(b)

Figure 16: Prediction accuracy of hindcast studies S2, S3, and S4 for dUT1, in terms of MAE [ms]. Only the ResLearner is used. (a) displays the results for the days -31 to 0, while (b) for the days 1 to 31.

4.3 Analysis of ESA EOP data: a hindcast study

This analysis corresponds to the last study (S5) in Figure 2, the role of which is to validate our approach against an independent dataset of EOPs. The following points are important regarding this study.

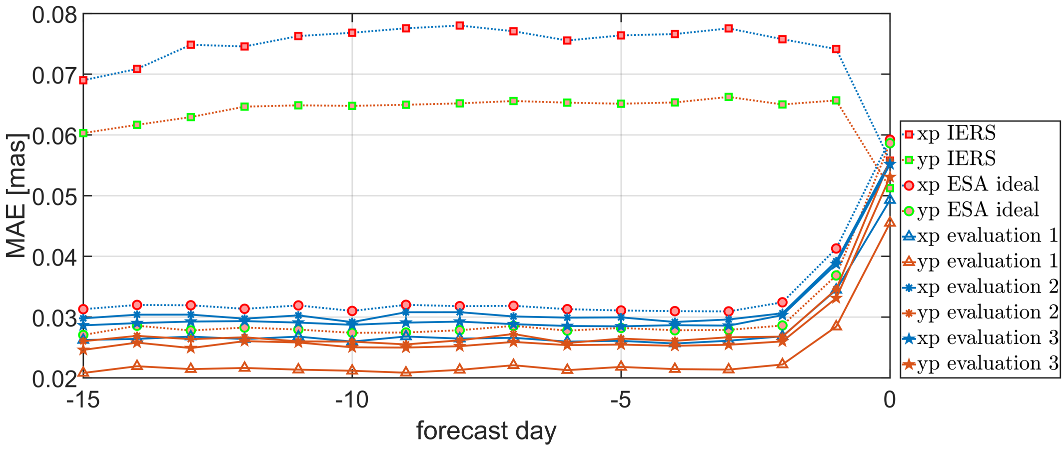
- The prediction horizon is 31 days, i.e., days -15 to 15
- Two baselines are considered: the rapid EOPs as provided by IERS and by ESA EOPs
- The final ESA EOPs are used for evaluation
- Validation is done against both the ideal and realistic ESA hindcast scenarios described in Section 3

We perform three different evaluations, namely:

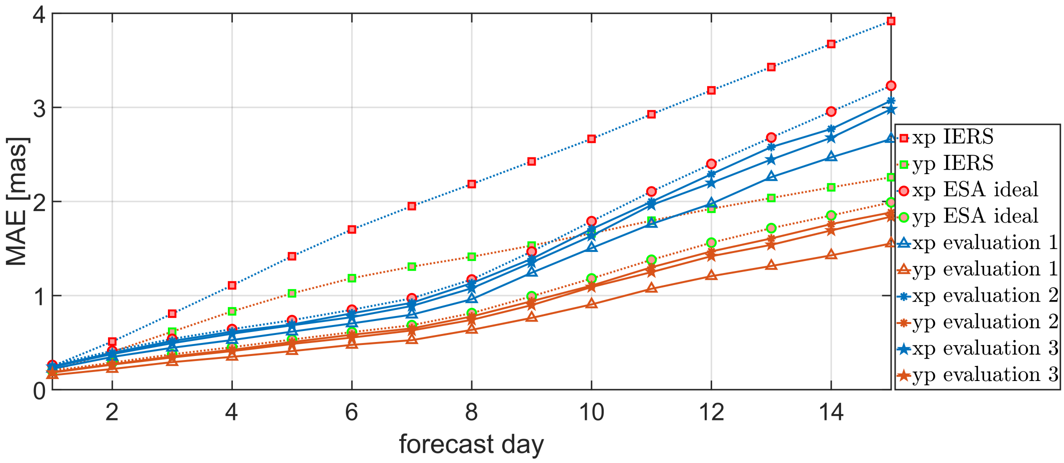
- evaluation 1: training only on IERS final EOPs up to the end of 2022,
- evaluation 2: training only on IERS final EOPs up to the respective time of ESA EOPs, using retraining at each epoch,
- evaluation 3: training on a combination of IERS and ESA EOPs, similarly with retaining.

The first evaluation is a hindcast study based on the pre-trained models. This means that no retraining is needed and predictions are made all at once. The second evaluation is more of operational nature, although in the past. The training period is thereby assumed to extend from 2015 up to the prediction day. In the third evaluation, IERS data from 2015 up to the end of 2017 are used for the training and first prediction. For each subsequent prediction, the ESA final data are added day-by-day to the training.

We analyze both the ideal and realistic scenarios mentioned in Section 3. First, we discuss the ideal case. The results of these evaluations are shown in Figures 17-18. Considering these results, we would like to highlight the following points: First, ResLearner is able to further improve the prediction accuracy based on ESA data, confirming its flexibility for different datasets. Second, there is not much difference between the results of the three evaluations. Only evaluation 1 presents minor superiority over the other evaluations. This is expected, however, as in this case, the model has seen not only the past but also the future final IERS EOPs. Third, all evaluations, as well as the ideal ESA baseline, show a significant improvement compared to the IERS baseline. Moreover, they show a more realistic behavior of the error of day 0, omitting the anomalous behavior seen in the IERS baseline (the error of day 0 being smaller than that of day -1). Application of ResLearner unmixer here points mostly again towards the EAM as the culprit. Furthermore, it shows that ESA and IERS data are slightly inconsistent at day 0, with the rapid IERS baseline accuracy being better when evaluated against IERS 14 C04. This, however, does not have an impact on the high prediction accuracy of both ESA baseline scenarios, which is close to that achieved with ResLearner.

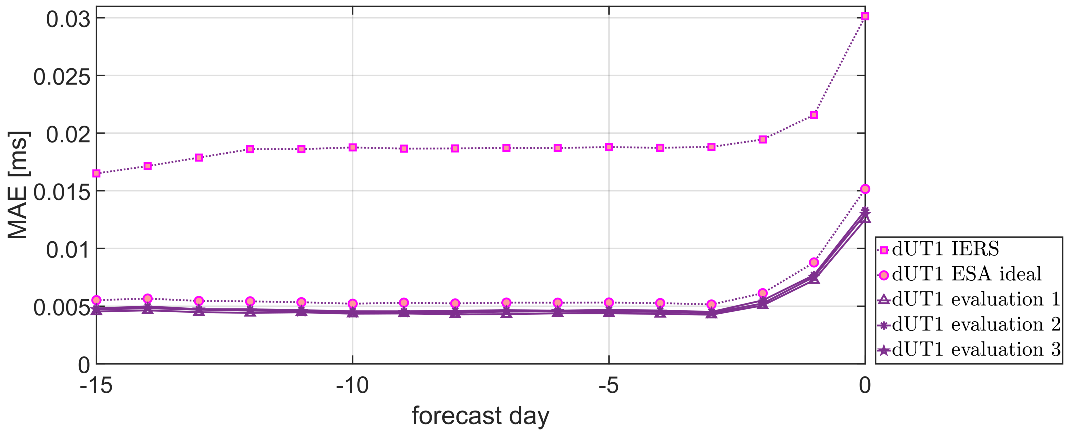


(a)

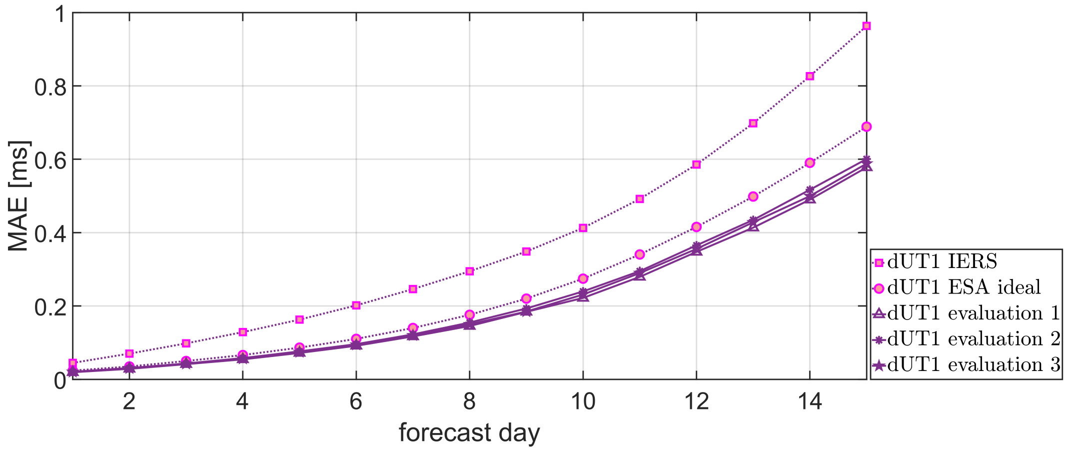


(b)

Figure 17: Prediction accuracy of the ResLearner algorithm for polar motion components xp and yp, based on study 5 (S5) and for three different evaluations: 1) training only IERS final EOPs up to the end of 2022, 2) evaluation 2: training only IERS final EOPs up to the respective time of ESA, 3) evaluation 3: training on a combination of IERS and ESA ideal data. Two baselines are presented: rapid IERS and rapid ESA ideal scenario. The data are evaluated against the final ESA ideal data. (a) shows the results for prediction days -31 to 0, while (b) for days 1 to 31.



(a)



(b)

Figure 18: Prediction accuracy of the ResLearner algorithm for dUT1, based on study 5 (S5) and for three different evaluations: 1) training only IERS final EOPs up to the end of 2022, 2) evaluation 2: training only IERS final EOPs up to the respective time of ESA, 3) evaluation 3: training on a combination of IERS and ESA ideal data. Two baselines are presented: rapid IERS and rapid ESA ideal scenario. The data are evaluated against the final ESA ideal data. (a) shows the results for prediction days -31 to 0, while (b) for days 1 to 31.

Figure 19 presents the results of the ESA realistic scenario for dUT1. While there is no significant difference between the ESA ideal and realistic scenarios for polar motion, dUT1 shows a clear reduction in prediction accuracy for days -15 to 0 compared to the ESA ideal scenario. This can be related to the missing of VLBI 24-hour data on these days, as the ESA realistic scenario only considers VLBI intensive sessions and GNSS rapids in the rapid combination. However, the change in prediction accuracy from days 1 to 15 is insignificant.

For ResLearner trained on the ESA realistic data, the prediction horizons between -15 and 0 days show a significant improvement compared to the ESA realistic scenario. This is in contrast to the results achieved by training on the ESA ideal scenario, where the additional improvement achieved by ResLearner is only minor. Thus, the results suggest that ResLearner can contribute to mitigating the effect of the processing latency of 24-hour VLBI sessions, which are crucial for a reliable determination of dUT1.

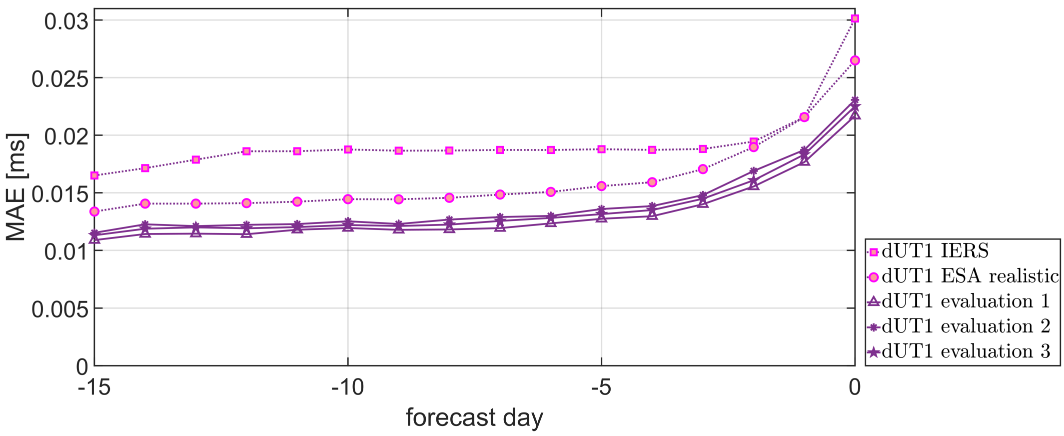


Figure 19: Prediction accuracy of the ResLearner algorithm for dUT1, based on study 5 (S5) and for three different evaluations: 1) training only IERS final EOPs up to the end of 2022, 2) evaluation 2: training only IERS final EOPs up to the respective time of ESA, 3) evaluation 3: training on a combination of IERS and ESA realistic data. Two baselines are presented: rapid IERS and rapid ESA realistic scenario. The data are evaluated against the final ESA realistic data.

4.4 Further discussions and recommendations

Several consequences arise from the results presented above. First, in order to analyze the sensitivity of the anomalous behavior at day 0 between the rapid and final IERS EOP series for evaluation, we evaluate the results of ResLearner and ResLearner PhycoRNN against the IERS 20 C04 series. This is similar to what is presented in Figure 5, but the reference EOP series is different. The results are shown in Figure 20. Comparing Figures 5 and 20, we observe that the anomalous behavior at day 0 is less severe. This further shows the dependence of the results on the version of IERS final and confirms that the choice of reference evaluation series is important when evaluating in general, and in this case especially for day 0. Note that we also trained the algorithms based on the IERS 20 C04 series and observed that the anomalous behavior at day 0 is less severe. This attests to the suitability of IERS 20 C04 to address this problem to a certain extent.

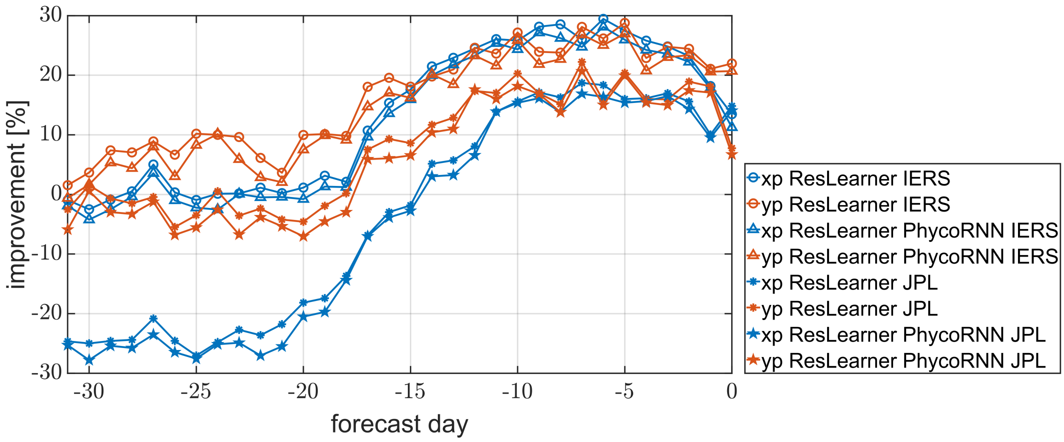


Figure 20: Improvement in prediction accuracy of polar motion components xp, yp for the first study presented in Figure 2, in terms of percentage. This is similar to Figure 5, but evaluated against the IERS 20 C04 instead of IERS 14 C04. Only the days -31 to 0 are shown to check for anomalous behavior at day 0.

In addition, since there are several types of the ResLearner method, we can compute an ensemble of all types based on IERS 20 C04 as target EOPs. A simple weighted ensemble is used, with the weights computed based on the overall prediction performance of individual types of ResLearner. We call this type of ResLearner the full ensemble ResLearner. The results of improvement for the full ensemble ResLearner are shown in Figure 21. The problem at day 0 is almost eliminated and we achieve up to 50% improvement in accuracy compared to the IERS rapid data. Note, however, that the improvements for days 1 to 31 are smaller compared to those presented in Figure 5, thereby suggesting that using the full ensemble approach is only beneficial in days -31 to 0. Crucial to note is that training a similar full ensemble based on IERS 14 C04 is not beneficial as the error would still persist.

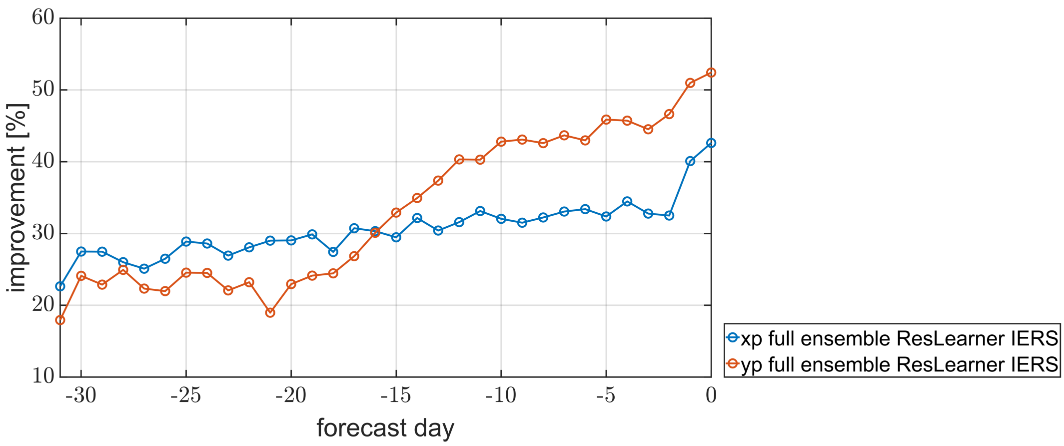


Figure 21: Improvement in prediction accuracy of polar motion components xp, yp for the first study presented in Figure 2, in terms of percentage. This is similar to Figures 5 and 20, but a weighted ensemble of the types of ResLearner algorithm is used. Only the days -31 to 0 are shown to check for anomalous behavior at day 0.

Based on our thorough analyses, we present some recommendations regarding the improvement of rapid EOP data, summarized in Table 3.

Table 3: Recommendations based on the numerical results presented in Section 4.

characteristics	recommendation
type of ResLearner	non-linear ResLearner with self-calibration
most relevant features	EAM, semi-diurnal, diurnal, zonal tides, MEI
EOP series for training and evaluation	IERS 20 C04

5 Conclusions

We devised a new machine learning method called ResLearner for the purpose of reducing errors in rapid EOPs w.r.t. final EOPs. The method is essentially non-linear and has a physically-constrained form called ResLearner PhycoRNN based on coupled oscillatory recurrent neural networks. Additionally, we also investigated the linear form of the method. Unmixing and self-calibration problems are analyzed as well, used for finding the causes of discrepancies between rapid and final EOPs, and calibrating the errors in the input features. Extensive numerical investigations are performed on both IERS and JPL final data, as well as validations against independent series of ESA hindcast experiments. The results show the superiority of non-linear ResLearner compared to the linear methods. Furthermore, ResLearner PhycoRNN can outperform ResLearner in the yp component of polar motion, while ResLearner is better in the xp component. Generally, the improvement in the accuracy of both polar motion components is over 40% across a large portion of the prediction horizon and can reach up to 60%. For dUT1, the improvement in prediction accuracy is smaller, but becomes larger for later prediction days, reaching up to 25%. In this context, validation against the ESA hindcast experiments demonstrates the capability of ResLearner to partially compensate for quality limitations in rapid dUT1 determination that are related to the latency of 24-hour VLBI data. As technical limitations will not allow for a faster availability of these data in the foreseeable future, ResLearner could become a valuable component in enhancing the quality of this parameter crucial for low-latency and real-time applications.

There is an anomalous behavior in the IERS rapid EOP data at day 0, where the consistency with the IERS finals appears to be better than at day -1. The unmixing algorithm suggests that errors in EAM, dominance of GNSS-derived polar motion, and tides are the main causes of this behavior. By applying the ResLearner self-calibration to the data, the errors are reduced and further improvement is achieved. Furthermore, using the IERS 20 C04 series either as the target in the training phase or as reference series for evaluation reduces this anomalous behavior, which suggests the superiority of the IERS 20 C04 over the 14 C04 EOP series. This is further justified when an ensemble of all types of ResLearner methods is used, in which case we no longer observe this anomalous behavior.

We further discussed the importance of geophysical information and found that besides EAM functions, tidal corrections and CI contribute to the prediction performance. Subdiurnal, diurnal, and long-period (zonal) tides in the oceans are all found to be relevant. Furthermore, the multivariate ENSO index is found to be the most relevant CI. Further investigation in this context should focus on each individual component in order to judge whether errors assigned to a certain part of a (conventional) model are actually to be related to it. In this context, feature importance can give hints on where model deficiencies might have an impact on the quality of current EOP determination.

Up to now, the ResLearner-based EOP determination realises a rapid EOP product that does not have a seamless transition from the corresponding final EOPs. This is in contrast to the EOP series realised by the ESA approach, where final and rapid EOPs combined from space-geodetic observations are directly complemented by a prediction that uses the last set of rapid (combined) EOPs as initial values. Further investigation might put focus on incorporating ML-based features already as conditions into the combination of the space-geodetic techniques, thereby realising a seamless EOP time series from the past into the future.

Since the method developed in this paper is based on the concept of physically-constrained neural networks, by modifying the geophysical constraints it can be used for other adjustment and prediction problems as well. One such problem in the field of Earth rotation is the long-term prediction of changes in the length-of-day. We hope that the results presented in this paper stimulate further research in this direction to combine the mathematical rigor of neural networks and the strength of geophysical information.

Acknowledgments
The authors acknowledge the European Space Agency (ESA) for providing series of hindcast experiments derived within the ESA project on “Independent Generation of Earth Orientation Parameters” (ESA-EOP; ESA Contract 4000120430/17/D/SR).

Declarations

Conflict of interest: None

Data availability

The improved rapid EOPs based on the methodology presented in this paper are operationally available on the ETH Zurich Geodetic Prediction Center (GPC) website at `https://gpc.ethz.ch/EOP/Rapid/`. The 14-day forecasts of EAM functions can be accessed at the ETH Zurich GPC website at `https://gpc.ethz.ch/EAM/`. EAM analysis products of GFZ German Research Center for Geosciences are available for download at `http://rz-vm115.gfz-potsdam.de:8080/repository`. IERS rapid and final EOPs (series 14 C04) are available at `https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html`. EOP series 20 C04, consistent with ITRF 2020, can be accessed via `https://hpiers.obspm.fr/iers/eop/eopc04_20/eopc04.1962-now`. The JPL final EOP series can be obtained via `https://eop2-external.jpl.nasa.gov/`. ESA data used in the study has been provided on request for this study (cf. Kehm et al., 2023). The developed software is available at `https://doi.org/10.5281/zenodo.7712379`. Information regarding the rapid files processing strategy can be accessed at `https://maia.usno.navy.mil/ser7/archive.notes` and `https://maia.usno.navy.mil/information/iers-gaz13.txt`. The multivariate ENSO index can be accessed via `https://psl.noaa.gov/enso/mei/` and the MJI data via `https://www.psl.noaa.gov/mjo/mjoindex/`. Data regarding NAI are available at `https://www.ncei.noaa.gov/access/monitoring/nao/`.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... Zheng, X. (2016). Tensorflow: A system for large-scale machine learning. In *12th usenix symposium on operating systems design and implementation (OSDI*

16), usenix association (p. 265-283). doi: https://dl.acm.org/doi/10.5555/3026877.3026899

Barnes, R., Hide, R., White, A., & Wilson, C. (1983). Atmospheric angular momentum fluctuations, length-of-day changes and polar motion. *Proceedings of the Royal Society A*, *387*. doi: https://doi.org/10.1098/rspa.1983.0050

Bengio, Y., Lodi, A., & Prouvost, A. (2021). Machine learning for combinatorial optimization: A methodological tour d'horizon. *European Journal of Operational Research*, *290*(2), 405-421. doi: https://doi.org/10.1016/j.ejor.2020.07.063

Bishop, C. (2006). *Pattern recognition and machine learning*. Springer Information Science and Statistics.

Bizouard, C., Lambert, S., Gattano, C., Becker, O., & Richard, J. (2019). The IERS EOP 14C04 solution for Earth orientation parameters consistent with itrf 2014. *Journal of Geodesy*, *93*, 621–633. doi: https://doi.org/10.1007/s00190-018-1186-3

Bizouard, C., & Seoane, L. (2010). Atmospheric and oceanic forcing of the rapid polar motion. *Journal of Geodesy*, *84*, 19–30. doi: https://doi.org/10.1007/s00190-009-0341-2

Brzezinski, A., & Nastula, J. (2002). Oceanic excitation of the Chandler wobble. *Advances in Space Research*, *30*, 195–200. doi: https://doi.org/10.1016/S0273-1177(02)00284-3

Chao, B. (1984). Interannual length-of-day variation with relation to the southern oscillation/El Nino. *Geophysical Research Letters*, *11*, 541–544. doi: https://doi.org/10.1029/GL011i005p00541

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*, *78*, 343–353. doi: https://doi.org/10.1007/s00190-004-0411-4

Dahlen, F. (1976). The passive influence of the oceans upon the rotation of the Earth. *Geophysical Journal International*, *46*, 363–406. doi: https://doi.org/j.1365-246X.1976.tb04163.x

Dick, R., & Thaller, D. (2018). *IERS Annual Report 2018*. IERS.

Dickman, S. (2003). Evaluation of "effective angular momentum function" formulations with respect to core-mantle coupling. *Journal of Geophysical Research: Solid Earth*, *108*. doi: https://doi.org/10.1029/2001JB001603

Dill, R., & Dobslaw, H. (2010). Short-term polar motion forecasts from Earth system modeling data. *Journal of Geodesy*, *84*, 529–536. doi: https://doi.org/10.1007/s00190-010-0391-5

Dill, R., Dobslaw, H., Hellmers, H., Kehm, A., Bloßfeld, M., Thomas, M., ... Schönemann, E. (2020). Evaluating processing choices for the geodetic estimation of Earth orientation parameters with numerical models of global geophysical fluids. *Journal of Geophysical Research: Solid Earth*, *125*. doi: https://doi.org/10.1029/2020JB020025

Dill, R., Dobslaw, H., & Thomas, M. (2019a). Improved 90-day Earth orientation predictions from angular momentum forecasts of atmosphere, ocean, and terrestrial hydrosphere. *Journal of Geodesy*, *93*, 287–295. doi: https://doi.org/10.1007/s00190-018-1158-7

Dill, R., Saynisch-Wagner, J., Irrgang, C., & Thomas, M. (2021). Improving atmospheric angular momentum forecasts by machine learning. *Earth and Space Science*, *8*(12). doi: https://doi.org/10.1029/2021EA002070

Di Lorenzo, E., Xu, T., Zhao, Y., Newman, M., Capotondi, A., Stevenson, S., ... Zhang, H. (2023). Modes and mechanisms of pacific decadal-scale variability. *Annual Review of Marine Science*, *15*(1), null. doi: https://doi.org/10.1146/annurev-marine-040422-084555

Dobslaw, H., & Dill. (2018). Predicting Earth orientation changes from global forecasts of atmosphere-hydrosphere dynamics. *Advances in Space Research*, *61*, 1047–1054. doi: https://doi.org/10.1016/j.asr.2017.11.044

Dobslaw, H., & Dill, R. (2019b). Effective angular momentum functions from Earth system modelling at GeoForschungsZentrum in Potsdam. In *Product description document: GFZ German Research Centre for Geosciences Department 1: Geodesy Section 1.3: Earth System Modelling.* Retrieved from `http://rz-vm115.gfz-potsdam.de:8080/repository`

Dobslaw, H., Dill, R., Grötzsch, A., Brzezinski, A., & Thomas, M. (2010). Seasonal polar motion excitation from numerical models of atmosphere, ocean, and continental hydrosphere. *Journal of Geophysical Research: Solid Earth*, *115*. doi: https://doi.org/10.1029/2009JB007127

Fischler, M., & Bolles, R. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, *24*, 381-395. doi: https://doi.org/10.1145/358669.358692

Ganaie, M., Hu, M., Malik, A., Tanveer, M., & Suganthan, P. (2022). Ensemble deep learning: A review. *Engineering Applications of Artificial Intelligence*, *115*. doi: https://doi.org/10.1016/j.engappai.2022.105151

Geneva, N., & Zabaras, N. (2020). Modeling the dynamics of PDE systems with physics-constrained deep auto-regressive networks. *Journal of Computational Physics*, *403*, 109056. doi: https://doi.org/10.1016/j.jcp.2019.109056

Gross, R. (1997). Earth rotation variations-long period. *Treatise on Geophysics*, *3*, 239–294. doi: https://doi.org/10.1016/B978-044452748-6.00057-2

Gross, R. (2008). An improved empirical model for the effect of long-period ocean tides on polar motion. *Journal of Geodesy*, *83*, 635–644. doi: https://doi.org/10.1007/s00190-008-0277-y

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE conference on computer vision and pattern recognition (CVPR).*

Hendon, H. (1995). Length of day changes associated with the Madden-Julian Oscillation. *Journal of the Atmospheric Sciences*, *52*, 2373–2383. doi: https://doi.org/10.1175/1520-0469(1995)052⟨2373:LODCAW⟩2.0.CO;2

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*, 1735–1780. doi: https://doi.org/10.1162/neco.1997.9.8.1735

Huber, P. (1964). Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, *35*, 73-101. doi: https://doi.org/10.1214/aoms/1177703732

Huber, P. (1973). Robust regression: Asymptotics, conjectures and Monte Carlo. *The Annals of Statistics*, *1*, 799-821. doi: https://doi.org/10.1214/aos/1176342503

Kalarus, M., Schuh, H., Kosek, W., Akyilmaz, O., Bizouard, C., Gambis, D., ... Zotov, L. (2010). Achievements of the Earth orientation parameters prediction comparison campaign. *Journal of Geodesy*, *84*, 587–596. doi: https://doi.org/10.1007/s00190-010-0387-1

Kehm, A., Hellmers, H., Bloßfeld, M., Dill, R., Angermann, D., Seitz, F., ... Enderle, W. (2023). Combination strategy for consistent final, rapid and predicted Earth rotation parameters. *Journal of Geodesy*, *97*. doi: https://doi.org/10.1007/s00190-022-01695-w

Kiani Shahvandi, M., Gou, J., Schartner, M., & Soja, B. (2022). Data driven approaches for the prediction of Earth's effective angular momentum functions. *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, *127*, 6550–6553. doi: https://doi.org/10.1109/IGARSS46834.2022.9883545

Kiani-Shahvandi, M., Schartner, M., & Soja, B. (2022). Neural ODE differential learning and its application in polar motion prediction. *Journal of Geophysical Research: Solid Earth*, *127*. doi: https://doi.org/10.1029/2022JB024775

Kiani-Shahvandi, M., & Soja, B. (2021). Small geodetic datasets and deep networks: Attention-based residual LSTM autoencoder stacking for geodetic time series.

In *7th international conference on machine learning, optimization, and data science.* doi: https://doi.org/10.1007/978-3-030-95467-3_22

Kiani-Shahvandi, M., & Soja, B. (2022). 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. *Advances in Space Research*, *70*. doi: https://doi.org/10.1016/j.asr.2022.05.042

Kiladis, G., Dias, J., Straub, K., Wheeler, M., Tulich, S., Kikuchi, K., ... Ventrice, M. (2014). A comparison of OLR and circulation-based indices for tracking the MJO. *Monthly Weather Review*, *142*(5), 1697–1715. doi: https://doi.org/10.1175/MWR-D-13-00301.1

Kingma, D., & Ba, J. (2015). Adam: A method for stochastic optimization. In *International conference on learning representations (iclr).*

Kur, T., Dobslaw, H., Śliwińska, J., Nastula, J., Wińska, M., & Partyka, A. (2022). Evaluation of selected short-term predictions of UT1-UTC and LOD collected in the second Earth orientation parameters prediction comparison campaign. *Earth, Planets and Space*, *74*. doi: https://doi.org/10.1186/s40623-022-01753-9

Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2016). Simple and scalable predictive uncertainty estimation using deep ensembles. In *Advances in neural information processing systems.* Curran Associates, Inc. Retrieved from https://doi.org/10.48550/arXiv.1612.01474

Lambeck, K. (1980). The Earth's variable rotation: Geophysical causes and consequences. *Cambridge University Press*. doi: https://doi.org/10.1017/CBO9780511569579

Liu, D., & Nocedal, J. (1989). On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, *45*, 503–528. doi: https://doi.org/10.1007/BF01589116

Liu, S., & Dobriban, E. (2020). Ridge regression: Structure, cross-validation, and sketching. In *International conference on learning representations.* Retrieved from https://openreview.net/forum?id=HklRwaEKwB

Luo, J., Chen, W., Ray, J., & Li, J. (2022). Short-term polar motion forecast based on the Holt-Winters algorithm and angular momenta of global surficial geophysical fluids. *Surveys in Geophysics*. doi: https://doi.org/10.1007/s10712-022-09733-0

Maksymilian, W., & Chen, K. (2020). Feature importance ranking for deep learning. In *Advances in neural information processing systems 33.* Morgan Kaufmann Publishers.

Marquardt, D., & Snee, R. (1975). Ridge regression in practice. *The American Statistician*, *29*, 3-20.

Mazzarella, A. (2007). The 60-year solar modulation of global air temperature: the Earth's rotation and atmospheric circulation connection. *Theoretical and Applied Climatology*, 193–199. doi: https://doi.org/10.1007/s00704-005-0219-z

Minderer, M., Djolonga, J., Romijnders, R., Hubis, F., Zhai, X., Houlsby, N., ... Lucic, M. (2021). Revisiting the calibration of modern neural networks. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, & J. W. Vaughan (Eds.), *Advances in neural information processing systems* (Vol. 34, pp. 15682–15694). Curran Associates, Inc. Retrieved from https://proceedings.neurips.cc/paper/2021/file/8420d359404024567b5aefda1231af24-Paper.pdf

Modiri, S., Belda, S., Heinkelmann, R., Hoseini, M., Ferrandiz, J., & Schuh, H. (2018). Polar motion prediction using the combination of SSA and Copula-based analysis. *Earth, Planets and Space*, *70*. doi: https://doi.org/10.1186/s40623-018-0888-3

Modiri, S., Belda, S., Hoseini, M., Heinkelmann, R., Ferrandiz, J., & Schuh, H. (2020). A new hybrid method to improve the ultra-short-term prediction of LOD. *Journal of Geodesy*, *94*. doi: https://doi.org/10.1007/s00190-020-01354-y

Nastula, J., & Ponte, R. (1999). Further evidence for oceanic excitation of polar motion. *Geophysical Journal International*, *139*, 123–130. doi: https://doi.org/10.1046/j.1365-246X.1999.00930.x

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... Chintala, S. (2019). Pytorch: An imperative style, high-performance deep learning library. In *Advances in neural information processing systems 32* (pp. 8024–8035). Curran Associates, Inc. Retrieved from `http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf`

Petit, G., & Luzum, B. (2010). *IERS Technical Note 36*. IERS.

Ratcliff, J., & Gross, R. (2022). Combinations of Earth orientation measurements: SPACE2021, COMB2021, and POLE2021. *JPL Publications*.

Raut, S., Modiri, S., Heinkelmann, R., Balidakis, K., Belda, S., Kitpracha, C., & Schuh, H. (2022). Investigating the relationship between length of day and El-Niño using wavelet coherence method. In (pp. 1–6). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: https://doi.org/10.1007/1345_2022_167

Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning internal representations by error propagation. In *Parallel distributed processing* (pp. 318–362). MIT Press. doi: https://dl.acm.org/doi/10.5555/104279.104293

Rusch, T., & Mishra, S. (2021). Coupled oscillatory recurrent neural network (coRNN): An accurate and (gradient) stable architecture for learning long time dependencies. In *International conference on learning representations.*

Rusch, T., Mishra, S., Erichson, N., & Mahoney, M. (2022). Long expressive memory for sequence modeling. In *International conference on learning representations.*

Sullivan, T. (2015). *Introduction to uncertainty quantification* (Vol. 63). Springer, Texts in Applied Mathematics.

Sun, Q., Zhou, W., & Fan, J. (2020). Adaptive Huber regression. *Journal of the American Statistical Association*, *115*. doi: https://doi.org/10.1080/01621459.2018.1543124

Szandała, T. (2021). Review and comparison of commonly used activation functions for deep neural networks. *Bio-inspired Neurocomputing*, 203–224. doi: https://doi.org/10.1007/978-981-15-5495-7_11

Teunissen, P. (2003). *Adjustment theory: an introduction.* VSSD Series on Mathematical Geodesy and Positioning.

Timmermann, A., An, S., Kug, J., Jin, F., Cai, W., Capotondi, A., ... Zhang, X. (2018). El Niño–Southern Oscillation complexity. *Nature*, *559*(7715), 535–545. doi: https://doi.org/10.1038/s41586-018-0252-6

Visbeck, M., Hurrell, J., Polvani, L., & Cullen, H. (2001). The North Atlantic Oscillation: Past, present, and future. *Proceedings of the Natural Academy of Sciences*, *98*, 12876–12877. doi: https://doi.org/10.1073/pnas.231391598

Volland, H. (1996). Atmosphere and Earth's rotation. *Surveys in Geophysics*, *17*, 101–144. doi: https://doi.org/10.1007/BF01904476

Wolter, K., & Timlin, M. (1993). Monitoring ENSO in COADS with a seasonally adjusted principal component index. Norman: Proceedings of the 17th Climate Diagnostics Workshop.

Xu, X., Zhou, Y., Duan, P., Fang, M., Kong, Z., Xu, C., & An, X. (2022). Contributions of oceanic and continental AAM to interannual variation in ΔLOD with the detection of 2020–2021 La Nina event. *Journal of Geodesy*, *96*. doi: https://doi.org/10.1007/s00190-022-01632-x

Zhang, R., Liu, Y., & Sun, H. (2020). Physics-informed multi-LSTM networks for metamodeling of nonlinear structures. *Computer Methods in Applied Mechanics and Engineering*, *369*, 113226. doi: https://doi.org/10.1016/

j.cma.2020.113226