



HAL
open science

Assessment of length-of-day and universal time predictions based on the results of the Second Earth Orientation Parameters Prediction Comparison Campaign

Justyna Śliwińska-Bronowicz, Tomasz Kur, Malgorzata Wińska, Henryk Dobslaw, Jolanta Nastula, Aleksander Partyka, Santiago Belda, Christian Bizouard, Dale Boggs, Sara Bruni, et al.

► To cite this version:

Justyna Śliwińska-Bronowicz, Tomasz Kur, Malgorzata Wińska, Henryk Dobslaw, Jolanta Nastula, et al.. Assessment of length-of-day and universal time predictions based on the results of the Second Earth Orientation Parameters Prediction Comparison Campaign. *Journal of Geodesy*, 2024, 98, 10.1007/s00190-024-01824-7 . insu-04858857

HAL Id: insu-04858857

<https://insu.hal.science/insu-04858857v1>

Submitted on 30 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Assessment of length-of-day and universal time predictions based on the results of the Second Earth Orientation Parameters Prediction Comparison Campaign

Justyna Śliwińska-Bronowicz¹ · Tomasz Kur² · Małgorzata Wińska³ · Henryk Dobslaw⁴ · Jolanta Nastula¹ · Aleksander Partyka¹ · Santiago Belda⁵ · Christian Bizouard⁶ · Dale Boggs⁷ · Sara Bruni⁸ · Lue Chen⁹ · Mike Chin⁷ · Sujata Dhar⁴ · Robert Dill⁴ · Jose Manuel Ferrandiz⁵ · Junyang Gou¹⁰ · Richard Gross⁷ · Sonia Guessoum⁵ · Songtao Han⁹ · Robert Heinkelmann⁴ · Christopher Irrgang⁴ · Mostafa Kiani Shahvandi¹⁰ · Jia Li¹¹ · Marcin Ligas¹² · Lintao Liu¹³ · Weitao Lu⁹ · Volker Mayer¹⁴ · Maciej Michalczak¹² · Sadegh Modiri¹⁵ · Michiel Otten⁸ · Todd Ratcliff⁷ · Shrishail Raut^{4,16} · Jan Saynisch-Wagner⁴ · Matthias Schartner¹⁰ · Erik Schoenemann¹⁷ · Harald Schuh^{4,16} · Benedikt Soja¹⁰ · Xiaoqing Su¹⁸ · Daniela Thaller¹⁵ · Maik Thomas⁴ · Guocheng Wang¹³ · Yuanwei Wu¹⁹ · Xueqing Xu²⁰ · Xinyu Yang¹⁹ · Xin Zhao¹⁹ · Zhijin Zhou⁹

Received: 31 October 2023 / Accepted: 25 January 2024 / Published online: 20 March 2024
© The Author(s) 2024

Abstract

Predicting Earth Orientation Parameters (EOP) is crucial for precise positioning and navigation both on the Earth's surface and in space. In recent years, many approaches have been developed to forecast EOP, incorporating observed EOP as well as information on the effective angular momentum (EAM) derived from numerical models of the atmosphere, oceans, and land-surface dynamics. The Second Earth Orientation Parameters Prediction Comparison Campaign (2nd EOP PCC) aimed to comprehensively evaluate EOP forecasts from many international participants and identify the most promising prediction methodologies. This paper presents the validation results of predictions for universal time and length-of-day variations submitted during the 2nd EOP PCC, providing an assessment of their accuracy and reliability. We conduct a detailed evaluation of all valid forecasts using the IERS 14 C04 solution provided by the International Earth Rotation and Reference Systems Service (IERS) as a reference and mean absolute error as the quality measure. Our analysis demonstrates that approaches based on machine learning or the combination of least squares and autoregression, with the use of EAM information as an additional input, provide the highest prediction accuracy for both investigated parameters. Utilizing precise EAM data and forecasts emerges as a pivotal factor in enhancing forecasting accuracy. Although several methods show some potential to outperform the IERS forecasts, the current standard predictions disseminated by IERS are highly reliable and can be fully recommended for operational purposes.

Keywords Earth Orientation Parameters (EOP) · Length-of-day (LOD) · UT1–UTC · Prediction

1 Introduction

Time-variable Earth Orientation Parameters (EOP) serve as an essential link between the celestial and terrestrial reference frames, allowing for the transformation of coordinates between them (Petit and Luzum 2010). This connection has a wide range of applications in modern geodesy (precise positioning and navigation on the Earth's surface and in space, and determining orbits of satellites), astronomy (astronomical instruments orientation), and in the operation

of space missions. EOP comprise polar motion (PM), differences between universal time and coordinated universal time (UT1–UTC), or its time-derivative length-of-day (LOD) variations, and corrections to the conventional precession-nutation model, i.e., celestial pole offsets (CPO). EOP are routinely determined with high accuracy with the means of Global Navigation Satellite Systems (GNSS, Ferland et al. 2009; Steigenberger et al. 2006; Zajdel et al. 2020), Satellite Laser Ranging (SLR, Bloßfeld et al. 2018; Glaser et al. 2015; Sośnica et al. 2019), Very Long Baseline Interferometry (VLBI, Karbon et al. 2017; Robertson et al. 1985; Schuh

Extended author information available on the last page of the article

and Böhm 2013; Sovers et al. 1998), and Doppler Orbitography and Radiopositioning Integrated by Satellite (DORIS, Moreaux et al. 2023).

The International Earth Rotation and Reference Systems Service (IERS) is a widely recognized institution that is responsible for the regular delivery of the EOP series (Bizouard and Gambis 2009). The monitoring and sharing of daily, monthly, and long-term EOP data, as well as leap second announcements, are the responsibilities of the Earth Orientation Center hosted by the Paris Observatory that acts under the auspices of IERS (Bizouard et al. 2019; Gambis 2004; Gambis and Luzum 2011). The former version of the final EOP series provided by IERS, i.e., IERS EOP 14 C04, was consistent with the conventional International Terrestrial Reference Frame 2014 (ITRF 2014) and the Second International Celestial Reference Frame (ICRF2) (Bizouard et al. 2019). The latest implementation of the C04 series, namely IERS EOP 20 C04, aligned with the most recent international terrestrial reference frame ITRF 2020 (Altamimi et al. 2023) and celestial reference frame ICRF3 (Charlot et al. 2020), was made available in February 2023 (IERS Message No. 471 distributed by the IERS Central Bureau: https://datacenter.iers.org/data/2/message_471.txt). The formal errors associated with the final IERS EOP 20 C04 data delivered by the IERS currently indicate an uncertainty level of approximately 10 microseconds (μs) for UT1–UTC, while rapidly processed observations may have errors several times higher.

Many operational applications require the knowledge of EOP in real time. However, the complexity of computations and the processing of data from various observational techniques, each characterized by different levels of accuracy, stability, availability, and temporal resolution, increase the processing time required for the EOP determination. As a result, rapidly processed but less precise EOP datasets are provided once per day, while the most accurate solutions are delivered with delays of 30 days and longer. Consequently, most real-time applications exploit short-term EOP predictions, which are currently processed by institutions around the world. For example, the U.S. Naval Observatory (USNO) routinely issues forecasts that are subsequently officially disseminated by the IERS (Luzum et al. 2001). In addition, Deutsches GeoForschungsZentrum (GFZ; Dill et al. 2019), Eidgenössische Technische Hochschule (ETH, Kiani Shahvandi et al. 2022a; Soja et al. 2022), Jet Propulsion Laboratory (JPL; Gross et al. 1998), and the European Space Agency (ESA; Bruni et al. 2021) deliver EOP forecasts on a regular basis. The latter two institutes provide truly independently processed series of EOP based on data from various observational techniques (Bruni et al. 2021; Ratcliff and Gross 2019), whereas other groups such as ETH and GFZ partly rely on IERS input data (Kiani Shahvandi et al. 2022a; Dill et al. 2019).

For most real-time applications, such as precise positioning and navigation, it is sufficient to have an EOP forecast for the next few days. EOP predictions for a duration of at least 1 year into the future might be utilized for climate forecasting and long-term satellite orbit prediction (Lei et al. 2023). In terms of required EOP prediction accuracy, the current accuracy of predictions (below 1 ms for UT1–UTC ultra-short-term predictions disseminated by IERS) meets the needs for ephemerides or pointing astronomical instruments (Luzum 2010). For the purpose of tracking and navigating interplanetary spacecraft, forecast accuracy requirements of approximately 10 mas for PM and 0.65 ms for UT1–UTC have been formulated (Oliveau and Freedman 1997). Even higher accuracy of EOP predictions is essential for real-time satellite orbit determination and VLBI analysis. Given the broad range of possible future satellite applications related to, e.g., the monitoring of precipitable water vapor, tsunamis, and earthquakes, the demand for accurate estimates and predictions of EOP might even increase further.

In addition to the observed EOP values, contemporary Earth orientation predictions commonly incorporate analyses and forecasts of effective angular momentum (EAM). In theory, changes in the rotation of the solid Earth can be examined by employing the principle of conserving angular momentum within the Earth system, including the surrounding fluid layers (atmosphere, oceans, continental hydrosphere) (Gross 2007). According to this principle, the rotation of the solid Earth is altered by external torques, internal redistribution of mass, and the exchange of angular momentum between the solid Earth and the adjacent fluid layers. EAM functions describe the excitation of Earth orientation changes caused by the sum of atmospheric, oceanic, and hydrospheric mass redistributions (Barnes 1983; Brzeziński 1992). In particular, the axial component (χ_3) of EAM excites UT1–UTC. Previous works (e.g., Gross et al. 2004) have shown that atmospheric and oceanic effects can explain up to 90% of observed UT1–UTC variation. Such relationships between modelled EAM and the EOP have prompted the adoption of EAM forecasts for predicting Earth rotation especially for shorter time horizons (Dobslaw and Dill 2018; Freedman et al. 1994).

Between 2006 and 2008, a comprehensive comparison and evaluation of different EOP forecasts was carried out as part of the EOP Prediction Comparison Campaign (EOP PCC) organized by Vienna University of Technology and Centrum Badań Kosmicznych Polskiej Akademii Nauk (CBK PAN) with the support of the IERS (Kalarus et al. 2010). The primary objectives of this initiative were to identify the optimal methods for EOP forecasting and to eventually develop a combined series of EOP predictions. The EOP PCC yielded valuable insights into various prediction techniques under consistent rules and conditions. The benefits of combining submitted solutions were demonstrated, and it was shown

that incorporating atmospheric angular momentum (AAM) forecast data as an input improves EOP prediction accuracy. However, the choice of the best prediction technique was shown to be dependent on the selected EOP and the targeted prediction horizon, indicating that no single technique outperformed all the others (Kalarus et al. 2010).

Since the completion of the EOP PCC in 2008, considerable progress has been made in the development of observational data, the advancement of new forecasting methods, and the understanding of the role of EAM in EOP variations. There has also been a substantial increase in the number of teams involved in EOP prediction, with different teams applying various inputs, forecasting algorithms, and prediction horizons. Consequently, there are clear differences in the accuracy of the resultant individual predictions. As a result of the progress in the field of EOP forecasting, the IERS has established a dedicated working group (Working Group on the 2nd EOP PCC, for details see <https://www.iers.org/IERS/EN/Organization/WorkingGroups/PredictionComparison/predictionComparison.html>) with the primary objective of conducting a new comprehensive reassessment of the current capabilities of EOP forecasting. The 2nd EOP PCC was handled in collaboration between CBK PAN and GFZ. The EOP PCC Office, established at CBK PAN, was responsible for routinely collecting and evaluating forecasts submitted by registered campaign participants (Śliwińska et al. 2022). The operational phase of the 2nd EOP PCC lasted 70 weeks from September 1, 2021, until December 31, 2022, involving 18 active teams from 23 institutes around the world, who routinely provided predictions of all EOP based on 50 different methods distinguished with individual IDs. During the campaign, 7327 individual predictions of any EOP were collected. A summary of the most relevant events related to the 2nd EOP PCC and some technical aspects, such as file format requirements and data submission rules, as well as more detailed statistics, can be found in Śliwińska et al. (2022).

Although only the values of PM, UT1–UTC, and precession-nutation are sufficient to perform coordinate transformations between the celestial and terrestrial reference frames, the 2nd EOP PCC Office also collected and analyzed LOD predictions, even though LOD is not directly included in the transformation matrix. LOD can be defined as the first negative derivative of UT1–UTC with respect to time, which means LOD is equal to the rate of change of UT1–UTC over time, expressed as:

$$\text{LOD} = -\frac{d(\text{UT1} - \text{UTC})}{dt} \quad (1)$$

In turn, UT1–UTC can be estimated from LOD by integrating LOD values and adding the value of UT1–UTC at zero epoch (Miksch et al. 2019). Estimation of LOD is

important because of the inherent disturbances arising from GNSS data in the case of UT1–UTC, as it is challenging to reliably distinguish between linear drifts in the satellite constellation and changes in the Earth's rotational angle. LOD remains unaffected by this issue, making it feasible to estimate accurately using GNSS alone, without the requirement of a consistent combination with VLBI. LOD can be therefore used to densify and predict UT1–UTC (Senior et al. 2010).

In the current study, we present a scientific summary of the 2nd EOP PCC results focusing on a thorough assessment of predictions of UT1–UTC and LOD submitted by the campaign participants. Our goal is to provide an objective analysis of a wide variety of UT1–UTC and LOD forecasts developed using diverse approaches, prediction methods, and input data. We intend to determine the current forecasting capabilities for these parameters and offer recommendations on the most effective methodologies for this purpose. No such comprehensive analysis of multiple UT1–UTC and LOD forecasts has been conducted since the conclusion of the 1st EOP PCC. Section 2 of this paper provides a brief overview of prediction methodologies and input data used by contributors, along with statistics regarding the number of valid files received and the most popular prediction horizons. Section 3 describes the methodology for prediction evaluation. Section 4 presents a thorough evaluation of the predictions, divided into a general overview of predictions without distinguishing specific IDs (Sect. 4.1) and a detailed assessment of all predictions (Sect. 4.2). Section 5.1 presents an analysis of the dependence of prediction accuracy on the considered time period, whereas Sect. 5.2 deals with the transformation between LOD and UT1–UTC. Section 6 includes a ranking of all methods based on previously applied criteria, and finally, Sect. 7 summarizes all the results and provides concluding remarks.

2 Overview of UT1–UTC and LOD predictions

2.1 Prediction methods and input data exploited by participants

During the 2nd EOP PCC, UT1–UTC was predicted by 15 participants using 25 different combinations of methods and inputs (indicated with individual IDs), while LOD predictions were performed by 12 teams with the use of 25 approaches (Table 1). 15 of the IDs provided forecasts for both parameters, which makes a total of 35 IDs providing UT1–UTC and/or LOD predictions. In total, the EOP PCC Office received 1399 files for UT1–UTC and 1226 files for LOD predictions. A summary of the prediction methods and input data used by the participants to predict UT1–UTC and LOD is presented in Table 2 and a full description of the

Table 1 Summary of the number of all registered participants, methods, and prediction files involved in the 2nd EOP PCC compared with the 1st EOP PCC.

	1st EOP PCC ^a			2nd EOP PCC		
	All parameters	UT1–UTC	LOD	All parameters	UT1–UTC	LOD
Number of registered participants	11	7	8	18	15	12
Number of institutes	11	7	8	23	19	16
Number of countries of participant origin	8	6	7	8	8	7
Total number of team members	12	7	9	55	41	34
Number of registered methods (IDs)	20 (+ 1 combined prediction series)	13 (+ 1 combined prediction series)	14 (+ 1 combined prediction series)	50	25	25
Number of predictions	6176	861 (ultra-short-term) 493 (short-term) 238 (medium-term) ^b	808 (ultra-short-term) 548 (short-term) 268 (medium-term)	7327	1399	1226

The table only contains information relating to active participants of the campaign (i.e., those who have submitted their forecasts at least once). A participant is an institute or a consortium of several institutes that work together on a given forecast and have a common ID or several common IDs. An institute is a single research center that either acts as a participant in developing its own forecasts or cooperates with other institutes

^aBased on Kalarus et al. (2010)

^bUltra-short-term < 10 days, short-term < 30 days, medium-term < 500 days

groups is provided in “Appendix 1”. A wide variety of methods are exploited, but algorithms based on least squares (LS) and their modifications and approaches based on machine learning (ML), were most popular among all IDs. However, only two institutes used ML-based methods for UT1–UTC and LOD forecasting. ML has been declared in 14 different IDs, 13 of which were developed by one group (ETH) (see Table 2 and “Appendix 1” for more details). The LS methods are usually combined with autoregression (AR), autoregressive integrated moving average (ARIMA), convolution, local approximation (LA), or kriging methods. Of the methods that do not belong to the two most popular groups (ML and LS), the most noteworthy are Kalman filtering, adaptive polyharmonic models, normal time–frequency transform (NTFT), singular spectrum analysis (SSA), and Copula approaches (Table 2).

The input data exploited are more homogeneous than methods of prediction as almost all participants use the EOP 14 C04 series provided by the IERS, usually supplemented with daily datasets from the IERS Rapid Service/Prediction Center at USNO. Only two IDs use the EOP final series from other data sources. A total of 23 IDs declared the use of EAM data as an additional input and most of these exploit EAM series provided by GFZ (Dill et al. 2013, 2019, 2022). Although GFZ routinely delivers data and 6-day forecasts for atmospheric, oceanic, hydrological, and sea-level angular

momentum (AAM, OAM, HAM, and SLAM, respectively), not all participants used each of the four components.

2.2 Submissions of UT1–UTC and LOD predictions

Figure 1 presents statistics of submitted files for UT1–UTC and LOD predictions, specifically the number of files submitted by each ID throughout the entire campaign period (Fig. 1a), the total number of all prediction files received on individual submission days (Fig. 1b), and the most common prediction lengths (Fig. 1c). Figure 1a shows that only two IDs (IDs 100 and 126) provided valid predictions of UT1–UTC in all 70 weeks of the campaign, but another 10 out of the 25 total IDs involved in UT1–UTC prediction provided more than 60 UT1–UTC forecasts. The IDs for which we received the lowest number of predictions (IDs 146–149) were registered around halfway through the 2nd EOP PCC operation. Even though as many methods were used for LOD forecasting as for UT1–UTC prediction, there are over 170 fewer files submitted for LOD (Table 1, Fig. 1). This is partly because in the case of LOD there were more methods registered later in the campaign (IDs 141–145, 156, 157), and some participants stopped (e.g., ID 108) or started (ID 117) forecasting LOD a few weeks after initiating active participation. Over 60 LOD predictions were sent by 10 out of the 25 IDs, of which only 2 provided files in all weeks of the campaign duration. The least active participant (ID 157)

Table 2 Summary of prediction methods and input data used by participants to predict UT1–UTC and LOD

ID	Predicted parameters	Prediction method	Input data
100	UT1–UTC; LOD	LS + AR	IERS 14 C04; IERS finals.daily
101	UT1–UTC; LOD	LS + AR with piecewise parameter optimization	IERS 14 C04; IERS finals.daily; GFZ AAM + OAM + HAM + SLAM data and 6-day predictions
102	UT1–UTC; LOD	LS + AR	IERS 14 C04; GFZ AAM 6-day predictions
103	UT1–UTC	Adaptive polyharmonic models with infrequent components	IERS 14 C04
104	UT1–UTC; LOD	Kalman filter	EOP data from IGS, ILRS, JPL, GSFC; NCEP/NCAR (before 9/26/21) and GFZ (after 10/19/21) AAM data and 6-day predictions
105	UT1–UTC; LOD	LS + AR	IERS 14 C04; IERS finals.daily; GFZ AAM + OAM + HAM data and 6-day predictions
107	UT1–UTC	Differential LS + AR	IERS 14 C04; IERS finals.daily
108	UT1–UTC; LOD	LS + multi-elements AR (LS + MAR)	EOP 14 C04 IERS; GFZ AAM data
112	UT1–UTC; LOD	LS + Convolution	IERS 14 C04; IERS finals.daily; GFZ AAM + OAM + HAM + SLAM data
113	UT1–UTC; LOD	LS + AR	IERS 14 C04
114	UT1–UTC; LOD	LS collocation	IERS 14 C04
115	UT1–UTC; LOD	Neural networks	IERS 14 C04
116	UT1–UTC	LS + AR	ESA ERP estimates based on data from GNSS Rapid and Final, SLR, DORIS, VLBI intensive sessions, VLBI rapid turnaround sessions; GFZ AAM + OAM + HAM data and 6-day predictions
117	UT1–UTC; LOD	SSA + Copula	IERS 14 C04; IERS Bulletin A; GFZ AAM + OAM + HAM + SLAM data and 90-day predictions
118	UT1–UTC	Local approximation (LA) + LS	IERS 14 C04
121	UT1–UTC; LOD	Normal time–frequency transform (NTFT)	IERS 14 C04
122	UT1–UTC; LOD	Weighted LS + ARIMA	IERS 14 C04; IERS finals.daily
123	UT1–UTC; LOD	Unknown	IERS 14 C04; GFZ AAM + OAM data

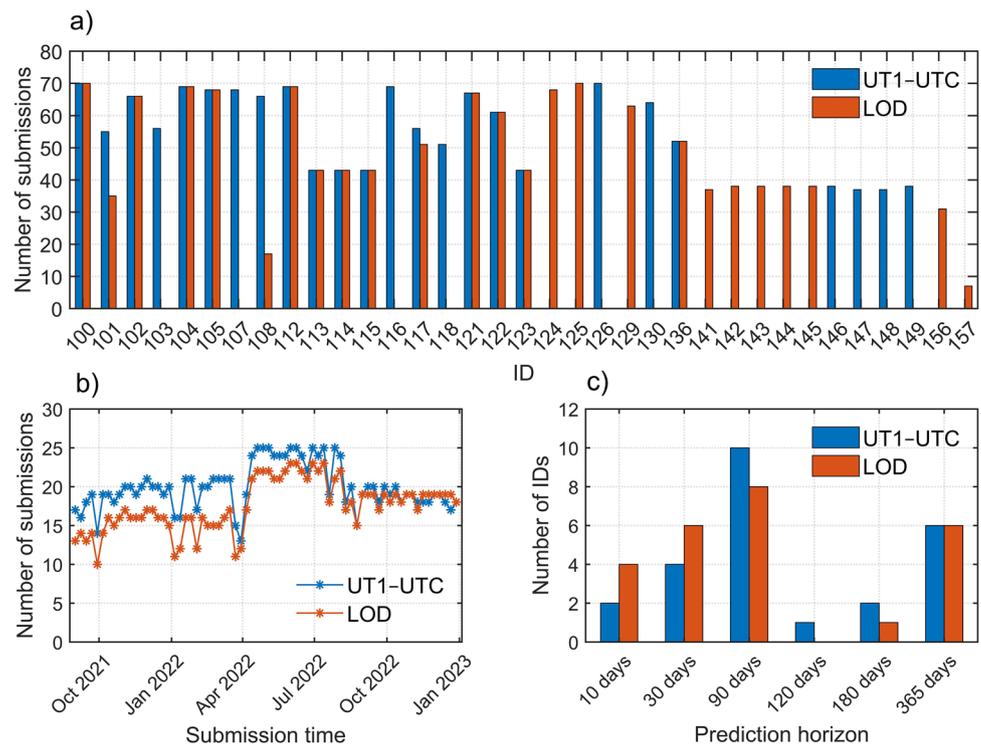
Table 2 (continued)

ID	Predicted parameters	Prediction method	Input data
124	LOD	Encoder–decoder long short-term memory (LSTM)	IERS finals.daily; GFZ AAM data and 6-day predictions
125	LOD	LSTM auto-encoder stacking augmented with residual learning and attention mechanism	SYRTE EOP series ^a ; GFZ AAM data and 6-day predictions
126	UT1–UTC	First-order neural ordinary differential equations (ODEs)	SYRTE EOP series; GFZ AAM data and 6-day predictions
129	LOD	First-order neural (ODEs)	SYRTE EOP series; GFZ AAM data and 6-day predictions
130	UT1–UTC	First-order neural (ODEs)	SYRTE EOP series; GFZ AAM + HAM + OAM + SLAM data and 6-day predictions
136	UT1–UTC; LOD	LS + AR	IERS 14 C04; IERS finals.daily; GFZ AAM + OAM + HAM data and 6-day predictions
141	LOD	LS + kriging	IERS 14 C04; IERS finals.daily
142	LOD	Encoder–decoder LSTM	IERS finals.daily; GFZ AAM + HAM + OAM + SLAM data and 6-day predictions
143	LOD	Revised multilayer perceptron	IERS finals.daily; GFZ AAM data and 6-day predictions
144	LOD	First-order neural (ODEs)	IERS finals.daily; GFZ AAM data and 6-day predictions
145	LOD	First-order neural (ODEs)	SYRTE EOP series; GFZ AAM data and 6-day predictions
146	UT1–UTC	First-order neural (ODEs)	SYRTE EOP series; GFZ AAM + HAM + OAM + SLAM data and 6-day predictions
147	UT1–UTC	First-order neural ODEs with Residual modelling	SYRTE EOP series; IERS finals.daily predictions; GFZ AAM + HAM + OAM + SLAM data and 6-day predictions
148	UT1–UTC	First-order neural ODEs with Residual modelling	IERS finals.daily data and predictions; GFZ AAM + HAM + OAM + SLAM data and 6-day predictions
149	UT1–UTC	First-order neural ODEs with Residual modelling	SYRTE EOP series; GFZ AAM + HAM + OAM + SLAM data and 6-day predictions
156	LOD	LS + ARIMA	IERS 14 C04; IERS finals.daily
157	LOD	Dynamic mode decomposition (DMD)	IERS 14 C04; IERS finals.daily

A full description of the groups is provided in “Appendix 1”

^aSYRTE EOP series are the eopc04_extended files provided by Systèmes de Référence Temps Espace (SYRTE) of Paris Observatory. SYRTE EOP are based on rapid data only and does not exploit the final post-processed solution which is typically available 4 weeks later

Fig. 1 **a** Number of UT1–UTC and LOD predictions submitted by each ID during the whole campaign duration, **b** number of all UT1–UTC and LOD predictions submitted on individual submission days, and **c** number of IDs providing UT1–UTC and LOD predictions for the specified prediction horizon



registered only in November 2022 and sent just 7 LOD forecasts.

For about the first 7 months of the campaign, apart from a few occasional drops, the number of valid UT1–UTC and LOD forecasts received each Wednesday by the EOP PCC Office was relatively stable (around 20 predictions of UT1–UTC and 16 predictions of LOD) (Fig. 1b). After a noticeable decrease at the end of March 2022, the quantity of submitted files increased (around 25 for UT1–UTC and 22 for LOD) because of the registration of several new methods by one of the participants. These numbers eventually diminished in September 2022 to reach 18 or 19 for both parameters.

The 2nd EOP PCC participants had full freedom in terms of the forecast horizon, with the only condition that the predictions could not be longer than 365 days into the future. A histogram of the prediction horizons used in the 2nd EOP PCC is shown in Fig. 1c. The most popular forecast length for both UT1–UTC and LOD was 90 days into the future and the second most popular was the prediction for 365 days ahead. In contrast, in the 1st EOP PCC, the most frequently submitted prediction lengths were ultra-short-term forecasts (predictions for up to 10 days into the future) (Kalarus et al. 2010).

3 Prediction evaluation methodology

Because of the large number of IDs predicting both parameters, we decided to present detailed results in groups of contributions. As stated in Sect. 2.1, the exploited input data were rather homogeneous, so we formed the groups manually based on the prediction method used. The most popular approach among participants was exploiting LS in combination with AR and with possible modification of the method, so we decided to distinguish two groups based on this group of methods. The first group, “LS + AR”, includes IDs that do not use EAM data, while the second, “LS + AR + EAM”, includes IDs that additionally use EAM data. The third group (“ML”) includes ML approaches, most of which use EAM data as well (only for one ID, i.e., ID 115, we do not have a clear information whether this type of data was used or not). The last group, “Other”, contains more unique methods that cannot be included in the previous three groups. In this group, 4 out of 8 IDs declared the use of EAM data (IDs 102, 104, 117, and 123). Every group contains between 5 and 8 IDs. A summary of the method clustering with information on the assignment of IDs to the groups is shown in Table 3.

Following the previous campaign, we use the mean absolute error (MAE) as a basic parameter for EOP prediction evaluation (Kalarus et al. 2010; Kur et al. 2022; Śliwińska et al. 2022):

Table 3 ID assignment to groups according to prediction method and EAM usage

	LS + AR	LS + AR + EAM	ML	Other
UT1–UTC	100, 107, 113, 114, 118, 122	101, 105, 108, 112, 116, 136	115, 126, 130, 146, 147, 148, 149	102, 103, 104, 117, 121, 123
LOD	100, 113, 114, 122, 156	101, 105, 108, 112, 136	115, 124, 125, 129, 142, 143, 144, 145	102, 104, 117, 121, 123, 141, 157

$$MAE_i = \frac{1}{n_p} \sum_{j=1}^{n_p} |x_i^{obs} - x_{i,j}^{pred}|, \quad i = 1, 2, \dots, I, \quad (2)$$

where n_p is the number of valid prediction files submitted by a campaign participant under a single ID, x_i^{obs} is the EOP reference data for the i th day of reference series, $x_{i,j}^{pred}$ is the predicted value for the i th day of the j th prediction, and I is the forecast horizon. In this study, we use $I = 10$ days and $I = 30$ days.

As reference data, we use the IERS 14 C04 series. However, some forecasts evaluated in this study are predicted from other input data (e.g., IDs 104, 116, see Table 2). Since prediction algorithms are typically optimized with respect to the underlying input data, this could potentially result in increased MAE for forecasts based on the data other than the official IERS solution. Some insights regarding the differences between individual reference EOP series are included in Śliwińska et al. (2022), where the IERS 14 C04 solution was compared with other combined EOP data (e.g., SPACE, Bulletin A), as well as with the single-technique series based on GNSS, SLR and VLBI measurements. It was found that for UT1–UTC, the RMS of differences between IERS 14 C04 and other solutions ranged from 0.019 to 0.212 ms, while for LOD, these values ranged from 0.010 and 0.098 ms. Hence, the choice of reference data for assessing EOP prediction may be important in individual cases.

The results for UT1–UTC from each of the group are also compared with predictions disseminated by the IERS, which are produced by USNO for 90 days into the future. The IERS forecasts for UT1–UTC rely on observations from VLBI, GNSS, and SLR, alongside AAM analysis and prediction data. The AAM data utilized for UT1–UTC predictions are sourced from a combination of the operational National Centers for Environmental Prediction (NCEP) and Navy Global Environment Model (NAVGEM) (IERS 2023). USNO employs AAM-based projections to determine UT1–UTC forecasts for a period extending up to 7.5 days ahead. For longer-term predictions, LOD excitations are combined smoothly with the longer-term UT1–UTC predictions. The method for predicting UT1–UTC beyond 7.5 days utilizes a straightforward approach known as differencing (McCarthy and Luzum 1991). Details on the computation are provided in IERS Annual Report (IERS 2023).

The IERS forecasts were taken from daily updated files `finals.2000A.daily` (<https://www.iers.org/iers/en/DataProducts/EarthOrientationData/eop.html>—accessed 2023-05-01). The ID 200 was assigned for these predictions. Note that LOD forecasts are not provided by IERS. However, to relate the LOD predictions provided by campaign participants to the forecasts delivered by IERS, we converted the UT1–UTC predictions from ID 200 into LOD values.

During routine checking of prediction files, the EOP PCC Office detected some erroneous predictions. Sometimes outliers were caused by problems in input preparation resulting from availability issues with the most recent EOP or EAM data, defects in the software, or data errors in the files (e.g., predictions provided for the wrong days) (personal communication with participants). Incorrect predictions may affect the objective assessment of the accuracy of a given forecasting method, which is undesirable; however, the EOP PCC Office had to avoid interfering with the supplied files or any manual modification of the submissions. To solve the problem of erroneous predictions, we decided to incorporate a two-step screening process.

In the first step of data selection, called “ σ criterion”, we computed the standard deviation S_j of the differences between reference and prediction ($x_i^{obs} - x_{i,j}^{pred}$) for all submitted predictions separately, and then, the standard deviation of differences was computed for all submissions (S_{total}). Individual predictions with $S_j > S_{total}$ were eliminated from further processing. This step removes highly inaccurate predictions whose values significantly differ from observational data and other predictions, which could have been caused by factors such as incorrect units.

In the second step, called “ β criterion”, we considered each ID individually to eliminate those single predictions of a given participant that noticeably differ from the other predictions by the same provider. To achieve this, following Kalarus et al. (2010), we determined a threshold independently for each ID using the β parameter, the value of which depends on the median of differences between an observation and a prediction (median absolute error, MDAE):

$$\beta_j = \sum_{i=1}^I \left(\alpha \cdot MDAE_i - |x_i^{obs} - x_{i,j}^{pred}| \right), \quad (3)$$

$$\text{MDAE}_i = \text{median}\left(\left|x_i^{\text{obs}} - x_{i,1}^{\text{pred}}\right|, \dots, \left|x_i^{\text{obs}} - x_{i,n_p}^{\text{pred}}\right|\right), \quad (4)$$

where the α value was chosen subjectively to preserve a representative number of predictions. In this study, we used $\alpha = 3$. All predictions with $\beta_j < 0$ were eliminated from further processing.

Table 4 provides the number of UT1–UTC predictions eliminated by the σ and β criteria for each ID for the 10- and 30-day prediction horizon as well as the ratio of rejected files to all submitted files. The use of the σ condition resulted in the elimination of nine files for ID 102 (for both prediction horizons), whereas no gross errors were detected for the other IDs. Indeed, for ID 102, we noticed some problems mainly in October 2021 and again between September and November 2022. The participant reported that gross errors in the first period might have resulted from the use of UT1–TAI difference instead of UT1–UTC, while the GFZ AAM forecasts (upon which those predictions are based for the first 7 days) were not downloaded correctly in the second period (personal correspondence with Christian Bizouard). Some of the participants with the largest number of rejections reported problems with the availability of the most recent EOP values from the IERS EOP Rapid solution, using observational data from wrong days, or unintentional lack of use of daily IERS EOP Rapid data. In general, after applying σ and β criteria, 6.0% of UT1–UTC predictions for the 10-day horizon and 4.8% of UT1–UTC predictions for the 30-day prediction horizon were eliminated.

The number of LOD predictions eliminated by the σ and β criteria are presented in Table 5. Again, the use of the σ criterion eliminated several files for ID 102, which was probably related to the problem with the download of the AAM predictions. In addition, two files for ID 129 were eliminated and this participant reported a few problems with the availability of EAM forecasts delivered by GFZ. The largest number of rejections for β was detected for ID 105. This was caused mainly by improper accounting for long-periodic ocean tides at the beginning of the campaign. This problem was then traced back and solved by the participant (the details of this issue are described in Kur et al. 2022). In general, after the applying σ and β criteria, 4.7% of LOD predictions for 10-day horizon and 2.2% of LOD predictions for 30-day prediction horizon were eliminated.

As mentioned above, one of the reasons for the reduction in the accuracy of 10-day forecasts that were eliminated from further processing may be problems with the availability of EAM data and their 6-day forecasts provided by GFZ, which are routinely used by participants, especially for the LOD forecasting. Therefore, in Tables 4 and 5, we also included the number of rejections for days in which the GFZ EAM predictions were unavailable. According to GFZ (personal

correspondence with Robert Dill), we learned that in the operational phase of the campaign, in 11 out of 70 weeks there were problems with the availability of EAM forecasts on the day of submitting the EOP forecasts to the EOP PCC Office (Wednesday) or on the day before. However, the elimination of outlier predictions coincides with the unavailability of EAM predictions only in some cases (for LOD 13 deletions out of a total of 52 occurred on days when GFZ EAM predictions were inaccessible) (Table 5), indicating that there is no single reason for all the identified problems. It is also worth mentioning that the lack of EAM forecasts not only on the day of submission or the day before may reduce the accuracy of the predictions. When participants use the full set of daily EAM forecasts in their algorithm, missing values on any other day are essential. The most critical effect of EAM data unavailability is for those predictions where the lack of EAM caused an error, but the error was not large enough so that the prediction was eliminated. Those submissions with errors due to a lack of EAM, but not eliminated, can cause the average performance to degrade significantly.

4 Evaluation of UT1–UTC and LOD predictions

4.1 General overview of predictions

We will begin the assessment of UT1–UTC and LOD predictions with a general overview of forecasts (without distinguishing individual IDs) to study the overall accuracy achieved during the campaign. Boxplots of differences between reference and predicted values for the 1st, 5th, 10th, 15th, 20th, 25th, and 30th day into the future are shown in Fig. 2 (for UT1–UTC) and Fig. 3 (for LOD). Statistics for the differences between reference and predicted values for the 10- and 30-day horizons are given in Table 6.

In the case of UT1–UTC, the lowest inter-quartile range (IQR, range between the first quartile (Q1) and the third quartile (Q3) of the data) of differences and the fewest outlier values were detected for the LS + AR + EAM and ML groups (Fig. 2). However, for the LS + AR + EAM group only, the most extreme data points that are not indicated as outliers (outliers defined as data points that fall below $Q1 - 1.5 \cdot \text{IQR}$ or above $Q3 + 1.5 \cdot \text{IQR}$), are within the range of ± 10 ms for the 30-day prediction horizon. A gradual increase in the range of differences can be observed as the prediction length increases. For all considered prediction days in the “Other” group and starting from the 20th prediction day for the LS + AR group, there is a clear asymmetry in the distribution of differences (there are more positive differences), indicating that the predictions underestimate the observed values in these cases.

Table 4 Number of UT1–UTC predictions rejected after applying σ and β criteria, number of σ and β rejections for days when EAM GFZ predictions were unavailable, and percentage (%) of rejected predictions in relation to the number of all submitted files

Group	ID	Number of predictions rejected by σ Number of σ rejections for days when EAM GFZ predictions were unavailable		Number of predictions rejected by β Number of β rejections for days when EAM GFZ predictions were unavailable		% of rejection for $\sigma + \beta$	
		10 days	30 days	10 days	30 days	10 days	30 days
LS + AR	100	–	–	711	510	10.0	7.1
	107	–	–	210	210	2.9	2.9
	114	–	–	710	410	16.3	9.3
	118	–	–	–	511	0.0	9.8
	122	–	–	210	210	3.3	3.3
LS + AR + EAM	101	–	–	511	–	9.1	0.0
	105	–	–	310	310	4.4	4.4
	108	–	–	610	210	9.1	3.0
	112	–	–	211	210	2.9	2.9
	116	–	–	–	211	0.0	2.9
	136	–	–	410	–	7.7	0.0
ML	126	–	N/A ^a	712	N/A	10.0	NA
	130	–	N/A	512	N/A	7.8	NA
	146	–	–	211	111	5.3	2.6
	147	–	–	311	111	8.1	2.7
	148	–	–	–	311	0.0	8.1
	149	–	–	211	310	5.3	7.9
Other	102	911	911	511	511	21.2	21.2
	103	–	–	210	110	3.6	1.8
	104	–	–	310	210	4.3	2.9
	117	–	–	610	410	13.0	8.9
	121	–	–	211	210	3.0	3.0
	123	–	–	110	–	2.3	0.0
IERS	200	–	–	210	510	2.9	7.2
Total		911	911	78112	5416	6.0	4.8

For IDs not listed (IDs 113, 115), none of the files was removed. Note that the prediction horizon for IDs 126 and 130 is shorter than 30 days, therefore the 30-day prediction horizon values are not provided for those IDs
^aN/A not applicable

Figure 3 shows that in the case of LOD predictions, except for the ML group, the highest number of outlier values is received for the initial prediction days (1st and 5th). Unlike UT1–UTC forecasts, there is no clear increase in the range of differences with increasing day of forecast.

The IQR for all UT1–UTC predictions for both the 10- and 30-day prediction horizons are 0.575 and 2.031 ms, respectively, which is around twice that of the ID 200 (Table 6). It is worth noting that despite eliminating outlier predictions, the range between the minimum and maximum values is prominent (83.523 and 106.902 ms for the 10- and 30-day prediction horizon, respectively), which was mainly influenced by outlier values of differences for just a few IDs (IDs

102, 122, and 126). Even though the difference values for EOP PCC contributors' predictions were generally considerably greater than those of the forecasts from ID 200, mean and median values for predictions from the campaign participants are lower than corresponding values for ID 200, especially for the 30-day prediction horizon. For both ID 200 and 2nd EOP PCC predictions, the differences between observed and predicted values as well as their statistics (min, max, median, mean, root-mean-square (RMS), range, IQR) are greater for the 30-day than for the 10-day prediction horizon, which shows that the error of UT1–UTC forecast grows as the prediction horizon increases.

Table 5 Number of LOD predictions rejected after applying σ and β criteria, number of σ and β rejections for days when EAM GFZ predictions were not available, and percentage (%) of rejected predictions in relation to the number of all submitted files

Group	ID	Number of predictions rejected by σ Number of σ rejections for days when EAM GFZ predictions were unavailable		Number of predictions rejected by β Number of β rejections for days when EAM GFZ predictions were unavailable		% of rejection for $\sigma + \beta$	
		10 days	30 days	10 days	30 days	10 days	30 days
LS + AR	100	–	–	110	110	1.4	1.4
	122	–	–	110	–	1.6	0.0
	156	–	–	210	110	6.5	3.2
LS + AR + EAM	105	–	–	1313	–	19.1	0.0
	108	–	–	110	–	5.9	0.0
	112	–	–	110	110	1.4	1.4
ML	115	–	–	410	510	9.3	11.6
	124	–	N/A	413	N/A	6.0	N/A
	125	–	N/A	412	N/A	5.7	N/A
	129	211	N/A	210	N/A	6.5	N/A
	142	–	N/A	211	N/A	5.3	N/A
	143	–	–	211	211	5.3	5.3
	144	–	–	311	211	7.9	5.3
	145	–	–	211	111	5.3	2.6
Other	102	611	611	–	–	9.1	9.1
	104	–	–	311	110	4.3	1.4
	117	–	–	310	–	7.0	0.0
	121	–	–	110	110	1.5	1.5
	123	–	–	210	–	4.7	0.0
	141	–	–	110	110	2.7	2.7
IERS	200	–	–	110	–	1.4	1.4
Total		812	611	53113	1713	4.7	2.2

For IDs not listed (IDs 101, 113, 114, 136, 157), none of the files was removed. Note that the prediction horizon for IDs 124, 125, 129 and 142 is shorter than 30 days, therefore the 30-day prediction horizon values are not provided for those IDs

For statistics of differences computed for LOD predictions, there is less difference between the 10-day and 30-day horizon than in the case of UT1–UTC (Table 6). The IQR for all LOD predictions for the 10- and 30-day prediction horizon are 0.138 and 0.245 ms, respectively, while the range between the minimum and maximum values are as high as 9.445 and 9.718 ms for 10- and 30-day predictions, respectively. The comparison of statistics of differences obtained for LOD between ID 200 and 2nd EOP PCC participants shows that in both cases, the IQR is at a similar level for both time horizons. In turn, the differences received in the campaign exhibit higher extreme values than in the case of ID 200, but the mean and median values are lower for the campaign participants than for ID 200.

4.2 Detailed assessment of predictions

We further assess the quality of UT1–UTC and LOD predictions by specifically studying MAE for the 10- and 30-day horizons. Analysis is performed for the groups described in Sect. 3. Figure 4 presents MAE plots for UT1–UTC for the 10-day prediction horizon for each group compared with the mean MAE for the group and the mean MAE for all IDs. The plots also contain MAE values for Day 0, which is the submission day (i.e., Wednesday, the last day for which observational data are available). By showing this value, we can verify whether a given participant had errors at the stage of preparing observational data, which could later reduce the accuracy of the forecast itself. Except for IDs 112 (Fig. 4b), 148 (Fig. 4c), 117 and 121 (Fig. 4d), there were no major issues at the data preparation stage for most participants, since MAE for Day 0 is usually very low. The providers of

Fig. 2 Boxplot of differences between reference and predicted values for UT1–UTC for 1st, 5th, 10th, 15th, 20th, 25th, and 30th day into the future for groups **a** LS + AR, **b** LS + AR + EAM, **c** ML, and **d** Other

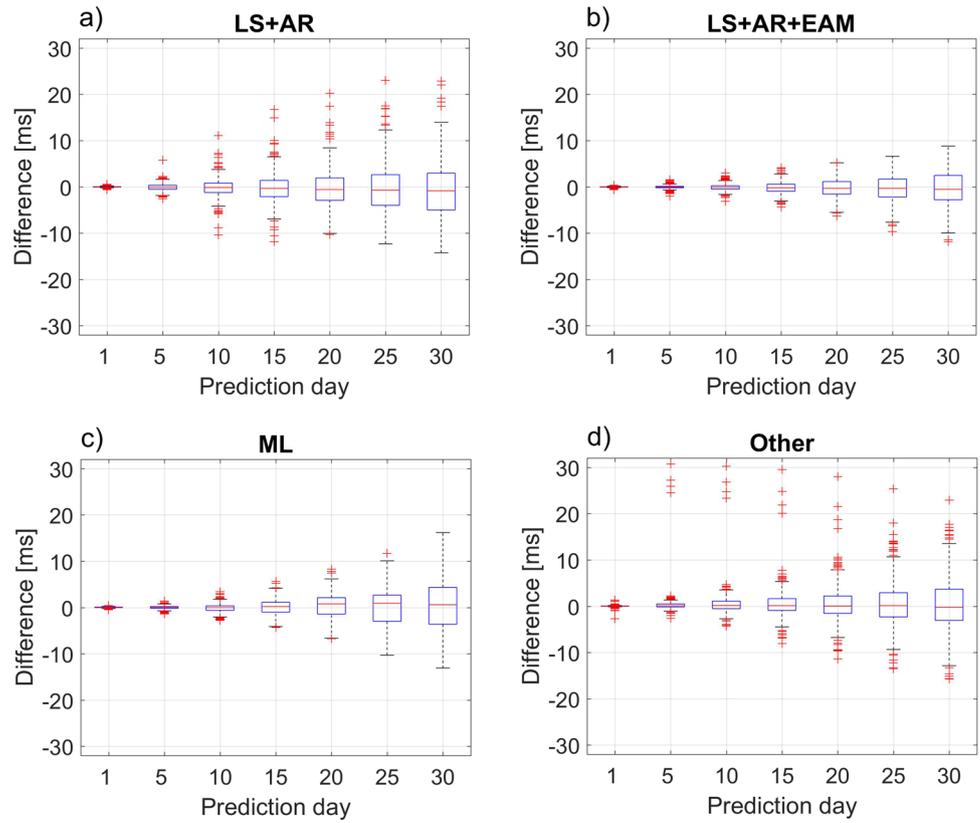


Fig. 3 Boxplot of differences between reference and predicted values for LOD for the 1st, 5th, 10th, 15th, 20th, 25th, and 30th day into the future for groups **a** LS + AR, **b** LS + AR + EAM, **c** ML, and **d** Other

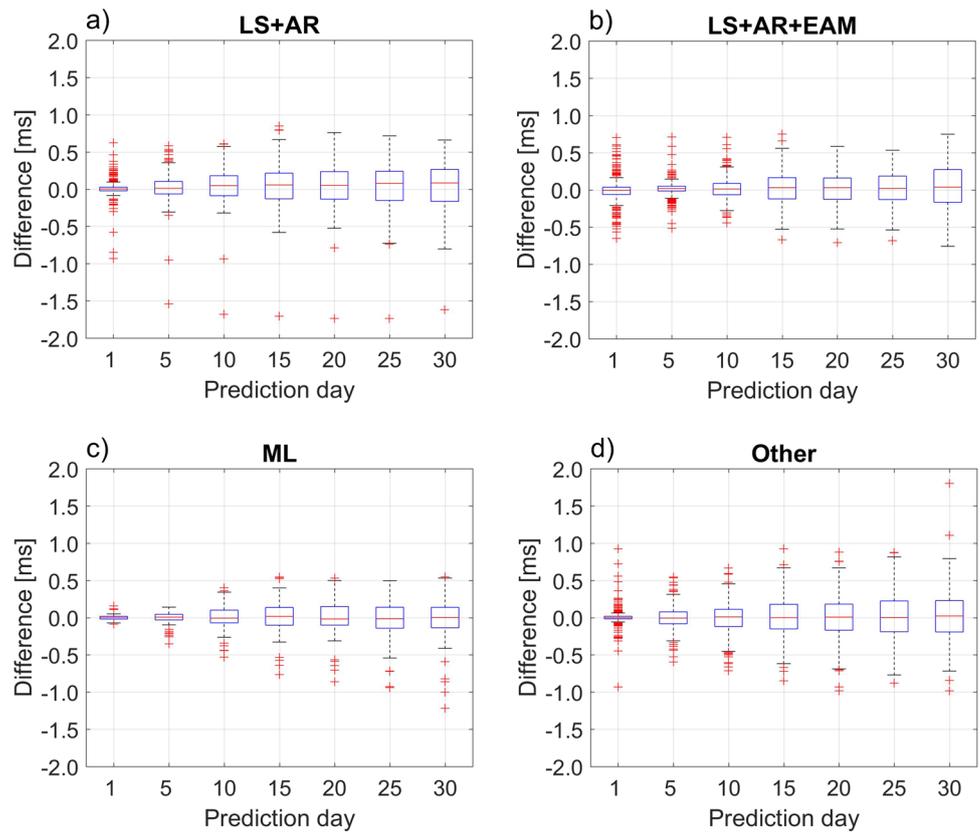


Table 6 Statistics (min, max, mean, root-mean-square (RMS), median, range, and inter-quartile range (IQR)) for the differences between reference and predicted values for UT1–UTC and LOD for predictions from all participants (Total) and for prediction from ID 200 for prediction horizons of 10 and 30 days into the future

	UT1–UTC [ms]				LOD [ms]			
	ID 200 (IERS)		Total		ID 200 (IERS)		Total	
	10 days	30 days	10 days	30 days	10 days	30 days	10 days	30 days
Min	− 1.444	− 12.084	− 11.309	− 16.523	− 0.204	− 0.627	− 7.857	− 7.857
Max	0.762	8.258	72.215	90.378	0.406	0.965	1.588	1.860
Mean	− 0.093	− 0.274	0.098	− 0.047	0.052	0.050	0.008	0.018
RMS	0.298	1.987	1.945	3.429	0.112	0.222	0.231	0.251
Median	− 0.045	− 0.111	− 0.012	− 0.032	0.049	0.047	0.008	0.014
Range	2.205	20.342	83.523	106.902	0.609	1.591	9.445	9.718
IQR	0.240	0.960	0.575	2.031	0.134	0.227	0.138	0.245

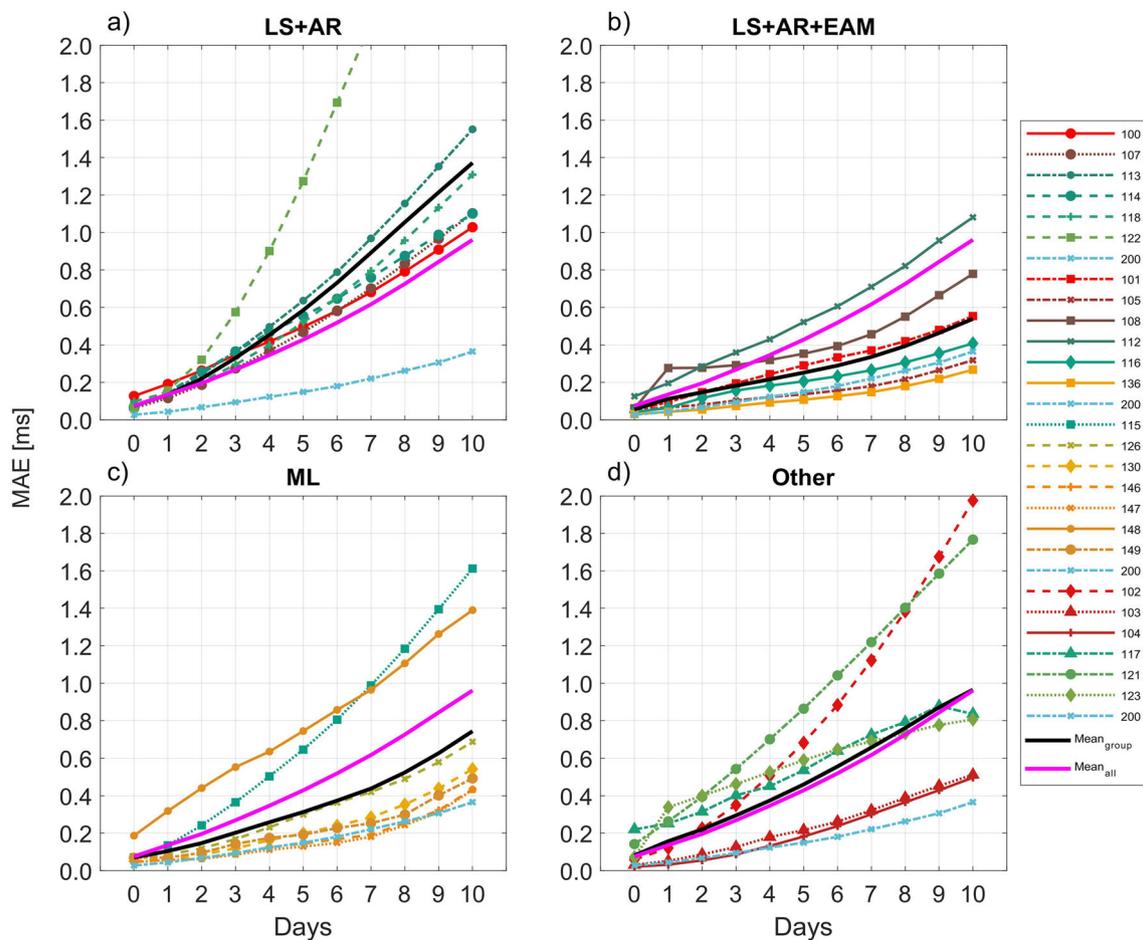


Fig. 4 MAE for UT1–UTC for up to 10 days into the future for groups **a** LS + AR, **b** LS + AR + EAM, **c** ML, and **d** Other. The thick black line represents the mean MAE for the group, and the thick magenta line represents the mean MAE for all IDs (the same for all subplots)

prediction with ID 117 performed a thorough analysis and validation of their methodology to uncover the source of the offset observed on Day 0. They discovered that the issue was caused by an error in input data for several submissions. Typically, ID 117 used a combination of IERS 14 C04 (with a 30-day delay) and Bulletin A (with a 2-day delay) as an input. However, it appeared that for a few submissions, inadvertently, only IERS 14 C04 was used (omitting Bulletin A). Due to the latency of IERS 14 C04, predictions for UT1–UTC (but also for LOD and PM) pertained to the last month rather than the current day and the subsequent days. These discrepancies ultimately impacted the MAE values. The forecast group identified sessions with incomplete input data that relied solely on C04 data with a 30-day lag as MJD submission dates, i.e. 59,528, 59,535, 59,542, 59,556, 59,591, and 59,899 (personal communication with Sadegh Modiri). The σ and β criteria utilized by the EOP PCC Office indicated only the dates 59,528, 59,535, and 59,556 as outliers. The participant also indicated the inconsistent combination of two distinct input time series (IERS 14 C04 and Bulletin A) as potential reason for the offset observed on Day 0.

For ID 108, the MAE for Day 0 is close to the rest of the participants, but starting from Day 1 of prediction, the MAE increases rapidly, which might suggest some weaknesses of the prediction method. The comparison of MAE plots for individual groups (Fig. 4a–d) shows that the IDs from the LS + AR + EAM group had the lowest forecasting errors (the mean for this group is lower than for the other groups, and as many as 5 out of 6 IDs in this group achieve results better than the mean for all methods). Most IDs from the ML group show a similar MAE change with increasing prediction day, with values congruous with those obtained in the LS + AR + EAM group (Fig. 4c). However, IDs 115 and 148 differ notably from the other participants in the ML group. Results for ID 115 are characterized by a much more rapid increase in the MAE with the length of prediction, and ID 148 suffers from high error at Day 0. Group LS + AR presents consistent results for all IDs (except for the more rapid MAE growth observed in ID 122); however, all submissions are characterized by the largest errors compared with participants from other groups (the mean MAE for this group is higher than for other groups and all IDs from the LS + AR group present results with lower accuracy than the mean for all IDs). This suggests that not including EAM data in the prediction procedure can deteriorate forecast quality. As the Other group contains several unique forecasting methods, results for individual IDs differ from each other both in the MAE values on particular prediction days and in the MAE increase rate.

To quantitatively assess the accuracy of UT1–UTC predictions for the 10-day prediction horizon, we show in Table 7 numerical values of MAE for two cases: MAE for the 10th day of prediction (MAE[10]) and mean MAE for the whole 10-day horizon, i.e., mean of MAE for the days from 0 to

10 (MAE[0–10]). While we initially focus only on errors for the last day of the considered prediction horizon, in the second case, we examine the performance of the method for the entire prediction horizon. Table 7 includes the highest, lowest, and mean values of MAE[10] and MAE[0–10] for each group in comparison with quantities received for ID 200. The results confirm the findings from Fig. 4 that methods from the LS + AR + EAM group are most reliable for UT1–UTC prediction as MAE[10] and MAE[0–10] for this group are 0.57 and 0.29 ms, respectively. The lowest MAE[10] and MAE[0–10] values were received for ID 136 (0.27 and 0.12 ms, respectively). It should be noted that IDs 136 and 105 are the only methods that perform better than ID 200 for the 10-day prediction horizon and both were submitted by the same team (GFZ). The only difference between the IDs is that in the case of ID 136, the last observed UT1–UTC value is taken from the Bulletin A only after its last update around 19:00 UTC, while the prediction from ID 105 is processed earlier in the day, which means that the initial value from Bulletin A of the day before is used (the value for Day 0 provided by ID 105 is actually a prediction and not an observed quantity). The analyses have shown that waiting for the most recent observed value can slightly improve the accuracy of the UT1–UTC prediction at the 10-day prediction horizon, which is plausible in view of the additional geodetic data incorporated. However, relying on additional input that is not available until late afternoon generally increases the risk that the forecast will not be provided in time.

MAE plots for UT1–UTC for the 30-day prediction horizon for each group are given in Fig. 5, while statistics of MAE for UT1–UTC for the Day 30 of prediction and for the whole 30-day forecast horizon are given in Table 8. Figure 5 shows that IDs 102, 113, 115, 121, and 122 suffer from an exceptionally sharp MAE increase starting from the first days of the prediction, while for many other methods (IDs 105, 116, 136, 146, 147, 149, 200) the MAE value starts to increase after about Day 10 into the future. Again, forecasts from the LS + AR + EAM group are characterized by the smallest prediction error (MAE[30] and MAE[0–30] of 3.21 and 1.29 ms, respectively), while the LS + AR group generally performs worst (MAE[30] and MAE[0–30] of 4.86 and 2.33 ms, respectively). This underlines the importance of using EAM information in improving the accuracy of UT1–UTC forecasting. Indeed, the best IDs in groups ML and Other are also those that exploit EAM data. Nevertheless, although several participants almost approach the accuracy of forecasts to that obtained for ID 200 (IDs 103, 104, 105, 116), only ID 105 achieves a MAE slightly lower than that obtained for forecasts from ID 200. Notably, starting from around Day 12, MAE for ID 136 becomes slightly higher than for ID 105, even though the former uses the most up-to-date observational data. It can also be observed from Fig. 5 that for the first days of prediction, the MAE values for most ML

Table 7 Statistics (maximum, minimum, and mean) of MAE for UT1–UTC for the 10th day of the prediction and for the whole 10-day forecast horizon

	LS + AR		LS + AR + EAM		ML		Other		ID 200 (IERS)
	ID	MAE [ms]	ID	MAE [ms]	ID	MAE [ms]	ID	MAE [ms]	MAE [ms]
<i>MAE for 10th day of forecast horizon (MAE [10])</i>									
Max	122	3.15	112	1.08	115	1.61	102	1.98	0.37
Min	100	1.03	136	0.27	146, 147	0.43	104	0.50	
Mean		1.54		0.57		0.80		1.07	
<i>Mean MAE for 10-day forecast horizon (MAE[0–10])</i>									
Max	122	1.42	112	0.55	148	0.77	121	0.90	0.17
Min	107	0.51	136	0.12	147	0.16	104	0.21	
Mean		0.72		0.29		0.37		0.54	

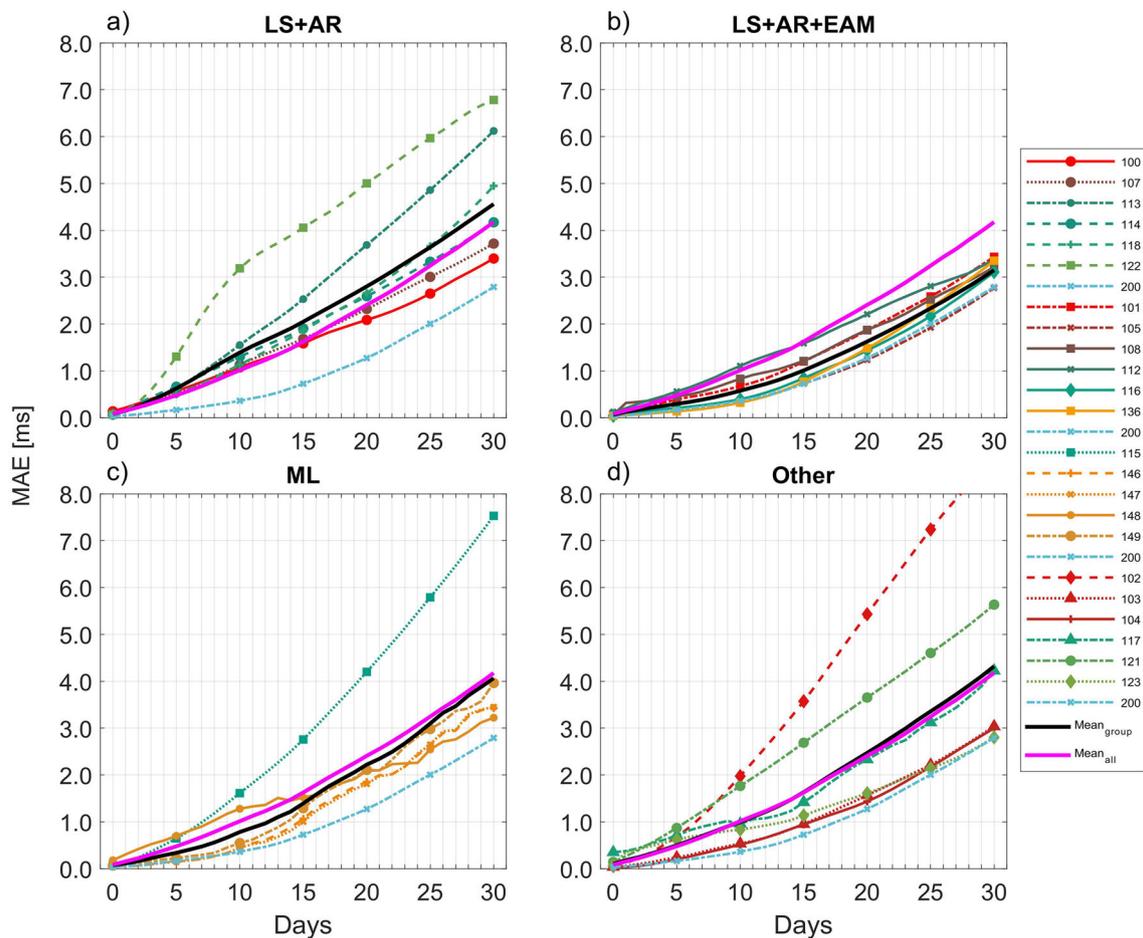


Fig. 5 MAE for UT1–UTC for up to 30 days into the future for groups **a** LS + AR, **b** LS + AR + EAM, **c** ML, and **d** Other. A thick black line represents the mean MAE for the group, whereas a thick magenta line represents the mean MAE for all IDs (the same for all subplots)

Table 8 Statistics (maximum, minimum, and mean) of MAE for UT1–UTC for the 30th day of the prediction and for the whole 30-day forecast horizon

	LS + AR		LS + AR + EAM		ML		Other		ID 200 (IERS)
	ID	MAE [ms]	ID	MAE [ms]	ID	MAE [ms]	ID	MAE [ms]	MAE [ms]
<i>MAE for the 30th day of forecast horizon (MAE[30])</i>									
Max	122	6.78	101	3.44	115	7.53	102	8.74	2.79
Min	100	3.40	105	2.77	148	3.22	123	2.80	
Mean		4.86		3.21		4.32		4.57	
<i>Mean MAE for the 30-day forecast horizon (MAE[0–30])</i>									
Max	122	3.81	112	1.67	115	3.14	102	3.89	1.00
Min	100	1.62	105	0.97	147	1.31	104	1.13	
Mean		2.33		1.29		1.78		2.01	

methods are close to the MAE observed for ID 200. However, starting from around the Day 10 of prediction, the MAE values for IDs 146, 147, and 149 begin to noticeably vary from the MAE observed for ID 200, which suggests that these submissions are most effective for ultra-short-term horizons.

Table 8 shows that for the group of most promising prediction methods for UT1–UTC (LS + AR + EAM), the mean MAE[30] and MAE[0–30] are 3.21 and 1.29 ms, respectively, while the corresponding values for ID 200 are 2.79 and 1.00 ms, which indicates there is still some room for improvement for groups dealing with UT1–UTC forecasting. For ID 105, which has been identified as the most promising method, these values are only slightly lower (2.77 and 0.97 ms for MAE[30] and MAE[0–30], respectively) than for the forecast disseminated by IERS.

The MAE plots for UT1–UTC predictions reveal a gradual increase in MAE along with the day of prediction, and the growth starts to accelerate within 10 days (Figs. 4, 5). Only ID 122 shows a reverse trend (after an initial rapid surge, the rate of MAE change slows after Day 10). An important indicator of the accuracy of a given forecasting algorithm is not only the error for a given day of prediction but also the rate of this error increase. Therefore, to study the growth rate of the MAE, we determined the mean MAE increase per day by differentiating the MAE values for consecutive prediction days within non-overlapping time intervals (i.e., 0–5 days, 6–10 days, 11–15 days, 16–30 days, and the 30-day prediction horizon (0–30 days)). These values were computed for each group along with ID 200 and all IDs together and are presented in Fig. 6. Methods from the LS + AR + EAM group show the least rapid increase in MAE. Apart from the LS + AR group, the mean MAE increase is lowest for the 0–5-day interval and highest for the 16–30 day-interval. Although the MAE increase for all groups is generally greater than that for ID 200, if we take into account individual submissions, some

of them (IDs 105 and 136) are characterized by a slightly lower MAE growth rate than for ID 200.

We now move on to a detailed assessment of the LOD predictions. The MAE plots for a 10-day horizon (Fig. 7) indicate certain issues with the MAE on Day 0 for several submissions (IDs 100, 108, 112, and 117). It is worth noting that while we observe a significant error on Day 0 for ID 101, the MAE values are much lower for the following prediction days. This discrepancy could suggest that the incorrect observed LOD value on Day 0 was included only in the file sent to the EOP PCC Office, rather than the utilization of erroneous observations in the forecasting process itself. The MAE results are most uniform for the ML group, with mean MAE[10] and MAE[0–10] of 0.102 and 0.064 ms, respectively (Table 9). Similar to UT1–UTC predictions, for LOD, the most promising methodology for a 10-day prediction horizon is that utilized by ID 136 (MAE[10] and MAE[0–10] are equal to 0.072 and 0.040 ms, respectively). The other submission from the same participant (ID 105) exhibits only slightly higher errors. The corresponding values received for ID 200 (LOD predictions transformed from UT1–UTC predictions) are 0.104 and 0.083 ms for MAE[10] and MAE[0–10], respectively. Notably, the mean MAE for the LS + AR and LS + AR + EAM groups is distorted by a few individual submissions, where high errors are observed starting from the submission day. This naturally impacts the overall performance of these groups. If we take into account the mean MAE for the whole 10-day prediction horizon (MAE[0–10]) for the entire group, ML methods turn out to provide lowest errors. However, it is worth mentioning that the use of EAM data may be important in this case, not only the method itself. Indeed, in the ML group almost all participants use EAM data (note that we have not obtained information about whether EAM was used for ID 115). The LS + AR group is the only one in which none of the methods use EAM data, which may be the reason for the highest mean

Fig. 6 Mean MAE increase per day for UT1–UTC prediction for each group, all predictions together (Total), and ID 200

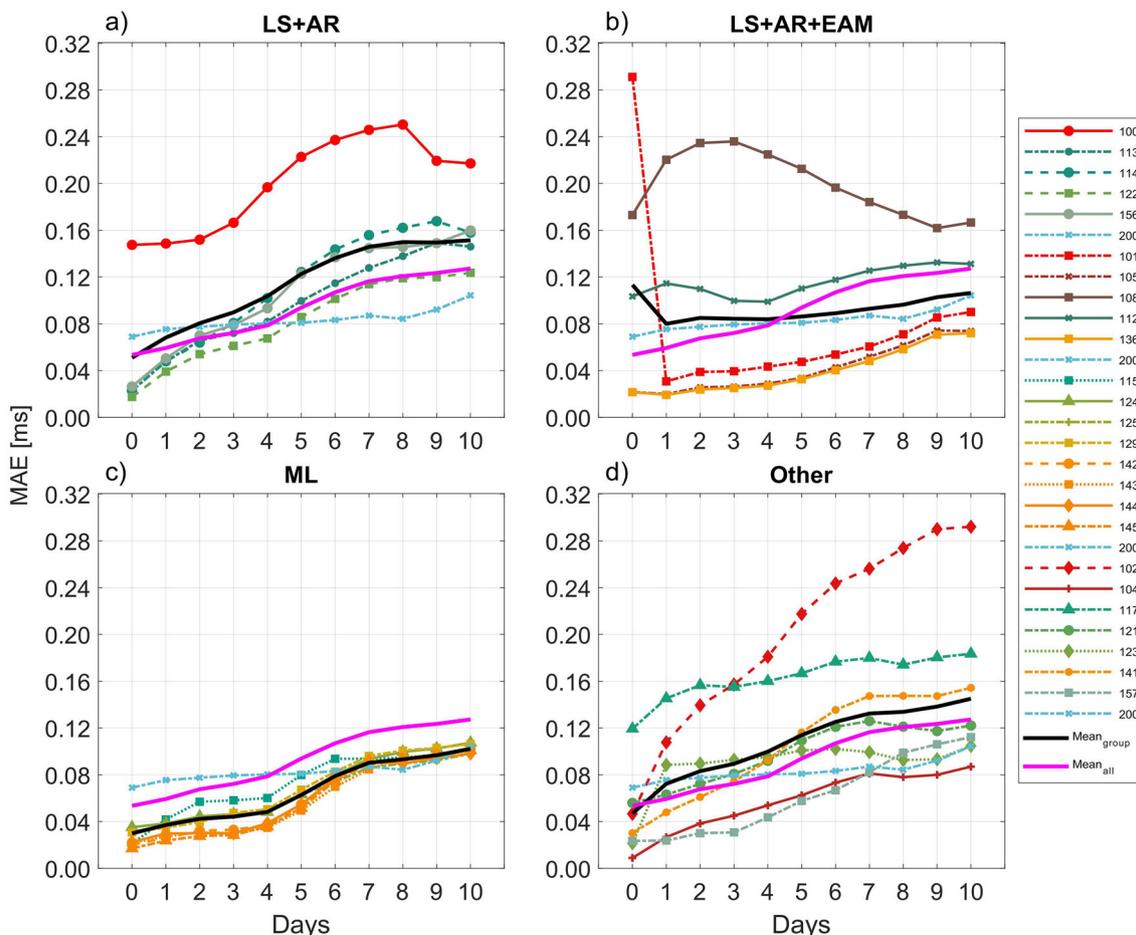
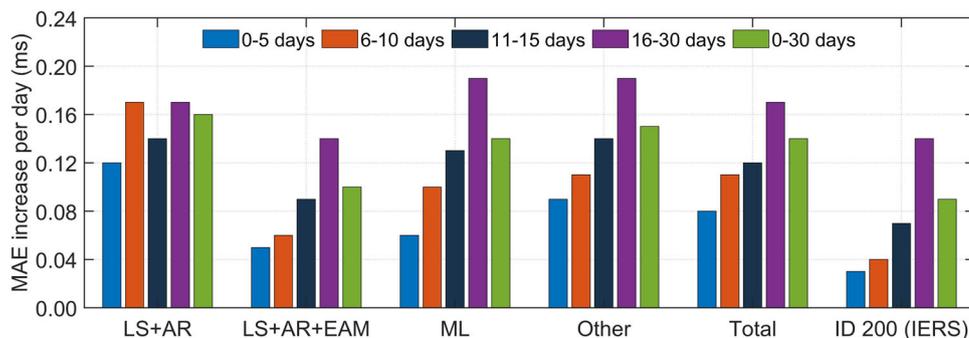


Fig. 7 MAE for LOD for up to 10 days into the future for groups **a** LS + AR, **b** LS + AR + EAM, **c** ML, and **d** Other. The thick black line represents the mean MAE for the group, and the thick magenta line represents the mean MAE for all IDs (the same for all subplots)

MAE values in this group. The impact of EAM inclusion on prediction accuracy improvement should be dominant for 10-day predictions, since their forecasts are only available for 6 days into the future.

The results presented in Fig. 8 and Table 10 indicate that the ML methods perform best in the case of LOD for a 30-day horizon (mean MAE[30] and MAE[0–30] of 0.161 and 0.119 ms, respectively). However, if we take into account individual approaches, it turns out that the most promising

results were recorded for ID 157 (Table 10). Nevertheless, this finding should be treated with caution, as this participant registered relatively late (November 2022) in the campaign, and therefore the MAE was determined on the basis of only seven submissions. To evaluate this method objectively, more predictions from this participant should be assessed.

Table 11 summarizes the performance of methods developed by the EOP PCC participants compared with the algorithm exploited for ID 200, giving the number of IDs with

Table 9 Statistics (maximum, minimum, and mean) of MAE for LOD for the 10th day of the prediction and for the whole 10-day forecast horizon

	LS + AR		LS + AR + EAM		ML		Other		ID 200 (IERS)
	ID	MAE [ms]	ID	MAE [ms]	ID	MAE [ms]	ID	MAE [ms]	MAE [ms]
<i>MAE for 10th day of forecast horizon (MAE[10])</i>									
Max	100	0.217	108	0.167	124	0.107	102	0.292	0.104
Min	122	0.124	136	0.072	125, 144	0.098	104	0.087	
Mean		0.161		0.107		0.102		0.151	
<i>Mean MAE for 10-day forecast horizon (MAE[0-10])</i>									
Max	100	0.200	108	0.198	115	0.072	102	0.200	0.083
Min	122	0.082	136	0.040	143, 145	0.058	104	0.058	
Mean		0.120		0.095		0.064		0.111	

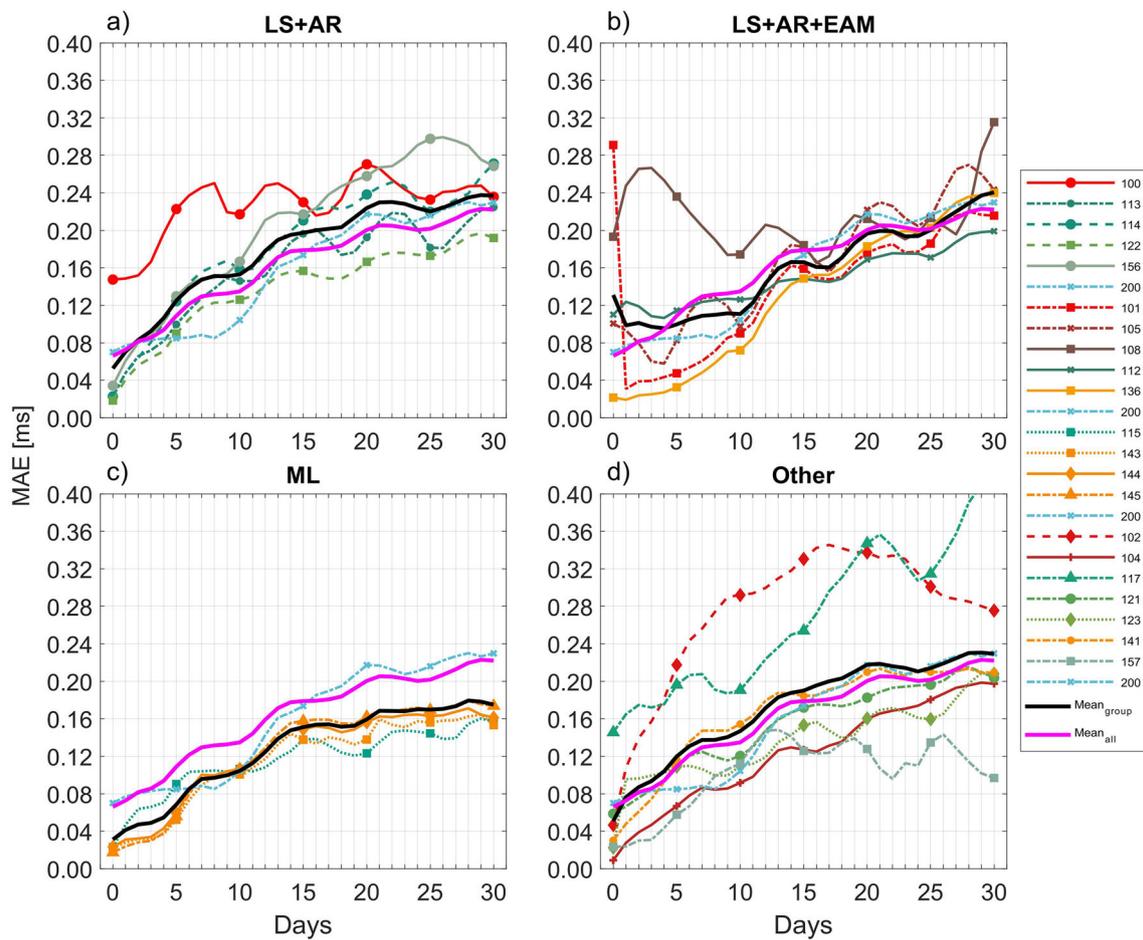


Fig. 8 MAE for LOD for up to 30 days into the future for groups **a** LS + AR, **b** LS + AR + EAM, **c** ML, and **d** Other. The thick black line represents the mean MAE for the group and the thick magenta line represents the mean MAE for all IDs (the same for all subplots)

Table 10 Statistics (maximum, minimum, and mean) of MAE for LOD for the 30th day of prediction and for the whole 30-day forecast horizon

	LS + AR		LS + AR + EAM		ML		Other		ID 200 (IERS)
	ID	MAE [ms]	ID	MAE [ms]	ID	MAE [ms]	ID	MAE [ms]	MAE [ms]
<i>MAE for 30th day of forecast horizon (MAE[30])</i>									
Max	114	0.271	108	0.315	145	0.174	117	0.414	0.230
Min	122	0.192	112	0.199	143	0.153	157	0.097	
Mean		0.238		0.243		0.161		0.229	
<i>Mean MAE for 30-day forecast horizon (MAE[0–30])</i>									
Max	100	0.227	108	0.213	145	0.125	102	0.274	0.158
Min	122	0.137	136	0.130	143	0.115	157	0.102	
Mean		0.183		0.159		0.119		0.175	

Table 11 Number of IDs in each group with MAE[10], MAE[0–10], MAE[30], and MAE[0–30] of UT1–UTC and LOD prediction equal to or lower than corresponding MAE values for ID 200 in relation to the number of all IDs in the group that predict UT1–UTC and LOD

	UT1–UTC					LOD				
	LS + AR	LS + AR + EAM	ML	Other	All groups	LS + AR	LS + AR + EAM	ML	Other	All groups
MAE[10]	0/6	2/6	0/7	0/6	2/25	0/5	3/5	6/8	2/7	11/25
MAE[0–10]	0/6	2/6	2/7	0/6	4/25	1/5	3/5	8/8	2/7	14/25
MAE[30]	0/6	1/6	0/5	1/6	2/23	2/5	2/5	4/4	5/7	13/21
MAE[0–30]	0/6	1/6	0/5	0/6	1/23	1/5	3/5	4/4	4/7	12/21

MAE[10], MAE[0–10], MAE[30], and MAE[0–30] equal to or below the corresponding MAE values obtained from the IERS forecasts. Of the 25 different approaches that apply UT1–UTC, only a few achieve accuracies that match ID 200. In contrast, for LOD predictions, about half of the IDs forecast with an error lower than that for ID 200.

5 Dependence of prediction accuracy on selected factors

5.1 Evolution of prediction accuracy

This part of our analysis focuses on how the accuracy of individual predictions has evolved over time and aims to verify whether the improvements made by the participants during the campaign were effective. To do so, we analyze MAE for a 10-day horizon in eight consecutive two-month periods named P1–P8, whose start and end dates are shown in Table 12.

Figure 9 shows MAE for UT1–UTC for up to 10 days into the future for consecutive 2-month periods. It can be seen that for many participants prediction errors were not stable throughout the duration of the campaign. For example, ID 117 experienced some difficulties on Day 0 between the

Table 12 Two-month periods (P1–P8) considered in the analysis of forecast evolution over time and their start and end dates

Period	Start date	End date
P1	September 1, 2021	October 31, 2021
P2	November 1, 2021	December 31, 2021
P3	January 1, 2022	February 28, 2022
P4	March 1, 2022	April 30, 2022
P5	May 1, 2022	June 30, 2022
P6	July 1, 2022	August 31, 2022
P7	September 1, 2022	October 31, 2022
P8	November 1, 2022	December 31, 2022

beginning of November 2021 and the end of February 2022, which affected the overall accuracy of the forecast. Visible offsets at Day 0 for this submission were primarily related to problems at the data preparation stage (see Sect. 4.2. for more details). However, starting in March 2022, the issue was resolved, and the participant visibly reduced the errors in the forecast. Periodic reduction in forecast accuracy was also observed for several IDs. During the initial period of the campaign, the accuracy of the method exploited by ID 102 was high but notably larger errors were recorded for some months, especially between September 1 and October

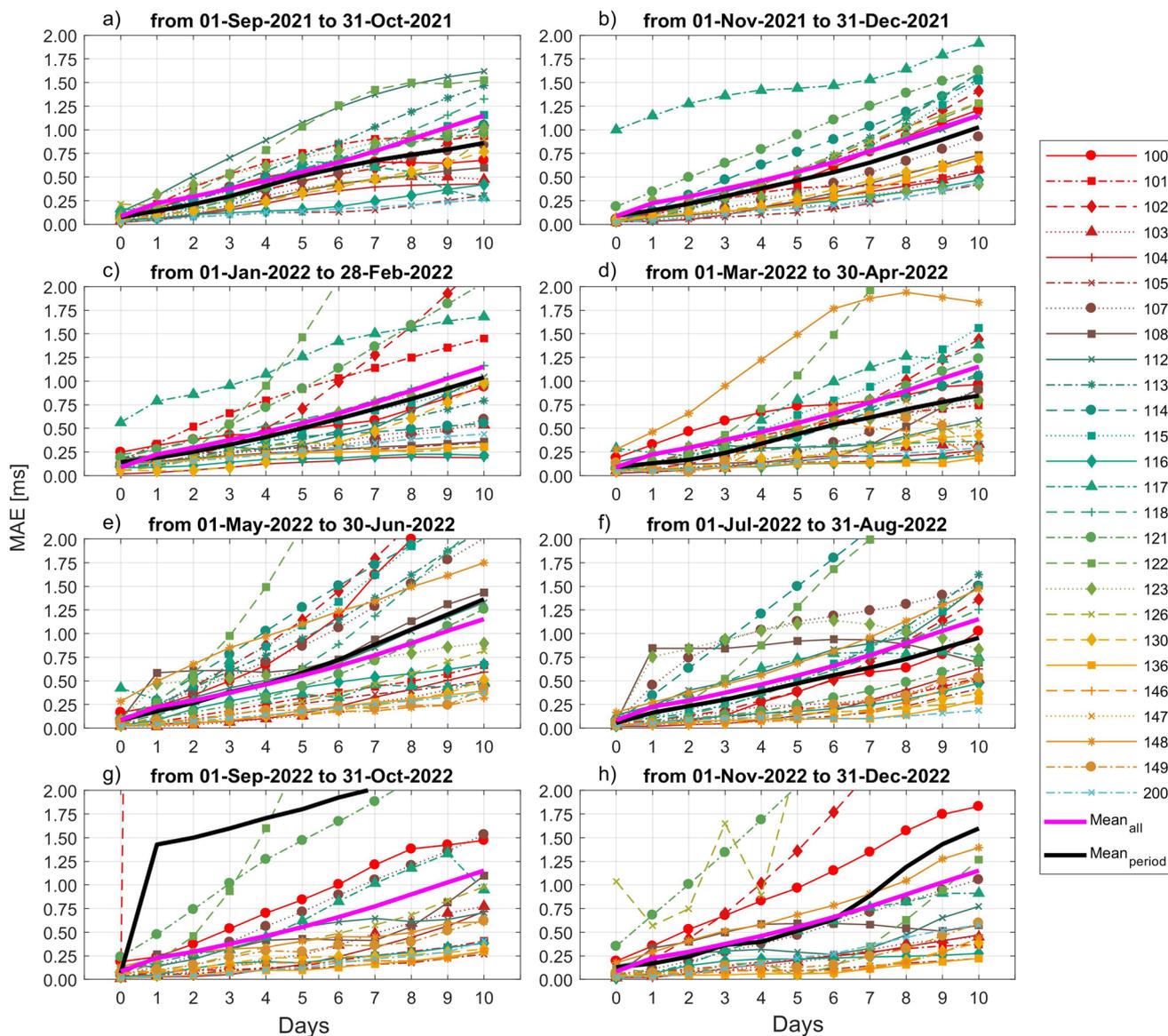


Fig. 9 MAE for UT1–UTC for up to 10 days into the future for consecutive 2-month periods (a–h). The thick black line represents the mean MAE for the period and the thick magenta line represents the mean MAE for the whole campaign (the same for all subplots)

31, 2022. Another participant with ID 148 started submitting forecasts in April 2022 with relatively high MAE, but then managed to reduce the errors in the following months.

We are unable to identify a single ID for which errors consistently decreased from period to period, which would indicate continuous improvement of the method. Instead, we observe periodic increases and decreases in MAE for individual participants. In general, the methods that exhibited the lowest errors throughout the entire campaign period (IDs 104, 105, 116, 136, 146, 147, and 149, see Fig. 4) demonstrate visibly greater stability in terms of MAE across individual periods than other submissions. Comparing the average MAE

for a given period with the average MAE for the entire operational phase of the campaign allows us to conclude that the period between March and May 2022 had the highest number of forecasts with errors below the overall average. It is worth noting that at the end of the campaign (in the four last months), the majority of methods exhibited low errors, and the mean MAE for those periods were only disrupted by a few individual submissions.

When considering LOD, it becomes apparent that during the first 2 months, the MAE values for IDs 105 and 117 exhibit a distinct sinusoidal fluctuation that is not present in any of the other submissions (Fig. 10). This problem was thoroughly discussed in Kur et al. (2022), who provided a

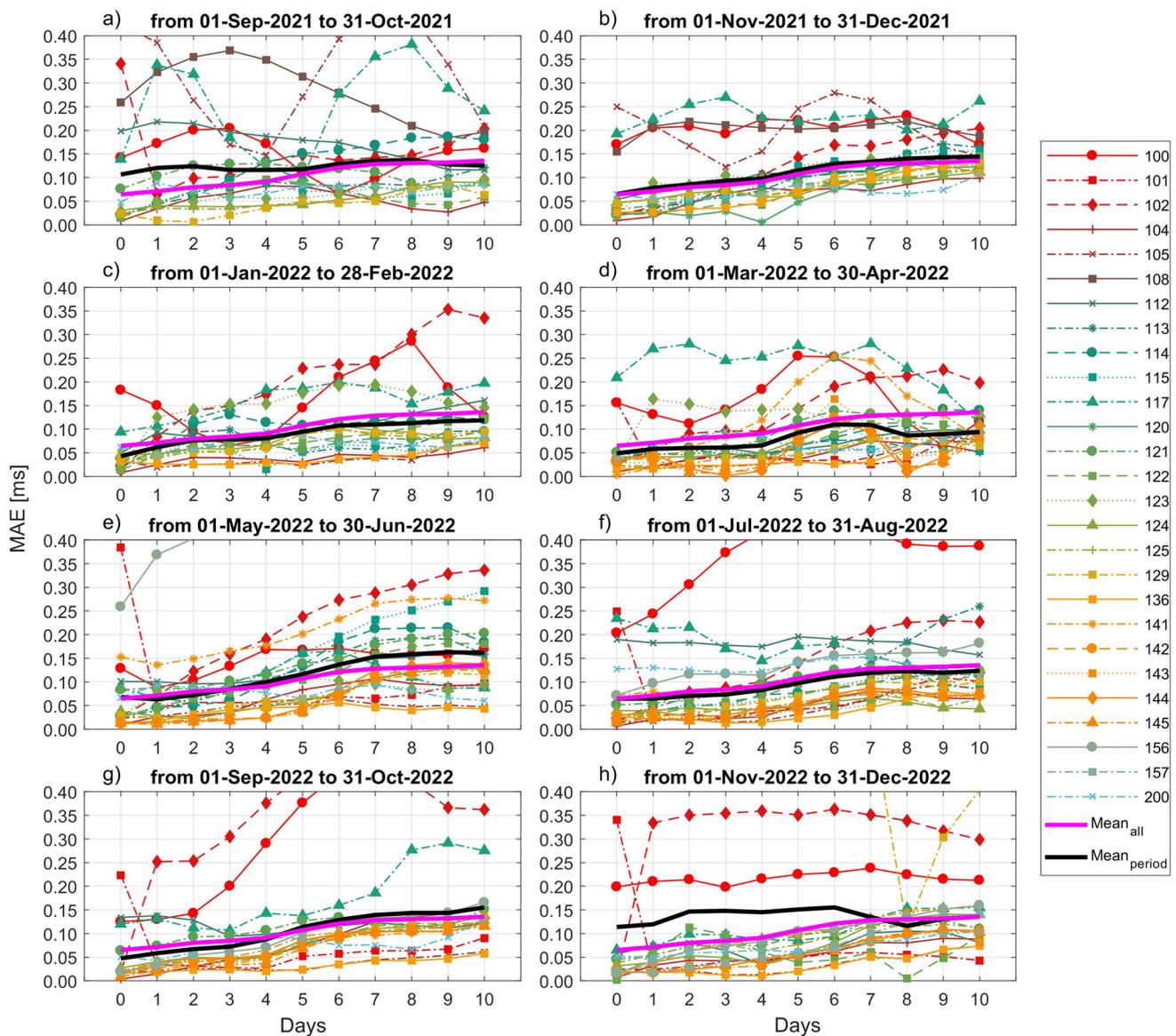


Fig. 10 MAE for LOD for up to 10 days into the future for consecutive two-month periods (a–h). The thick black line represents the mean MAE for the period and the thick magenta line represents the mean MAE for the whole campaign (the same for all subplots)

preliminary evaluation of LOD predictions collected during the 2nd EOP PCC. It was concluded that the issue was caused by incorrect handling of impact from long-period ocean tides (Bizouard et al. 2022). In subsequent periods, the oscillation vanished, which is considered an effective intermediate result of the preliminary evaluations by the EOP PCCO Office (Kur et al. 2022). There was a notable enhancement of accuracy when comparing MAE results for ID 108 for Period P1 with MAE for Period P2; however, starting from December 22, 2021, this participant stopped providing LOD forecasts while continuing to predict UT1–UTC and PM. Figure 10 shows that the issue with incorrect value at Day 0 for ID 101 discussed in the previous section (Fig. 7) became apparent from May 2022 and was not solved until the end of the 2nd EOP

PCC. Similar to the results for UT1–UTC, MAE for almost all predictions after July 2022 are below the mean for the whole campaign duration and only individual methods (IDs 100 and 102) have MAE visibly higher than the mean MAE computed for the whole 2nd EOP PCC period.

To study the time evolution of prediction accuracy in more detail, for each ID, we computed a percentage change (PCh) of MAE for a given 2-month period compared with the preceding period:

$$PCh = \text{mean} \left(\frac{MAE_i(n) - MAE_i(n+1)}{MAE_i(n)} \right) \cdot 100\%, \quad (5)$$

where MAE_i is the value for i th point of the prediction computed for the n th period ($n = 1, \dots, 7$). If $PCh > 0$, the preceding period has lower MAE (prediction accuracy has improved), if $PCh < 0$, the preceding period has higher MAE (prediction accuracy has deteriorated).

Values of PCh for UT1–UTC and LOD prediction from each participant are displayed in Fig. 11. The figure confirms that practically no method exhibits a continuous improvement in accuracy, but there are also no instances in which the quality consistently declines. Rather, we observe alternating periods of better and worse prediction performance. Forecasts from ID 200 also demonstrate such tendencies. This may not necessarily be linked to modifications of the method itself but perhaps to the temporal occurrence of certain phenomena affecting LOD and UT1–UTC, which may be more challenging to forecast.

The PCh statistics presented in Table 13 indicate a noticeable decrease in accuracy for several campaign participants and for ID 200 as well (minimum PCh values), which could have been caused by periodic errors in data preparation, lack of observational EOP data and EAM predictions or delays in access to the data, as discussed before. Both for UT1–UTC and LOD, the number of positive and negative values of PCh is similar, which, combined with a median value close to zero, allows us to conclude that the majority of predictions on a global scale were rather stable over time.

5.2 Transformation between LOD and UT1–UTC

In this section, we analyze MAE values received for the original UT1–UTC and LOD predictions, as well as the MAE of these parameters transformed from the respective LOD and UT1–UTC forecasts. Figure 12 presents MAE for up to 10 days into the future for original and transformed UT1–UTC and LOD predictions as well as the differences in MAE between the original and transformed predictions. The results are provided only for IDs who forecast both UT1–UTC and LOD. In general, for the methods that performed equally well (or poorly) for both UT1–UTC and LOD predictions (e.g., IDs 104, 105, 112, 117, and 136), the differences in MAE before and after transformation are relatively small. However, in cases where there are large prediction errors for one parameter and small errors for the second parameter, the corresponding difference becomes more pronounced. For example, ID 108 predicts LOD with a relatively large error compared with other participants, while UT1–UTC forecasts provided by this participant are accurate. Consequently, it turns out that transformed UT1–UTC predictions from ID 108 exhibit higher MAE than the original UT1–UTC predictions, but the transformed LOD shows a lower MAE than the original LOD forecasts. In other words, for this method, it is better not to forecast LOD but to transform it from the UT1–UTC predictions. The reverse applies to the method from

ID 122, which poorly predicts UT1–UTC but has a more accurate prediction of LOD. This analysis demonstrates the influence of differences in accuracy between UT1–UTC and LOD predictions on the results of the transformation between these parameters, rather than the impact of the transformation itself on the accuracy of the transformed predictions. The forecasting method also seems to have little influence, as the IDs with the smallest differences in MAE before and after transformation belong to different groups.

6 Ranking of prediction approaches

Although it is not possible to identify a single ID that would provide the highest prediction accuracy for all EOP and for different forecast horizons, we attempted to find the most universal and reliable combination of prediction methodology and input data. To do so, we have developed a ranking of all IDs based on the following criteria:

- (A) percentage of rejected submissions—to assess the credibility of predictions for a given algorithm;
- (B) range of differences between prediction and reference—to evaluate predicting repeatability (accurate predictions with high stability over time should be characterized by a small range of differences);
- (C) values of $MAE[1]$, $MAE[6]$, $MAE[7]$, and $MAE[10]$ —to check the quality of predictions at the beginning, middle, and end of the 10-day prediction horizon. To include all IDs in the ranking, we do not consider the prediction for 30 days into the future;
- (D) median of PCh—to assess the stability of the accuracy of the method.

Each ID has been assigned points equal to its position in the classification, assuming that the lower the number of points, the higher place in the classification is reached. The ranking for UT1–UTC is shown in Table 14, and the ranking for LOD is in Table 15. Overall, the prediction from ID 200 (IERS) is placed 2nd for UT1–UTC and 10th for LOD predictions transformed from UT1–UTC predictions, which confirms the reliability of the algorithm performed at USNO. For both UT1–UTC and LOD, the highest places are dominated by IDs exploiting EAM data mostly from groups LS + AR + EAM and ML. The leader for both parameters turned out to be ID 136 with the LS + AR + EAM method exploited by GFZ.

7 Summary and conclusions

The main objective of this research was to conduct a thorough evaluation of multiple predictions of UT1–UTC and

Fig. 11 Percentage change (PCh) of MAE of **a** UT1–UTC and **b** LOD predictions in individual analysis periods (P2–P8) in relation to the previous periods (P1–P7)

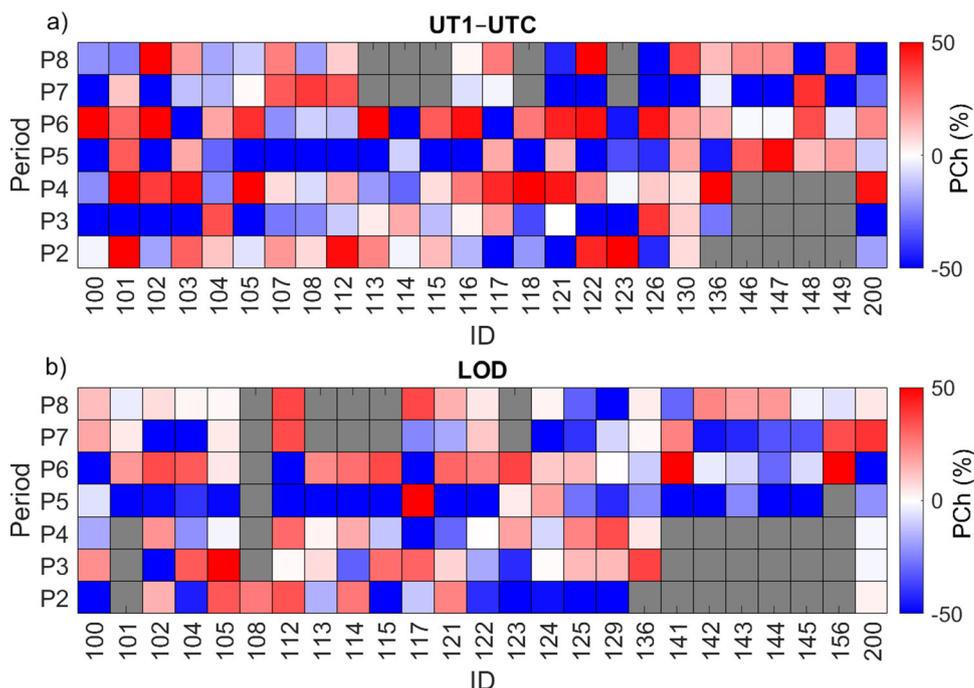


Table 13 Statistics (min, max, mean, RMS, median) for PCh for predictions of UT1–UTC and LOD for ID 200 and predictions from all participants together (Total), and number of positive and negative values of PCh

	UT1–UTC		LOD	
	ID 200 (IERS)	Total	ID 200 (IERS)	Total
Min (%)	– 101	– 9465	– 69	– 2266
Max (%)	46	74	40	80
Mean (%)	– 21	– 87	– 7	– 35
RMS (%)	50	775	31	212
Median (%)	– 19	– 1	– 2	0
Number of PCh > 0	2	75	3	66
Number of PCh < 0	5	77	4	65

LOD obtained during the 2nd EOP PCC, using the IERS 14 C04 solution as a reference. The primary goal of the campaign was to assess the current potential of EOP prediction, which encompassed exploring new methodologies such as ML methods that have been rapidly evolving in recent years, along with studying the contribution of input data (both EOP observations and EAM data and predictions) on forecast performance. The 2nd EOP PCC provided an objective evaluation platform for scientists from different countries and institutions to collaborate and compete in enhancing EOP predictions. Thanks to the effort and experience of the 23 participating institutions from around the world, an unprecedented set of EOP forecasts was gathered during an operational phase spanning 70 weeks.

Since the conclusion of the 1st EOP PCC, there has been an increased interest in EOP forecasts, which was evident in the considerably larger number of teams participating

in the most recent EOP PCC. Lessons from the first campaign have been learned in the sense that, as recommended in the conclusions of the 1st EOP PCC, there has been an increased interest in utilizing EAM when predicting EOP in order to enhance forecasting algorithms. In the case of UT1–UTC and LOD, 23 out of 35 IDs exploited such data. While the focus in the first campaign was primarily on utilizing AAM, in the current campaign, OAM, HAM, and SLAM data and their predictions were also incorporated. However, at present, EAM predictions developed by GFZ are accessible for a maximum of 6 days ahead, and this notably influences the outcomes obtained by 2nd EOP PCC participants who utilized these data. Extending the length of EAM predictions could potentially help to reduce the prediction errors for longer prediction horizons. The only center, apart from GFZ, that currently provides EAM predictions is ETH Zurich (Kiani Shahvandi et al. 2022a, 2023; Soja et al. 2022). Daily

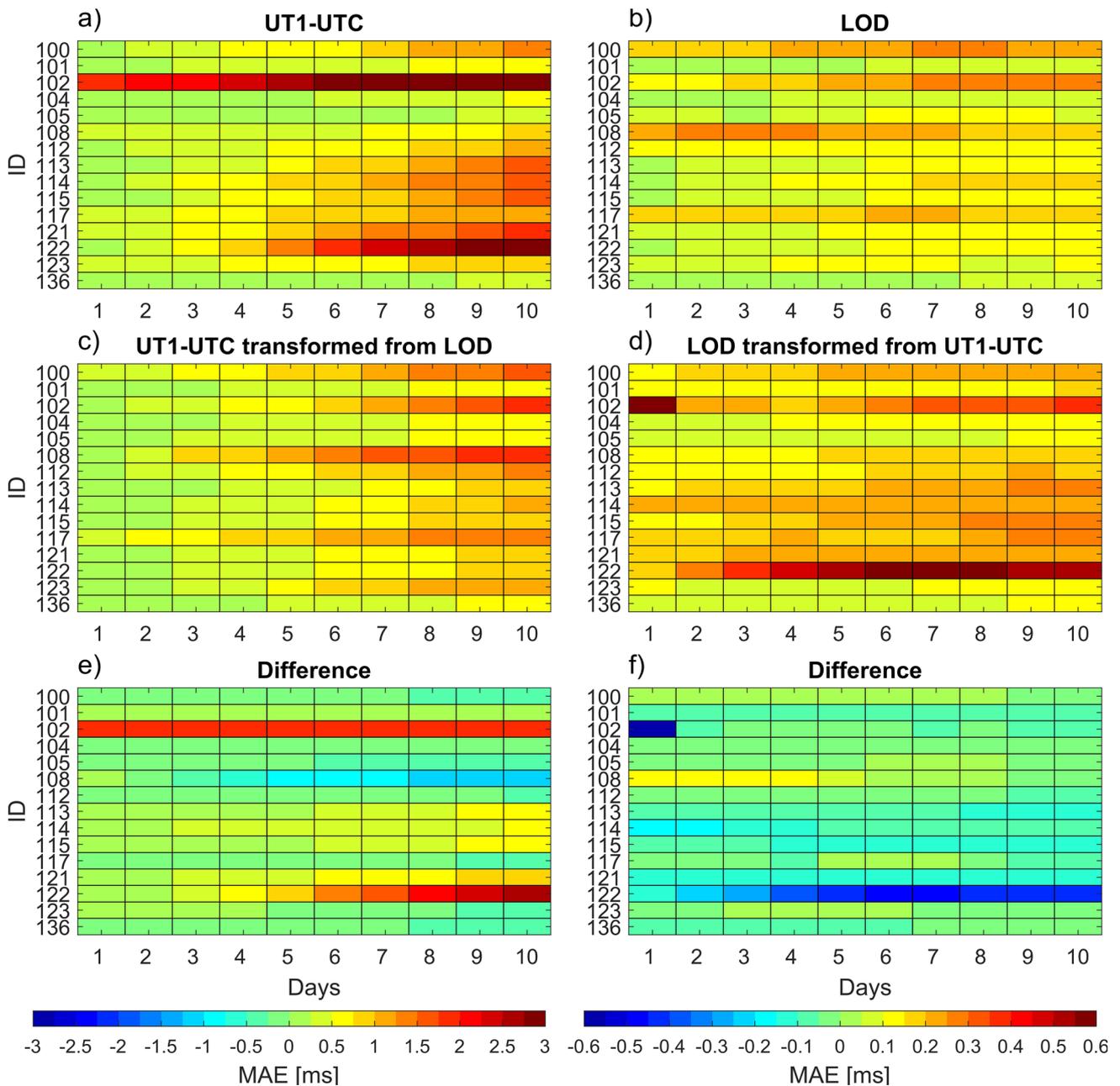


Fig. 12 MAE for up to 10 days into the future for **a** UT1–UTC, **b** LOD, **c** UT1–UTC transformed from LOD, **d** LOD transformed from UT1–UTC, **e** difference between MAE of UT1–UTC and MAE of UT1–UTC

transformed from LOD, and **f** difference between MAE of LOD and MAE of LOD transformed from UT1–UTC

forecasts of all four EAM components from ETH are available for up to 14 days ahead; however, these predictions are generated through the utilization of ML techniques applied to data sourced from GFZ, so they are not entirely independent. Nevertheless, their utilization has the potential to enhance the accuracy of EOP predictions for a maximum forecast horizon of about two weeks. Future research aiming to improve EOP predictions should encompass the development of new EAM data and longer-term EAM forecasts, by making use of

even longer prediction runs that are now being performed by various numerical weather prediction centers (Scaife et al. 2022).

Regarding the best methods for predicting UT1–UTC and LOD, the Kalman filter (with AAM forecast from NCEP used), wavelet decomposition + autocovariance, and adaptive transformation from AAM to LOD residuals (LODR) were identified as the most effective approaches in the 1st EOP PCC (Kalarus et al. 2010). In the current campaign, LS +

Table 14 Ranking of IDs according to the adopted criteria and the number of points awarded to each ID in individual categories for UT1–UTC

Rank	ID	Group	Total	% of rejected predictions	Range	MAE[1]	MAE[6]	MAE[7]	MAE[10]	Median PCh
1	136	LS + AR + EAM	23	10	1	2	1	1	1	7
2	200	IERS	26	2	4	3	5	5	3	4
3	116	LS + AR + EAM	39	0	2	7	7	7	4	12
4	105	LS + AR + EAM	41	8	3	9	3	2	2	14
5	146	ML	43	9	5	6	4	4	6	9
5	147	ML	43	12	6	5	2	3	5	10
7	104	Other	57	7	11	1	8	9	8	13
8	149	ML	63	9	9	10	6	6	7	16
9	103	Other	70	6	7	4	10	10	9	24
10	130	ML	79	11	10	8	9	8	10	23
11	123	Other	85	1	8	26	19	15	14	2
12	101	LS + AR + EAM	88	13	12	12	11	11	11	18
13	126	ML	93	14	26	11	12	12	12	6
14	112	LS + AR + EAM	102	2	14	21	16	17	17	15
14	115	ML	102	0	18	15	22	23	23	1
16	113	LS + AR	105	0	16	16	21	22	22	8
17	107	LS + AR	108	3	22	13	14	16	19	21
18	148	ML	112	0	17	25	23	21	21	5
19	108	LS + AR + EAM	115	13	13	24	13	13	13	26
20	118	LS + AR	117	0	15	17	20	20	20	25
21	117	Other	119	15	21	22	17	18	15	11
22	100	LS + AR	120	14	23	20	15	14	16	18
23	114	LS + AR	128	16	19	18	18	19	18	20
24	102	Other	132	17	25	14	24	24	25	3
25	121	Other	143	4	20	23	25	25	24	22
25	122	LS + AR	143	5	24	19	26	26	26	17

The result for prediction from IERS is shown in bold

AR with EAM data and predictions and ML with EAM data and predictions were found to achieve the highest accuracy. It should be noted that although ML methods have been rapidly developing in recent years, and the ML groups considered in this article were the largest in terms of the number of IDs, the majority of these methodologies were developed by a single participant (ETH). Therefore, it would be beneficial for other teams to join in the exploitation of this promising new technology. All ML methods that achieved high accuracy utilized EAM as an additional data source and the usage of these data seems to be crucial in improving the accuracy of UT1–UTC and LOD predictions. The source of the EOP observations used appears to play a secondary role. However, proper implementation of input data is crucial, as errors for the submission day resulting from issues such as the limited availability of EOP and/or EAM data, or internal problems with the data retrieval algorithm, contribute to a bias specific to each method.

When it comes to the numbers, for the best prediction methodologies chosen for UT1–UTC, MAE[10] was 0.27 ms, while MAE[30] was 2.77 ms. For forecasts from ID 200 those quantities were only slightly larger, i.e., 0.37 ms and 2.79 ms for MAE[10] and MAE[30], respectively. In turn, during the 1st EOP PCC, the best UT1–UTC prediction methods ensured MAE[10] of around 0.60 ms and MAE[30] as high as 3.80 ms. In the case of LOD, the best achievements from the 2nd EOP PCC were 0.072 ms for MAE[10] and 0.097 ms for MAE[30], while in the 1st EOP PCC, optimal values were around 0.130 ms and 0.220 ms for MAE[10] and MAE[30], respectively. There has clearly been considerable progress made in EOP forecasting over the past years. Nevertheless, there remains room for improvement for teams predicting UT1–UTC as only 2 (in the case of MAE [10] and MAE[30]) out of the 25 IDs revealed slightly higher prediction accuracy than official forecasts disseminated by IERS (ID 200).

Table 15 Ranking of IDs according to the adopted criteria and the number of points awarded to each ID in individual categories for LOD

Rank	ID	Group	Total	% of rejected predictions	Range	MAE[1]	MAE[6]	MAE[7]	MAE[10]	Median PCh
1	136	LS + AR + EAM	18	0	3	1	1	1	1	11
2	101	LS + AR + EAM	35	0	2	9	3	3	4	14
3	105	LS + AR + EAM	51	17	22	2	2	2	2	4
4	145	ML	54	8	6	3	8	11	8	10
5	104	Other	55	6	11	6	6	4	3	19
5	157 ^a	Other	55	0	1	4	4	5	15	26
7	143	ML	56	8	8	7	5	6	7	15
8	125	ML	61	9	9	11	7	7	5	13
9	142	ML	63	8	4	5	9	9	10	18
10	200	IERS	70	2	7	20	13	8	11	9
11	144	ML	77	14	12	8	10	10	6	17
12	122	LS + AR	90	4	13	13	15	16	17	12
12	124	ML	90	11	10	12	11	12	14	20
14	113	LS + AR	94	0	16	15	17	19	19	8
15	121	Other	95	3	15	19	19	18	16	5
16	123	Other	101	7	5	21	16	15	12	25
17	114	LS + AR	102	0	14	17	22	22	21	6
18	112	LS + AR + EAM	103	2	18	23	18	17	18	7
19	115	ML	107	16	17	14	14	13	9	24
20	129	ML	109	12	26	10	12	14	13	22
21	156	LS + AR	121	12	25	18	21	20	22	3
22	141	Other	122	5	19	16	20	21	20	21
23	108	LS + AR + EAM	129	10	20	26	24	24	23	2
24	102	Other	137	15	21	22	26	26	26	1
25	100	LS + AR	141	1	24	25	25	25	25	16
26	117	Other	153	13	23	24	23	23	24	23

The result for prediction from IERS is shown in bold

^aID 157 submitted only seven predictions

To summarize the achievements of the 2nd EOP PCC and provide some perspectives, it should be stated that currently, the most important factor in improving the accuracy of UT1–UTC and LOD forecasts is the use of precise and reliable EAM data and predictions. Therefore, the first step in this regard should be the development of EAM forecasts for longer horizons. It has also been demonstrated that modern ML-based algorithms have tremendous potential and should continue to be developed, although classical approaches based on LS + AR (with added EAM information) have also proven to be reliable.

Acknowledgements The EOP PCC Office would like to acknowledge the efforts of all participants of the 2nd EOP PCC for their unquestioned contributions to the campaign.

Author's contributions TK, JŚB, and MW proposed the general idea of this contribution. JŚB and TK did the computations. JŚB wrote the draft version of the manuscript. JN and HD commented regularly on

the results and gave suggestions. JN, HD, JŚB, TK, MW, AP, and 2nd EOP PCC participants reviewed and revised the manuscript.

Funding This study was funded by the National Science Centre, Poland under the OPUS call in the Weave programme, Grant number 2021/43/I/ST10/01738. H. Dobslaw is supported by the project DISCLOSE, funded by the German Research Foundation (DO 1311/6-1). The work of D. Boggs, M. Chin, R. Gross, and T. Ratcliff described in this paper was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. S. Belda was partially supported by Generalitat Valenciana (SEJIGENT/2021/001) and the European Union—NextGenerationEU (ZAMBRANO 21-04). J. M. Ferrandiz was partially supported by Spanish Project PID2020-119383 GB-I00 funded by Ministerio de Ciencia e Innovación (MCIN/AEI/10.13039/501100011033/) and PROMETEO/2021/030 funded by Generalitat Valenciana.

Availability of data and materials All EOP predictions analyzed in this study were submitted to the EOP PCC Office by registered participants in the frame of the 2nd EOP PCC. The data can be accessed from the GFZ Data Services under the following link: <https://doi.org/10.5880/>

GFZ.1.3.2023.001. Predictions developed by IERS/USNO, as well as the IERS 14 C04 solution used in this study to validate EOP predictions, are available at <https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html>.

Code availability Not applicable.

Declarations

Conflict of interest The authors declare no conflict of interests.

Consent to participate Not applicable.

Consent for publication Not applicable.

Ethical approval Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix 1

Description of each ID predicting UT1–UTC and/or LOD as provided by 2nd EOP PCC participants.

100

Authors with affiliations:

Xueqing Xu

CAS Key Laboratory of Planetary Sciences, Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China

Method description:

As complex variations of the Earth's rotation, there are commonly relative regular and irregular signals coupling in EOP data series, such as the trend, annual, Chandler terms, and high frequency trembles in polar motion; and the trend, interannual, seasonal and sub-seasonal oscillations in LOD changes. For the predictions of these stable signals, we adopt the LS model expressed by polynomial trend and harmonic oscillations; and a stochastic process AR model can be employed for the predictions of irregular variations (Xu et al. 2015, 2012).

101

Authors with affiliations:

Yuanwei Wu, Xin Zhao, Xinyu Yang

National Time Service Center of Chinese Academy of Sciences, Xi'an, China

Method description:

The C04 14 and the latest IERS daily files are combined as EOP inputs.

The GFZ EAM products, including 6 days predictions are used as EAM inputs.

The method we used to predict PM coordinates (PM_x, PM_y), LOD, and UT1 is similar to Dill et al. (2019) but with some revisions:

- (a) given the 1 day delay to GFZ's EAM prediction, the 6 days of prediction is adjust of 5 days.
- (b) in the step of LS and AR, the parameter is optimized but evaluations day for different parameters at different time scale, details of the optimization are given in the proceeding paper Wu et al. (2022).

102

Authors with affiliations:

Christian Bizouard

Observatoire de Paris, SYRTE, Paris, France

Method description:

For all predicted EOP parameters, the past data allowing to build the prediction are the daily operational IERS 14 C04 series up to the current date. The predicted values are given for each day of the interval [MJD0, MJD0 + 365] at 0 h UTC. Notation: MJD0 is a current date at 0 h UTC.

UT1–UTC & LOD: Considering UTR–TAI and LOD residuals (LODR): (1) LS fit of a degree 2 polynomial trend, 365.242 day and 182.26 day terms over the last 433 days. (2) Autoregressive modelling of the residuals (3) The 6 first days of the UT1 prediction from MJD0 to MJD0 + 6 are calculated by integrating the atmospheric angular momentum forecast of the GFZ (the LOD prediction is directly given by the AAM forecasted values to a constant factor and a bias) (4) The LS + AR model is extrapolated from the last day (MJD0 + 6) of the predicted values from AAM (5) Zonal tide effects are added back.

104

Authors with affiliations:

Richard Gross, Dale Boggs, Mike Chin, Todd Ratcliff

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, USA

Method description:

A Kalman filter is used to combine independent measurements of the Earth's orientation taken by the space-geodetic observing techniques of Satellite Laser Ranging, Very Long Baseline Interferometry, and the Global Navigation Satellite System. In order to improve the predicted EOPs, atmospheric and oceanic angular momenta analyses, and forecasts are used as proxy length-of-day and polar motion excitation measurements. Prior to their combination, the data series are

adjusted to have the same bias and rate, the stated uncertainties of the measurements are adjusted, and data points considered to be outliers are deleted (Freedman et al. 1994; Gross et al. 1998).

105

Authors with affiliations:

Robert Dill, Henryk Dobslaw, Jan Saynisch-Wagner, Christopher Irrgang, Maik Thomas

Department 1: Geodesy, Deutsches GeoForschungsZentrum GFZ, Potsdam, Germany

Method description:

EOP prediction is based on the GFZ EAM Predictor (Dill et al. 2019). The sum of EAM (4 years of model-based effective angular momentum functions including EAM 6-day forecasts) and the residual of geodetic angular momentum (GAM) (4 years of geodetic angular momentum derived from IERS 14 C04), extrapolated for the last ~ 30 days up to the end of the EAM 6-day forecasts by a 1st LS + AR step, is predicted into the future by a 2nd LS + AR step. GFZ uses this two-step GFZ EAM Predictor to provide daily updated EAM predictions from -90 days in the past to + 90 days into the future with 3-hourly sampling (<http://esmdata.gfz-potsdam.de:8080/>). The submitted EOP prediction #105 is generated as soon as the EAM prediction is available (~ 11 UTC). Using the latest available EOP coordinates from IERS rapid EOPs (finals.daily) from the day before as initial values for the Liouville equation a 90-day EOP prediction is derived from the EAM prediction. The predicted 3-hourly EOPs are re-sampled to daily time intervals and the first day is cut off to start the time series at the actual day of submission.

107

Authors with affiliations:

Zhijin Zhou, Lue Chen, Weitao Lu, Songtao Han

National Key Laboratory of Science and Technology on Aerospace Flight Dynamic, Beijing Aerospace Control Center, Beijing, China

Method description:

First determine the length of the training data and reads the raw data, preprocessing according to its type, then the least square fitting, calculating the residuals between the data and the LS model, the residual prediction data is obtained by autoregressive modeling, and the least square model is extrapolated. The two are added together and post-processed to obtain the prediction product (Chen et al. 2014).

108

Authors with affiliations:

Weitao Lu, Lue Chen, Zhijin Zhou, Songtao Han

National Key Laboratory of Science and Technology on Aerospace Flight Dynamic, Beijing Aerospace Control Center, Beijing, China

Method description:

The prediction method is LS + MAR, in which LS means difference LS, and MAR means Multi-elements AR. The inputs including EOP data released by IERS and AAM data released by GFZ. The prediction parameters include both PM components and UT1-UTC; the longest prediction day is 365.

112

Authors with affiliations:

Xueqing Xu

CAS Key Laboratory of Planetary Sciences, Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China

Method description:

In this method, the EAM from the Earth System Modelling Group of GeoForschungsZentrum Potsdam (ESMGFZ) are selected as the input excitation series. Afterward, the interannual, seasonal and sub-seasonal terms of EOP are calculated from the EAM predictions by the Liouville convolution equation. Meanwhile, the rest of the EOP trend terms are extrapolated by the polynomial LS model. Finally, the total EOP predictions are combined with the excitation calculations and trend extensions (Xu et al. 2023).

116

Authors with affiliations:

Erik Schoenemann (1), Sara Bruni (2), Michiel Otten (2), Volker Mayer (3)

- (1) ESA/ESOC, Robert-Bosch-Straße 5, Darmstadt, Germany
- (2) PosiTim UG at ESA/ESOC, Robert-Bosch-Straße 5, Darmstadt, Germany
- (3) LSE Space GmbH at ESA/ESOC, Robert-Bosch-Straße 5, Darmstadt, Germany

Method description:

ESA's contribution to the second IERS EOP PCC was based on the output of the ESA ERP Service (<http://navigation-office.esa.int/products/erp/>) generated from September 2021 to December 2022. The Service provides daily updates of ERP estimates and the relevant predictions for 119 days in the future. The estimation phase is based on a rigorous combination at the normal equation level of different geodetic solutions. The combination takes into account the full correlation matrices, and realizes a seamless transition between ERP estimates based on final and rapid input products. For the IERS EOP PCC, the combination included ESA's GNSS, SLR and DORIS official products submitted to the relevant International Association of Geodesy (IAG) Services, as well as BKG and Deutsches Geodätisches Forschungsinstitut (DGFI) solutions for intensive and rapid-turnaround VLBI sessions, respectively.

Concerning the prediction phase, the software implements a combination of least-square fitting and autoregressive modelling based on the whole history of ESA ERP estimates to characterize the deterministic part of the ERP variability and the high-frequency variability induced by non-tidal atmospheric and oceanic dynamics. Then, the short-range (6 days) EAM forecasts provided by GFZ are used to predict the irregular variations generated by the atmospheric, oceanic and hydrological dynamics. In order to stabilize the short-term predictions, EAM forecasts are also combined in the excitation domain with an additional signal extrapolating the difference between the geodetic excitation necessary to generate the observed history of ERP variations and the corresponding excitation extracted from geophysical models.

Additional details on the ESA ERP Service can be found in (Bruni et al. 2021); the ESA ERP Software is described in (Kehm et al. 2023).

117

Authors with affiliations:

Sadegh Modiri (1), Daniela Thaller (1), Santiago Belda (2), Sonia Guessoum (2), Jose M Ferrandiz (2), Shrishail Raut (3, 4), Sujata Dhar (3), Robert Heinkelmann (3), Harald Schuh (3, 4)

- (1) Federal Agency for Cartography and Geodesy BKG, Frankfurt am Main, Germany
- (2) UAVAC, University of Alicante, Alicante, Spain
- (3) Department 1: Geodesy, Deutsches Geo-Forschungszentrum GFZ, Potsdam, Germany
- (4) Institute for Geodesy and Geoinformation Science, Technische Universität Berlin, Berlin, Germany

Method description:

We used two different approaches for Earth Rotation Parameters (ERP) and celestial pole offsets (CPO) prediction. The proposed algorithm for predicting the ERP is called SSA + Copula-based analysis (Modiri et al. 2018, 2020; Modiri 2021). The algorithm splits the observed PM time series into periodic terms and anomalies, which are modelled using SSA and Copula-based analysis, respectively. The SSA periodic terms estimation involves selecting a window parameter, forming a trajectory matrix, performing singular value decomposition, selecting a proper group of singular values and corresponding singular vectors, and calculating the trend. The Copula anomaly modelling involves forming the trajectory matrix of residual time series, computing the marginal distribution, transforming data to the rank space, computing the empirical and conditional Copula, and sampling random data from the conditional Copula cumulative distribution function (CDF). The final predicted PM data is

the sum of the predicted periodic terms using SSA and the predicted anomaly using the Copula-based model.

124

Authors with affiliations:

Mostafa Kiani Shahvandi, Matthias Schartner, Junyang Gou, Benedikt Soja

Institute of Geodesy and Photogrammetry, ETH Zurich, Zurich, Switzerland

Method description:

This method focuses on the 10-day prediction of LOD using a hybrid modeling approach, which combines physical, statistical, and encoder-decoder LSTM (EDLSTM) network while considering geophysical excitations. We first remove the secular trend and known signals from the LOD time series by combining the Savitzky–Golay filter (Savitzky and Golay 1964), tidal corrections (Petit and Luzum 2010) and least-squares adjustment (Brockwell and Davis 2002) to generate the LOD residuals (LODR). The GFZ AAM data (Dobslaw and Dill 2018) and corresponding 6-day forecasts are also preprocessed using least-squares adjustment to obtain AAM residuals (AAMR).

Then, the LODR, AAMR, and 6-day forecasts of AAMR are concatenated and considered as input features. The final dimension of the input tensors is 30×8 since we consider the previous 30 days as the input sequence. Then we feed this tensor into the EDLSTM network (Hochreiter and Schmidhuber 1997, Nayak and Ng 2020) to predict LODR for the next ten days. In order to restore the full LOD, we should also predict the previously removed components. The tidal and seasonal signals can be easily predicted since they are estimated using deterministic models. The long-term trend will be predicted using PCHIP (Piecewise Cubic Hermite Interpolating Polynomial) extrapolation (Fritsch and Carlson 1980).

For more details, please refer to Gou et al. (2023).

125

Authors with affiliations:

Mostafa Kiani Shahvandi, Matthias Schartner, Junyang Gou, Benedikt Soja

Institute of Geodesy and Photogrammetry, ETH Zurich, Zurich, Switzerland

Method description:

This architecture is designed (Kiani Shahvandi and Soja 2022a) to extract meaningful information even from the scarce data. The elements of this architecture are as follows: (1) residual learning; (2) attention mechanism; (3) long short-term memory (LSTM). Several consecutive blocks of the mentioned elements are stacked together in an auto-encoder manner. Each of these blocks are pretrained using the so-called greedy-layer wise pretraining in order to facilitate the main training phase. Similar to the studies (Kiani Shahvandi and Soja 2021) on Transformer architecture, number of the blocks depends on the accuracy obtained on the validation

set during training, but for the prediction of LOD there are usually two blocks in the architecture. It is also important to note that the architecture of LSTM element is similar to the one used in Gou et al. (2023). The input and output sequence lengths to this architecture are 30 and 12, respectively. The loss function is mean absolute error. The algorithm is retrained at each prediction epoch to take advantage of the new EOP and EAM data being available. This architecture is trained on LOD residuals (i.e., after the removal of secular trends, tidal effects, and seasonal signals) having the AAM and its 6-day forecasts as additional features (seasonal signals are removed from AAM). The final prediction is the summation of output of the architecture and the prediction of secular trends, tides, and seasonal signals of the LOD.

126, 129, 130, 144–146

Authors with affiliations:

Mostafa Kiani Shahvandi, Matthias Schartner, Junyang Gou, Benedikt Soja

Institute of Geodesy and Photogrammetry, ETH Zurich, Zurich, Switzerland

Method description:

The architecture used is based on the first-order neural ordinary differential equations (Neural ODEs). In this architecture the hidden state in the hidden layer should follow a differential equation. To apply this concept to the EOPs, it is assumed that EOPs follow first-order differential equations the exact form of which should be determined by fitting neural networks to the observations. The general approach of Neural ODE differential learning (Kiani Shahvandi et al. 2022a) is modified (i.e., in a way that does not require using the rates of EOPs) and used as the primary architecture. A variation of this architecture is the so-called simple recursive method (Kiani Shahvandi et al. 2022b), in which an attempt is made to incorporate the uncertainties in the observational data in the training for a more reliable estimation of parameters of the neural networks (Kiani Shahvandi and Soja 2022b). As the result, the loss function here is the mean squared error. The architecture does not require any pre-processing of the input features. However, in case of LOD prediction it is used on the LOD residuals after the removal of secular trends, tides, and seasonal signals (Gou et al. 2023). The forecasting horizon of the architecture contains both 10 and 30 days. The input sequence length to the architectures is 10 days. The architectures are trained at each prediction epoch to take advantage of the most recent available EOP and EAM data.

136

Authors with affiliations:

Robert Dill, Henryk Dobslaw, Jan Saynisch-Wagner, Christopher Irrgang, Maik Thomas

Department 1: Geodesy, Deutsches GeoForschungsZentrum GFZ, Potsdam, Germany

Method description:

EOP prediction is based on the GFZ EAM Predictor (Dill et al. 2019). The sum of EAM (4 years of model-based effective angular momentum functions including EAM 6-day forecasts) and the residual of GAM (4 years of geodetic angular momentum derived from IERS 14 C04), extrapolated for the last ~ 30 days up to the end of the EAM 6-day forecasts by a 1st LS + AR step, is predicted into the future by a 2nd LS + AR step. GFZ uses this two-step GFZ EAM Predictor to provide daily updated EAM predictions from – 90 days in the past to + 90 days into the future with 3-hourly sampling (<http://esmdata.gfz-potsdam.de:8080/>). The submitted EOP prediction #136 is generated as soon as the rapid EOP solution for the actual day is available (~ 17:15 UTC). The latest non-predicted EOP coordinates from IERS rapid EOPs (finals.daily) are taken as initial values for the Liouville equation to derive a 90-day EOP prediction from the EAM prediction. The predicted 3-hourly EOPs are re-sampled to daily time intervals.

141

Authors with affiliations:

Maciej Michalczyk, Marcin Ligas

AGH University of Krakow, Adama Mickiewicza 30, 30–059 Kraków, Poland

Method description:

Ordinary kriging belongs to a broad family of geostatistical methods of prediction, and is optimal in the sense of Best Linear Unbiased Prediction (BLUP) if the mean value of a random function is an unknown constant. Prediction by means of kriging requires a structure function that describes continuity and variability of a random process. A semivariogram, due to its advantage over a covariance function, is used to describe a structure hidden in the residual series, that is obtained after the removal of a linear trend, periodic components and periodicity associated with solid Earth tides. Among theoretical semivariogram models the best performing one turned out to be exponential model. The final forecast consists of extrapolated deterministic part combined with the predicted (kriged) stochastic part. The entire process is presented in Michalczyk and Ligas (2021, 2022).

142

Authors with affiliations:

Mostafa Kiani Shahvandi, Matthias Schartner, Junyang Gou, Benedikt Soja

Institute of Geodesy and Photogrammetry, ETH Zurich, Zurich, Switzerland

Method description:

This method focuses on the 10-day prediction of LOD using a hybrid modeling approach, which combines physical, statistical, and encoder-decoder LSTM (EDLSTM) network while considering geophysical excitations. We first remove the secular trend and known signals from the LOD time

series by combining the Savitzky–Golay filter (Savitzky and Golay 1964), tidal corrections (Petit and Luzum 2010), and least-squares adjustment (Brockwell and Davis 2002) to generate the LOD residuals (LODR). The GFZ EAM (AAM + OAM + HAM) data (Dobslaw and Dill 2018) and corresponding 6-day forecasts are also preprocessed using least-squares adjustment to obtain EAM residuals (EAMR).

Then, the LODR, EAMR, and 6-day forecasts of EAMR are concatenated and considered as input features. The final dimension of the input tensors is 30×8 since we consider the previous 30 days as the input sequence. Then we feed this tensor into the EDLSTM network (Hochreiter and Schmidhuber 1997, Nayak and Ng 2020) to predict LODR for the next ten days. In order to restore the full LOD, we should also predict the previously removed components. The tidal and seasonal signals can be easily predicted since they are estimated using deterministic models. The long-term trend will be predicted using PCHIP (Piecewise Cubic Hermite Interpolating Polynomial) extrapolation (Fritsch and Carlson 1980).

For more details, please refer to Gou et al. (2023).

143

Authors with affiliations:

Mostafa Kiani Shahvandi, Matthias Schartner, Junyang Gou, Benedikt Soja

Institute of Geodesy and Photogrammetry, ETH Zurich, Zurich, Switzerland

Method description:

The basis of this architecture is the multilayer perceptron. However, in order to use several hidden layers and overcoming the problems with deep neural networks, we have added the residual learning blocks to this architecture, similar to Kiani Shahvandi and Soja (2022a). Therefore, we name this architecture the revised multilayer perceptron. The activation functions for the hidden layers are tangent hyperbolic, except for the last layer which is linear. Number of hidden layers is three. The input sequence length in this algorithm is 10. However, the output sequence length is either 12 or 32 depending on the forecasting horizon, both of which are provided. The architecture is used for LOD prediction in the same manner as Gou et al. (2023). The inputs to this algorithm are preprocessed, i.e., in the case of LOD the trends, tides, and seasonal signals are removed, while for EAM functions the seasonal signals are subtracted from the observations and forecasts. The mentioned subtracted signals are subsequently added to the predictions of the architecture in order to give the final value of LOD prediction.

147, 148, 149

Authors with affiliations:

Mostafa Kiani Shahvandi, Matthias Schartner, Junyang Gou, Benedikt Soja

Institute of Geodesy and Photogrammetry, ETH Zurich, Zurich, Switzerland

Method description:

This architecture is based on the first-order neural ordinary differential equations (Neural ODEs). The general Neural ODE differential learning architecture (Kiani Shahvandi et al. 2022a) is modified in a way that does not incorporate the rates of EOPs into the architecture. A simplified form of this architecture is also used (referred to as simple recursive) (Kiani Shahvandi et al. 2022b) in which the uncertainties in the input data are used to weigh the loss function (Kiani Shahvandi and Soja 2022b) (the loss function here is the mean squared error) for a more reliable estimation of the parameters of the neural networks. However, investigating the residuals of training phase reveals that some signals in the observations are not well captured by the Neural ODEs. Therefore, an attempt is made to model these residuals by Long Short-Term Memory (LSTM) neural networks in the same manner suggested by Gou et al. (2023). First, Neural ODEs are trained and then the fitted values are subtracted from the observations in order to compute the residuals. Subsequently these residuals are modelled by LSTM. The input to the LSTM architecture is only the past values of residuals of training of the Neural ODEs architecture. The input sequence length here is 10 and the loss function is mean absolute error. For this purpose, the predictions of IERS are also incorporated into the algorithm in order to predict the residuals at each training epoch (retraining is required).

156

Authors with affiliations:

Maciej Michalczyk, Marcin Ligas

AGH University of Krakow, Adama Mickiewicza 30, 30–059 Kraków, Poland

Method description:

The deterministic part, i.e., estimated linear trend, periodic components and tidal effects, is first removed from raw times series. ARIMA model is then used to predict the residual part. Autoregressive integrated moving average ARIMA(p, d, q) model is a combination of autoregressive model (p), moving average model (q) and differencing process (integrated part; d) that accounts for a potential non-stationarity of a residual process. The deterministic component is extrapolated for future time instances and then combined with ARIMA-based predicted stochastic part. The best set of parameters p and q is selected by means of corrected Akaike Information Criterion (AIC). Parameter d determines a degree of differencing to be applied in order to transform a non-stationary time series into a stationary one in the mean sense. Stationarity of the residual process is checked using Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. The entire prediction procedure is described in Michalczyk et al. (2022).

157

Authors with affiliations:

Maciej Michalczak, Marcin Ligas
AGH University of Krakow, Adama Mickiewicza 30,
30–059 Kraków, Poland

Method description:

Dynamic mode decomposition (DMD) is a relatively new technique of data decomposition that emerged in the field of fluid dynamics due to work by Schmid (2010). It is described as an ideal marriage of data decomposition methods (e.g., principal component analysis) and Fourier transform (Kutz et al. 2016). It is a data-driven, equation-free technique with the only assumption that some dynamics is present in data. It has the ability to reconstruct and forecast data in a single numerical procedure. Despite DMD is of spatio-temporal origin with a slight modification it can be used to univariate time series by splitting an input time series of length T into L sub-series (shifted by one time step ahead) of length K (Tirunagari et al. 2017). This generates a trajectory matrix known from, e.g., singular spectrum analysis (SSA). The main goal of the method is to capture a low rank structure of the analyzed dynamical system, i.e., to decompose it into the most dominant components (trends, harmonics) that may be used for a future state prediction later on. Due to numerical feasibility, during the 2nd EOP PCC the method was used without dimensionality reduction step (no low rank approximation involved).

References

- Altamimi Z, Rebischung P, Collilieux X et al (2023) ITRF2020: an augmented reference frame refining the modeling of nonlinear station motions. *J Geod* 97:47. <https://doi.org/10.1007/s00190-023-01738-w>
- Barnes RTH, Hide R, White AA, Wilson CA (1983) Atmospheric angular momentum fluctuations, length-of-day changes and polar motion. *Proc R Soc Lond* 387:31–73. <https://doi.org/10.1098/rspa.1983.0050>
- Bizouard C, Fernández LI, Zotov L (2022) Admittance of the earth rotational response to zonal tide potential. *J Geophys Res Solid Earth*. <https://doi.org/10.1029/2021JB022962>
- Bizouard C, Gambis D (2009) The combined solution C04 for earth orientation parameters consistent with international terrestrial reference frame 2005. In: International association of geodesy symposia, vol 134. https://doi.org/10.1007/978-3-642-00860-3_41
- Bizouard C, Lambert S, Gattano C, Becker O, Richard JY (2019) The IERS EOP 14 C04 solution for earth orientation parameters consistent with ITRF 2014. *J Geodesy*. <https://doi.org/10.1007/s00190-018-1186-3>
- Bloßfeld M, Rudenko S, Kehm A, Panadina N, Müller H, Angermann D, Hugentobler U, Seitz M (2018) Consistent estimation of geodetic parameters from SLR satellite constellation measurements. *J Geod*. <https://doi.org/10.1007/s00190-018-1166-7>
- Brockwell PJ, Davis RA (2002) Introduction to time series and forecasting. Springer, Berlin
- Bruni S, Schoenemann E, Mayer V, Otten M, Springer T, Dilssner F, Enderle W, Zandbergen R (2021) ESA'S earth orientation parameter product, Austria, Vienna: EGU General Assembly 2021, online, 19–30 Apr 2021. <https://doi.org/10.5194/egusphere-egu21-12989>
- Brzeziński A (1992) Polar motion excitation by variations of the effective angular momentum function: considerations concerning deconvolution problem. *Manuscr Geod* 17:3–20
- Charlot P, Jacobs CS, Gordon D, Lambert S, De Witt A, Böhm J, Fey AL, Heinkelmann R, Skurikhina E, Titov O, Arias EF, Bolotin S, Bourda G, Ma C, Malkin Z, Nothnagel A, Mayer D, Macmillan DS, Nilsson T, Gaume R (2020) The third realization of the international celestial reference frame by very long baseline interferometry. *Astron Astrophys* 644:1–28. <https://doi.org/10.1051/0004-6361/202038368>
- Chen L, Tang G, Hu S, Ping J, Xu X, Xia J (2014) High accuracy differential prediction of UT1-UTC. *J Deep Space Explor* 1(3):230–235. <https://doi.org/10.15982/j.issn.2095-7777.2014.03.012>
- Dill R, Dobsław H, Thomas M (2013) Combination of modeled short-term angular momentum function forecasts from atmosphere, ocean, and hydrology with 90-day EOP predictions. *J Geodesy* 87(6):567–577. <https://doi.org/10.1007/s00190-013-0631-6>
- Dill R, Dobsław H, Thomas M (2019) Improved 90-day Earth orientation predictions from angular momentum forecasts of atmosphere, ocean, and terrestrial hydrosphere. *J Geodesy* 93(3):287–295. <https://doi.org/10.1007/s00190-018-1158-7>
- Dill R, Dobsław H, Thomas M (2022) ESMGFZ products for earth rotation prediction. *Artif Satell* 57(s1):254–261. <https://doi.org/10.2478/arsa-2022-0022>
- Dobsław H, Dill R (2018) Predicting earth orientation changes from global forecasts of atmosphere-hydrosphere dynamics. *Adv Space Res* 61(104):47–1054. <https://doi.org/10.1016/j.asr.2017.11.044>
- Ferland R, Piraszewski M (2009) The IGS-combined station coordinates, earth rotation parameters and apparent geocenter. *J Geod* 83(3–4):385–392. <https://doi.org/10.1007/s00190-008-0295-9>
- Freedman AP, Steppe JA, Dickey JO, Eubanks TM, Sung LY (1994) The short-term prediction of universal time and length of day using atmospheric angular momentum. *J Geophys Res* 99:6981–6996. <https://doi.org/10.1029/93JB02976>
- Fritsch FN, Carlson RE (1980) Monotone piecewise cubic interpolation. *SIAM J Numer Anal* 17:238–246. <https://doi.org/10.1137/0717021>
- Gambis D (2004) Monitoring earth orientation using space-geodetic techniques: state-of-the-art and prospective. *J Geodesy*. <https://doi.org/10.1007/s00190-004-0394-1>
- Gambis D, Luzum B (2011) Earth rotation monitoring, UT1 determination and prediction. *Metrologia*. <https://doi.org/10.1088/0026-1394/48/4/S06>
- Glaser S, Fritsche M, Sošnica K, Rodriguez-Solano C, Wang K, Dach R, Hugentobler U, Rothacher M, Dietrich R (2015) A consistent combination of GNSS and SLR with minimum constraints. *J Geod* 89(12):1165–1180. <https://doi.org/10.1007/s00190-015-0842-0>
- Gou J, Kiani Shahvandi M, Hohensinn R, Soja B (2023) Ultra-short-term prediction of LOD using LSTM neural networks. *J Geod* 97:52. <https://doi.org/10.1007/s00190-023-01745-x>
- Gross RS (2007) Earth rotation variations—long period. In: Schubert G (ed) Treatise on geophysics. Elsevier, Amsterdam, pp 239–294. <https://doi.org/10.1016/B978-0-444-52748-6.00057-2>
- Gross RS, Eubanks TM, Steppe JA, Freedman AP, Dickey JO, Runge TF (1998) A Kalman-filter-based approach to combining independent earth-orientation series. *J Geodesy*. <https://doi.org/10.1007/s001900050162>
- Gross RS, Fukumori I, Menemenlis D, Gegout P (2004) Atmospheric and oceanic excitation of length-of-day variations during 1980–2000. *J Geophys Res* 109:B01406. <https://doi.org/10.1029/2003JB002432>
- Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9:1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- IERS Annual Report 2019, Dick WR, Thaller D (eds) (2023) International earth rotation and reference systems service, Central Bureau.

- Frankfurt am Main: Verlag des Bundesamts für Kartographie und Geodäsie, ISBN 978-3-86482-136-3. <http://www.iers.org/IERS/AR2019>
- Kalarus M, Schuh H, Kosek W, Akyilmaz O, Bizouard C, Gambis D, Gross R, Jovanović B, Kumakshev S, Kutterer H, Mendes Cerveira PJ, Pasynok S, Zotov L (2010) Achievements of the Earth orientation parameters prediction comparison campaign. *J Geodesy* 84(10):587–596. <https://doi.org/10.1007/s00190-010-0387-1>
- Karbon M, Soja B, Nilsson T, Deng Z, Heinkelmann R, Schuh H (2017) Earth orientation parameters from VLBI determined with a Kalman filter. *Geodesy Geodyn* 8(6):396–407. <https://doi.org/10.1016/j.geog.2017.05.006>
- Kehm A, Hellmers H, Bloßfeld M, Dill R, Angermann D, Seitz F et al (2023) Combination strategy for consistent final, rapid and predicted earth rotation parameters. *J Geodesy* 97:1. <https://doi.org/10.1007/s00190-022-01695-w>
- Kiani Shahvandi M, Soja B (2021) Modified deep transformers for GNSS time series prediction. In: 2021 IEEE international geoscience and remote sensing symposium IGARSS, Brussels, Belgium, pp 8313–8316. <https://doi.org/10.1109/IGARSS47720.2021.9554764>
- Kiani Shahvandi M, Schartner M, Soja B (2022a) Neural ODE differential learning and its application in polar motion prediction. *J Geophys Res Solid Earth* 127:e2022JB024775. <https://doi.org/10.1029/2022JB024775>
- Kiani Shahvandi M, Gou J, Schartner M, Soja B (2022b) Data driven approaches for the prediction of Earth's effective angular momentum functions. In: IGARSS 2022–2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, pp 6550–6553. <https://doi.org/10.1109/IGARSS46834.2022.9883545>
- Kiani Shahvandi M, Schartner M, Gou J, Soja B (2023) Operational 14-day-ahead prediction of earth's effective angular momentum functions with machine learning. *IUGG Berl.* <https://doi.org/10.57757/IUGG23-0346>
- Kiani Shahvandi M, Soja B (2022b) Inclusion of data uncertainty in machine learning and its application in geodetic data science, with case studies for the prediction of earth orientation parameters and GNSS station coordinate time series. *Adv Space Res* 70(3):563–575. <https://doi.org/10.1016/j.asr.2022.05.042>
- Kiani Shahvandi M, Soja B (2022a) Small geodetic datasets and deep networks: attention-based residual LSTM autoencoder stacking for geodetic time series. In: Machine learning, optimization, and data science. LOD 2021. Lecture notes in computer science, vol 13163. Springer, Cham. https://doi.org/10.1007/978-3-030-95467-3_22
- Kur T, Dobsław H, Śliwińska J, Nastula J, Wińska M (2022) Evaluation of selected short-term predictions of UT1 - UTC and LOD collected in the second earth orientation parameters prediction comparison campaign. *Earth Planets Space.* <https://doi.org/10.1186/s40623-022-01753-9>
- Kutz JN, Brunton SL, Brunton BW, Proctor J (2016) Dynamic mode decomposition: data-driven modeling of complex systems. *Soc Ind Appl Math Phila.* <https://doi.org/10.1137/1.9781611974508>
- Lei Y, Zhao D, Guo M (2023) Medium- and long-term prediction of polar motion using weighted least squares extrapolation and vector autoregressive modeling. *Artif Satell* 58(2):42–55. <https://doi.org/10.2478/arsa-2023-0004>
- Luzum B (2010) Future of earth orientation predictions. *Artif Satell* 45(2):107–110. <https://doi.org/10.2478/v10018-010-0011-x>
- Luzum BJ, Ray JR, Carter MS, Josties FJ (2001) Recent improvements to IERS bulletin A combination and prediction. *GPS Solut* 4(3):34–40. <https://doi.org/10.1007/PL00012853>
- McCarthy DD, Luzum BJ (1991) Prediction of earth orientation. *Bull Géodésique* 65:18–21. <https://doi.org/10.1007/BF00806338>
- Michalczyk M, Ligas M (2021) Kriging-based prediction of the earth's pole coordinates. *J Appl Geodesy.* <https://doi.org/10.1515/jag-2021-0007>
- Michalczyk M, Ligas M (2022) The (ultra) short term prediction of length-of-day using kriging. *Adv Space Res.* <https://doi.org/10.1016/j.asr.2022.05.007>
- Michalczyk M, Ligas M, Kudryś J (2022) Prediction of earth rotation parameters with the use of rapid products from IGS, code and GFZ data centres using Arima and Kriging—a comparison. *Artif Satell* 57(s1):275–289. <https://doi.org/10.2478/arsa-2022-0024>
- Mikschi M, Böhm J, Böhm S, Horozovic D (2019) Comparison of integrated GNSS LOD to dUT1. In: Proceedings of the 24th European VLBI group for geodesy and astrometry working meeting, vol 4. Chalmers University of Technology, pp 247–251
- Modiri S (2021) On the improvement of earth orientation parameters estimation: using modern space geodetic techniques, PhD Thesis, Technische Universität Berlin (Germany). <https://doi.org/10.48440/gfz.b103-21107>
- Modiri S, Belda S, Heinkelmann R, Hoseini M, Ferrándiz JM, Schuh H (2018) Polar motion prediction using the combination of SSA and Copula-based analysis. *Earth Planets Space* 70:115. <https://doi.org/10.1186/s40623-018-0888-3>
- Modiri S, Belda S, Hoseini M, Heinkelmann R, Ferrándiz JM, Schuh H (2020) A new hybrid method to improve the ultra-short-term prediction of LOD. *J Geod* 94:23. <https://doi.org/10.1007/s00190-020-01354-y>
- Moreaux G, Lemoine FG, Capdeville H, Otten M, Štěpánek P, Saunier J, Ferrage P (2023) The international DORIS service contribution to ITRF2020. *Adv Space Res* 72(1):65–91. <https://doi.org/10.1016/j.asr.2022.07.012>
- Nayak T, Ng HT (2020) Effective modeling of encoder-decoder architecture for joint entity and relation extraction. *Proc AAAI Conf Artif Intell* 34:8528–8535. <https://doi.org/10.1609/aaai.v34i05.6374>
- Oliveau SH, Freedman AP (1997) Accuracy of earth orientation parameter estimates and short-term predictions generated by the Kalman earth orientation filter. The Telecommunications and Data Acquisition Progress Report, pp 1–10. Accessed 20 Aug 2023 https://tda.jpl.nasa.gov/progress_report/42-129/129C.pdf
- Petit G, Luzum B (2010) IERS conventions (2010). Verlag des Bundesamts für Kartographie und Geodäsie, Frankfurt am Main
- Ratcliff JT, Gross RS (2019) Combinations of earth orientation measurements: SPACE2018, COMB2018, and POLE2018. Jet Propulsion Laboratory, California Institute of Technology, Publication 19–7. Accessed 20 Aug 2023 <https://hdl.handle.net/2014/45971>
- Robertson D, Carter W, Campbell J et al (1985) Daily earth rotation determinations from IRIS very long baseline interferometry. *Nature* 316:424–427. <https://doi.org/10.1038/316424a0>
- Savitzky A, Golay MJ (1964) Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem* 36:1627–1639. <https://doi.org/10.1021/ac60214a047>
- Scaife AA, Hermanson L, van Niekerk A et al (2022) Long-range predictability of extratropical climate and the length of day. *Nat Geosci* 15:789–793. <https://doi.org/10.1038/s41561-022-01037-7>
- Schmid PJ (2010) Dynamic mode decomposition of numerical and experimental data. *J Fluid Mech* 656:5–28. <https://doi.org/10.1017/S0022112010001217>
- Schuh H, Böhm J (2013) Very long baseline interferometry for geodesy and astrometry. In: Xu G (ed) Sciences of geodesy—II. Springer, Berlin. https://doi.org/10.1007/978-3-642-28000-9_7
- Senior K, Kouba J, Ray J (2010) Status and prospects for combined GPS LOD and VLBI UT1 measurements. *Artif Satell* 45:57–73. <https://doi.org/10.2478/v10018-010-0006-7>
- Soja B, Kiani Shahvandi M, Schartner M, Gou J, Kłopotek G, Crocetti L, Awadaljeed M (2022) The new geodetic prediction center at ETH Zurich, EGU General Assembly 2022, Vienna, Austria,

- 23–27 May 2022, EGU22-9285. <https://doi.org/10.5194/egusphere-egu22-9285>
- Sośnica K, Bury G, Zajdel R, Strugarek D, Drożdżewski M, Kazmierski K (2019) Estimating global geodetic parameters using SLR observations to Galileo, GLONASS, BeiDou, GPS, and QZSS. *Earth Planets Space* 71:20. <https://doi.org/10.1186/s40623-019-1000-3>
- Sovers OJ, Fanselow JL, Jacobs CS (1998) Astrometry and geodesy with radio interferometry: experiments, models, results. *Rev Mod Phys*. <https://doi.org/10.1103/RevModPhys.70.1393>
- Steigenberger P, Rothacher M, Dietrich R, Fritsche M, Rülke A, Vey S (2006) Reprocessing of a global GPS network. *J Geophys Res Solid Earth*. <https://doi.org/10.1029/2005JB003747>
- Śliwińska J, Kur T, Wińska M, Nastula J, Dobslaw H, Partyka A (2022) Second earth orientation parameters prediction comparison campaign (2nd EOP PCC): overview. *Artif Satell* 57(S1):237–253. <https://doi.org/10.2478/arsa-2022-0021>
- Tirunagari S, Kouchaki S, Poh N, Bober M, Windridge D (2017) Dynamic mode decomposition for univariate time series: analysing trends and forecasting. <https://hal.science/hal-01463744>
- Wu Y, Zhao X, Yang X (2022) Improved prediction of polar motions by piecewise parameterization. *Artif Satell J Planet Geodesy* 57(S11):290–299. <https://doi.org/10.2478/arsa-2022-0025>
- Xu XQ, Zhou YH, Xu CC (2023) Earth rotation parameters prediction and climate change indicators in it. *Artif Satell J Planet Geodesy* 57(S11):262–273. <https://doi.org/10.2478/arsa-2022-0023>
- Xu XQ, Zhou YH (2015) EOP prediction using least square fit in and autoregressive filter over optimized data intervals. *Adv Space Res* 56:2248–2253. <https://doi.org/10.1016/j.asr.2015.08.007>
- Xu XQ, Zhou YH, Liao XH (2012) Short-term earth orientation parameters predictions by combination of the least squares, AR model and Kalman filter. *J Geodyn* 62:83–86. <https://doi.org/10.1016/j.jog.2011.12.001>
- Zajdel R, Sośnica K, Bury G (2020) System-specific systematic errors in earth rotation parameters derived from GPS, GLONASS, and Galileo. *GPS Solut* 24:74. <https://doi.org/10.1007/s10291-020-00989-w>

Authors and Affiliations

Justyna Śliwińska-Bronowicz¹  · Tomasz Kur² · Małgorzata Wińska³ · Henryk Dobslaw⁴ · Jolanta Nastula¹ · Aleksander Partyka¹ · Santiago Belda⁵ · Christian Bizouard⁶ · Dale Boggs⁷ · Sara Bruni⁸ · Lue Chen⁹ · Mike Chin⁷ · Sujata Dhar⁴ · Robert Dill⁴ · Jose Manuel Ferrandiz⁵ · Junyang Gou¹⁰ · Richard Gross⁷ · Sonia Guessoum⁵ · Songtao Han⁹ · Robert Heinkelmann⁴ · Christopher Irrgang⁴ · Mostafa Kiani Shahvandi¹⁰ · Jia Li¹¹ · Marcin Ligas¹² · Lintao Liu¹³ · Weitao Lu⁹ · Volker Mayer¹⁴ · Maciej Michalczak¹² · Sadegh Modiri¹⁵ · Michiel Otten⁸ · Todd Ratcliff⁷ · Shrishail Raut^{4,16} · Jan Saynisch-Wagner⁴ · Matthias Schartner¹⁰ · Erik Schoenemann¹⁷ · Harald Schuh^{4,16} · Benedikt Soja¹⁰ · Xiaoqing Su¹⁸ · Daniela Thaller¹⁵ · Maik Thomas⁴ · Guocheng Wang¹³ · Yuanwei Wu¹⁹ · Xueqing Xu²⁰ · Xinyu Yang¹⁹ · Xin Zhao¹⁹ · Zhijin Zhou⁹

✉ Justyna Śliwińska-Bronowicz
jsliwinska@cbk.waw.pl

¹ Centrum Badań Kosmicznych Polskiej Akademii Nauk, Bartycka 18A, 00-716 Warsaw, Poland

² Institute of Geodesy and Geoinformatics, Wrocław University of Environmental and Life Sciences, Norwida 25, 50-375 Wrocław, Poland

³ Faculty of Civil Engineering, Warsaw University of Technology, Armii Ludowej 16, 00-637 Warsaw, Poland

⁴ Department 1: Geodesy, Deutsches GeoForschungsZentrum GFZ, Potsdam, Germany

⁵ UAVAC, University of Alicante, Alicante, Spain

⁶ Observatoire de Paris, SYRTE, Paris, France

⁷ Jet Propulsion Laboratory, California Institute of Technology, Pasadena, USA

⁸ PosiTim UG at ESA/ESOC, Robert-Bosch-Straße 5, Darmstadt, Germany

⁹ National Key Laboratory of Science and Technology on Aerospace Flight Dynamic, Beijing Aerospace Control Center, Beijing, China

¹⁰ Institute of Geodesy and Photogrammetry, ETH Zurich, Zurich, Switzerland

¹¹ Xi'an Surveying and Mapping Center, Xi'an, China

¹² AGH University of Krakow, Adama Mickiewicza 30, 30-059 Kraków, Poland

¹³ Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Beijing, China

¹⁴ LSE Space GmbH at ESA/ESOC, Robert-Bosch-Straße 5, Darmstadt, Germany

¹⁵ Federal Agency for Cartography and Geodesy BKG, Frankfurt am Main, Germany

¹⁶ Institute for Geodesy and Geoinformation Science, Technische Universität Berlin, Berlin, Germany

¹⁷ ESA/ESOC, Robert-Bosch-Straße 5, Darmstadt, Germany

¹⁸ Shandong University of Technology, Zibo, China

¹⁹ National Time Service Center of Chinese Academy of Sciences, Xi'an, China

²⁰ CAS Key Laboratory of Planetary Sciences, Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China