Prospects of Predicting the Polar Motion Based on the Results of the Second Earth Orientation Parameters Prediction Comparison Campaign

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

AGU ADVANCING EARTH AND SPACE SCIENCES

Earth and Space Science

RESEARCH ARTICLE
10.1029/2023EA003278

Key Points:

- Polar motion predictions for the *y* component are more accurate than for the *x* component
- Least squares and auto regression with effective angular momentum data provides the best results for polar motion prediction
- Some of the submitted predictions show higher accuracy than the IERS solution

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

T. Kur, J. Śliwińska-Bronowicz, M. Wińska, H. Dobslaw and J. Nastula, tomasz.kur@upwr.edu.pl; jsliwinska@cbk.waw.pl; malgorzata.winska@pw.edu.pl; dobslaw@gfz-potsdam.de; nastula@cbk.waw.pl

Citation:

Kur, T., Śliwińska-Bronowicz, J., Wińska, M., Dobslaw, H., Nastula, J., Partyka, A., et al. (2024). Prospects of predicting the polar motion based on the results of the second earth orientation parameters prediction comparison campaign. *Earth and Space Science*, *11*, e2023EA003278. https://doi.org/10.1029/2023EA003278

Received 12 SEP 2023
Accepted 26 SEP 2024

Author Contributions:

Conceptualization: Tomasz Kur, Justyna Śliwińska-Bronowicz, Malgorzata Wińska, Henryk Dobslaw, Jolanta Nastula, Aleksander Partyka
Data curation: Tomasz Kur, Justyna Śliwińska-Bronowicz, Malgorzata Wińska, Henryk Dobslaw, Jolanta Nastula, Aleksander Partyka, Santiago Belda, Christian Bizouard, Dale Boggs, Sara Bruni, Lue Chen, Mike Chin, Sujata Dhar, Robert Dill, PengShuo Duan, Jose M. Ferrandiz, Junyang Gou, Richard Gross,

Prospects of Predicting the Polar Motion Based on the Results of the Second Earth Orientation Parameters Prediction Comparison Campaign

Tomasz Kur[1,2], Justyna Śliwińska-Bronowicz[2], Malgorzata Wińska[3], Henryk Dobslaw[4], Jolanta Nastula[2], Aleksander Partyka[2], Santiago Belda[5], Christian Bizouard[6], Dale Boggs[7], Sara Bruni[8], Lue Chen[9], Mike Chin[7], Sujata Dhar[4], Robert Dill[4], PengShuo Duan[10], Jose M. Ferrandiz[5], Junyang Gou[11], Richard Gross[7], Sonia Guessoum[5], Songtao Han[9], Robert Heinkelmann[4], ChengLi Huang[10,12], Christopher Irrgang[4], Jacek Kudrys[13], Jia Li[14], Marcin Ligas[13], Lintao Liu[15], Weitao Lu[9], Volker Mayer[16], Wei Miao[17], Maciej Michalczak[13], Sadegh Modiri[18], Michiel Otten[8], Todd Ratcliff[7], Shrishail Raut[4,19], Jan Saynisch-Wagner[4], Matthias Schartner[11], Erik Schoenemann[20], Harald Schuh[4,19], M. Kiani Shahvandi[11], Benedikt Soja[11], Xiaoqing Su[21], Daniela Thaller[18], Maik Thomas[4], Guocheng Wang[15], Yuanwei Wu[22], CanCan Xu[12,17], Xueqing Xu[10], Xinyu Yang[22], Xin Zhao[22], and Zhijin Zhou[9]

[1]Wrocław University of Environmental and Life Sciences, Institute of Geodesy and Geoinformatics, Wrocław, Poland, [2]Centrum Badań Kosmicznych Polskiej Akademii Nauk, Warsaw, Poland, [3]Faculty of Civil Engineering, Warsaw University of Technology, 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, CA, USA, [8]PosiTim UG at ESA/ESOC, Darmstadt, Germany, [9]National Key Laboratory of Science and Technology on Aerospace Flight Dynamic, Beijing Aerospace Control Center, Beijing, China, [10]CAS Key Laboratory of Planetary Sciences, Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai, China, [11]ETH Zurich, Institute of Geodesy and Photogrammetry, Zurich, Switzerland, [12]School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing, China, [13]AGH University of Krakow, Krakow, Poland, [14]Xi'an Surveying and Mapping Center, Xi'an, China, [15]Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan, China, [16]LSE Space GmbH at ESA/ESOC, Darmstadt, Germany, [17]Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai, China, [18]Federal Agency for Cartography and Geodesy BKG, Frankfurt am Main, Germany, [19]Technische Universität Berlin, Institute for Geodesy and Geoinformation Science, Berlin, Germany, [20]ESA/ESOC, Darmstadt, Germany, [21]Shandong University of Technology, Shandong, China, [22]National Time Service Center of Chinese Academy of Sciences, Shaanxi, China

Abstract Growing interest in Earth Orientation Parameters (EOP) resulted in various approaches to the EOP prediction algorithms, as well as in the exploitation of distinct input data, including the observed EOP values from various operational data centers and modeled effective angular momentum functions. Considering these developments and recently emerged new methodologies, the Second Earth Orientation Parameters Prediction Comparison Campaign (2nd EOP PCC) was pursued in 2021–2022. The campaign was led by Centrum Badań Kosmicznych Polskiej Akademii Nauk in cooperation with Deutsches GeoForschungsZentrum and under the auspices of the International Earth Rotation and Reference Systems Service. This paper provides the analysis and evaluation of the polar motion predictions submitted during the 2nd EOP PCC with the prediction horizons between 10 and 30 days. Our analysis shows that predictions are highly reliable with only a few occasional discrepancies identified in the submitted files. We demonstrate the accuracy of EOP predictions by (a) calculating the mean absolute error relative to polar motion observations from September 2021 through December 2022 and (b) assessing the stability of the predictions in time. The analysis shows unequal results for the *x* and *y* components of polar motion (PMx and PMy, respectively). Predictions of PMy are usually more accurate and have a smaller spread across all submitted files when compared to PMx. We present an analysis of similarity between the participants to indicate what methods and input data give comparable output. We also prepared the ranking of prediction methods for polar motion summarizing the achievements of the campaign.

Plain Language Summary Polar motion consists of two time-variable angles that characterize the orientation of the Earth's rotational axis with respect to a terrestrial reference frame attached to the surface of the solid Earth. It can be measured by space geodetic techniques, like Global Navigation Satellite Systems or Very Long Baseline Interferometry (VLBI). However, the final VLBI solutions used by geodetic processing centers to provide the values of polar motion have a latency of around 1 month. Therefore, predicted values are necessary for

Sonia Guessoum, Songtao Han, Robert Heinkelmann, ChengLi Huang, Christopher Irrgang, Jacek Kudrys, Jia Li, Marcin Ligas, Lintao Liu, Weitao Lu, Volker Mayer, Wei Miao, Maciej Michalczak, Sadegh Modiri, Michiel Otten, Todd Ratcliff, Shrishail Raut, Jan Saynisch-Wagner, Matthias Schartner, Erik Schoenemann, Harald Schuh, M. Kiani Shahvandi, Benedikt Soja, Xiaoqing Su, Daniela Thaller, Maik Thomas, Guocheng Wang, Yuanwei Wu, CanCan Xu, Xueqing Xu, Xinyu Yang, Xin Zhao, Zhijin Zhou
Formal analysis: Tomasz Kur, Justyna Śliwińska-Bronowicz, Malgorzata Wińska
Investigation: Tomasz Kur, Justyna Śliwińska-Bronowicz, Malgorzata Wińska
Methodology: Tomasz Kur, Justyna Śliwińska-Bronowicz, Malgorzata Wińska, Jolanta Nastula, Aleksander Partyka
Resources: Henryk Dobslaw, Santiago Belda, Christian Bizouard, Dale Boggs, Sara Bruni, Lue Chen, Mike Chin, Sujata Dhar, Robert Dill, PengShuo Duan, Jose M. Ferrandiz, Junyang Gou, Richard Gross, Sonia Guessoum, Songtao Han, Robert Heinkelmann, ChengLi Huang, Christopher Irrgang, Jacek Kudrys, Jia Li, Marcin Ligas, Lintao Liu, Weitao Lu, Volker Mayer, Wei Miao, Maciej Michalczak, Sadegh Modiri, Michiel Otten, Todd Ratcliff, Shrishail Raut, Jan Saynisch-Wagner, Matthias Schartner, Erik Schoenemann, Harald Schuh, M. Kiani Shahvandi, Benedikt Soja, Xiaoqing Su, Daniela Thaller, Maik Thomas, Guocheng Wang, Yuanwei Wu, CanCan Xu, Xueqing Xu, Xinyu Yang, Xin Zhao, Zhijin Zhou
Software: Tomasz Kur
Supervision: Henryk Dobslaw, Jolanta Nastula
Validation: Tomasz Kur, Justyna Śliwińska-Bronowicz, Malgorzata Wińska
Visualization: Tomasz Kur
Writing – original draft: Tomasz Kur
Writing – review & editing: Justyna Śliwińska-Bronowicz, Malgorzata Wińska, Henryk Dobslaw, Jolanta Nastula, Aleksander Partyka, Santiago Belda, Christian Bizouard, Dale Boggs, Sara Bruni, Lue Chen, Mike Chin, Sujata Dhar, Robert Dill, PengShuo Duan, Jose M. Ferrandiz, Junyang Gou, Richard Gross, Sonia Guessoum, Songtao Han, Robert Heinkelmann, ChengLi Huang, Christopher Irrgang, Jacek Kudrys, Jia Li, Marcin Ligas, Lintao Liu, Weitao Lu, Volker Mayer, Wei Miao, Maciej Michalczak, Sadegh Modiri, Michiel Otten, Todd Ratcliff, Shrishail Raut, Jan Saynisch-Wagner, Matthias Schartner, Erik Schoenemann,

operational applications such as spacecraft navigation. To assess current methods of predicting polar motion time series, the Second EOP Prediction Comparison Campaign was pursued under the auspices of the International Earth Rotation and Reference Systems Service. The campaign aimed to test current achievements in polar motion predictions obtained with a variety of computational methods (including least squares, machine learning, and a Kalman filter) under realistic conditions. By evaluating the results of the campaign, we show that some of the prediction methods utilized do indeed reduce prediction errors and enhance prediction accuracy by using geophysical information from the fluid Earth's layers: Atmosphere, oceans, and terrestrial hydrosphere.

1. Introduction

Earth Orientation Parameters (EOP) comprise polar motion, differences between universal time and coordinated universal time (UT1-UTC), its time-derivative Length-of-Day (LOD) change, and nutation. EOP are routinely determined by advanced space geodetic techniques, that is, Global Navigation Satellite System (GNSS, Byram & Hackman, 2012), Very Long Baseline Interferometry (VLBI, Schuh & Böhm, 2014), and Satellite Laser Ranging (SLR, Coulot et al., 2010). EOP are computed sequentially, using the next available solutions from each space geodetic technique. All geodetic techniques provide data with different accuracies, stabilities, and latencies. Due to the delay caused by the time needed to collect and process input data from various sources, the final estimated EOP parameters cannot be made available in real-time. Instead, short-term predictions are utilized for real-time applications like tracking and navigating interplanetary spacecrafts. EOP predictions are computed by several prediction centers for example, the U.S. Naval Observatory (USNO, Luzum et al., 2001), Jet Propulsion Laboratory (JPL) of the National Aeronautics and Space Administration (NASA, Gross et al., 1998), and the European Space Agency (ESA, Bruni et al., 2021; Kehm et al., 2023).

The International Earth Rotation and Reference Systems Service (IERS) regularly provides EOP in the form of time series derived from a combination of various measurement techniques (Bizouard & Gambis, 2009). The Earth Orientation Center, established by the IERS and hosted by the Paris Observatory, is responsible for operational EOP monitoring and the delivery of daily, monthly, and long-term EOP data, the release of time dissemination (UT1-UTC), and leap second announcements (Bizouard et al., 2019; Gambis, 2004; Gambis & Luzum, 2011). The IERS 14 C04 series is aligned to conventional reference frames (International Terrestrial Reference Frame ITRF 2014 and International Celestial Reference Frame ICRF2) and has been the de-facto standard for EOP solutions for many years (Bizouard et al., 2019). The newest EOP IERS 20 C04 series, available since 14 February 2023, is aligned to the ITRF 2020 (Altamimi et al., 2023, IERS Message No. 471 distributed by the IERS Central Bureau).

The Earth Orientation Department at the USNO is hosting the IERS Rapid Service/Prediction Center, which operates as a product center within the IERS. Its main task is to deliver EOP by gathering relevant geodetic observations and conducting data analyses to determine and predict the continuously changing orientation of the terrestrial reference frame within the quasi-inertial celestial reference frame. The service is primarily utilized by real-time users requiring high-quality EOP information within 30 days, which is the nominal latency of the IERS final series provided by the Paris Observatory. EOP is essential in applications such as precise navigation and positioning, particularly for real-time purposes. The service provides daily Bulletin A EOP solutions and other Earth orientation results. Daily EOP solutions (including finals.daily, finals2000A.daily, and gpsrapid.daily) are usually uploaded at 18:00 UTC each day. Bulletin A EOP data is available every Thursday by 20:00 UTC.

Interest from various scientific institutions concerned with EOP forecasting led to the initiation of the first Earth Orientation Parameters Prediction Comparison Campaign (1st EOP PCC), which was carried out between 2005 and 2008 to compare various prediction methodologies and to identify the most promising strategies for EOP predictions (Kalarus et al., 2010). The campaign was prepared jointly by Vienna University of Technology and Centrum Badań Kosmicznych Polskiej Akademii Nauk (CBK PAN). One of the main conclusions from the 1st EOP PCC was that no single prediction technique worked best for all EOP components and every prediction horizon. It has also been demonstrated that incorporating data on Atmospheric Angular Momentum (AAM) into the forecasting process generally enhances the accuracy of EOP predictions.

Since that time, scientists have been constantly seeking new forecasting methods to ensure the best possible EOP prediction accuracy. This growing interest eventually resulted in the establishment of the 2nd EOP PCC by a

Harald Schuh, M. Kiani Shahvandi, Benedikt Soja, Xiaoqing Su, Daniela Thaller, Maik Thomas, Guocheng Wang, Yuanwei Wu, CanCan Xu, Xueqing Xu, Xinyu Yang, Xin Zhao, Zhijin Zhou

working group of IERS. The 2nd EOP PCC was organized in 2021 to specifically assess machine learning (ML) approaches for EOP prediction that have matured substantially in the most recent past. In addition, the knowledge about the role of geophysical fluid layers in exciting Earth rotation variations relevant to short-term predictions has improved after the first campaign (e.g., Afroosa et al., 2021; Bizouard & Seoane, 2010; Dill et al., 2019; Harker et al., 2021; Kiani Shahvandi, Gou, et al., 2022). A growing number of teams have begun to utilize Effective Angular Momentum (EAM) data for forecasting EOP, which is likely one of the factors contributing to the improvement in the accuracy of EOP forecasting in recent years. The second campaign was again run by CBK PAN in close cooperation with Deutsches GeoForschungsZentrum (GFZ). The operational part of the 2nd EOP PCC ended on 31 December 2022. The EOP PCC Office established at CBK PAN coordinated participants' efforts and compared their predictions using well-defined rules. The campaign is an international initiative with 23 registered institutions from 8 different countries, involving a total of over 50 people who regularly delivered predictions based on sometimes very different methods, each assigned with an individual ID number (Kur et al., 2022; Śliwińska et al., 2022; Śliwińska-Bronowicz et al., 2024). Forecasts of all EOP were submitted by participants every Wednesday, and predictions were evaluated as soon as the geodetic final EOP observations of the forecasted period eventually became available. A more comprehensive description of the EOP PCC Office's preparations for the campaign and the activities undertaken throughout is provided in Śliwińska et al. (2022). Data from the 2nd EOP PCC are publicly available at GFZ Data Services (Śliwińska et al., 2023).

This paper gives an insight into the polar motion predictions submitted during the 2nd EOP PCC with a focus on forecast horizons at 10 and 30 days. Short descriptions of the different prediction methods, together with evaluation methodology and a summary of submissions, are provided in Section 2. In Section 3, the accuracy of the predictions and performance of various combinations of methods and input data are investigated. In Section 4, we assess the relationship between various prediction approaches. Finally, Section 5 summarizes all the results and provides concluding remarks.

2. Overview of Submissions and Analysis Approaches

2.1. Description of Registered Prediction Methods

During the 2nd EOP PCC, 33 participant IDs were registered for the purposes of predicting polar motion. Essential information about each ID's prediction method is presented in Table 1. The full description of the IDs' prediction methodology given by participant is attached in Appendix A. The most common prediction methods were LS and autoregression methods (LS + AR) as well as ML, which should be understood here in the widest possible sense. These prediction methods were based typically on IERS 14 C04. Additionally, 19 IDs use EAM analysis and forecast data as provided by GFZ (Dill et al., 2019, 2022; Dobslaw & Dill, 2018). Effective Angular Momentum consists of AAM, Oceanic Angular Momentum (OAM), Hydrospheric Angular Momentum (HAM), and SLAM (Sea-Level Angular Momentum). However, some IDs use only selected components of EAM (e.g., AAM and OAM) as input to their prediction algorithm.

The relevance of different excitation mechanisms for polar motion prediction was recently assessed by Kiani Shahvandi, Gou, et al. (2022), Kiani Shahvandi, Gou, et al. (2022). Those authors exploited Neural Ordinary Differential Equations (ODE) Differential Learning to predict polar motion and demonstrated that using OAM, HAM, and SLAM improves the prediction accuracy more than twice compared to the case of not using EAM. Atmospheric Angular Momentum contributes the least to the prediction and reduces the forecasting error by about 20%. The lowest prediction errors in Kiani Shahvandi, Gou, et al. (2022), Kiani Shahvandi, Gou, et al. (2022) were achieved with the use of AAM + OAM + HAM and AAM + OAM + HAM + SLAM—mean sum squared (mean sum squared (MSS)) equal to 1.19 and 1.21 mas^2, respectively compared to a reference solution without EAM with MSS = 4.81 mas^2. The research suggests that the optimal combinations of EAM for linear models consist of AAM + OAM + HAM or AAM + OAM + HAM + SLAM. Notably, these combinations are closely comparable. Atmospheric Angular Momentum alone does not contribute substantially to performance improvement, underscoring the limited utility of AAM as a standalone feature for predicting polar motion. HAM and SLAM exhibit lower importance compared to OAM. Overall, utilizing EAM enhances prediction performance compared to models without EAM.

Table 1
Summary of Prediction Methods and Input Data Used for Polar Motion Prediction

Group members	ID	Prediction method	Input data	References
Xueqing Xu	100	LS + AR	IERS 14 C04; IERS finals.daily	X. Xu et al. (2022), X. Q. Xu et al. (2012), X. Xu and Zhou (2015)
	112	LS + convolution	IERS 14 C04; IERS finals.daily; GFZ AAM, OAM, HAM, SLAM	
Yuanwei Wu, Xin Zhao, Xinyu Yang	101	LS + AR with piecewise parameter optimization	IERS 14 C04; IERS daily; GFZ AAM, OAM, HAM, SLAM (data + predictions)	Wu et al. (2022)
Christian Bizouard	102	LS + AR	IERS 14 C04	X
Richard Gross, Dale Boggs, Mike Chin, Todd Ratcliff	104	Kalman filter	EOP data from IGS, ILRS, JPL, GSFC; GFZ AAM and OAM forecasts	Freedman et al. (1994), Gross et al. (1998)
Robert Dill, Henryk Dobslaw, Maik Thomas, Jan Saynisch-Wagner, Christopher Irrgang	105	LS + AR	IERS 14 C04; IERS finals.daily; GFZ AAM, OAM, HAM (data + predictions)	Dill et al. (2019)
	136	LS + AR	IERS 14 C04; IERS finals.daily; GFZ AAM, OAM, HAM (data + predictions)	Dill et al. (2019)
Weitao Lu, Lue Chen, Zhijin Zhou, Songtao Han	107	LS + AR	IERS 14 C04; IERS finals.daily	Chen et al. (2014)
	108	LS + multiscale AR	IERS 14 C04; GFZ AAM	X
	137	Artificial neural network + AR	IERS 14 C04	X
Erik Schoenemann, Sara Bruni, Michiel Otten, Volker Mayer	116	LS + AR	ESA GNSS rapid; ESA GNSS final; ESA SLR; ESA DORIS; BKG VLBI intensive sessions; DGFI VLBI rapid turnaround sessions; GFZ AAM, OAM, HAM (data + predictions)	Bruni et al. (2021), Kehm et al. (2023)
Sadegh Modiri, Daniela Thaller, Shrishail Raut, Sujata Dhar, Robert Heinkelmann, Harald Schuh, Santiago Belda, Sonia Guessoum, Jose M. Ferrandiz	117	singular spectrum analysis (SSA) + copula	IERS 14 C04; IERS Bulletin A	Belda et al. (2016, 2018), Modiri (2021), Modiri et al. (2018, 2020)
Lintao Liu, Xiaoqing Su, Guocheng Wang	121	Normal time-frequency transform (NTFT)	IERS 14 C04	X
Jia Li	122	Weighted least squares + autoregressive integrated moving average (WLS + ARIMA)	IERS 14 C04; IERS finals.daily	X
Matthias Schartner, Mostafa Kiani Shahvandi, Junyang Gou, Benedikt Soja	128	First order neural ordinary differential equations	JPL EOP2 series; GFZ AAM (data + predictions)	Gou et al. (2023), Kiani Shahvandi, Gou, et al. (2022), Kiani Shahvandi, Schartner, and Soja (2022), Kiani Shahvandi and Soja (2022)
	131		IERS finals.daily; GFZ AAM, OAM, HAM, SLAM (data + predictions)	
	132		JPL EOP2 series; GFZ AAM, OAM, HAM, SLAM (data + predictions)	
	133		SYRTE EOP series; GFZ AAM, OAM, HAM, SLAM (data + predictions)	
	150	First-order neural ordinary differential equations with residual modeling	IERS finals.daily; GFZ AAM, OAM, HAM, SLAM (data + predictions)	Gou et al. (2023), Kiani Shahvandi, Gou, et al. (2022), Kiani Shahvandi, Schartner, and Soja (2022), Kiani Shahvandi and Soja (2022)
	151			
	152			
	153	Physics-constrained neural networks	IERS finals.daily; GFZ AAM (data + predictions)	Gou et al. (2023), Kiani Shahvandi, Dill, Dobslaw, Kehm, et al. (2023), Kiani Shahvandi, Dill, Dobslaw, Mishra, and Soja (2023), Kiani Shahvandi, Schartner, and Soja (2022)
ChengLi Huang, PengShuo Duan, CanCan Xu	138	Deconvolution + LS, AR, fast Fourier transform + convolution	IERS finals.daily; GFZ AAM, OAM, HAM, SLAM (data + predictions)	Dill et al. (2019), Dobslaw and Dill (2018)

Table 1
Continued

Group members	ID	Prediction method	Input data	References
Maciej Michalczak, Marcin Ligas, Jacek Kudrys	141	LS + ordinary kriging prediction	IERS 14 C04; IERS finals.daily	Michalczak and Ligas (2021, 2022)
	156	LS + ARIMA	IERS 14 C04; IERS finals.daily	Michalczak et al. (2022)
	157	LS + vector autoregressive prediction	IERS 14 C04; IERS finals.daily	X
Anonymous	103	adaptive polyharmonic models with infrequent components	IERS	X
	113	LS + AR	IERS 14 C04	
	114	Least squares collocation	IERS 14 C04	
	115	Neural networks	IERS 14 C04	
	118	Local approximation technique (LA) + LS	IERS	
	123	Undisclosed	GFZ AAM, OAM	
	135	LS + AR	IERS	

Note. Full descriptions provided by campaign participants can be found in Appendix A.

2.2. Statistics on Polar Motion Predictions Submission

During 70 weeks of the campaign duration, the EOP PCC Office received 3,414 valid files with predictions for both components of polar motion. Figure 1a shows the number of files submitted by each campaign participant during the 2nd EOP PCC. Only two participants (with ID 100 and 128) provided their predictions every week throughout the campaign duration, while the last registered method (ID 157) only delivered 7 predictions. Figure 1b shows the number of prediction files with PMx and PMy submitted every submission day (Wednesday). We can distinguish three periods with visibly varying numbers of submitted files: the initial stage from September 2021 to April 2022, where the submission numbers were slowly increasing, a middle period from April 2022 to September 2022, with the highest number of submitted files, and a slight decline in participation toward the end of the campaign from September 2022 to December 2022. Figure 1c demonstrates the continuity of submission for all IDs predicting polar motion. The EOP PCC Office received almost 74% of all potential files, with a mean of 24 ± 4 files per submission day.

In Table 2, we report the number of files submitted with the maximum prediction forecast lengths equal to 10, 30, 90, 120, 180, and 365 days. The most popular forecast horizons were 90 and 365 days (1,052 and 964 files, respectively). That makes short-term predictions from 10 to 30 days the least chosen length of the horizon, as only 986 files were submitted for both. During the 1st EOP PCC, which lasted over 2 years, the campaign office received 2240 ultra short-term predictions (<10 days, 21 predictions were rejected), 1,840 short-term predictions (<30 days, 17 predictions were rejected), and 1,350 for medium-term (<500 days, 60 predictions were rejected) from 13 participants for both PMx and PMy. The relation of the number of files to the length of prediction is therefore reversed between the campaigns.

2.3. Clustering of IDs

According to the large number of IDs predicting polar motion, we manually divided campaign participants' approaches into groups based on methodology and the use of EAM input data. We do not consider geodetic input data explicitly because it turned out this was too uniform among the participants to be a distinguishing feature. We created four groups: (a) *LS + AR* contains prediction methods based on the least squares with autoregression only; (b) *LS + AR + EAM* consists of IDs with LS and autoregression, but with the use of EAM as an additional data set; (c) *ML + EAM* contains ML-based approaches where most of the participants also use EAM (only ID 137 do not exploit it, but here ML was the key parameter). The final group (d) *Other* is formed from all other IDs, which mostly use methods unique in the campaign, that is, none of the other participants decided to exploit it, for example, the Kalman filter. Complete information on the IDs' assignment to the groups is shown in Table 3. Each group is additionally compared with ID 200, which is the official IERS prediction from the finals.2000A.daily file

10.1029/2023EA003278

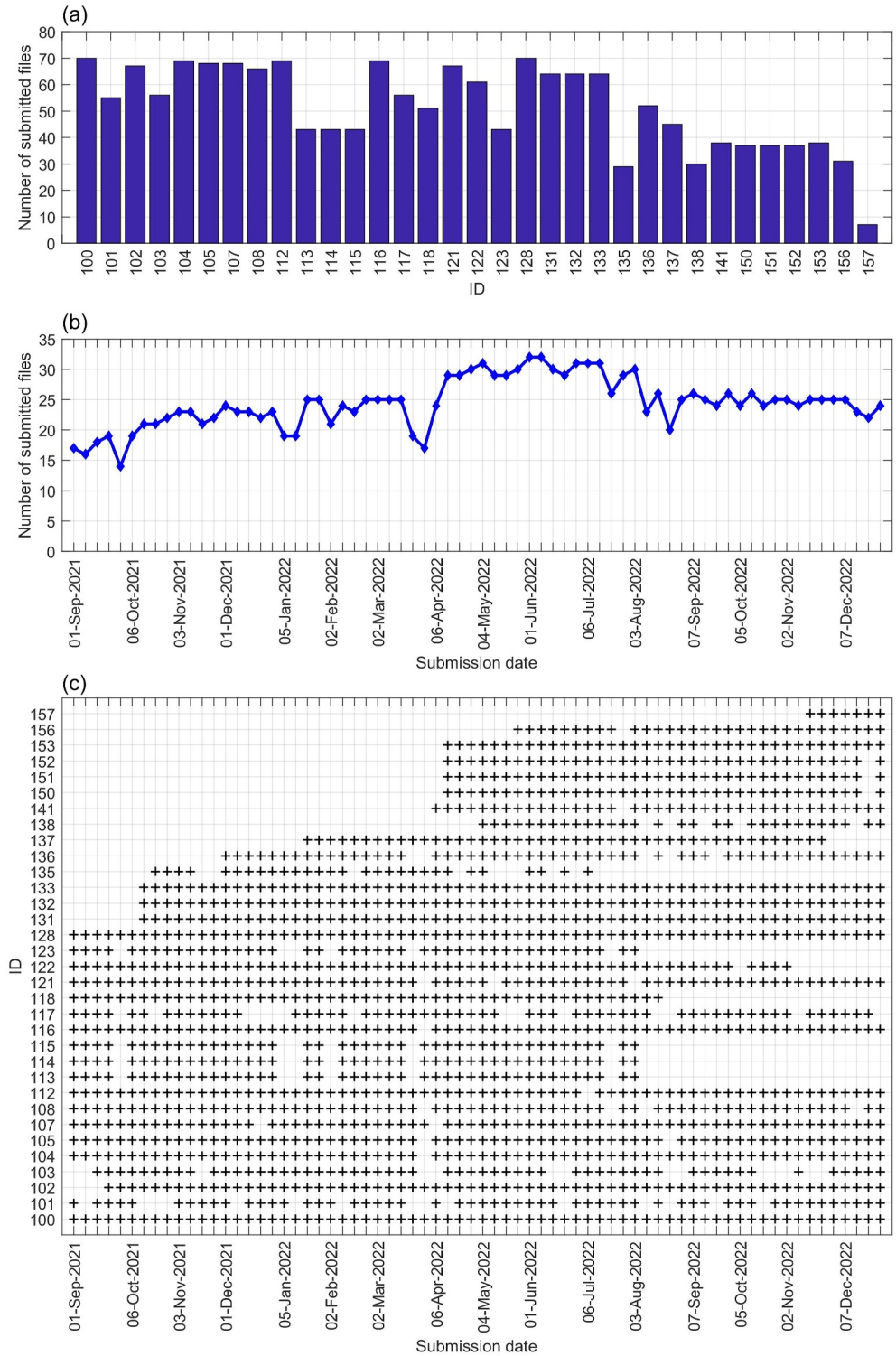

Figure 1. (a) Number of files submitted by campaign participants between 1 September 2021 and 28 December 2022. (b) Number of submitted files from all IDs containing PMx and PMy predictions between 1 September 2021 and 28 December 2022. (c) PMx and PMy submissions made by participants every week during the campaign.

Table 2
Number Of Files According to the Prediction Horizon

Prediction horizon	10 days	30 days	90 days	120 days	180 days	365 days
Number of files	536	450	1,052	138	274	964

saved by the EOP PCC Office upon the deadline set for the weekly submissions (i.e., on Wednesdays at 20:00 UTC) (https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html—accessed 1.05.2023).

2.4. General Evaluation Methodology

The mean absolute error (MAE, Kalarus et al., 2010; Kur et al., 2022; Śliwińska et al., 2022) was used as an objective parameter to evaluate predictions with respect to the IERS 14 C04 series:

$$MAE_i = \frac{1}{n_p}\sum_{j=1}^{n_p}\left|x_i^{\text{obs}} - x_{i,j}^{\text{pred}}\right|, i = 1,2,...,I \tag{1}$$

where n_p is the number of predictions related to the same ID, x_i^{obs} is the observed EOP data for the i th day, $x_{i,j}^{\text{pred}}$ is the value for the i th day of the j th prediction, and I is the forecast horizon (i.e., the time period predicted into the future). The evaluation of the predictions is performed for forecast horizons equal to 10 and 30 days with differentiation into groups as defined in Section 2.3.

A few predictions submitted to the campaign are affected by gross errors, which cannot be entirely avoided in an operational setting. Since gross errors adversely affect the objective assessment of the accuracy of a given forecasting method, we attempted to identify and remove them. This was achieved by incorporating a two-step algorithm to reject prediction outliers based on statistical values computed with the same rules for all participants. The first step of data selection, called σ-criterion, includes computation of standard deviation S_j for $x_i^{\text{obs}} - x_{i,j}^{\text{pred}}$ for all individual predictions submitted. The values obtained are compared with the standard deviation S_{total} for all forecasts. All predictions with $S_j > S_{\text{total}}$ are rejected. This method is a simple and efficient way to remove predictions with highly inaccurate values introduced during data preparation considering the general accuracy of all predictions. However, each ID is characterized with different accuracy during the campaign. Thus, in the second step of data selection, we also evaluated predictions by using thresholds computed individually for each ID by determining a β-parameter:

$$\beta_j = \sum_{i=1}^{I}\left(\alpha \cdot MDAE_i - \left|x_i^{\text{obs}} - x_{i,j}^{\text{pred}}\right|\right), \tag{2}$$

$$MDAE_i = \text{median}\left(\left|x_i^{\text{obs}} - x_{i,1}^{\text{pred}}\right|, \left|x_i^{\text{obs}} - x_{i,2}^{\text{pred}}\right|, \ldots, \left|x_i^{\text{obs}} - x_{i,n_p-1}^{\text{pred}}\right|, \left|x_i^{\text{obs}} - x_{i,n_p}^{\text{pred}}\right|\right). \tag{3}$$

If $\beta_j < 0$, the prediction is not considered further. The α-value is deduced empirically to preserve a representative set of data (Kalarus et al., 2010). We applied the β-parameter with an α-value equal to 4. Table 4 summarizes the number of rejected files at each step of the data selection process. We can notice that (a) $LS + AR$ and (b)

Table 3
ID Assignment to Groups According to Main Prediction Method and Effective Angular Momentum Usage

Group	IDs
(1) *LS + AR*	100; 107; 113; 114; 118; 122; 135; 156; 157
(2) *LS + AR + EAM*	101; 105; 108; 112; 116; 136; 138
(3) *ML + EAM*	115; 123[a]; 128; 131; 132; 133; 137; 150; 151; 152; 153
(4) *Other*	102; 103; 104; 117; 121; 141

[a]Method description was not provided, assignment made by the EOP PCC Office based on evaluation results (see details in Sections 3 and 4).

Table 4
Summary of the Number of Files Rejected in Each Step of Data Selection

		(1) *LS + AR*	(2) *LS + AR + EAM*	(3) *ML + EAM*	(4) *Other*	Total
PMx						
10 days	σ-criterion	0	0	3	0	3
	β-parameter	4	2	18	2	26
30 days	σ-criterion	2	0	1	8	11
	β-parameter	5	4	2	1	12
PMy						
10 days	σ-criterion	1	0	3	1	5
	β-parameter	3	2	11	0	16
30 days	σ-criterion	1	0	1	15	17
	β-parameter	1	1	1	2	5

LS + AR + EAM have the lowest number of rejected predictions, while (c) *ML + EAM* and (d) *Other* have relatively more predictions with gross errors. In the case of *ML + EAM*, most IDs delivered predictions up to 10 days. These groups have the highest number of identified instances. In the case of (d) Other, most erroneous predictions were detected for the 30-day forecast horizon with the σ-criterion. The detailed description of the data selection step is given in the Supporting Information S1.

3. Predictions Quality Assessment

Figure 2 shows the boxplots of differences between IERS 14 C04 and all predictions fulfilling the accuracy requirements for 10- and 30-day prediction horizons. The plots indicate that all predictions with huge errors were successfully rejected. Boxplots are ordered according to the clusters presented in Table 3. For both prediction horizons considered, PMy predictions are more consistent than PMx forecasts across all participants. The range of differences for PMx is also visibly higher than for PMy. Only a few participants obtained results that were better or comparable to those processed at the USNO and published by IERS (ID 200) for PMx for 10- and 30-day horizons. For PMy, this number is substantially higher, indicating room for improvement in the USNO approach, particularly for this component. Data points beyond the whiskers are outliers computed as 3 times interquartile range (IQR) obtained for data after two-step selection. From Figure 2, we can observe that results for some IDs are notably asymmetric, for example, for IDs 131, 132, and 133 for PMx for 10 days, where outliers appear mostly below the minimum value. This might suggest that forecasts are severely underestimated for individual submission days. On the contrary, outliers for ID 104 for both parameters and forecast horizons are almost perfectly symmetric. This indicates a lack of long-term stability of differences between reference and forecast. For most participants, however, values of IQR are relatively small, as the mean difference between the 3rd and 1st quartile for PMx (10 days) is equal to 2.7 mas, while for PMy (10 days) it is 1.8 mas. For the 30-day forecast horizon, these values are equal to 6.7 and 4.3 mas, respectively. On the other hand, minimum and maximum values of differences can sometimes exceed ±10 mas and ±20 mas for 10- and 30-day horizons, respectively (e.g., for IDs 121 and 137).

We further analyzed the quality of the predictions by reporting MAE values for 10 and 30 days into the future in reference to IERS 14 C04. Results are grouped into four clusters as introduced in Section 2.3. In addition, we always show the mean value computed for each group, the mean value of all predictions from all groups, and the results for ID 200 to also allow for cross-comparisons. In Figure 3, we compare the values of MAE with respect to the mean value for all IDs and the mean value for each cluster for a 10-day prediction horizon. Day 0 in each plot represents submission day (the last day for which observational data is available). Higher discrepancies between IDs are obtained for PMx, while MAE for PMy is more consistent across participants, as already recognized in the previous section. IDs 121 and 135 visibly suffer from poor data preparation as the temporal variation of the MAE is substantially biased. For ID 137, MAE for day zero is close to the rest of the participants in the group, but it seems that predictions from the 2nd day onwards are notably different from predictions of other IDs, which suggests some kind of systematic error in the case of ID 137. For both PMx and PMy in the groups (a) *LS + AR*

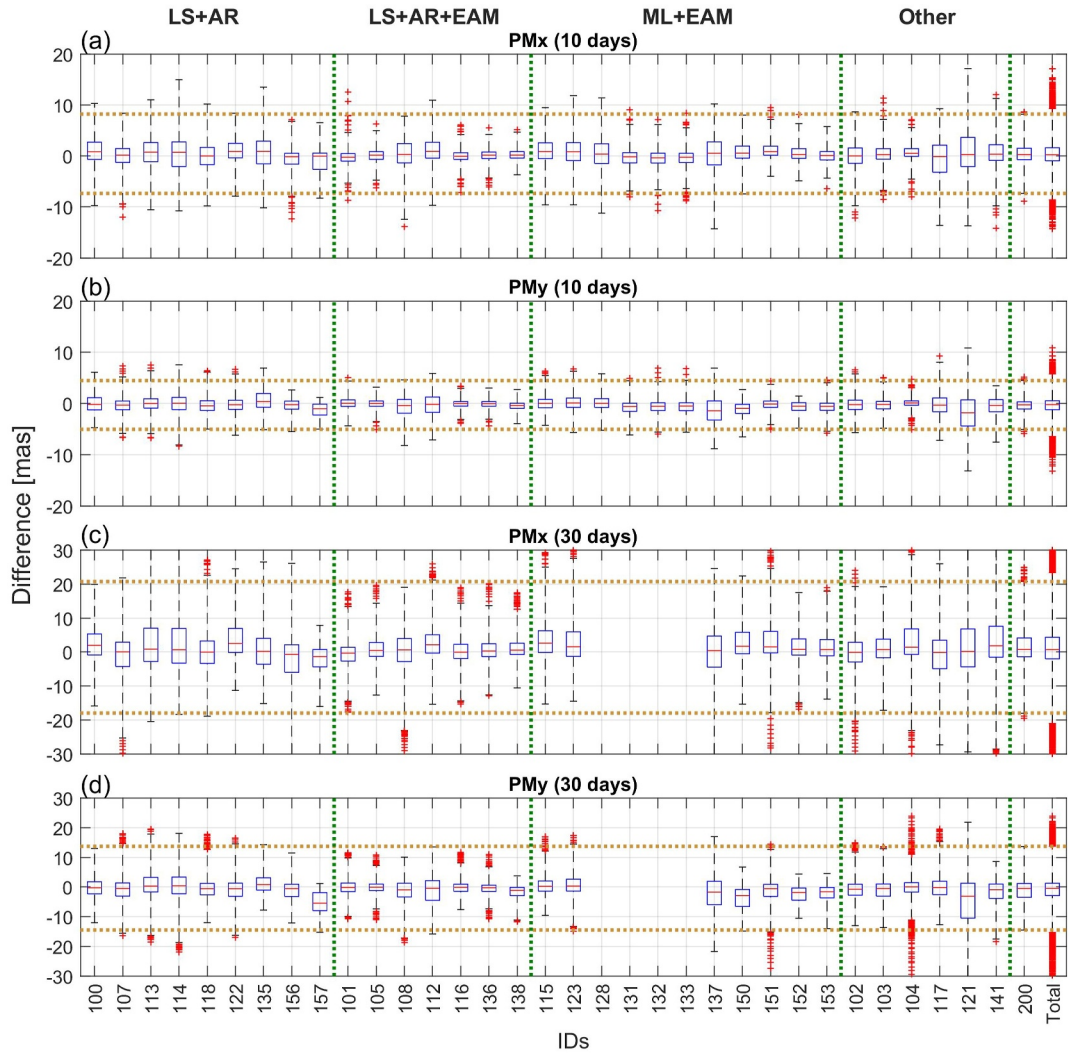

Figure 2. Boxplot for differences between IERS 14 C04 and predictions for each ID: (a) PMx with the 10-day horizon, (b) PMy with the 10-day horizon, (c) PMx with the 30-day horizon, (d) PMy with the 30-day horizon. The vertical dotted line represents each cluster. The horizontal dotted line represents the maximum and minimum values of prediction released by IERS. Data points beyond the whiskers are outliers computed as 3 times IQR.

and (d) *Other*, most IDs do not perform better than the predictions provided by IERS. On the contrary, groups (b) *LS + AR + EAM* and (c) *ML + EAM*, which use EAM, mostly provide better results than the IERS. However, group (b) *LS + AR + EAM* is also characterized by the largest difference between mean values for the group (black thick line in Figure 3c) and for all (magenta thick line in Figure 3c) for PMx caused by two IDs performing worse than ID 200. It suggests that participants may not always incorporate EAM forecasts in an optimal manner. Groups (b) *LS + AR + EAM* and (c) *ML + EAM* consist of IDs that exploit EAM predictions until day 6, which becomes somewhat apparent for PMx between 7th and 10th day when MAE grows more rapidly than for initial days (see the "elbow" characteristic in the course of the line, which is more visible for (b) *LS + AR + EAM* group but also for ID 104 using Kalman filter with EAM). PMy is characterized by smaller values of MAE on the 10th day than PMx for all the clusters. Comparison with IERS predictions for the 10-day forecast horizon shows that 13 IDs for PMx and 13 IDs for PMy (out of 33 in total) have a lower MAE than ID 200. This clearly indicates room for improvements within the polar motion prediction algorithm implemented at the USNO.

Figure 4 demonstrates MAE values for 30-day predictions. IDs with shorter prediction horizons are not included. For most of the predictions MAE linearly increases, but in group (d) *Other* we can see an inflation of error for IDs 104 and 141 for PMx. Generally, group (d) *Other* is characterized by rather more dispersed results for 30 days

10.1029/2023EA003278

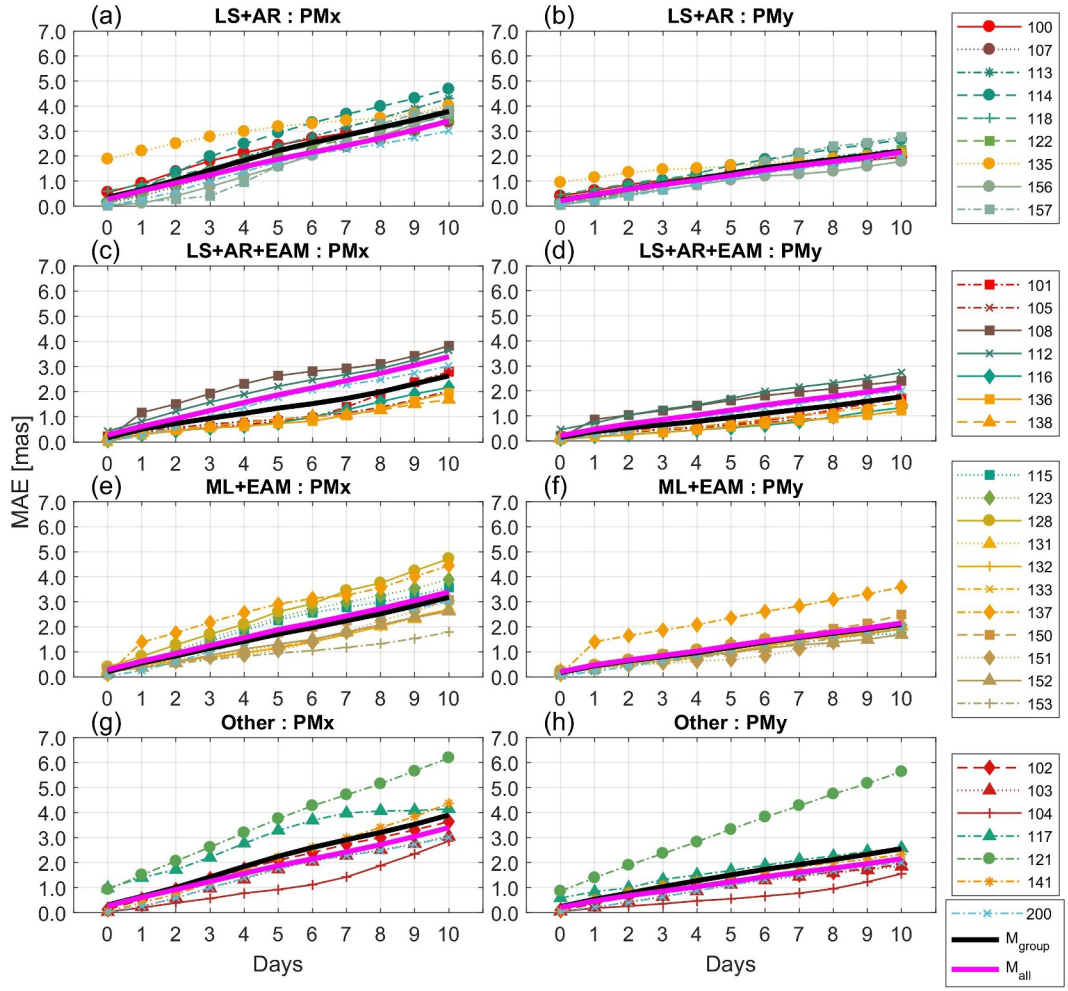

Figure 3. Mean absolute error for PMx and PMy computed for 10 days into the future relative to IERS C04: (a, b) for group (1) $LS + AR$, (c, d) for group (2) $LS + AR + EAM$, (e, f) for group (3) $ML + EAM$, (g, h) for group (4) *Other*. M_{group} means the mean for each cluster and M_{all} represents the mean for all predictions.

than it was found for MAE for 10 days. Groups (a) $LS + AR$, (b) $LS + AR + EAM$, and (c) $ML + EAM$ are rather stable on consecutive days during a forecast. Generally, most predictions for PMy are more accurate than forecasts from ID 200, but for PMx there are only a few that perform better than IERS and most of them are part of the group (b) $LS + AR + EAM$. Comparison with IERS predictions for the 30-day forecast horizon shows that 9 IDs for PMx and 20 IDs for PMy out of 29 have lower MAE than ID 200. Comparing the number of IDs with mean values of MAE lower than the mean for IERS, as reported in Table 5, we can clearly see the advantage of using EAM functions in predictions. The exceptions are predictions for PMx for 30 days computed with ML (group (c) $ML + EAM$), where only two IDs are better than the official IERS product.

To summarize MAE evaluation, we show minimum, maximum, and mean values of MAE for both PMx and PMy for the 1st, 10th and 30th day of prediction (MAE[1], MAE[10] and MAE[30], respectively) in Table 6. Additionally, values for ID 200 are reported to compare campaign results with the IERS predictions. It can be noticed that most of the IDs are repeated between PMx and PMy. Usually, IDs from the (b) LS + AR + EAM group have lower MAE values than the IERS predictions. The highest MAE ratio obtained for participant to MAE for IERS prediction is almost 9 for PMx and ID 135 in the case of MAE[1]—2.2 mas (MAE[1] for ID 135) to 0.25 mas (MAE[1] for ID 200). (b) LS + AR + EAM group provides the lowest mean values of MAE in all cases, performing better than IERS, especially for MAE[10] and MAE[30]. For each group, in most cases, the maximum values of MAE for respective days are higher for PMx than for PMy, while for IERS, this difference is smaller than for campaign participants.

10.1029/2023EA003278

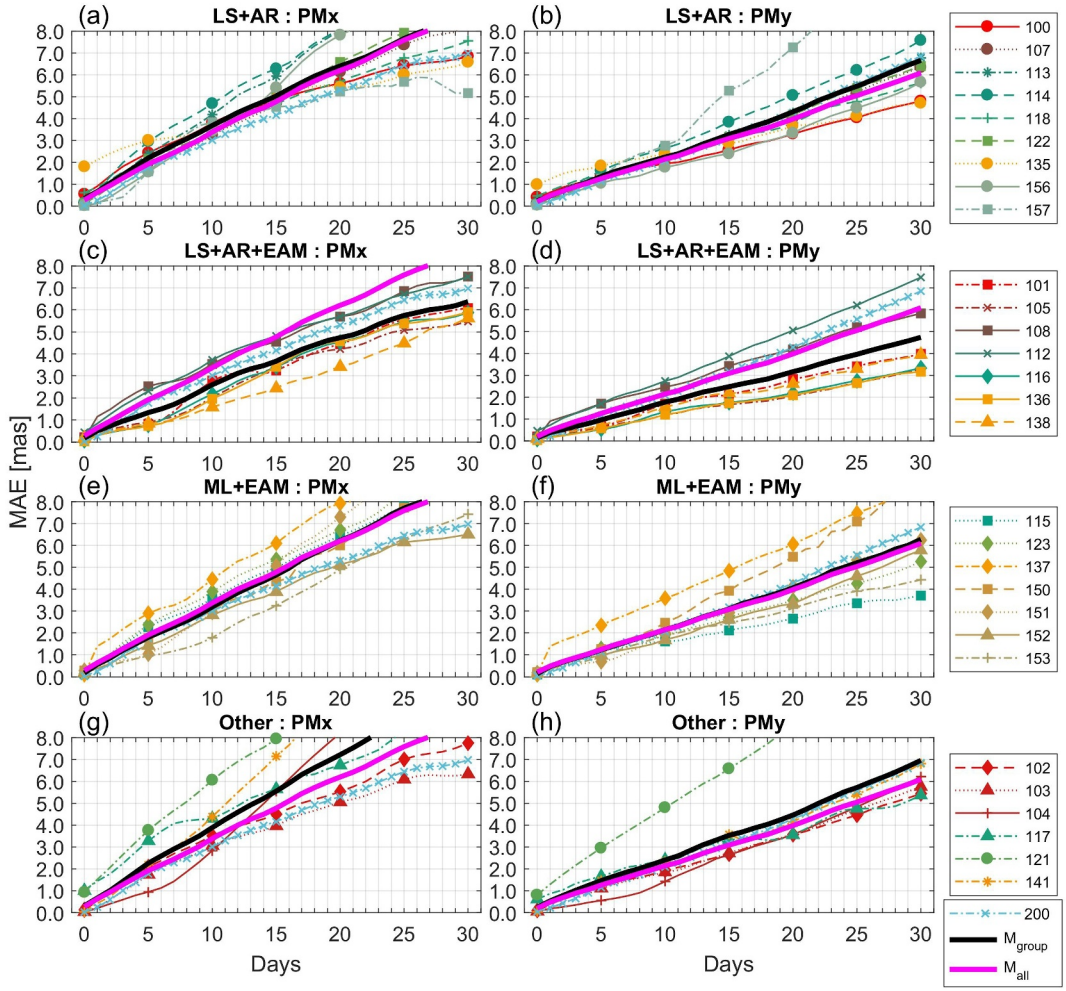

Figure 4. Mean absolute error for PMx and PMy computed for 30 days into the future relative to IERS C04: (a, b) for group (1) *LS + AR*, (c, d) for group (2) *LS + AR + EAM*, (e, f) for group (3) *ML + EAM*, (g, h) for group (4) *Other*. M_{group} indicates the mean for each cluster and M_{all} represents the mean for all predictions.

One rule adopted for the 2nd EOP PCC was the free choice of input data, allowing each participant to decide about the use of EOP from different sources and epochs. Consequently, participants did not always utilize the latest solutions for polar motion. For instance, longer processing times for specific methods could cause delays, preventing timely submission by the ID. These inhomogeneities in the input data (in particular, the date of the last observed value) may partly explain discrepancies between submissions from different IDs that declared the use of the same data for their predictions. Conversely, a crucial aspect of the campaign was identifying the most accurate prediction methodology under realistic conditions. This involved testing not only the algorithm but also the

Table 5
The Number of IDs in Each Group With Mean Mean Absolute Error (MAE) From 10 to 30 Days of Prediction Lower Than the Mean MAE Received for IERS Prediction (ID 200), and Respective Mean MAE Values for ID 200, for Example, 1/9 Means That One ID Out of Nine in the Group Meets the Condition

	Group (1) *LS + AR*	Group (2) *LS + AR + EAM*	Group (3) *ML + EAM*	Group (4) *Other*	Total	Mean MAE for ID 200
PMx (10 days)	0/9	5/7	6/11	2/6	13/33	1.6 mas
PMy (10 days)	1/9	5/7	5/11	2/6	13/33	1.1 mas
PMx (30 days)	1/9	5/7	2/7	1/6	9/29	4.0 mas
PMy (30 days)	6/9	5/7	5/7	4/6	20/29	3.3 mas

Table 6
Statistics (Minimum, Maximum, Mean) of Mean Absolute Error for PMx and PMy in Each Group for the 1st, 10th, and 30th Day of the Prediction

	LS + AR		LS + AR + EAM		ML + EAM		Other			
	ID	MAE [mas]	ID	MAE [mas]	ID	MAE [mas]	ID	MAE [mas]		MAE [mas]
PMx										
MAE for 1st day of forecast horizon MAE[1]										
Min	156	0.12	136	0.29	152	0.36	104	0.20	200 (IERS)	0.25
Max	135	2.20	108	1.17	137	1.39	121	1.51		
Mean	–	0.71	–	0.54	–	0.59	–	0.71		
MAE for 10th day of forecast horizon MAE[10]										
Min	107	3.35	138	1.68	153	1.79	104	2.86	200 (IERS)	3.01
Max	114	4.68	108	3.83	128	4.72	121	6.19		
Mean	–	3.86	–	2.58	–	3.20	–	4.04		
MAE for 30th day of forecast horizon MAE[30]										
Min	157	5.15	105	5.46	152	6.50	103	6.33	200 (IERS)	6.97
Max	156	12.25	108	7.52	151	11.91	141	15.58		
Mean	–	8.76	–	6.28	–	9.00	–	10.96		
PMy										
MAE for 1st day of forecast horizon MAE[1]										
Min	157	0.23	136	0.17	115	0.31	104	0.18	200 (IERS)	0.22
Max	135	1.14	108	0.86	137	1.40	121	1.39		
Mean	–	0.48	–	0.40	–	0.47	–	0.59		
MAE for 10th day of forecast horizon MAE[10]										
Min	156	1.78	105	1.19	152	1.69	104	1.56	200 (IERS)	2.00
Max	157	2.76	112	2.73	137	3.39	121	5.63		
Mean	–	2.22	–	1.73	–	2.11	–	2.65		
MAE for 30th day of forecast horizon MAE[30]										
Min	135	4.70	136	3.17	115	3.72	117	5.37	200 (IERS)	6.84
Max	157	11.87	112	7.48	150	8.98	121	12.24		
Mean	–	6.63	–	4.43	–	6.20	–	6.97		

selection of data and the necessary preparations to ensure a reliable prediction. From our analysis, we conclude that the pure choice of input data from available EOP series is not the most important variable determining the accuracy of the prediction. The combined information in Table 1 and related figures of MAE (Figures 3 and 4) fails to pinpoint any particular data as an essential choice for highly accurate EOP forecasting. However, most methods utilized past epochs of the IERS 14 C04 data, which were also employed as the reference series in our study. All methods also used good quality geodetic data, either from IERS or official analysis centers of the different services of the International Association of Geodesy (IAG) like IDs 104 and 116. Our results suggest that the choice of the geodetic input data does not matter as long as certain quality levels are reached.

The last analysis in this section concerns changes of MAE in separate time periods to investigate the stability of predictions. This analysis was done for the 10-day forecast horizon. First, we computed MAE for data in 5 periods for visualization purposes (4 periods are 12 weeks long: September 2021–November 2021; December 2021–February 2022; March 2022–May 2022 June 2022–August 2022; the last period covers 16 weeks: September 2022–December 2022). Then, in a separate analysis, we computed percentage change (PCh) for each month of the 2nd EOP PCC campaign to analyze how prediction certainty was changing during the campaign:

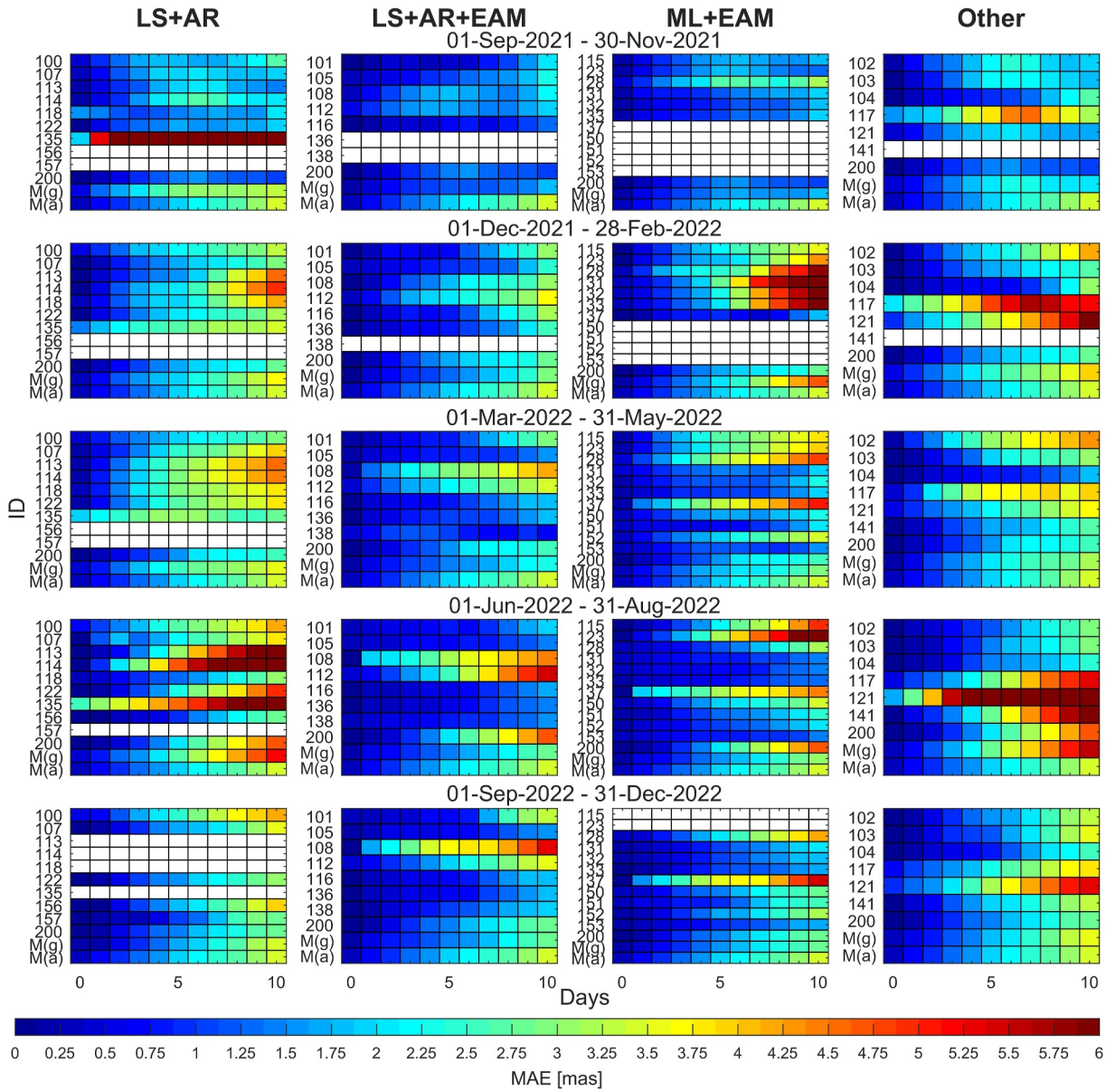

Figure 5. Change of mean absolute error for PMx in consecutive periods. $M(g)$ means the mean for each cluster and investigated period and $M(a)$ represents the mean for all predictions from four clusters for the whole campaign duration. White means no solution was provided in the period.

$$PCh = \text{mean}\left(\frac{MAE_i(n) - MAE_i(n+1)}{MAE_i(n)}\right) \cdot 100\%, \tag{4}$$

where MAE_i(n) is the value for ith day of the prediction computed for nth group. While $PCh > 0$ it means that the next period has a lower MAE. When $PCh < 0$ it means the next period has higher MAE and predictions are worse.

Figures 5 and 6 show values of MAE for the previously selected periods for PMx and PMy, respectively. For the PMx, changes of MAE values between consecutive periods are not dependent across groups or time. For instance, the initial prediction submitted by ID 137 has better accuracy than others from the group (c) $ML + EAM$, but since March 2022 results are worse than for other IDs. For ID 128, 131, 132, and 133 which were submitted by the same institution, all predictions have problems after the 5th day in the time frame December 2021–February 2022. Since March, only ID 128 still had a visibly lower accuracy than other IDs from the group, whereas the other IDs

10.1029/2023EA003278

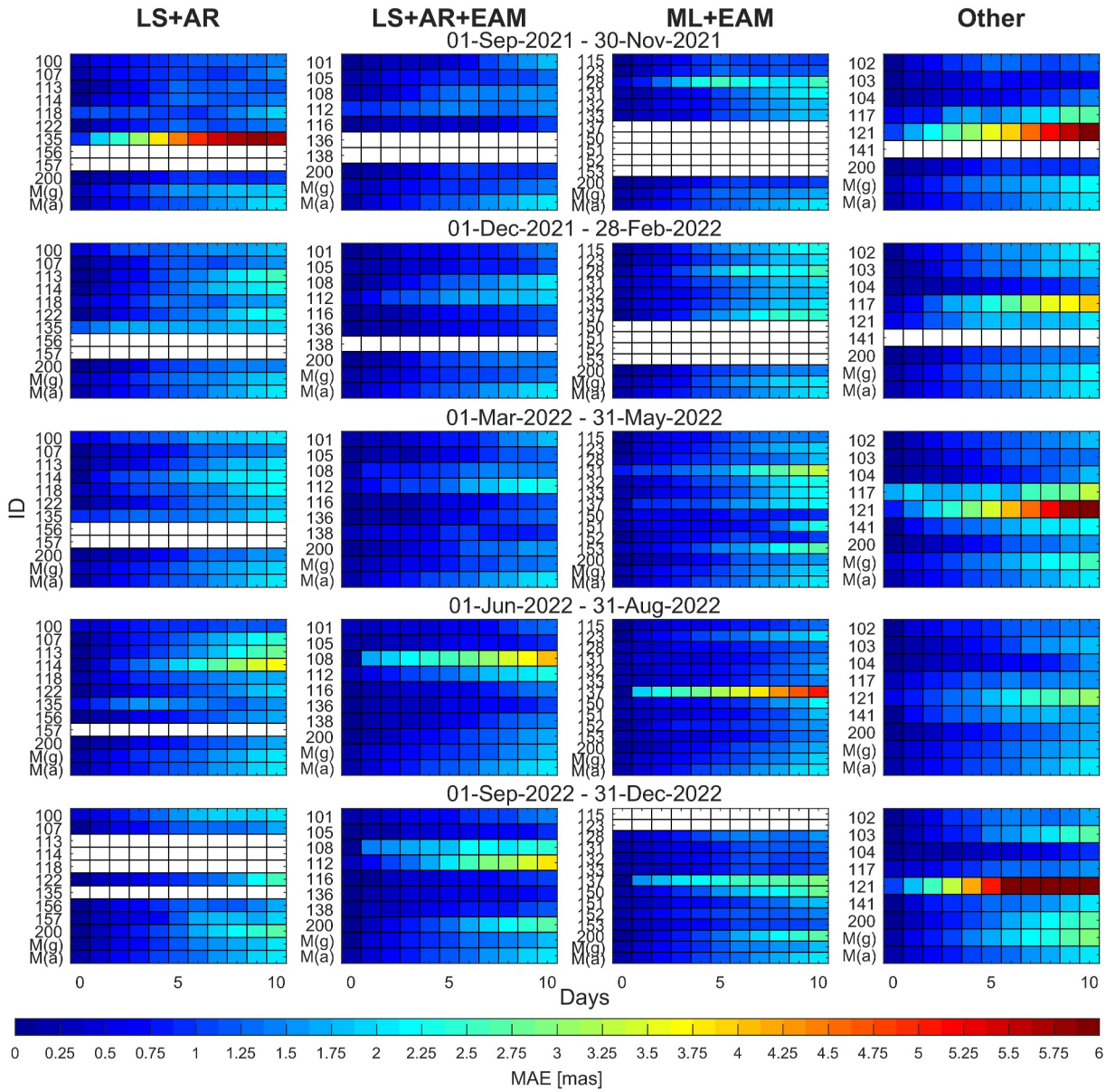

Figure 6. Change of mean absolute error for PMy in consecutive periods. $M(g)$ means the mean for each cluster and investigated period and $M(a)$ represents the mean for all predictions from four clusters for the whole campaign duration. White means no solution was provided in the period.

improved remarkably. Generally, more predictions sustain MAE less than 1 mas for the first 2–3 days, and below 2 mas for the next 3 days. In some cases, using EAM helps achieve more days with better accuracy, but that is not always the case. Results of MAE for consecutive periods for PMy shown in Figure 6 are much more stable for most IDs. ID 121 has an error pattern not observed elsewhere, that is, MAE is growing significantly at about the 5th day, which might suggest systematic errors in this implementation. The error is growing temporarily for other IDs, probably caused by some internal issues. We would like also to highlight ID 135, which started with very erroneous predictions for both PMx and PMy but the participant was able to substantially improve the predictions during the course of the campaign.

To investigate the change of prediction accuracy (per ID) in time, we also computed MAE for each month and then we calculated PCh for each day of prediction. Statistical results for the 1st, 6th, 7th, and 10th days of prediction are displayed in Table 7. The minimum values of PCh for PMx show some severe accuracy degradation for several IDs. According to information provided to the EOP PCC Office by participants, most problems were

Table 7
Statistics and Number of Positive and Negative Values of Percentage Change Computed for 1st, 6th, 7th, and 10th Day of Prediction for All IDs Together

	Mean [%]	Std [%]	Min [%]	Max [%]	Median [%]	PCh > 0	PCh < 0
PMx							
Day 1	−48.4	170.4	−1,456.2	96.7	1.1	154	148
Day 6	−47.1	182.7	−1,350.6	95.9	3.3	158	144
Day 7	−60.1	269.0	−2961.5	94.6	0.7	152	149
Day 10	−40.3	159.9	−1,699.3	85.4	−1.7	149	153
PMy							
Day 1	−87.1	557.7	−6,633.4	95.8	2.7	156	146
Day 6	−30.3	133.7	−1,536.7	94.6	5.0	159	143
Day 7	−30.1	121.7	−893.2	90.7	4.2	162	140
Day 10	−40.3	134.7	−751.6	95.3	2.1	155	147

caused in the data preparation step by utilizing wrong input or sometimes due to lack of external data used for algorithms (or delays in its delivery). Similar values are also noticed for PMy. In the case of PMx, the number of negative PCh is slightly higher or almost equivalent to the number of negative PCh, while for PMy positive PCh is more frequent, but still the difference in the count is negligible. More promising evidence is found in values of median values, which are in the range of −1.7%–5.0%. This suggests that predictions are generally relatively stable in time, as neither high degradations nor relevant improvements are apparent.

4. Relations Between IDs

We also investigate the relations between the 2nd EOP PCC participants' prediction methodology used for individual PM forecasts. In the previous sections, we used clusters of participants selected a priori based on prediction algorithm characteristics. In the following, we also use empirical evidence from the submissions themselves to quantify levels of similarity across submissions from different participants.

We start with an analysis of the correlation for the 5th, 6th, and 7th day of predictions between all participants (Figure 7). One of the most distinctive characteristics between different IDs is the usage of EAM predictions, which are provided up to 6 days into the future (Dill et al., 2019, 2022). By computing correlations, we aimed to check potential dependencies one day before and after the EAM forecast cutoff. For unrelated IDs, that is, computed by different teams, correlation coefficients are usually between −0.2 and 0.2. Higher correlation coefficients are found for predictions from the same a priori group. For most participants, correlation values thus depend on the institution rather than on the day of forecast or EAM data utilization. However, for quite a number of pairs, we note a visible change of prediction characteristics, for example, in PMx the correlations of ID 157 with IDs 105, 135, and 138 for the day 5th and 6th are higher than 0.2, while for the day 7th correlation coefficient are < -0.7. Interestingly, the feature does not emerge for PMy for this ID. Comparing the correlation in each cluster of IDs, there is no pattern in any of them that might suggest a specific impact of the method. Moreover, between PMx and PMy there is some agreement in the values for most pairs of IDs, which probably points to the fact that both polar motion components are typically predicted in combination.

To further extend the analysis related to the similarity of different forecasting methods we adopted unsupervised ML algorithms dedicated to clustering samples to classify the IDs. It allows us to make a more objective comparison of the results obtained by each ID, deprived of the cognitive bias that might be inherent in our manual grouping, which was applied before. Also, it might help to reveal connections between IDs that have been missed in the analysis presented so far. To perform clustering, we exploited the scikit-learn package for Python (Pedregosa et al., 2011). Before starting, we define parameters describing each ID. For this, we used the prediction method, input data, values of RMS and bias computed against EOP 14 C04, correlation with IERS predictions, and finally MAE for 0th, 1st, 7th, and 10th day of prediction for both PMx and PMy. That gives 16 parameters to describe each ID, which is a more detailed parametrization than the one used for manual grouping based only on

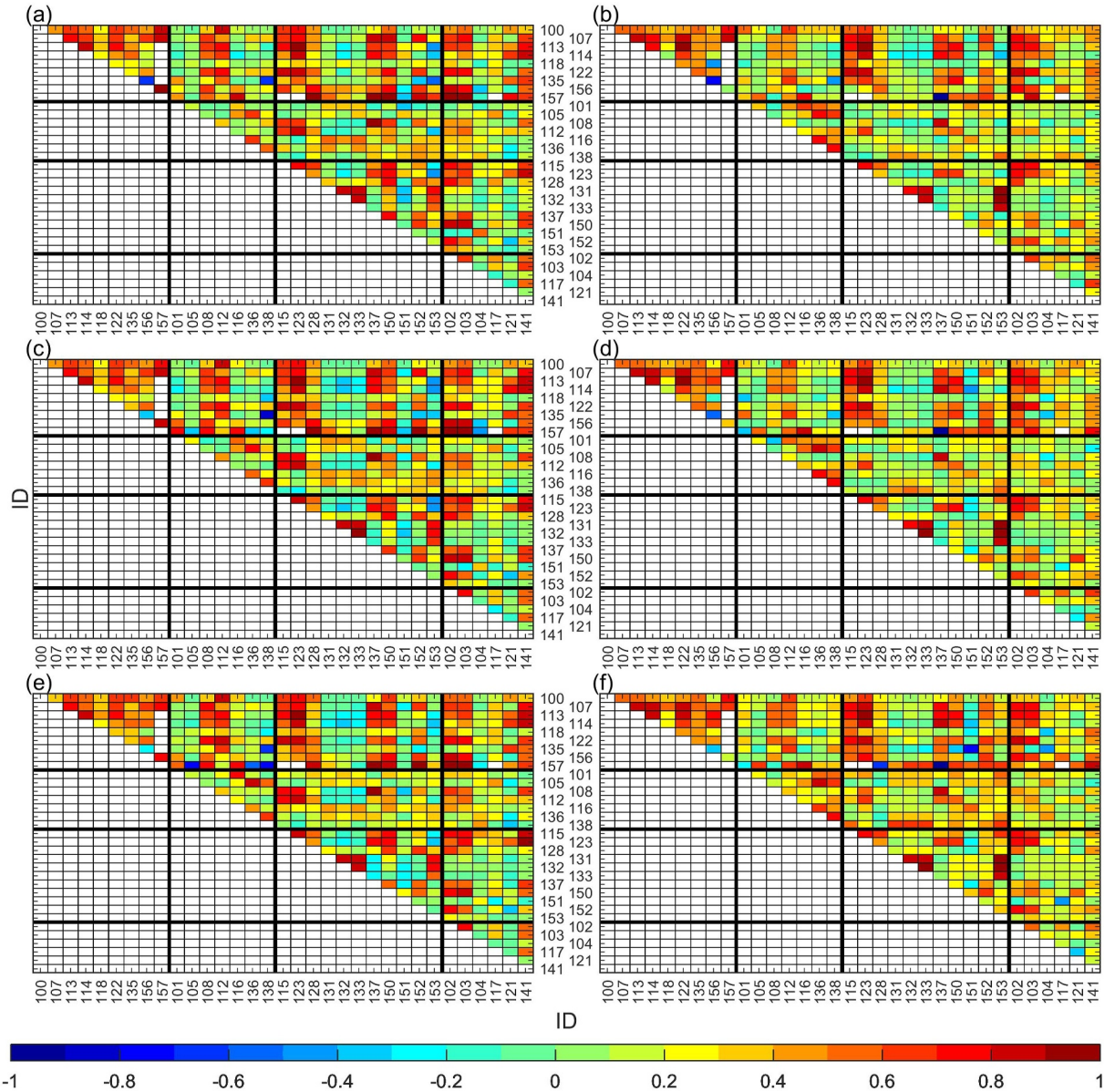

Figure 7. Correlation between participants for (a) PMx and 5th day, (b) PMy and 5th day, (c) PMx and 6th day, (d) PMy and 6th day, (e) PMx and 7th day, (f) PMy and 7th day. Thick black lines represent clusters on the *x* and *y* axis.

method and input data. According to the clustering algorithms requirements for non-numerical input, we applied a one-hot encoder (OHE) and term-frequency times inverse document-frequency (tf-idf). One-hot encoder is the process by which categorical data are converted into numerical as binary features, that is, a feature is represented by a column and receives a "1" if used by ID otherwise it receives a "0". Tf-idf works in a similar manner to OHE but additionally re-weights the count feature into float values. For more details, we refer to Pedregosa et al. (2011) and the scikit-learn webpage (scikit-learn.org—accessed on 14.04.2023). Out of the many methods implemented in the package we performed Monte Carlo analysis, usually with their initial settings, to find out the most representative results. We applied only two criteria: the number of determined groups should be between 3 and 6 and the number of IDs assigned to each cluster should not be lower than 3. These conditions allowed for direct comparison with manual grouping. Eventually, we arrived at 10 clustering methods (CM, Table 8) that met the requirements. 6 of them returned three groups, 3 of them had four groups, and only one method led to five groups.

Table 8
Clustering Algorithms That Fulfilled the Requirements (i.e., the Number of Determined Groups Should Be Between 3 and 6 and the Number of IDs Assigned to Each Cluster Should Not Be Lower Than 3) Which Were Then Used to Classify Prediction Approaches

	Clustering method	Non-numerical conversion type	Number of groups
CM-1	Affinity propagation	OHE	3
CM-2	Affinity propagation	Tf-idf	3
CM-3	Agglomerative clustering	OHE	3
CM-4	Agglomerative clustering	Tf-idf	3
CM-5	MiniBatch KMeans	Tf-idf	3
CM-6	K-means	Tf-idf	3
CM-7	K-means	Tf-idf	4
CM-8	Affinity propagation	OHE	4
CM-9	Agglomerative clustering	OHE	4
CM-10	K-means	Tf-idf	5

Note. CM, cluster method; OHE, a one-hot encoder; TF-IDF, term-frequency times inverse document-frequency.

Figure 8 provides the number of common occurrences for each pair of IDs in the same group across different clustering algorithms including manual clustering. The highest value of common assignment is seen for the group (c) *ML + EAM* consisting of ML with EAM, which makes this group the most comprehensive cluster of individual prediction approaches. Instances that are part of this group have visibly less in common with other clusters (i.e., more white fields in Figure 8 when compared with other groups). Predictions from the group (a) *LS + AR* are the

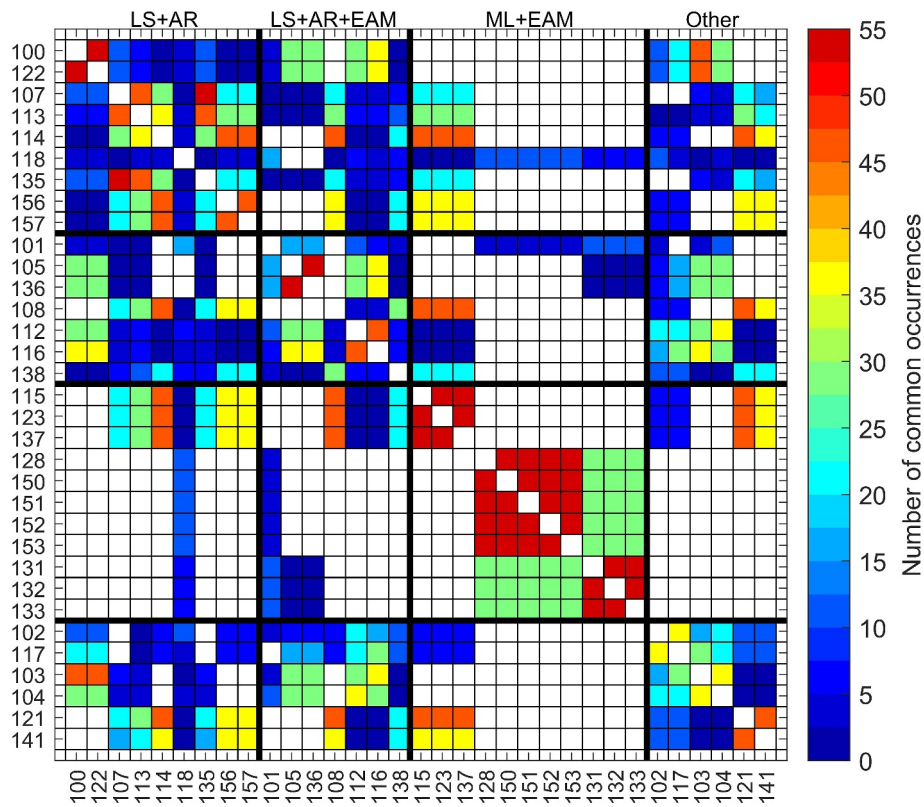

Figure 8. The number of common occurrences of IDs across all clustering approaches (including manual clustering). Black thick lines delineate groups obtained from manual allocation. IDs are grouped by the predicting team and are not sorted in ascending order. White color means no common occurrences.

most variable within the clusters, that is, the prediction method is not shaping the outcome as much as for example, ML. Other criteria like input data or algorithm settings appear to be of minor influence only. Groups based on (a) *LS + AR* have varying internal consistency as most of the IDs from these clusters are not assigned to the same group in this evaluation. Interesting results were obtained for the group (d) *Other*, which collects all remaining IDs that do not suit groups (a) *LS + AR*, (b) *LS + AR + EAM*, and (c) *ML + EAM*. In fact, these IDs could be split into two groups consisting of IDs 102, 103, 104, and 117 (which have much resemblance to group (b) *LS + AR*), and the second cluster could be composed of IDs 121 and 141 (which behave more like group (c) *ML + EAM* based methods). Figure 8 also reveals that IDs 101, 102 and 118 do not have many connections with other IDs, which makes these approaches somewhat one-of-a-kind. On the other hand, IDs 108, 114, or 137 have much in common with many other IDs. IDs 115, 123, and 137 are the only ML-based IDs that are more like (a) *LS + AR*-based approaches than other IDs from the group (c) *ML + EAM*.

5. Discussion and Conclusions

After completion of the operational phase of the 2nd EOP PCC, we evaluated predictions of the *x* and *y* components of polar motion submitted by all participants against the IERS 14 C04 solution. The main aim of the campaign was to analyze current prospects of EOP predictions, including new methods, that is, dynamically evolving ML with its wide range of approaches, as well as input data such as EAM predictions. New computation algorithms might help to fill gaps when modeling of physical phenomena is not readily available. After the end of the first campaign, essential improvements were also made in space geodetic techniques, including new GNSS constellations like Galileo and BeiDou. More advanced observation and computation possibilities result in data processing capabilities with higher accuracy. The conclusions from the 1st EOP PCC indicated the necessity of better modeling EAM, particularly AAM, OAM, and HAM. Currently, predictions of EAM are available for up to 6 days into the future, which clearly impacts the results obtained for 2nd EOP PCC participants using these data, as evident from the increase of MAE after the 7th day of the prediction. Thus, extension of the EAM prediction lengths would potentially allow for lowering polar motion prediction errors for forecast horizons longer than a week. Another point that should be mentioned is the dependence of campaign participants on GFZ EAM data and predictions. During the course of the campaign, ETH also launched an EAM prediction service (https://gpc.ethz.ch/EAM/ - accessed 19.04.2023 (Kiani Shahvandi, Gou, et al., 2022; Soja et al., 2022), but again relying on GFZ EAM. It is thus obvious that any disruption or even discontinuation of the GFZ service might critically impact the community's ability to predict polar motion, particularly at forecast horizons between two and 7 days.

In total 33 different methods were used for polar motion prediction by 18 campaign participants, which gave us an opportunity to evaluate a rich ensemble of different approaches. To indicate the most valuable connection of prediction methodology and input data we prepared a ranking of the IDs based on the following criteria: (a) Percentage of rejected submissions; (b) range of differences between predicted and observed values; (c) MAE on 1st, 6th, 7th, 10th forecast day; and (d) median of PCh. We assigned points equal to the place in the classification for each ID, which means that the lower the number of points, the higher the place reached in the classification. Criterium (A) assesses the reliability of the prediction of the algorithm. Criterium (B) helps to evaluate predicting repeatability—small values of range characterize accurate prediction with high stability in time. Criterium (C) intends to check the quality of predictions on the 1st and 10th day, while MAE for the 6th and 7th might help to assess or justify the use of the EAM predictions. Criterium (D) judges the tolerance of the method on the external data sets and their current quality with respect to the official IERS reference series. These criteria do not evaluate just prediction accuracy. Some IDs submit highly accurate predictions for selected periods rather than for the whole course of the campaign. We also limited ourselves to a 10-day forecast horizon for the ranking, as the short-term predictions are the most needed for many practical applications.

Ranking for PMx is shown in Table 9, and for PMy in Table 10. Overall, the official IERS prediction processed by USNO is placed 11th and 13th for PMx and PMy, respectively. The results also confirm the previous understanding that PMy predictions are generally more accurate than PMx forecasts. Thus, it might be useful to predict both parameters separately with differently configured methods for the highest accuracy. Another option is that PMx requires more accurate determinations of geophysical properties as the change of MAE shown in Figure 5 varies more than for PMy as given in Figure 6. This deviation for PMx is rather independent from the group and is present across various methods, so it is likely that the land-sea distribution (and its varying impact on PMx and PMy) is more important than the actual method employed. Atmospheric and oceanic effects also contribute differently to PMx and PMy changes, just as to the polar motion excitation (Harker et al., 2021). In addition,

Table 9
Ranking of IDs Sorted by Number of Points and Number of Points Reached for Each Criterium (the Less Points the Better the Method) for PMx on 10-Day Forecast Horizon

Rank	ID	Group	Total	% Of rejected predictions	Range	MAE[1]	MAE[6]	MAE[7]	MAE[10]	Median PCh
1	136	LS + AR + EAM	24	0	2	6	1	1	3	11
2	138	LS + AR + EAM	29	0	1	8	4	2	1	13
3	153	ML + EAM	36	3	3	10	6	4	2	8
4	105	LS + AR + EAM	45	0	4	14	5	3	4	15
5	104	Other	52	0	9	3	7	7	8	18
6	116	LS + AR + EAM	52	0	6	7	3	5	5	26
7	152	ML + EAM	59	7	5	11	11	11	11	3
8	151	ML + EAM	66	0	7	17	8	8	12	14
9	157[a]	LS + AR	72	0	8	2	17	19	24	2
10	101	LS + AR + EAM	74	0	24	9	2	6	6	27
11	**200**	**IERS**	**80**	**0**	**14**	**5**	**16**	**14**	**10**	**21**
12	156	LS + AR	81	0	17	1	14	16	21	12
13	150	ML + EAM	87	0	10	18	13	13	14	19
14	133	ML + EAM	89	8	13	13	9	9	7	30
15	103	Other	98	0	18	4	15	15	13	33
16	132	ML + EAM	101	10	15	16	10	10	9	31
17	115	ML + EAM	102	0	16	20	22	22	18	4
18	122	LS + AR	103	0	11	15	20	18	16	23
19	107	LS + AR	108	4	21	22	18	17	17	9
20	118	LS + AR	109	0	19	28	21	21	19	1
21	131	ML + EAM	117	11	12	23	12	12	15	32
22	102	Other	118	0	23	19	19	20	20	17
23	141	Other	130	0	33	12	24	25	29	7
24	113	LS + AR	135	0	26	21	27	28	28	5
25	112	LS + AR + EAM	146	1	22	27	23	23	22	28
26	123	ML + EAM	146	0	25	24	26	26	25	20
27	100	LS + AR	150	0	20	29	25	24	23	29
28	137	ML + EAM	161	0	31	31	30	29	30	10
29	108	LS + AR + EAM	163	2	27	30	29	27	26	22
30	128	ML + EAM	166	6	28	26	28	30	32	16
31	114	LS + AR	175	0	32	25	31	31	31	25
32	117	Other	176	0	29	32	32	32	27	24
33	121	Other	179	5	34	33	33	34	34	6
34	135	LS + AR	207	9	30	34	34	33	33	34

[a]Only seven predictions were submitted by ID 157. The result for prediction from IERS is shown in bold.

estimates for AAM wind term contributions over oceanic areas are generally more uncertain than those over land (e.g., Masaki, 2008).

Figure 9 displays the distribution of total points for each ID in both polar motion components, including group assignment. The figure shows differences in prediction methods efficiency used for PMx and PMy, also indicating some inclination of certain methods toward either PMx or PMy. A visible inclination (difference between number of points for PMx and PMy larger than 10) is found for almost 50% of IDs, reaching up to 60 points

Table 10
Ranking of IDs Sorted by Number of Points and Number of Points Reached for Each Criterium (the Less Points the Better the Method) for PMy on 10-Day Forecast Horizon

Rank	ID	Group	Total	% Of rejected predictions	Range	MAE[1]	MAE[6]	MAE[7]	MAE[10]	Median PCh
1	136	LS + AR + EAM	17	0	5	1	3	3	1	4
2	116	LS + AR + EAM	23	0	4	3	1	1	3	11
3	105	LS + AR + EAM	34	2	7	8	4	4	2	7
4	104	Other	45	0	12	2	2	2	4	23
5	101	LS + AR + EAM	53	0	9	9	5	5	7	18
6	152	ML + EAM	58	0	2	14	8	8	6	20
7	103	Other	61	0	11	4	13	11	10	12
8	151	ML + EAM	61	0	10	17	6	7	18	3
9	138	LS + AR + EAM	64	0	3	12	7	6	5	31
10	156	LS + AR	66	0	6	7	10	9	9	25
11	115	ML + EAM	73	0	14	11	16	14	8	10
12	153	ML + EAM	90	0	13	21	9	10	11	26
13	**200**	**IERS**	**104**	**1**	**19**	**5**	**14**	**15**	**17**	**33**
14	132	ML + EAM	110	7	27	20	11	12	12	21
15	133	ML + EAM	110	8	24	16	12	13	13	24
16	100	LS + AR	111	0	15	27	23	22	16	8
17	102	Other	115	0	22	18	18	16	14	27
18	131	ML + EAM	115	9	18	24	15	20	20	9
19	123	ML + EAM	118	0	23	19	20	19	15	22
20	107	LS + AR	119	3	29	15	17	17	19	19
21	150	ML + EAM	119	0	8	23	22	23	27	16
22	141	Other	120	0	16	26	25	26	25	2
23	113	LS + AR	126	0	30	10	24	25	24	13
24	122	LS + AR	127	0	26	13	19	18	23	28
25	157[a]	LS + AR	130	0	1	6	27	30	32	34
26	118	LS + AR	146	5	20	28	21	21	22	29
27	108	LS + AR + EAM	147	4	25	31	29	27	26	5
28	128	ML + EAM	154	11	17	25	26	24	21	30
29	135	LS + AR	158	10	21	32	32	32	30	1
30	112	LS + AR + EAM	165	0	28	29	31	31	31	15
31	137	ML + EAM	170	0	31	34	33	33	33	6
32	114	LS + AR	171	0	32	22	28	28	29	32
33	117	Other	173	6	33	30	30	29	28	17
34	121	Other	183	0	34	33	34	34	34	14

[a]Only seven predictions were submitted by ID 157. The result for prediction from IERS is shown in bold.

collected during the assessment, as can be deduced from the deviation from the dashed line that represents a perfectly balanced method for both polar motion components.

To conclude this research on polar motion predictions from the 2nd EOP PCC, the most promising methods typically exploit LS + AR or ML both in combination with EAM forecasts. The EOP input source seems to be of lesser importance as long as certain quality levels are reached. All methods of prediction evaluation demonstrate that many prediction approaches share common features, and the decisive factor is the proper method parameterization, which needs to be compatible with the input data. In some respects, however, ML-based computations

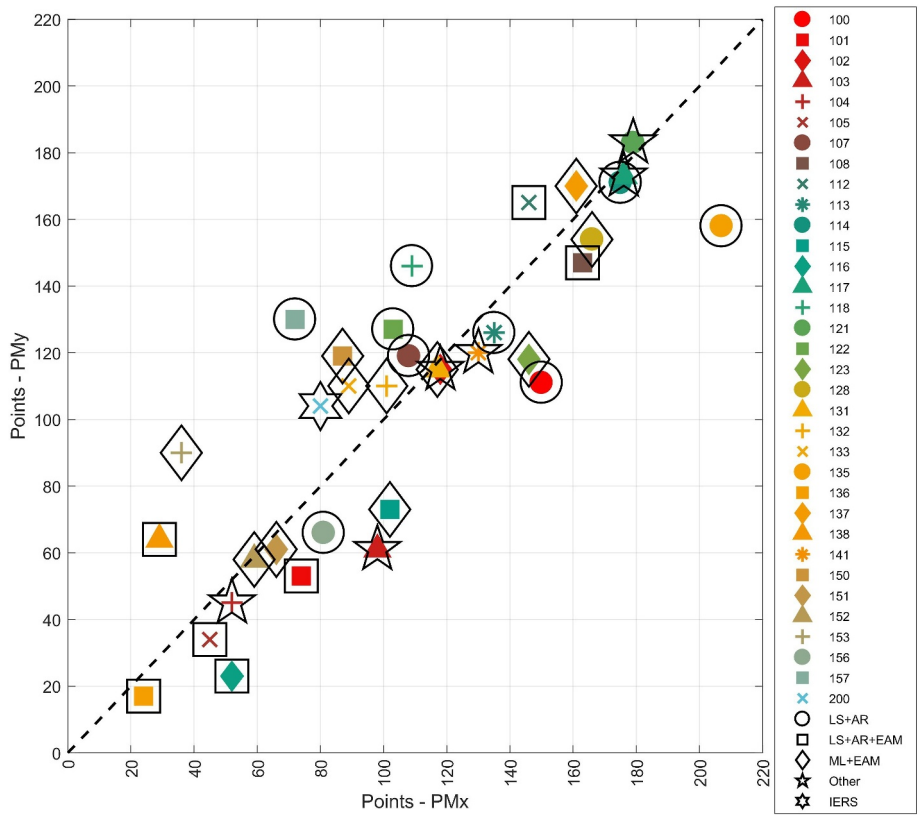

Figure 9. The distribution of total points for PMx and PMy for each ID, also considering the group assignment presented in Section 2.3.

are considered as unique approaches in comparison to the LS-based methods. There are promising ready-to-use prediction methods already operating more accurately than the official IERS predictions. On the other hand, still more analysis concerning the use of a full set of EOP should be done for example, in orbit determination of GNSS satellites or in VLBI solution. On another note, predictions submitted to the 2nd EOP PCC are more accurate in the short term than those published by the IERS, while IERS performs with higher quality in mid-term forecasts. It remains an open question how longer EAM prediction might affect EOP prediction, or how EAM computed by different institutes would improve forecasts. It might therefore be advantageous to continue the evaluation of operational polar motion predictions in the future after the end of the 2nd EOP PCC in order to continue fostering exchange among the various international research groups involved, and maybe even arrive at some combination product with value to routine users of EOP forecasts.

Appendix A: ID Descriptions

Description of Each ID as Provided by 2nd EOP PCC Participants.

A1. ID 100

There are both regular and irregular signals in EOP data series, such as the trend, annual, Chandler terms, and high frequency trembles in polar motion. For the predictions of the regular signals, we adopt the LS model expressed by polynomial trend and harmonic oscillations. For the irregular part, a stochastic autoregressive (AR) process model is employed (X. Q. Xu et al., 2012; X. Xu & Zhou, 2015).

A2. ID 101

The method utilizes EOP 14 C04 and the latest IERS finals.daily files are combined as the EOP inputs. Additionally, the GFZ EAM products, including 6 days predictions are utilized. The method we used to predict PMx and PMy is based on (Dill et al., 2019) with some revisions: given the 1-day delay to GFZ's EAM prediction, the 6 days of prediction is adjusted to 5 days. In the step of LS and AR, the parameter is optimized but the evaluation day for different parameters at different time scales.

A3. ID 102

Past data allowing to build the prediction are the EOP 14 C04 and finals.daily series up to the current date. The predicted values are given for 365 days into the future. The polar motion components are predicted individually using a least squares (LS) method to determine a harmonic model spanning the last 59 years. This model incorporates a second-degree polynomial trend, seasonal terms of 365 and 182 days, and the Chandler wobble of 433 days. Additionally, it includes side oscillations to account for the long-term modulation of the Chandler wobble over 23-year and 46-year periods. Residuals are modeled by an auto-regressive (AR) model of which the degree is calculated according to the values of the partial autocorrelation coefficients. LS and the AR results are extrapolated for up to 365 days.

A4. ID 104

A Kalman filter is used to combine independent measurements of the Earth's orientation taken by the space-geodetic observing techniques of SLR, VLBI, and the GNSS. In order to improve the predicted EOPs, AAM, and OAM analyses and forecasts are used as proxy 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).

A5. ID 105

EOP prediction is based on the GFZ EAM Predictor (Dill et al., 2019). The sum of EAM (4 years of model-based EAM 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 first LS + AR step, is predicted into the future by a second 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 on the actual day of submission.

A6. ID 107

First, we determine the length of the training data and read the raw data, preprocessing according to its type. Then we perform the least square fitting, calculate 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 signals are added together and post-processed to obtain the prediction product (Chen et al., 2014).

A7. ID 108

Our prediction method is LS + MAR, in which LS means difference least-squares adjustment, and MAR means Multi-elements autoregressive modeling. The inputs include EOP 14 C04 data released by the IERS and AAM data released by GFZ. The prediction parameters include both PM components and the longest prediction day is 365.

A8. ID 112

For our LS + Convolution method, the EAM from 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 equation. Meanwhile, the rest of the EOP trend terms are extrapolated by the polynomial least-square (LS) model. Finally, the total EOP predictions are combined with the excitation calculations and trend extensions (X. Xu et al., 2022).

A9. ID 116

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 IAG Services, as well as BKG and 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 modeling 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 that extrapolates 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 core routines of the ESA ERP Software are described in Kehm et al. (2023).

A10. ID 117

Our algorithm for predicting the ERP is called singular spectrum analysis (SSA) + Copula-based analysis (Modiri, 2021; Modiri et al., 2018, 2020). The algorithm splits the observed PM time series into periodic terms and anomalies, which are modeled 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 modeling 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 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.

A11. ID 128, ID 131, ID 132, ID 133

Our methods are based on first-order neural ODE (Neural ODEs) that implicitly assumes that 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, Schartner, & Soja, 2022) 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, Gou, et al., 2022), 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 & Soja, 2022). As a result, the loss function here is the mean squared error. The architecture does not require any preprocessing 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 contains both 10 and 30 days. The input

sequence length is 10 days. The architectures are trained at each prediction epoch to take advantage of the most recently available EOP and EAM data.

A12. ID 135

We utilize here PMx, PMy and also PMy-PMx series, and the final forecasts of each method were chosen to be combined according to the mathematical relationship between the PMy-PMx, PMy and PMx series. The combination that minimizes the forecast error, that is, the optimal forecast of the PMx series is obtained by combining the PMy-PMx forecast of the traditional method and the PMy forecast of the first-order difference method; the optimal forecast of the PMy series is obtained by combining the PMx forecast of the traditional method and the PMy-PMx forecast of the first-order difference method.

A13. ID 136

EOP prediction is based on the GFZ EAM Predictor (Dill et al., 2019). The sum of EAM (4 years of model-based EAM functions including EAM 6-day forecasts) and the residual of GAM (4 years of geodetic angular momentum derived from IERS 14C04), extrapolated for the last ∼30 days up to the end of the EAM 6-day forecasts by a first LS + AR step, is predicted into the future by a second 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 only generated once the rapid EOP solution for the actual day is available (typically around ∼17:15 UTC), and it thereby sometimes deviates from submission ID 105 by utilizing more recent (and thus more accurate) EOP data for the day zero. 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 (Dill et al., 2022). The Liouville equation utilized at GFZ follows Brzeziński (1992) with the numerical values as reported in Dill and Dobslaw (2010). The predicted 3-hourly EOPs are re-sampled to daily time intervals.

A14. ID 137

The prediction method is ANN (Artificial neural network) and autoregressive (AR) modeling. We adopt the wavelet function in ANN to predict the polar motion. A three-layer network is constructed, and the wavelet function is utilized in the mid-layer to approximate the non-linear relationship between the input and output, and finally to make a high-resolution prediction of polar motion. The inputs just contain the EOP 14 C04 data released by IERS, and the prediction parameters include both PM components, the longest prediction day is 365.

A15. ID 138

Our method for predicting polar motion was inspired by Dill et al. (2019) and Dobslaw and Dill (2018), which we enhance here in two ways: (a) Improve the deconvolution and convolution techniques to recalculate the geodetic residuals and lower the PM errors that were reproduced, (b) Develop some new algorithms that utilize the LS, AR, Fast Fourier Transform (FFT), and other methods to further reduce the EAM prediction errors.

A16. ID 141

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 the 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 residual series. Residual series are obtained after the removal of a linear trend and periodic components. Among theoretical semivariogram models, the best performing one turned out to be the Radon transform of the exponential model of order 2 (Markov 2nd order model). The final forecast consists of an extrapolated deterministic part combined with the predicted (kriged) stochastic part. The entire process is presented in Michalczak and Ligas (2021, 2022).

A17. ID 150, ID 151, ID 152

This method is based on the first order neural ODE (Neural ODEs). The general Neural ODE differential learning architecture (Kiani Shahvandi, Schartner, & Soja, 2022) is modified in a way that does not consider the rates of EOPs. A simplified form of this architecture is also used (referred to as simple recursive) (Kiani Shahvandi & Soja, 2022) in which the uncertainties in the input data are used to weigh the loss function (Kiani Shahvandi & Soja, 2022) (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 the 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 to compute the residuals. Subsequently, these residuals are modeled by LSTM. The input to the LSTM architecture is only the past values of residuals of training of the Neural ODEs. The input sequence length here is 10 and the loss function is MAE. For this purpose, the predictions of IERS are also incorporated into the algorithm for predicting the residuals at each training epoch (retraining is required).

A18. ID 153

This method is based on a modification of the differential form of neural ODE (Kiani Shahvandi, Schartner, & Soja, 2022) in which the Liouville equation for polar motion is incorporated into the algorithm as a geophysical constraint (Kiani Shahvandi, Dill, Dobslaw, Mishra, & Soja, 2023). This means that the input, that is, polar motion observations and EAM functions, are primarily used as features, while simultaneously being connected to each other via the Liouville differential equation to account for the rotational dynamics of the Earth. As such, this method is a simple physics-constrained neural network, introduced in a more general and rigorous form for the prediction of EOPs in Kiani Shahvandi, Dill, Dobslaw, Kehm et al. (2023). An attempt is also made to model the residuals of training using the LSTM neural networks similar to Gou et al. (2023). The inputs to the LSTM architecture are the residuals of the physics-constrained neural network during the training phase, with the input sequence length being 10.

A19. ID 156

The deterministic part, that is, estimated linear trend and periodic components are first removed from the raw times series. An autoregressive integrated moving average (ARIMA) model is then used to predict the residual part of the time series. 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 part is extrapolated for future time instances and then combined with the ARIMA-based predicted stochastic part. The best set of parameters p and q is selected by means of the 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. The stationarity of each residual ERP time series is checked using the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. The entire process is described in Michalczak et al. (2022).

A20. ID 157

Vector autoregression of order p VAR(p) is a multivariate counterpart of an autoregressive model AR(p) that describes evolution and coevolution of variables in time. It was applied to a joint prediction of residual polar motion series (PMx-res, PMy-res) that were obtained after removing linear trend and periodic components (fitted by LS to the input polar motion series; separately for x and y). Finally, all components, that is, extrapolated linear trends, periodicities and predicted residuals were added together to generate the final forecast. The prediction procedure depends on a set of parameters involving input time series length for trend and periodic components estimation, length of subseries for autoregression parameters estimation, order of autoregression, and number of periodic components. Since one of the variables in a joint forecast is usually predicted better, two separate bivariate predictions were applied, one for x and one for y component of polar motion.

Data Availability Statement

All predictions submitted to the EOP PCC Office in the frame of the 2nd EOP PCC can be accessed from the GFZ Data Services (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.

Acknowledgments
T. Kur, J. Śliwińska-Bronowicz, M. Wińska, J. Nastula, and A. Partyka representing the EOP PCC Office gratefully acknowledge the efforts of all participants of the 2nd EOP PCC for their invaluable contributions to the campaign. This study was funded by the National Science Centre, Poland under the OPUS call in the Weave programme, Grant 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 PID2020-119383GB-I00 funded by Ministerio de Ciencia e Innovación (MCIN/AEI/10.13039/501100011033/) and PROMETEO/2021/030 funded by Generalitat Valenciana.

References

Afroosa, M., Rohith, B., Paul, A., Durand, F., Bourdallé-Badie, R., Sreedevi, P. V., et al. (2021). Madden-Julian Oscillation winds excite an intraseasonal see-saw of ocean mass that affects Earth's polar motion. *Communications Earth & Environment*, *2*(1), 1–8. https://doi.org/10.1038/s43247-021-00210-x

Altamimi, Z., Rebischung, P., Collilieux, X., Métivier, L., & Chanard, K. (2023). ITRF2020: An augmented reference frame refining the modeling of nonlinear station motions. *Journal of Geodesy*, *97*(5), 47. https://doi.org/10.1007/s00190-023-01738-w

Belda, S., Ferrándiz, J. M., Heinkelmann, R., Nilsson, T., & Schuh, H. (2016). Testing a new free core nutation empirical model. *Journal of Geodynamics*, *94–95*, 94–95. https://doi.org/10.1016/j.jog.2016.02.002

Belda, S., Ferrándiz, J. M., Heinkelmann, R., & Schuh, H. (2018). A new method to improve the prediction of the celestial pole offsets. *Scientific Reports*, *8*(1), 13861. https://doi.org/10.1038/s41598-018-32082-1

Bizouard, C., & Gambis, D. (2009). *The combined solution C04 for Earth orientation parameters consistent with international terrestrial reference frame 2005* (Vol. 134, pp. 265–270). International Association of Geodesy Symposia. https://doi.org/10.1007/978-3-642-00860-3_41

Bizouard, C., Lambert, S., Gattano, C., Becker, O., & Richard, J. Y. (2019). The IERS EOP 14C04 solution for Earth orientation parameters consistent with ITRF 2014. *Journal of Geodesy*, *93*(5), 621–633. https://doi.org/10.1007/s00190-018-1186-3

Bizouard, C., & Seoane, L. (2010). Atmospheric and oceanic forcing of the rapid polar motion. *Journal of Geodesy*, *84*(1), 19–30. https://doi.org/10.1007/s00190-009-0341-2

Bruni, S., Schoenemann, E., Mayer, V., Otten, M., Springer, T., Dilssner, F., et al. (2021). ESA's Earth Orientation Parameter product. In *EGU General Assembly 2021 online, 19–30 Apilr 2021* (p. EGU21-12989). 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. *Manuscripta Geodaetica*, *17*(1), 3–20.

Byram, S., & Hackman, C. (2012). High-precision GNSS orbit, clock and EOP estimation at the United States naval observatory. In *Record - IEEE PLANS, position location and navigation symposium*. https://doi.org/10.1109/PLANS.2012.6236940

Chen, L., Tang, G., Hu, S., Ping, J., Xu, X., & Xia, J. (2014). High accuracy differential prediction of UT1-UTC. *Journal of Deep Space Exploration*, *1*(3), 230–235.

Coulot, D., Pollet, A., Collilieux, X., & Berio, P. (2010). Global optimization of core station networks for space geodesy: Application to the referencing of the SLR EOP with respect to ITRF. *Journal of Geodesy*, *84*(1), 31–50. https://doi.org/10.1007/s00190-009-0342-1

Dill, R., & Dobslaw, H. (2010). Short-term polar motion forecasts from earth system modeling data. *Journal of Geodesy*, *84*(9), 529–536. https://doi.org/10.1007/s00190-010-0391-5

Dill, R., Dobslaw, H., & Thomas, M. (2019). Improved 90-day Earth orientation predictions from angular momentum forecasts of atmosphere, ocean, and terrestrial hydrosphere. *Journal of Geodesy*, *93*(3), 287–295. https://doi.org/10.1007/s00190-018-1158-7

Dill, R., Dobslaw, H., & Thomas, M. (2022). ESMGFZ products for Earth rotation prediction. *Artificial Satellites*, *57*(s1), 254–261. https://doi.org/10.2478/arsa-2022-0022

Dobslaw, H., & Dill, R. (2018). Predicting Earth orientation changes from global forecasts of atmosphere-hydrosphere dynamics. *Advances in Space Research*, *61*(4), 1047–1054. https://doi.org/10.1016/j.asr.2017.11.044

Freedman, A. P., Steppe, J. A., Dickey, J. O., Eubanks, T. M., & Sung, L. Y. (1994). The short-term prediction of universal time and length of day using atmospheric angular momentum. *Journal of Geophysical Research*, *99*(B4), 6981–6996. https://doi.org/10.1029/93JB02976

Gambis, D. (2004). Monitoring Earth orientation using space-geodetic techniques: State-of-the-art and prospective. *Journal of Geodesy*, *78*(4–5), 295–303. https://doi.org/10.1007/s00190-004-0394-1

Gambis, D., & Luzum, B. (2011). Earth rotation monitoring, UT1 determination and prediction. *Metrologia*, *48*(4), S165–S170. https://doi.org/10.1088/0026-1394/48/4/S06

Gou, J., Kiani Shahvandi, M., Hohensinn, R., & Soja, B. (2023). Ultra-short-term prediction of LOD using LSTM neural networks. *Journal of Geodesy*, *97*(5), 52. https://doi.org/10.1007/s00190-023-01745-x

Gross, R. S., Eubanks, T. M., Steppe, J. A., Freedman, A. P., Dickey, J. O., & Runge, T. F. (1998). A kalman-filter-based approach to combining independent Earth-orientation series. *Journal of Geodesy*, *72*(4), 215–235. https://doi.org/10.1007/s001900050162

Harker, A. A., Schindelegger, M., Ponte, R. M., & Salstein, D. A. (2021). Modeling ocean-Induced rapid earth rotation variations: An update. *Journal of Geodesy*, *95*(9), 1–13. https://doi.org/10.1007/s00190-021-01555-z

Kalarus, M., Schuh, H., Kosek, W., Akyilmaz, O., Bizouard, C., Gambis, D., et al. (2010). Achievements of the Earth orientation parameters prediction comparison campaign. *Journal of Geodesy*, *84*(10), 587–596. https://doi.org/10.1007/s00190-010-0387-1

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. *Journal of Geodesy*, *97*(1), 3. https://doi.org/10.1007/s00190-022-01695-w

Kiani Shahvandi, M., Dill, R., Dobslaw, H., Kehm, A., Bloßfeld, M., Schartner, M., et al. (2023). Geophysically informed machine learning for improving rapid estimation and short-term prediction of earth orientation parameters. *Journal of Geophysical Research: Solid Earth*, *128*(10). https://doi.org/10.1029/2023JB026720

Kiani Shahvandi, M., Dill, R., Dobslaw, H., Mishra, S., & Soja, B. (2023). Improving the accuracy of rapid Earth Orientation Parameters with the "ResLearner" machine learning method. In *EGU General Assembly 2023, Vienna, Austria, 24–28 April 2023* (p. EGU23-4162). https://doi.org/10.5194/egusphere-egu23-4162

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

Kiani Shahvandi, M., Schartner, M., & Soja, B. (2022). Neural ODE differential learning and its application in polar motion prediction. *Journal of Geophysical Research: Solid Earth*, *127*(11). https://doi.org/10.1029/2022JB024775

Kiani Shahvandi, M., & Soja, B. (2022). Inclusion of data uncertainty in machine learning and its application in geodetic data science, with case studies for the prediction of Earth orientation parameters and GNSS station coordinate time series. *Advances in Space Research*, *70*(3), 563–575. https://doi.org/10.1016/j.asr.2022.05.042

Kur, T., Dobslaw, H., Śliwińska, J., Nastula, J., Wińska, M., & Partyka, A. (2022). Evaluation of selected short-term predictions of UT1-UTC and LOD collected in the second earth orientation parameters prediction comparison campaign. *Earth Planets and Space*, *74*(1), 191. https://doi.org/10.1186/s40623-022-01753-9

Luzum, B. J., Ray, J. R., Carter, M. S., & Josties, F. J. (2001). Recent improvements to IERS Bulletin A combination and prediction. *GPS Solutions*, *4*(3), 34–40. https://doi.org/10.1007/PL00012853

Masaki, Y. (2008). Wind field differences between three meteorological reanalysis data sets detected by evaluating atmospheric excitation of Earth rotation. *Journal of Geophysical Research*, *113*(D7), 7110. https://doi.org/10.1029/2007JD008893

Michalczak, M., & Ligas, M. (2021). Kriging-based prediction of the Earth's pole coordinates. *Journal of Applied Geodesy*, *15*(3), 233–241. https://doi.org/10.1515/jag-2021-0007

Michalczak, M., & Ligas, M. (2022). The (ultra) short term prediction of length-of-day using kriging. *Advances in Space Research*, *70*(3), 610–620. https://doi.org/10.1016/j.asr.2022.05.007

Michalczak, M., Ligas, M., & Kudrys, 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. *Artificial Satellites*, *57*(s1), 275–289. https://doi.org/10.2478/arsa-2022-0024

Modiri, S. (2021). On the improvement of Earth orientation parameters estimation using modern space geodetic techniques.

Modiri, S., Belda, S., Heinkelmann, R., Hoseini, M., Ferrándiz, J. M., & Schuh, H. (2018). Polar motion prediction using the combination of SSA and copula-based analysis. *Earth Planets and Space*, *70*(1), 115. https://doi.org/10.1186/s40623-018-0888-3

Modiri, S., Belda, S., Hoseini, M., Heinkelmann, R., Ferrándiz, J. M., & Schuh, H. (2020). A new hybrid method to improve the ultra-short-term prediction of LOD. *Journal of Geodesy*, *94*(2), 23. https://doi.org/10.1007/s00190-020-01354-y

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*.

Schuh, H., & Böhm, J. (2014). Very long baseline interferometry for geodesy and astrometry. In *Sciences of Geodesy - II: Innovations and future developments*. https://doi.org/10.1007/978-3-642-28000-9_7

Śliwińska, J., Dobslaw, H., Kur, T., Nastula, J., Wińska, M., Partyka, A., et al. (2023). EOP predictions collected during the operational phase of the second earth orientation parameters prediction comparison campaign [Dataset]. *GFZ Data Services*. https://doi.org/10.5880/GFZ.1.3.2023.001

Śliwińska, J., Kur, T., Wińska, M., Nastula, J., Dobslaw, H., & Partyka, A. (2022). Second earth orientation parameters prediction comparison campaign (2nd EOP PCC): Overview. *Artificial Satellites*, *57*(s1), 237–253. https://doi.org/10.2478/arsa-2022-0021

Śliwińska-Bronowicz, J., Kur, T., Wińska, M., Dobslaw, H., Nastula, J., Partyka, A., et al. (2024). 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*, *98*(3), 22. https://doi.org/10.1007/s00190-024-01824-7

Soja, B., Kiani Shahvandi, M., Schartner, M., Gou, J., Kłopotek, G., Crocetti, L., & Awadaljeed, M. (2022). The New Geodetic Prediction Center at ETH Zurich. In *EGU General Assembly 2022, Vienna, Austria, 23–27 May 2022* (p. EGU22-9285). https://doi.org/10.5194/egusphere-egu22-9285

Wu, Y., Zhao, X., & Yang, X. (2022). Improved prediction of polar motions by piecewise parameterization. *Artificial Satellites*, *57*(s1), 290–299. https://doi.org/10.2478/arsa-2022-0025

Xu, X., & Zhou, Y. (2015). EOP prediction using least square fitting and autoregressive filter over optimized data intervals. *Advances in Space Research*, *56*(10), 2248–2253. https://doi.org/10.1016/j.asr.2015.08.007

Xu, X., Zhou, Y., & Xu, C. (2022). Earth rotation parameters prediction and climate change indicators in it. *Artificial Satellites*, *57*(s1), 262–273. https://doi.org/10.2478/arsa-2022-0023

Xu, X. Q., Zhou, Y. H., & Liao, X. H. (2012). Short-term earth orientation parameters predictions by combination of the least-squares, AR model and Kalman filter. *Journal of Geodynamics*, *62*, 83–86. https://doi.org/10.1016/j.jog.2011.12.001