AGH

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie

DZIEDZINA NAUK INŻYNIERYJNO-TECHNICZNYCH

DYSCYPLINA INŻYNIERIA LĄDOWA, GEODEZJA I TRANSPORT

ROZPRAWA DOKTORSKA

Krótkoterminowa prognoza Parametrów Rotacji Ziemi

Autor: mgr inż. Maciej Patryk Michalczak

Promotor rozprawy: dr hab. inż. Marcin Ligas

Praca wykonana: Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie, Wydział Geodezji Górniczej i Inżynierii Środowiska

Kraków, 2024

Streszczenie

Parametry Orientacji Ziemi (EOP) to zbiór parametrów opisujących nieregularności w orientacji i ruchu obrotowym Ziemi. Wspomniane parametry składają się z: współrzędnych bieguna ziemskiego (PMx, PMy), różnicy czasów UT1-UTC oraz przesunięcia bieguna niebieskiego (dX, dY). Dodatkowym parametrem, który również opisuje zmiany w ruchu obrotowym Ziemi jest pierwsza ujemna pochodna po czasie UT1-UTC – długość doby (LOD). Podzbiorem Parametrów Orientacji Ziemi są Parametry Rotacji Ziemi (ERP), do których zalicza się współrzędne bieguna ziemskiego oraz UT1-UTC (bądź LOD). Zmienność wspomnianych parametrów spowodowana jest między innymi przez oddziaływanie Ziemi z ciałami niebieskimi takimi jak Słońce, Księżyc i innymi obiektami w Układzie Słonecznym, ruchami w jądrze i płaszczu Ziemi czy topnieniem pokrywy lodowej.

Parametry Orientacji Ziemi wyznaczane są za pomocą kombinacji różnych satelitarnych i kosmicznych technik geodezyjnych, takich jak: DORIS (Doppler Orbitography and Radiopositioning Integrated by Satellite), VLBI (Very Long Baseline Interferometry), GNSS (Global Navigation Satellite System) i SLR (Satellite Laser Ranging). Ich znajomość jest niezbędna do transformacji między niebieskim (CRF) a ziemskim układem odniesienia (TRF). Ponadto, są one wykorzystywane do wielu zadań, zarówno geodezyjnych, jak i astronomicznych, do których należą: precyzyjne pozycjonowanie w przestrzeni i na powierzchni Ziemi, nawigacja przez Deep Space Network (DSN) czy realizacja systemów czasu.

Złożoność technik pomiarowych oraz przetwarzania danych związanych z wyznaczaniem Parametrów Orientacji Ziemi uniemożliwia uzyskanie ich w czasie rzeczywistym. Rozwiązaniem tego problemu jest predykcja wspomnianych parametrów. Na przestrzeni lat stosowano wiele metod prognozy, jednak z biegiem lat wymagana jest coraz wyższa dokładność prognozy.

Głównym celem przeprowadzonych badań było opracowanie algorytmów predykcyjnych Parametrów Rotacji Ziemi z wykorzystaniem krigingu prostego i krigingu zwyczajnego, jedno- i wielozmiennych metod analizy szeregów czasowych (ARIMA, VAR) oraz dynamicznej dekompozycji modalnej (DMD).

W ramach badań przeprowadzono szereg analiz związanych z prognozą współrzędnych bieguna ziemskiego, różnicy czasów UT1-UTC i długości doby z wykorzystaniem różnych metod prognozy, danych wejściowych i schematów prognozy. W badaniach wykorzystano również dane o efektywnych momentach pędu (EAM), aby zwiększyć dokładność prognozy poprzez uwzględnienie dynamicznych zmian w atmosferze, hydrosferze lądowej i oceanach. Przeprowadzone badania pozwoliły również na porównanie uzyskanych rezultatów prognoz z wynikami uzyskanymi w ramach Drugiej Kampanii Porównawczej Prognozy Parametrów Orientacji Ziemi.

Wyniki uzyskane w ramach niniejszej pracy doktorskiej są potwierdzeniem postawionej tezy, że prognozowanie Parametrów Rotacji Ziemi za pomocą krigingu prostego i krigingu zwyczajnego, modelu ARIMA, VAR oraz dynamicznej dekompozycji modalnej umożliwia uzyskanie prognoz o dokładności wyższej lub porównywalnej do najdokładniejszych obecnie stosowanych metod. W szerszej perspektywie niezbędny jest ciągły rozwój zaprezentowanych algorytmów predykcyjnych, a także uwzględnianie dodatkowych informacji, pozwalających lepiej opisać i uwzględnić zjawiska fizyczne wpływające na ruch obrotowy Ziemi.

Summary

Earth Orientation Parameters (EOP) describe the irregularities in the Earth's orientation and rotation. These parameters include the Earth's pole coordinates (PMx, PMy), UT1-UTC and celestial pole offsets (dX, dY). Another parameter that reflects changes in Earth’s rotation is the negative first derivative of UT1-UTC, the Length of Day (LOD). A subset of the EOP, the Earth Rotation Parameters (ERP), comprises the Earth’s pole coordinates and UT1-UTC (or LOD). Variability in these parameters is driven by factors such as the Earth’s interaction with celestial objects like the Sun, Moon, and other objects in the Solar System, movements within the Earth’s core and mantle, and ice cap melting.

Earth Orientation Parameters are determined by combination of geodetic satellite and space techniques, such as: DORIS (Doppler Orbitography and Radiopositioning Integrated by Satellite), VLBI (Very Long Baseline Interferometry), GNSS (Global Navigation Satellite System), and SLR (Satellite Laser Ranging). These parameters are essential for transformations between the Celestial Reference Frame (CRF) and the Terrestrial Reference Frame (TRF). Additionally, EOP are needed for various geodetic and astronomical purposes, including precise spatial positioning, Deep Space Network (DSN) navigation and time keeping.

Complexity of measuring techniques and data processing involved in Earth Orientation Parameters determination makes it impossible to obtain them in real-time mode. The solution to this problem is the prediction of the mentioned parameters. Over the years many prediction methods have been used, but over time increasing prediction accuracy is needed.

The main objective of this research was to develop predictive algorithms for Earth Rotation Parameters using simple kriging and ordinary kriging, univariate and multivariate time series analysis methods (ARIMA, VAR), and Dynamic Mode Decomposition (DMD).

This study involved a series of analyses related to prediction of the Earth’s pole coordinates, UT1-UTC and Length of Day using various forecasting methods, input data, and prediction schemes. Effective Angular Momentum (EAM) data were also employed to improve forecast accuracy by considering dynamic changes in the atmosphere, land hydrosphere, and oceans. The conducted research also enabled a comparison of the obtained forecast results with the results achieved in the Second Earth Orientation Parameter Prediction Comparison Campaign.

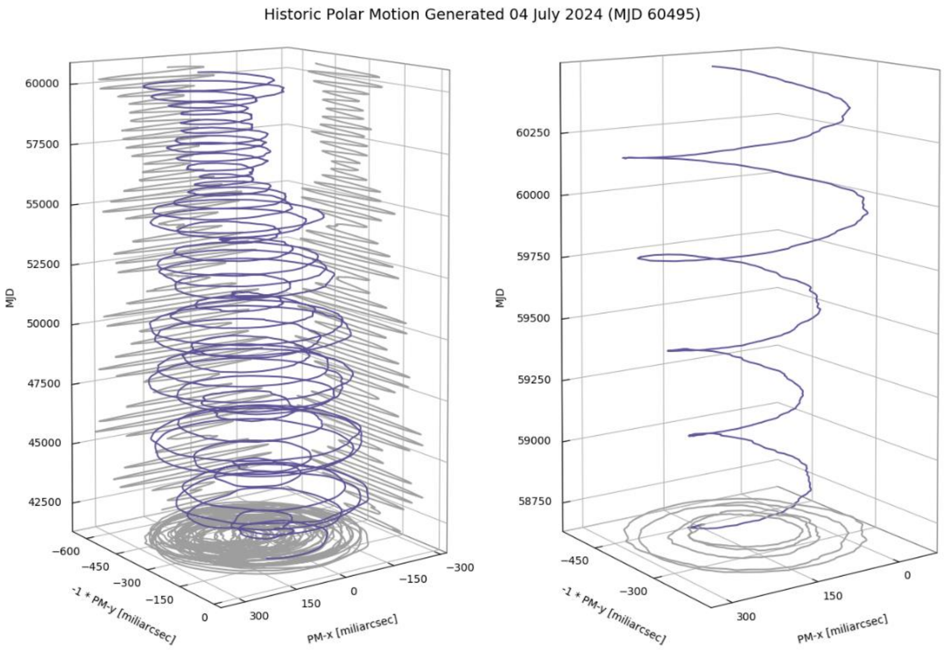
The results confirm the thesis that Earth Rotation Parameters prediction using simple kriging and ordinary kriging, ARIMA, VAR models, and Dynamic Mode Decomposition provides accuracy comparable to, or better than, the most accurate current methods. In a broader perspective, constant algorithms development and incorporation of additional information on physical phenomena affecting Earth’s rotation are essential.

Spis treści

1. Wstęp

Parametry Orientacji Ziemi (z ang. Earth Orientation Parameters – EOP) to zbiór parametrów opisujących nieregularności w orientacji i ruchu obrotowym Ziemi (Bizouard i in., 2019). Na Parametry Orientacji Ziemi składają się: współrzędne bieguna ziemskiego (PMx, PMy), różnica czasów UT1-UTC oraz przesunięcie bieguna niebieskiego (dX, dY). Dodatkowym parametrem, który również opisuje zmiany w ruchu obrotowym Ziemi jest pierwsza ujemna pochodna po czasie UT1-UTC – długość doby (z ang. Length of Day – LOD) (Modiri i in., 2020). Podzbiorem Parametrów Orientacji Ziemi są Parametry Rotacji Ziemi (z ang. Earth Rotation Parameters – ERP), do których zalicza się współrzędne bieguna ziemskiego oraz różnicę czasów UT1-UTC (bądź LOD).

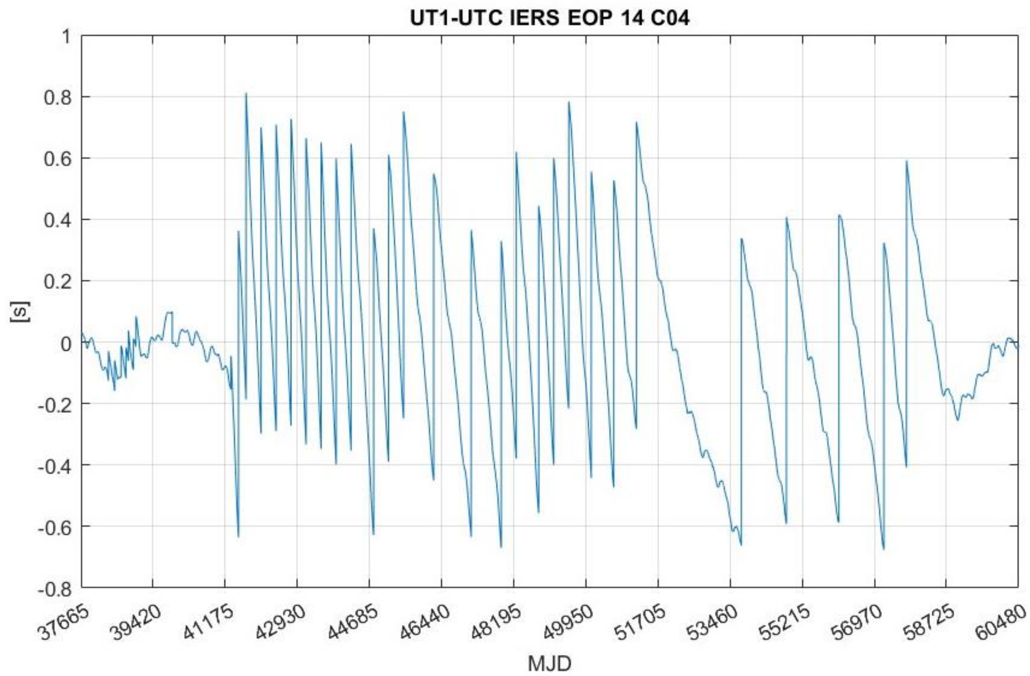
Współrzędne bieguna ziemskiego opisują położenie osi obrotu Ziemi względem skorupy ziemskiej. PMx definiowany jest wzdłuż południka 0° (Greenwich), zaś PMy wzdłuż długości geograficznej 90° zachodniej. Ruch bieguna ziemskiego, czyli zmiana jego współrzędnych w czasie, charakteryzuje się dwoma głównymi składowymi: oscylacją Chandlera (z okresem około 435 dni) oraz oscylacją roczną (Rysunek 1) (Gross, 2007). Zmiana współrzędnych spowodowana jest między innymi ruchami w jądrze i płaszczu Ziemi, redystrybucją masy wodnej, głównie topnieniem lądolodu Grenlandii, a także przez ruchy izostatyczne (Lambeck, 1989). Biegun ziemski porusza się przeciwnie do ruchu wskazówek zegara wokół swojej średniej pozycji, z nieregularnym dryfem w kierunku 80° na zachód (https://www.iers.org/IERS/EN/Science/EarthRotation/PolarMotion.html).



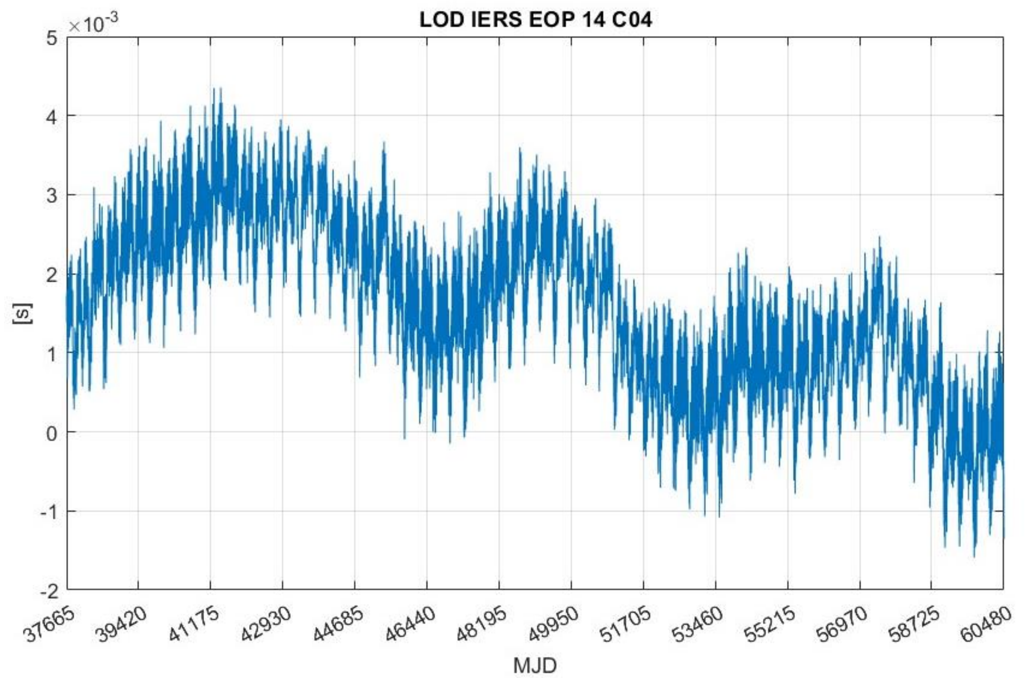
Rysunek 1 Zmiana współrzędnych bieguna ziemskiego w czasie. Rysunek po lewej stronie pokazuje ruch bieguna ziemskiego od 1973 roku, zaś rysunek po prawej przedstawia ruch bieguna podczas ostatnich pięciu lat (2019-2024). Źródło: https://maia.usno.navy.mil/information/what-is-eop.

Różnica czasów UT1-UTC (Rysunek 2) określa relację pomiędzy czasem astronomicznym opartym na ruchu obrotowym Ziemi (UT1) i czasem atomowym, bazującym na precyzyjnych zegarach atomowych (UTC). Długość doby (Rysunek 3) to różnica między satelitarnie wyznaczanym czasem trwania obrotu Ziemi wokół własnej osi a nominalną wartością 86400 sekund. Zmiany w UT1-UTC oraz LOD wynikają z różnych procesów zachodzących na Ziemi, np. zmian pływowych związanych z grawitacyjnym oddziaływaniem Słońca i Księżyca

(McCarthy i Luzum, 1993), czy też czynników pozapływowych, takich jak wymiana momentu pędu między Ziemią a atmosferą (Hide i in., 1980). Dodatkowo, istotny wpływ na prędkość obrotu Ziemi mają tropikalne monsuny występujące podczas zjawiska El Niño (Gross i in., 1996; Niedzielski i Kosek, 2008).

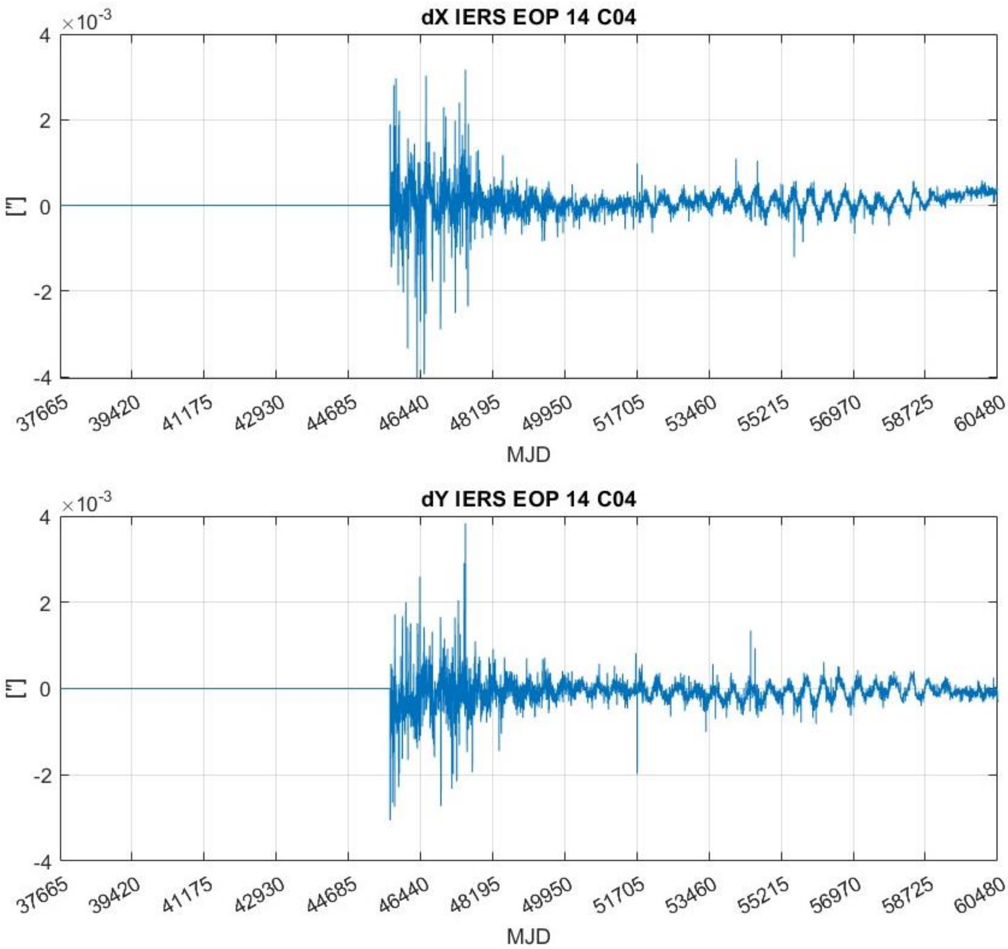


Rysunek 2 Zmiana różnicy czasów UT1-UTC w czasie (01.01.1962 – 19.06.2024). Na wykresie widoczne są charakterystyczne skoki w szeregu czasowym, kiedy to dodawana była do UT1-UTC sekunda przestępna. Sekunda przestępna, wprowadzana przez IERS (Bulletin C), ma zapewniać aby różnica między czasami UTC i UT1 nie przekroczyła 0.9 s (Nelson i in., 2001).



Rysunek 3 Zmiana długości doby (LOD) w czasie (01.01.1962 – 19.06.2024). Szereg czasowy LOD charakteryzuje się składowymi okresowymi, tj. składową roczną, półroczną czy składowymi związanymi z pływami strefowymi.

Różnica pomiędzy przewidywanymi współrzędnymi bieguna niebieskiego, wyznaczanymi przez modele precesji i nutacji, a jego rzeczywistym położeniem, określana jest mianem przesunięcia bieguna niebieskiego, dX i dY (z ang. Celestial Pole Offsets – CPO) (Nastula i in., 2020) (Rysunek 4). CPO reprezentuje dwuwymiarowy ruch, który w głównej mierze zdominowany jest przez wsteczny ruch oscylacyjny, określany jako swobodna nutacja jądra Ziemi (z ang. Free Core Nutation – FCN) (Kiani Shahvandi i in., 2024). Główną przyczyną występowania FCN jest brak współosiowości między osią obrotu płynnego rdzenia a płaszczem Ziemi, przy czym sugeruje się, iż ruch ten jest indukowany przez oddziaływanie procesów oceanicznych i atmosferycznych (Sasao i Wahr, 1981).



Rysunek 4 Zmiana przesunięcia bieguna niebieskiego, dX i dY w czasie (01.01.1962 – 19.06.2024).

Parametry Orientacji Ziemi wyznaczane są za pomocą kombinacji różnych satelitarnych i kosmicznych technik geodezyjnych, takich jak: DORIS (Doppler Orbitography and Radiopositioning Integrated by Satellite), VLBI (Very Long Baseline Interferometry), GNSS (Global Navigation Satellite System) i SLR (Satellite Laser Ranging) (Petit i Luzum, 2010). Połączenie powyższych technik daje możliwość wyznaczenia współrzędnych bieguna ziemskiego z dokładnością do około 50 μas (30 μas – w okresie ostatnich 365 dni; https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html; stan na 08.10.2024), w przypadku UT1-UTC z dokładnością około 30 μs (12 μs – w okresie ostatnich 365 dni; https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html; stan na 08.10.2024), zaś w przypadku LOD dokładność wyznaczenia oscyluje w granicach około 10 μs (8 μs – w okresie ostatnich 365 dni;

https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html; stan na 08.10.2024) (Dick i Thaller, 2020). Przesunięcie bieguna ziemskiego wyznaczane jest jedynie za pomocą techniki VLBI, a dokładność jego wyznaczania wynosi około 130 µas dla dX oraz 130 µas dla dY (https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html; stan na 08.10.2024).

Znajomość Parametrów Orientacji Ziemi jest niezbędna do transformacji między niebieskim (z ang. Celestial Reference Frame – CRF) a ziemskim układem odniesienia (z ang. Terrestial Reference Frame – TRF). Ponadto, są one wykorzystywane do wielu zadań, zarówno geodezyjnych, jak i astronomicznych, do których należą: precyzyjne pozycjonowanie w przestrzeni i na powierzchni Ziemi, nawigacja przez Deep Space Network (DSN) czy realizacja systemów czasu (Gambis i Luzum, 2011).

Złożoność technik pomiarowych oraz przetwarzania danych związanych z wyznaczaniem EOP uniemożliwia uzyskanie ich w czasie rzeczywistym. Finalne wersje szeregów czasowych zawierających Parametry Orientacji Ziemi publikowane są z około 30-dniowym opóźnieniem, między innymi przez Międzynarodową Służbę Ruchu Obrotowego Ziemi i Systemów Odniesienia (z ang. International Earth Rotation and Reference Systems Service – IERS) oraz wiele innych indywidualnych centrów i ośrodków analitycznych (IERS Annual Report 2018). Luka w danych, spowodowana wspomnianym opóźnieniem w ich publikacji, w podejściu operacyjnym musi zostać uzupełniona za pomocą tak zwanych szybkich produktów, publikowanych przez takie ośrodki jak: IERS, IGS (International GNSS Service), CODE (Center for Orbit Determination in Europe) czy GFZ (German Research Centre for Geosciences).

Na przestrzeni lat do prognozy Parametrów Orientacji Ziemi stosowano wiele metod, w tym: kolokację metodą najmniejszych kwadratów (Zotov, 2017), autoregresję (Kosek i in., 2007), sztuczne sieci neuronowe (Schuh i in., 2002; Gou i in., 2023), kombinację metody najmniejszych kwadratów i autoregresji z wykorzystaniem danych EAM (Dill in in., 2019), filtr Kalmana (Nastula i in., 2020) czy uczenie maszynowe (Kiani Shahvandi i in., 2022; Lei i in., 2017).

W latach 2006-2008, w celu zwiększenia dokładności prognozy EOP, przeprowadzona została, pod auspicjami IERS, Pierwsza Kampania Porównawcza Prognozy Parametrów Orientacji Ziemi (z ang. First Earth Orientation Parameter Prediction Comparison Campaign – 1st EOP PCC) (Kalarus i in., 2010). Kampania ta umożliwiła naukowcom z różnych krajów i instytutów badawczych porównanie dokładności prognoz EOP uzyskanych różnymi metodami predykcji. Dodatkowo dostarczyła ona wielu cennych spostrzeżeń, dotyczących trudności w prognozowaniu Parametrów Orientacji Ziemi. Metody prognozy wraz z upływem lat stawały się coraz dokładniejsze, dlatego we wrześniu 2021 roku rozpoczęta została Druga Kampania Porównawcza (2nd EOP PCC) (http://eoppcc.cbk.waw.pl/) i trwała do końca 2022 roku.

W wielu zadaniach geodezyjnych i astronomicznych, na przykład przy planowaniu misji kosmicznych, wymagana jest znajomość opisywanych parametrów zarówno w czasie rzeczywistym jak i w przyszłość (Schuh i in., 2002). Rozwiązaniem tego problemu jest ich prognoza, dlatego tak ważne są ciągłe badania związane z rozwojem metod prognozujących, a w związku z tym zwiększanie dokładności przewidywania EOP.

2. Wykaz publikacji będących podstawą rozprawy doktorskiej

Prezentowana rozprawa doktorska przygotowana została w formie spójnego tematycznie cyklu 6 publikacji naukowych. Zostały one opublikowane w czasopismach znajdujących się w wykazie czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych, prowadzonym przez Ministra Edukacji i Nauki.

Niniejsza rozprawa obejmuje następujące prace, które tworzą zbiór powiązanych tematycznie artykułów naukowych:

1. Michalczak, M., Ligas, M., (2021). Kriging-based prediction of the Earth's pole coordinates. Journal of Applied Geodesy, vol. 15, no. 3, pp. 233-241, https://doi.org/10.1515/jag-2021-0007
   - artykuł opublikowany 9 czerwca 2021 roku w czasopiśmie Journal of Applied Geodesy; wskaźnik cytowani czasopisma (ang. Impact Factor): 1.2; punktacja wg wykazu MNiSW: 100;

2. Michalczak, M., Ligas, M. (2022). The (ultra) short term prediction of length-of-day using kriging. Advances in Space Research, vol. 70, pp. 610-620, https://doi.org/10.1016/j.asr.2022.05.007
   - artykuł opublikowany 1 sierpnia 2022 roku w czasopiśmie Advances in Space Research; wskaźnik cytowani czasopisma (ang. Impact Factor): 2.8; punktacja wg wykazu MNiSW: 70;

3. Michalczak, M., Ligas, M., Kudrys, J. (2022). Prediction of Earth Rotation Parameters with the use of rapid products from IGS, CODE and GFZ data centers using ARIMA and kriging – a comparison. Artificial Satellites, Proceedings of the Second Earth Orientation Parameters Prediction Comparison Campaign (2nd EOP PCC) Workshop, vol.57, no.s1, 2022, pp. 274-289, https://doi.org/10.2478/arsa-2022-0024
   - artykuł opublikowany 5 stycznia 2023 roku w czasopiśmie Artificial Satellites; wskaźnik cytowani czasopisma (ang. Impact Factor): 0.7; punktacja wg wykazu MNiSW: 70;

4. Michalczak, M., Ligas, M. (2024). Short-term prediction of UT1-UTC and LOD via Dynamic Mode Decomposition and combination of least-squares and vector autoregressive model. Reports on Geodesy and Geoinformatics, vol. 117, pp. 45-54. https://doi.org/10.2478/rgg-2024-0006
   - artykuł opublikowany 6 marca 2024 roku w czasopiśmie Reports on Geodesy and Geoinformatics; wskaźnik cytowani czasopisma (ang. Impact Factor): 0.3; punktacja wg wykazu MNiSW: 70;

5. Ligas, M., Michalczak, M. (2024). Dynamic mode decomposition and bivariate autoregressive short-term prediction of Earth rotation parameters. Journal of Applied Geodesy, vol. 18, no. 2, pp. 211-221. https://doi.org/10.1515/jag-2023-0030

- artykuł opublikowany 5 września 2023 roku w czasopiśmie Journal of Applied Geodesy; wskaźnik cytowani czasopisma (ang. Impact Factor): 1.2; punktacja wg wykazu MNiSW: 100;

6. Michalczak, M., Ligas, M., Belda, S., Ferrándiz, J., Modiri, S. (2024). Advancing polar motion prediction with derivative information. Journal of Applied Geodesy. https://doi.org/10.1515/jag-2024-0046
   - artykuł opublikowany online 22 sierpnia 2024 roku w czasopiśmie Journal of Applied Geodesy; wskaźnik cytowani czasopisma (ang. Impact Factor): 1.2; punktacja wg wykazu MNiSW: 100;

W pierwszym artykule (Michalczak i Ligas, 2021) zaproponowano 2 nowe metody prognozy współrzędnych bieguna ziemskiego – kriging prosty i kriging zwyczajny. Bazując na dwóch referencyjnych szeregach czasowych, IERS EOP 05 C04 (IAU2000A) i IERS EOP 14 C04 (IAU2000A), przeprowadzono ultrakrótkoterminową prognozę (15 dni w przód) współrzędnych bieguna ziemskiego wykorzystując wcześniej wspomniane metody krigingu. Ponadto porównano uzyskane dokładności prognoz z wynikami osiągniętymi podczas Pierwszej Kampanii Porównawczej Prognoz Parametrów Orientacji Ziemi.

Drugi artykuł (Michalczak i Ligas, 2022) przedstawia wyniki ultrakrótkoterminowej (15 dni w przyszłość) i krótkoterminowej (30 dni w przyszłość) prognozy długości doby, z wykorzystaniem krigingu prostego i krigingu zwyczajnego. W badaniach wykorzystano dwie metody wyboru modelu semiwariogramu teoretycznego w procesie predykcyjnym. Pierwsza polegała na zastosowaniu jednego modelu podczas całego procesu predykcyjnego, podczas gdy w drugim podejściu przed każdą prognozą wybierany był semiwariogram (z wcześniej zadanej grupy) z najlepszym dopasowaniem do semiwariogramu empirycznego (z najmniejszą sumą kwadratów różnic pomiędzy semiwariogramem teoretycznym a empirycznym).

W trzeciej pracy (Michalczak i in., 2022) zaprezentowano wyniki prognozy trzech Parametrów Rotacji Ziemi – współrzędnych bieguna ziemskiego (PMx i PMy) oraz długości doby (LOD). W procesie predykcyjnym (na 15 i 30 dni w przyszłość) zastosowano metodę krigingu zwyczajnego oraz autoregresyjny zintegrowany model średniej ruchomej (z ang. AutoRegressive Integrated Moving Average – ARIMA). W pracy porównano również dokładność prognozy ERP z wykorzystaniem szybkich produktów z różnych ośrodków badawczych (IERS, CODE, IGS, GFZ).

Czwarty artykuł (Michalczak i Ligas, 2024) przedstawia dwie metody prognozy – kombinację wpasowania modelu trendu i składowych okresowych metodą najmniejszych kwadratów oraz wektorowej autoregresji na resztach modelu (z ang. combination of least-squares and vector autoregressive model – LS+VAR) a także, niewykorzystywaną wcześniej w tym celu, dynamiczną dekompozycję modalną (z ang. Dynamic Mode Decomposition – DMD). W pracy wykonano krótkoterminowe (30 dni w przyszłość) prognozy parametrów UT1-UTC i długości doby. Dodatkowo przy prognozie metodą LS+VAR wykorzystano szeregi czasowe wspomagające prognozę EOP, atmosferyczne momenty pędu (z ang. Atmospheric Angular Momentum - AAM). W badaniach zastosowano dwa podejścia wyboru parametrów sterujących. Podejście adaptacyjne polegało na wykonaniu wstępnej prognozy, mającej za zadanie wybrać zestaw parametrów do ostatecznej prognozy. W drugim podejściu, uśrednionym, wykonywane było kilka prognoz z różnymi zestawami parametrów sterujących, a następnie średnia z tych prognoz stanowiła ostateczną predykcję.

W piątej pracy (Ligas i Michalczak, 2024) wykorzystano wcześniej zaprezentowane metody, DMD i LS+VAR, do prognozy współrzędnych bieguna ziemskiego (DMD i LS+VAR) oraz długości doby (DMD). Badanie obejmowało horyzont prognozy 30-dniowej z przesunięciem 7-dniowym, przeprowadzonym osobno dla lat 2017–2022. Uzyskane wyniki porównano z wcześniej stosowanymi przez autorów metodami, krigingiem zwyczajnym i ARIMA, a także z dwoma najdokładniejszymi metodami zarejestrowanymi w Pierwszej Kampanii Porównawczej Prognozy Parametrów Orientacji Ziemi.

Szósty artykuł (Michalczak i in., 2024) prezentuje wykorzystanie kombinacji modelu wpasowania trendu i składowych okresowych metodą najmniejszych kwadratów i wektorowej autoregresji na resztach modelu (LS+VAR) do krótkoterminowej (30 dni w przód) prognozy współrzędnych bieguna Ziemskiego (PMx i PMy). Do procesu predykcyjnego wprowadzono dodatkowe szeregi czasowe zawierające informację o pochodnych obu współrzędnych, które potencjalnie powinny wnieść do procesu dodatkowe informacje wspomagające prognozę.

3. Cel i teza rozprawy doktorskiej

Głównym celem niniejszej rozprawy doktorskiej było zastosowanie wybranych metod geostatystycznych, takich jak kriging prosty i kriging zwyczajny, a także jedno- (ARIMA) i wielozmiennych (VAR) metod analizy szeregów czasowych do prognozy Parametrów Rotacji Ziemi.

W ramach przeprowadzonych prac badawczych zrealizowano również szereg dodatkowych celów, do których należą:

- zastosowanie dynamicznej dekompozycji modalnej (z ang. Dynamic Mode Decomposition – DMD) do prognozy Parametrów Rotacji Ziemi,
- wykorzystanie danych o Efektywnych Momentach Pędu (EAM) do prognozy parametrów ERP,
- opracowanie algorytmów predykcyjnych działających w trybie operacyjnym, umożliwiających uczestnictwo w Drugiej Kampanii Porównawczej Prognozy Parametrów Orientacji Ziemi,
- porównanie osiągniętych wyników z wynikami dostępnymi w literaturze oraz rezultatami Drugiej Kampanii Porównawczej Prognozy Parametrów Orientacji Ziemi.

Na podstawie badań literaturowych sformułowana została następująca teza:

„Prognoza Parametrów Rotacji Ziemi z wykorzystaniem krigingu prostego i krigingu zwyczajnego, jedno- i wielozmiennych metod analizy szeregów czasowych (ARIMA, VAR) oraz dynamicznej dekompozycji modalnej pozwala na uzyskanie prognoz o dokładności wyższej lub porównywalnej w stosunku do obecnie stosowanych, najdokładniejszych metod predykcyjnych.”

4. Metodyka badawcza

Rozdział ten został podzielony na dwie części, pierwszą zawierającą opisy wykorzystanych metod prognozy oraz drugą dotyczącą opisu danych i procesu predykcyjnego.

4.1. Wykorzystane metody prognozy

4.1.1. Procedura usunięcia trendu i składowych okresowych

W ogólnym założeniu całkowity model predykcji Parametrów Rotacji Ziemi składa się z dwóch części – deterministycznej i stochastycznej (za wyjątkiem DMD). Pierwsza z nich obejmuje estymowany trend liniowy oraz estymowane składowe okresowe, a w przypadku prognozy UT1-UTC oraz LOD są to również korekty związane z pływami strefowymi (Petit i Luzum, 2010). Składowe te są następnie ekstrapolowane w przyszłość i uwzględniane w procesie predykcyjnym (za wyjątkiem DMD). Część stochastyczna określa prognozowaną wartość residuów.

Ogólny model predykcji przedstawia się następująco:

$$\begin{gathered} Prognoza = trend\ liniowy + składowe\ okresowe + \\ \{korekty\ pływów\ strefowych\ (UT1 - UTC/LOD)\} + prognozowane\ „reszty” \end{gathered} \tag{1}$$

lub

$$v(t_p) = a_0^v + a_1^v t_p + \sum_{i=1}^{n} b_i^v \sin(c_i^v t_p + d_i^v) + \{dUT/dLOD\} + \varepsilon_{vv}(t_p) \tag{2}$$

gdzie: v to prognozowany parametr, t_p opisuje moment prognozy w czasie, a_0^v i a_1^v to współczynniki opisujące estymowany trend liniowy, b_i^v, c_i^v oraz d_i^v są estymowanymi wartościami amplitudy, częstotliwości i przesunięcia fazowego dla i-tej składowej okresowej, $\{dUT/dLOD\}$ to korekty związane z pływami strefowymi uwzględniane przy prognozie UT1-UTC i/lub LOD, zaś ε_{vv} jest prognozowaną wartością residuów.

Po usunięciu trendu liniowego i składowych okresowych, a w przypadku prognozy UT1-UTC oraz LOD także korekt wynikających z wpływu pływów strefowych, uzyskano szereg czasowy przedstawiający wartości residuów. Uzyskany w ten sposób szereg zostawał poddany prognozie za pomocą krigingu prostego, krigingu zwyczajnego, modelu ARIMA i VAR.

4.1.2. Kriging prosty i kriging zwyczajny

Kriging to grupa geostatystycznych metod prognozy, optymalnych w sensie nieobciążoności i minimalizacji średniokwadratowego błędu prognozy (Cressie, 1993; Stein, 1999). W badaniach wykorzystane zostały dwa liniowe warianty krigingu – kriging prosty i kriging zwyczajny.

Predyktor krigingu prostego w przypadku pojedynczej prognozy ma postać:

$$\hat{Z}_0 = \mu + \boldsymbol{\omega}^{\mathrm{T}} \boldsymbol{\Omega}^{-1} (\mathbf{Z} - \mathbf{u}\mu) \tag{3}$$

gdzie: $\hat{Z}_0$ określa prognozowaną wartość, μ jest stałą i znaną wartością oczekiwaną pola losowego, $\boldsymbol{\omega}$ to wektor kowariancji (obserwacje – prognoza), $\boldsymbol{\Omega}$ jest macierzą kowariancji (obserwacje – obserwacje), $\mathbf{Z}$ to wektor wartości obserwowanych, $\mathbf{u}$ to wektor jedynek.

W przypadku krigingu zwyczajnego postać predyktora przyjmuje poniższą postać:

$$\hat{Z}_0 = \hat{\mu} + \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}(\mathbf{Z} - \mathbf{u}\hat{\mu}) \tag{4}$$

,a stała i znana wartość oczekiwana pola losowego w krigingu prostym μ jest zastąpiona przez jej oszacowanie uogólnioną metodą najmniejszych kwadratów:

$$\hat{\mu} = (\mathbf{u}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{u})^{-1}\mathbf{u}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{Z} \tag{5}$$

Zarówno w przypadku prognozy współrzędnych bieguna ziemskiego (Michalczak i Ligas, 2021) oraz długości doby (Michalczak i Ligas, 2022), testowano różne modele semiwariogramów teoretycznych (sferyczny, kubiczny, eksponencjalny, gaussowski, sinusowy z „efektem dziury”, Markowa rzędu 2, Markowa rzędu 3 czy pentasferyczny). Ostatecznie dla prognozy PMx i PMy wybrany został model Markowa rzędu 2, zaś w przypadku prognozy długości doby wybrane zostały trzy modele, tj. kubiczny, eksponencjalny oraz sinusowy z „efektem dziury” (Olea, 1999). Poniżej przedstawiono modele wybranych semiwariogramów:

a) Markowa rzędu 2

$$\gamma(\tau, c_0, c, a) = \begin{cases} 0 & \tau = 0 \\ c_0 + c\left[1 - \left(1 + \frac{\tau}{a}\right)e^{-\left(\frac{\tau}{a}\right)}\right] & \tau > 0 \end{cases} \tag{6}$$

b) kubiczny

$$\gamma(\tau, c_0, c, a) = \begin{cases} 0 & \tau = 0 \\ c_0 + c\left[7\left(\frac{\tau}{a}\right)^2 - \frac{35}{4}\left(\frac{\tau}{a}\right)^3 + \frac{7}{2}\left(\frac{\tau}{a}\right)^5 - \frac{3}{4}\left(\frac{\tau}{a}\right)^7\right] & 0 < \tau \le a \\ c_0 + c & \tau > a \end{cases} \tag{7}$$

c) eksponencjalny

$$\gamma(\tau, c_0, c, a) = \begin{cases} 0 & \tau = 0 \\ c_0 + c\left[1 - e^{-\left(\frac{\tau}{a}\right)}\right] & \tau > 0 \end{cases} \tag{8}$$

d) sinusowy z „efektem dziury”

$$\gamma(\tau, c_0, c, a) = \begin{cases} 0 & \tau = 0 \\ c_0 + c\left[1 - \frac{sin\left(\frac{\pi\tau}{a}\right)}{\frac{\pi\tau}{a}}\right] & \tau > 0 \end{cases} \tag{9}$$

gdzie: τ to opóźnienie czasowe (odległość między obserwacjami), c_0 jest określane jako efekt samorodków (wariancja błędu pomiarowego), c to wartość częściowego progu (wariancja sygnału), zaś a to zakres autokorelacji.

Semiwariogram empiryczny dla szeregów czasowych (regularnie rozmieszczonych) prezentuje się następująco:

$$\hat{\gamma}(\tau) = \frac{1}{2(N-\tau)} \sum_{i=1}^{N-\tau} (Z_i - Z_{i+\tau})^2 \tag{10}$$

gdzie: N to liczba wartości obserwowanych.

Ostatecznie opisane modele teoretyczne zostały dopasowane do semiwariogramów empirycznych za pomocą nieliniowej metody najmniejszych kwadratów z ograniczeniami, spełniającymi wymaganie, aby wszystkie parametry (c_0, c, a) były nieujemne.

4.1.3. ARIMA

Autoregresyjny zintegrowany model średniej ruchomej to połączenie trzech elementów: procesu autoregresyjnego (AR), integracji procesu (I) oraz procesu średniej ruchomej (MA). Połączenie tych trzech procesów umożliwia uzyskanie większej elastyczności w dopasowywaniu modelu do rzeczywistych szeregów czasowych (Box i Jenkins, 1976).

a) Proces autoregresyjny (AR)

Proces autoregresyjny jest modelem, w którym każda obserwacja jest wyrażona jako liniowa kombinacja wcześniejszych obserwacji. Rząd procesu (p) określa, ile wcześniejszych wartości jest uwzględnianych przy wyznaczaniu bieżącej wartości. Proces AR(p) przedstawia się jako:

$$\hat{X}_t = \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t \tag{11}$$

gdzie: $\hat{X}_t$ to wartość szeregu w chwili t, p to rząd procesu, φ_i jest wartością określającą wpływ obserwacji poprzedniej na obserwację w chwili t, X_{t-i} to wartość szeregu w chwili $t - i$, zaś ε_t jest składnikiem losowym, zaburzeniem w chwili t.

b) Integracja procesu (I)

Integracja procesu jest odpowiedzialna za usunięcie niestacjonarności szeregu poprzez różnicowanie danych. Proces ten wykonywany jest do momentu, aż test KPSS (Kwiatkowski–Phillips–Schmidt–Shin) potwierdzi hipotezę o stacjonarności szeregu. Stopień integracji (d) określany jest jako krotność zastosowanych różnic.

$$\Delta X_t^i = X_t - X_{t-1} \tag{12}$$

gdzie: ΔX_t^i to i-ta różnica w chwili t, parametr i określa krotność różnicowania.

c) Proces średniej ruchomej (MA)

Proces średniej ruchomej opisuje związek między obserwacjami a wahaniami przypadkowymi (zaburzeniami) w chwili obecnej i wcześniejszych. Rząd procesu (q) określa, ile wcześniejszych zaburzeń jest uwzględnianych przy wyznaczaniu bieżącej wartości. Proces MA(q) przedstawia się jako:

$$\hat{X}_t = \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{13}$$

gdzie: θ_j to wartość określająca wpływ poprzedniego zaburzenia na obserwację w chwili t, q to rząd procesu.

Ostatecznie model ARIMA definiowany jest jako (Box i Jenkins, 1976):

$$\hat{X}_t^d = c + \sum_{i=1}^{p} \varphi_i X_{t-i}^d + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{14}$$

gdzie: $\hat{X}_t^d$ to prognozowana wartość szeregu, d to stopień integracji, c to stała.

Model ARIMA zapisywany jest jako ARIMA(p, d, q).

4.1.4. VAR

Model wektorowej autoregresji to wielozmienny odpowiednik jednozmiennego modelu autoregresji AR. Model łączy bieżące wartości zmiennych z ich przeszłymi wartościami oraz z przeszłymi wartościami innych zmiennych, które potencjalnie mogą oddziaływać na siebie w czasie. Rząd procesu (p) określa, ile wcześniejszych wartości jest uwzględnianych przy wyznaczaniu bieżącej wartości. VAR(p) definiowane jest jako (Hatemeni, 2004):

$$\hat{\mathbf{X}}_t = \mathbf{c} + \mathbf{A}_1 \mathbf{X}_{t-1} + \ldots + \mathbf{A}_p \mathbf{X}_{t-p} + \boldsymbol{\varepsilon}_t \tag{15}$$

gdzie: $\hat{\mathbf{X}}_t$ to wektor przyszłe wartości (n x 1, gdzie n to liczba zmiennych prognozowanych), $\boldsymbol{c}$ to wektor stałych (n x 1), $\mathbf{A}_1 \ldots \mathbf{A}_p$ to macierze współczynników (n x n), $\mathbf{X}_{t-1} \ldots \mathbf{X}_{t-p}$ to wektory szeregów czasowych biorących udział w procesie predykcyjnym (n x 1), $\boldsymbol{\varepsilon}_t$ to wektor zakłóceń (składniki losowe) (n x 1).

4.1.5. DMD

Dynamiczna dekompozycja modalna to metoda analizy dynamiki systemów nieliniowych, która pozwala na wyodrębnienie wzorców oraz prognozowanie przyszłych stanów systemu (Schmid, 2010). Rekonstrukcja i prognoza odbywa się w jednej procedurze obliczeniowej.

W DMD przyszłe stany systemu łączą się z przeszłymi stanami poprzez nieznany operator $\mathbf{A}$, który przechowuje dynamikę procesu:

$$\boldsymbol{x}_t = \mathbf{A}\boldsymbol{x}_{t-1} \tag{16}$$

gdzie: $\boldsymbol{x}_t$ to przyszły stan systemu w chwili t, $\boldsymbol{x}_{t-1}$ to przeszły stan w chwili $t-1$.

Wyznaczenie stanu przyszłego przyjmuje poniższą postać, gdy operator **A** opisywany jest przez wartości i wektory własne:

$$\boldsymbol{x}_t = \boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^{-1}\boldsymbol{x}_{t-1} \tag{17}$$

gdzie: $\boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^{-1}$ to elementy rozkładu macierzy $\mathbf{A}$ na wartości i wektory własne.

Macierz „klatek" (snapshot'ów) $\mathbf{X} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_m]$ przechowuje wszystkie zarejestrowane stany systemu w momentach czasu t = 1, 2, …, m i jest rozdzielana na dwie macierze $\mathbf{X_1} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_{m-1}]$ i $\mathbf{X_2} = [\mathbf{x}_2 \quad \mathbf{x}_3 \quad \cdots \quad \mathbf{x}_m]$, które są powiązane poprzez operator **A**:

$$\mathbf{X_2} = \mathbf{A}\mathbf{X_1} \tag{18}$$

Operator **A** jest następnie wyznaczany jako:

$$\mathbf{A} = \mathbf{X_2}\mathbf{X_1^+} \tag{19}$$

gdzie: $\mathbf{X_1^+} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^T$ to pseudoodwrotność Moore–Penrose'a wyrażona przez elementy rozkładu względem wartości szczególnych.

Niskowymiarowa reprezentacja operatora **A** za pomocą $\widetilde{\mathbf{A}}$ przedstawia się następująco:

$$\widetilde{\mathbf{A}} = \mathbf{U}^T\mathbf{A}\mathbf{U} = \mathbf{U}^T\mathbf{X_2}\mathbf{V}\mathbf{\Sigma}^{-1} \tag{20}$$

W pracy Tu i in. (2014) dowiedziono, że wartości własne macierzy $\widetilde{\mathbf{A}}$ są dokładnie takie same jak macierzy **A**:

$$\widetilde{\mathbf{A}}\mathbf{W} = \mathbf{W}\mathbf{\Lambda} \tag{21}$$

podczas gdy wektory własne macierzy **A** reprezentowane są przez:

$$\mathbf{\Phi} = \mathbf{X_2}\mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{W} \tag{22}$$

W przypadku analizy szeregów czasowych, wejściowe dane o długości T podzielone zostają na L podserii o długości K, które przesunięte są względem siebie o jeden krok czasowy do przodu, tworząc tym samym macierz **X**:

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & \cdots & x_L \\ x_2 & x_3 & \cdots & x_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_K & x_{K+1} & \cdots & x_{K+L-1} \end{bmatrix} = [\mathbf{x}'_1 \quad \mathbf{x}'_2 \quad \cdots \quad \mathbf{x}'_L] \tag{23}$$

gdzie: L jest wybierane tak aby $2 \leq L \leq \frac{T}{2}$, natomiast $K = T - L + 1$.

Następnie macierz **X** dzielona jest na dwie macierze, tak że $\mathbf{X_1} = [\mathbf{x}'_1 \quad \mathbf{x}'_2 \quad \cdots \quad \mathbf{x}'_{L-1}]$ i $\mathbf{X_2} = [\mathbf{x}'_2 \quad \mathbf{x}'_3 \quad \cdots \quad \mathbf{x}'_L]$, o wymiarach $K\ x\ (L-1)$.

Pozostałe etapy algorytmu DMD nie ulegają zmianie.

4.2. Opis danych i procesu predykcyjnego

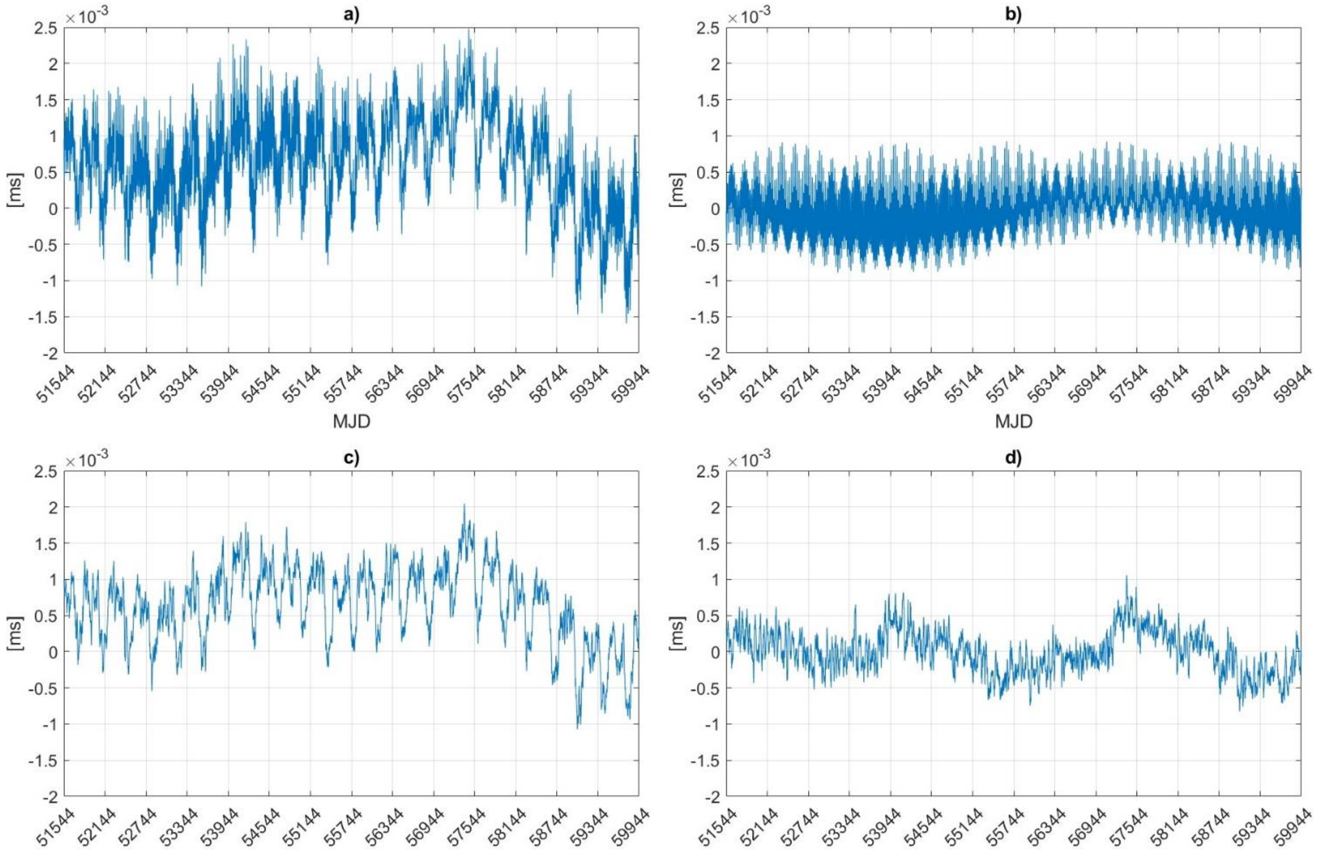
W przeprowadzonych badaniach, jako dane referencyjne, wykorzystano szeregi czasowe publikowane przez IERS, IERS EOP 05 C04 (IAU2000A), IERS EOP 14 C04 (IAU2000A) oraz IERS EOP 20 C04 (IAU2000A) dostępne na stronie https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html. Dane te są zgodne z Międzynarodowym Ziemskim Układem Odniesienia (ITRF2005, ITRF2014 oraz ITRF2020) w odniesieniu do modelu precesyjno-nutacyjnego IAU2000A. Zawierają one między innymi informacje o Parametrach Orientacji Ziemi, takich jak współrzędne bieguna ziemskiego (PMx, PMy), różnica czasów UT1-UTC, długość doby (LOD) oraz przesunięcie bieguna niebieskiego (dX, dY). Jak zostało opisane wcześniej, finalne dane EOP publikowane są z około 30-dniowym opóźnieniem, dlatego w trybie operacyjnym, niezbędne jest zastosowanie tak zwanych szybkich produktów. Do uzupełnienia luk w szeregach czasowych wykorzystano między innymi dane publikowane przez IERS (IERS Daily Rapid EOP Data (IAU2000)). Ponadto, w badaniach opisywanych w niniejszej rozprawie, zastosowano dodatkowe szeregi czasowe – efektywne momenty pędu, pochodzące z dwóch ośrodków obliczeniowych GFZ Potsdam i ETH Zurich (z niem. Eidgenössische Technische Hochschule Zürich). Dane te opisują wektor efektywnego momentu pędu Ziemi, uwzględniając takie czynniki jak ruchy mas powietrza, oceanów czy hydrosfery. Dane EAM składają się z czterech komponentów – Atmospheric Angular Momentum (AAM), Oceanic Angular Momentum (OAM), Barystatic Sea Level Angular Momentum (SLAM) oraz Hydrological Angular Momentum (HAM). Dane EAM zawierają: składowe równikowe (χ_1i χ_2), odpowiadające za ruch bieguna ziemskiego, oraz składową osiową (χ_3), odpowiadającą za zmiany w UT1-UTC, a w związku z tym również w LOD (Dill i in., 2019).

Pierwszym etapem procedury prognozy było znalezienie parametrów sterujących procesem predykcyjnym, które w zadanym oknie czasowym, pozwolą na wykonanie możliwie najdokładniejszej prognozy poszczególnych parametrów. Etap ten polegał na podzieleniu danych na dwa zestawy: treningowy (o różnej długości) i walidacyjny (zależny od długości prognozy). Następnie na tak podzielonych danych wykonywane były wstępne prognozy (wykorzystując metody opisane w rozdziale 4.1.), na podstawie których wybierano zestawy parametrów sterujących charakteryzujące się najmniejszymi wartościami błędów predykcji.
W tabeli 1 zaprezentowane zostały, w zależności od zastosowanej metody prognozy, poszukiwane parametry sterujące. Ze względu na bardzo dużą liczbę obliczeń, poszukiwanie wspomnianych parametrów obywało się z wykorzystaniem superkomputera ARES – ACK Cyfronet AGH za pośrednictwem infrastruktury PLGrid (https://plgrid.pl/).

Tabela 1 Zestawienie poszukiwanych parametrów sterujących w zależności od wykorzystanej metody prognozy.

Metoda/szukany parametr			
kriging prosty i kriging zwyczajny	**ARIMA(*p*,*d*,*q*)**	**VAR(*p*)**	**DMD**
T – długość wejściowego szeregu do estymacji trendu i składowych okresowych	*T* – długość wejściowego szeregu do estymacji trendu i składowych okresowych	*T* – długość wejściowego szeregu do estymacji trendu i składowych okresowych	*T* – długość wejściowego szeregu
NN – liczba najbliższych sąsiadów do konstrukcji układu równań	*p* – rząd procesu autoregresji	*p* – rząd procesu autoregresji	*K* – długość snapshot’u
num_waves – liczba składowych okresowych	*d* – stopień integracji	*subT* – długość podszeregu do estymacji VAR(*p*)	*L* – liczba snapshot’ów
	q – rząd procesu średniej ruchomej	*num_waves* – liczba składowych okresowych	
	num_waves – liczba składowych okresowych		

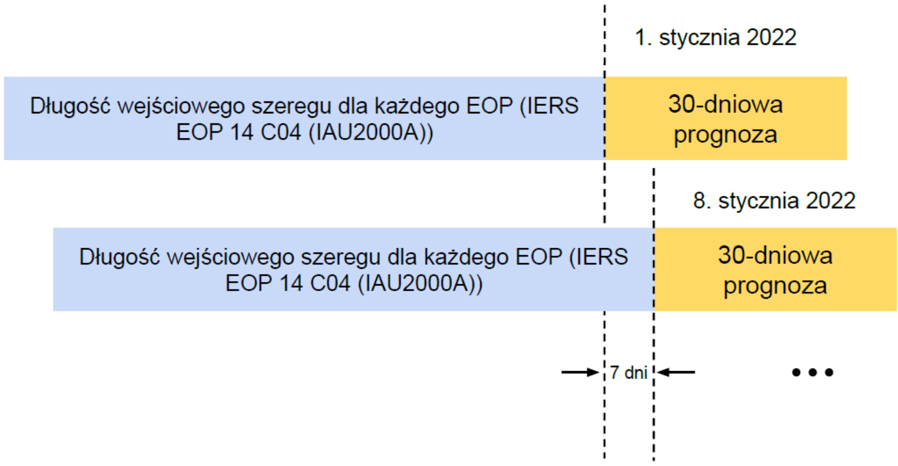
Kolejnym etapem było przeprowadzenie ostatecznej prognozy, bazującej na wcześniej wybranych parametrach sterujących. Ostatnią czynnością całego procesu było dodanie do prognozowanych reszt, ekstrapolowanych w przyszłość składowych (za wyjątkiem DMD), takich jak trend liniowy, składowe okresowe, czy w przypadku prognozy UT1-UTC i LOD, korekt związanych z pływami strefowymi. Schemat procedury przygotowania danych, na przykładzie LOD, został zaprezentowany na rysunku 5.



Rysunek 5 Szeregi czasowe przedstawiające LOD w zależności od etapu przygotowania danych wejściowych. Surowe dane LOD IERS EOP 14 C04 (a), korekty związane z pływami strefowymi (b), LOD z usuniętymi korektami związanymi z pływami strefowymi (c), szereg rezydualny LOD – po usunięciu korekt, trendu liniowego oraz składowych okresowych (d) – w przypadku LOD tak przygotowane dane stają się danymi wejściowymi do ostatecznej prognozy.

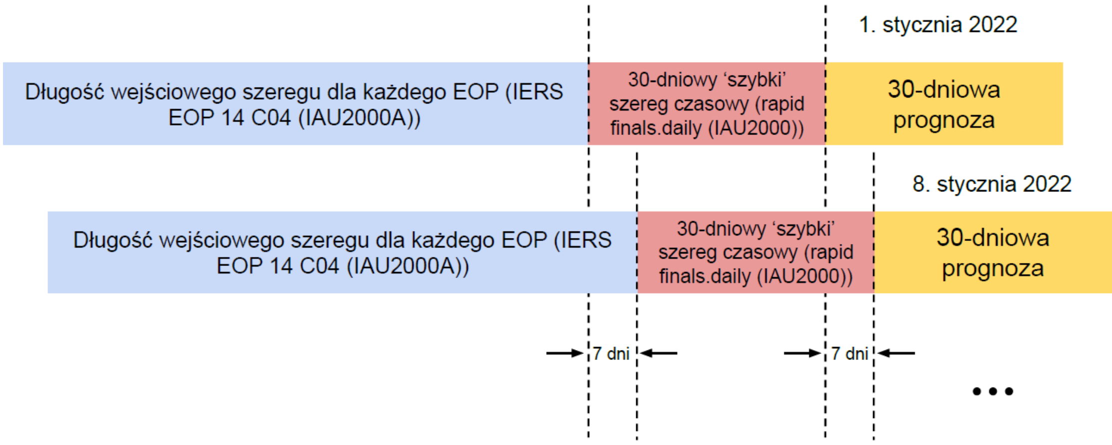
W przeprowadzonych badaniach zastosowano 2 rodzaje testów mających na celu sprawdzenie dokładności prognozy proponowanych metod:

- Offline – prognozy przeprowadzone były na danych archiwalnych/finalnych EOP (Rysunek 6).



Rysunek 6 Przykładowy schemat testu offline, bazującego na plikach archiwalnych/finalnych EOP.

- Quasi-offline – prognozy przeprowadzone były na kombinacji danych archiwalnych/finalnych i szybkich produktów EOP, mających za zadanie uzupełnić lukę w danych (Rysunek 7).



Rysunek 7 Przykładowy schemat testu quasi-offline, bazującego na kombinacji plików archiwalnych/finalnych i szybkich produktach EOP.

W przypadku wykorzystania danych EAM w procedurze predykcyjnej etapy postępowania uzupełnione są o odpowiednią transformację szeregu czasowego. W zależności od analizowanych parametrów EOP, od wcześniej przygotowanego szeregu odejmuje się odpowiednie dane EAM, uzyskując w ten sposób szereg rezydualny, który staje się szeregiem wejściowym do procedury prognozowania. Przygotowanie szeregu czasowego PMx i PMy, w celu wykorzystania danych EAM do prognozy, odbywa się poprzez zastosowanie równania Liouville'a, które opisuje relację pomiędzy geodezyjnym momentem pędu (z ang. Geodetic Angular Momentum – GAM) i ruchem bieguna ziemskiego (Gross i in., 2003; Xin i in., 2022):

$$\boldsymbol{p}(t) + \frac{i}{\sigma_c}\frac{d\boldsymbol{p}(t)}{dt} = \boldsymbol{\chi}(t) \tag{24}$$

zmienne w równaniu przedstawione są w zapisie zespolonym, tj. $\boldsymbol{\chi}(t) = \chi_1(t) + i\chi_2(t)$ oraz $\boldsymbol{p}(t) = p_1(t) - ip_2(t)$, gdzie: $\boldsymbol{\chi}$ to geodezyjny moment pędu, $\boldsymbol{p}$ opisuje ruch bieguna ziemskiego, σ_c to częstotliwość Chandlera wyznaczana jako: $\sigma_c = \frac{2\pi\left(1+\frac{i}{2Q}\right)}{T_{CW}}$, przy okresie

Chandlera $T_{CW} = 433$ dni i bezwymiarowym współczynniku dobroci (ang. quality factor) $Q = 179$ (Dill i Dobslaw, 2010).

Przekształcenie szeregów czasowych LOD oraz UT1-UTC, w celu zastosowania danych EAM, odbywa się poprzez poniższą transformację (Barnes i in., 1983):

$$LODR_{\chi_3} = -\frac{\Omega}{2\pi} LODR \tag{25}$$

gdzie: Ω to średnia prędkość kątowa Ziemi ($7.292115 \cdot 10^{-5}\ \frac{rad}{s}$), $LODR = LOD - pływy_{LOD}$ to szereg czasowy długości doby po usunięciu korekt związanych z pływami strefowymi w sekundach, $LODR_{\chi_3}$ to szereg przygotowany do dalszych analiz z uwzględnieniem EAM, w przypadku UT1-UTC szereg wejściowy $LODR = \frac{-d\,[(UT1-UTC)-pływy_{UT1-UTC}]}{dt}$.

Po przeprowadzeniu prognozy konieczne jest uzupełnienie prognozowanych wartości o uprzednio odjęte, zgodne w czasie, szeregi EAM.

5. Wyniki badań

Rozdział ten został podzielony na trzy części, w każdej z nich opisane zostały wyniki prognozy poszczególnych Parametrów Orientacji Ziemi.

Dokładność predykcji ERP została oceniona za pomocą średniego bezwzględnego błędu predykcji (z ang. Mean Absolute Prediction Error – MAPE) wyrażonego jako:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} |O_i - P_i| \tag{26}$$

gdzie: n to liczba prognoz, O_i to wartość obserwowana, P_i to prognozowana wartość.

5.1. Prognoza współrzędnych bieguna ziemskiego (PMx, PMy)

Wyniki prognozy offline współrzędnych bieguna ziemskiego na 15 dni w przód z wykorzystaniem dwóch wariantów krigingu zostały zaprezentowane w tabeli 2 (Michalczak i Ligas, 2021). Wyniki ujawniają, iż dokładność prognozy, zarówno PMx jak i PMy, z wykorzystaniem krigingu zwyczajnego, charakteryzuje się mniejszymi wartościami MAPE niż przy użyciu krigingu prostego. Błędy predykcji dla prognozy PMx z wykorzystaniem krigingu prostego są od 5 do 16 procent większe niż dla krigingu zwyczajnego, a w przypadku PMy jest to przedział od 2 do 9 procent. Kriging zwyczajny, w przypadku prognozy PMx na 15. dzień osiąga MAPE o 0.86 mas mniejsze niż kriging prosty, zaś korespondująca wartość dla PMy to 0.33 mas. Uzyskane wyniki potwierdzają znany z literatury fakt (Kiani Shahvandi i in., 2022; Xin i in., 2022), iż współrzędna PMy prognozowana jest dokładniej niż współrzędna PMx. Większa dokładność prognozy przy użyciu krigingu zwyczajnego wynika prawdopodobnie z faktu przyjęcia w procesie predykcyjnym krigingu prostego zerowej wartości oczekiwanej rezydualnego procesu losowego, podczas gdy w krigingu zwyczajnym jest ona każdorazowo szacowana.

Tabela 2 Dokładność prognozy PMx i PMy z wykorzystaniem krigingu prostego oraz zwyczajnego.

Dzień	MAPE [mas]			
	PMx		**PMy**	
	kriging zwyczajny	**kriging prosty**	**kriging zwyczajny**	**kriging prosty**
1	0.67	0.71	0.44	0.45
2	1.02	1.09	0.69	0.70
3	1.37	1.49	0.94	0.97
4	1.72	1.90	1.19	1.23
5	2.06	2.31	1.43	1.49
6	2.39	2.71	1.66	1.74
7	2.71	3.10	1.88	1.98
8	3.03	3.49	2.09	2.23
9	3.35	3.87	2.31	2.48
10	3.66	4.25	2.53	2.73
11	3.97	4.62	2.77	2.99
12	4.27	4.98	3.00	3.26
13	4.58	5.34	3.25	3.53
14	4.89	5.70	3.49	3.80
15	5.19	6.05	3.74	4.07

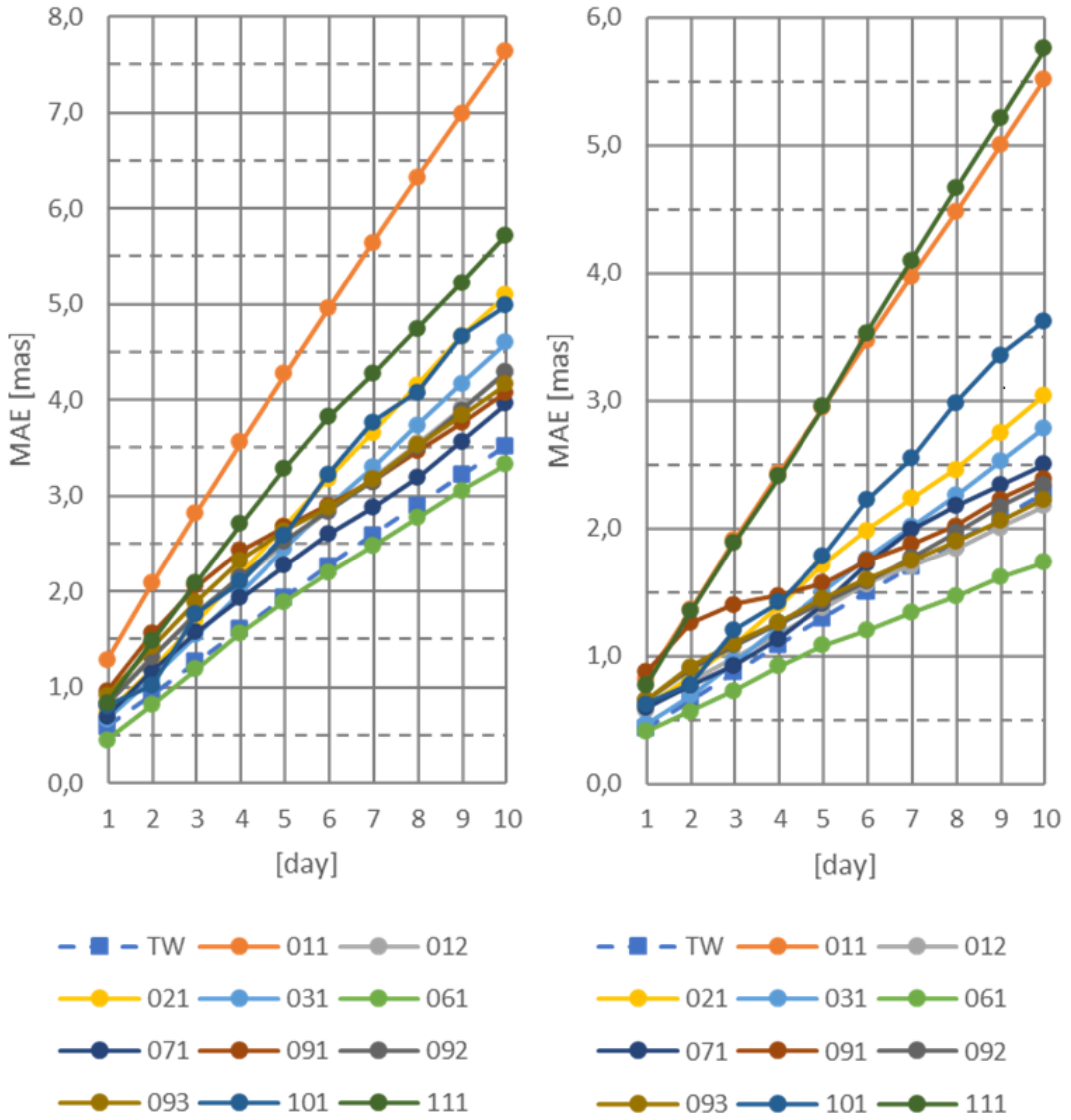
Tabela 3 przedstawia porównanie MAPE dla PMx i PMy z wykorzystaniem krigingu zwyczajnego z różnymi danymi wejściowymi, tj. IERS EOP 14 C04 i IERS EOP 05 C04. Porównanie zostało wykonane w oknie czasowym od 51544 MJD do 55929 MJD, co odpowiada dwóm tysiącom trzydziestu czterem 15-dniowych prognoz. Na podstawie poniższych wyników można zauważyć, że prognozy wykonane na szeregu IERS EOP 05 C04 charakteryzują się mniejszymi wartościami MAPE. W przypadku prognozy PMx, dokładność predykcji jest od 0.06 mas do 0.30 mas większa dla szeregu IERS EOP 05 C04, zaś w przypadku PMy wartości te mieszczą się w przedziale 0.03 mas – 0.30 mas.

Tabela 3 Porównanie MAPE dla PMx i PMy z wykorzystaniem krigingu zwyczajnego dla różnych danych wejściowych, tj. IERS EOP 14 C04 i IERS EOP 05 C04, oznaczonymi w tabeli odpowiednio jako (14) i (05).

	PMx			**PMy**		
Dzień	MAPE (14) [mas]	MAPE (05) [mas]	Różnica [mas]	MAPE (14) [mas]	MAPE (05) [mas]	Różnica [mas]
1	0.66	0.60	0.06	0.47	0.44	0.03
2	1.01	0.93	0.09	0.71	0.66	0.05
3	1.38	1.27	0.11	0.96	0.88	0.08
4	1.74	1.61	0.13	1.20	1.09	0.10
5	2.10	1.95	0.16	1.42	1.30	0.12
6	2.45	2.27	0.18	1.64	1.51	0.14
7	2.79	2.59	0.20	1.85	1.70	0.15
8	3.12	2.91	0.21	2.06	1.89	0.16
9	3.45	3.22	0.23	2.27	2.08	0.18
10	3.76	3.52	0.24	2.48	2.28	0.20
11	4.05	3.81	0.25	2.70	2.48	0.22
12	4.34	4.09	0.26	2.92	2.68	0.24

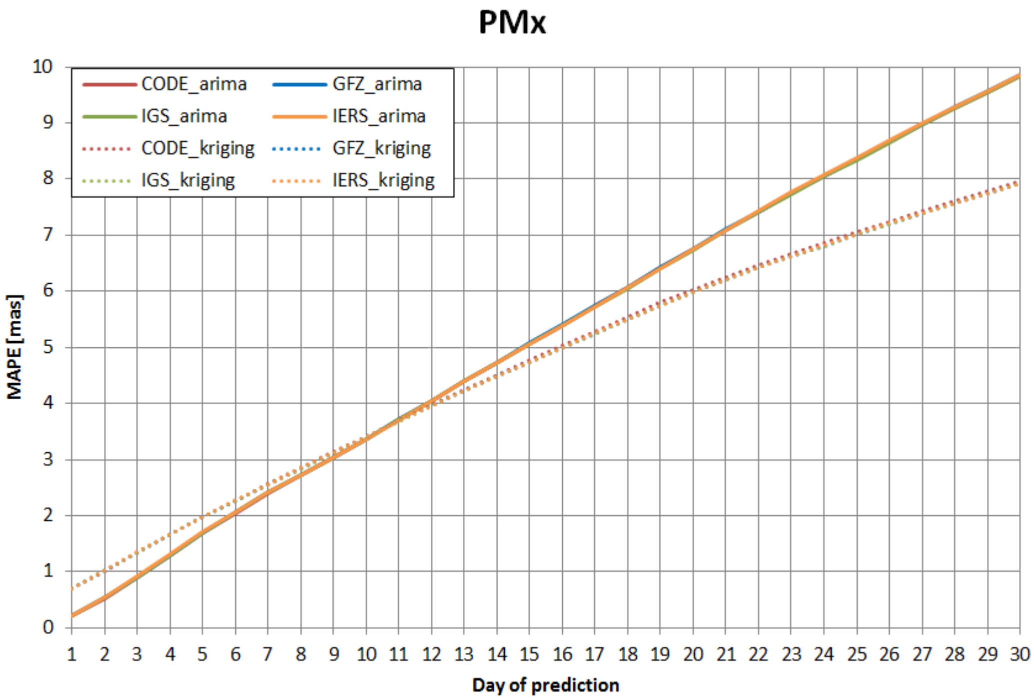
13	4.63	4.37	0.27	3.14	2.88	0.26
14	4.94	4.66	0.28	3.37	3.09	0.28
15	5.25	4.95	0.30	3.59	3.29	0.30

Na rysunku 8 zaprezentowane zostały porównania wyników prognozy offline dla 10-dniowej prognozy PMx i PMy z wykorzystaniem krigingu zwyczajnego (zaznaczone na wykresie jako TW), z wynikami uzyskanymi podczas Pierwszej Kampanii Porównawczej Prognozy Parametrów Orientacji Ziemi (oznaczenia na wykresie takie jak w Kalarus i in., 2010; w dalszej części pracy określanej jako 1st EOP PCC). Do porównania wykorzystano wyniki zaprezentowane w tabeli 3, dla szeregu IERS EOP 05 C04. Wyniki pokazują, iż dokładność prognozy współrzędnych bieguna ziemskiego z wykorzystaniem krigingu zwyczajnego jest porównywalna do najdokładniejszych metod biorących udział w 1st EOP PCC, takich jak ekstrapolacja metodą najmniejszych kwadratów połączona z prognozą modelem AR (LS extrapolation + AR), kolokacja metodą najmniejszych kwadratów czy sieci neuronowe.

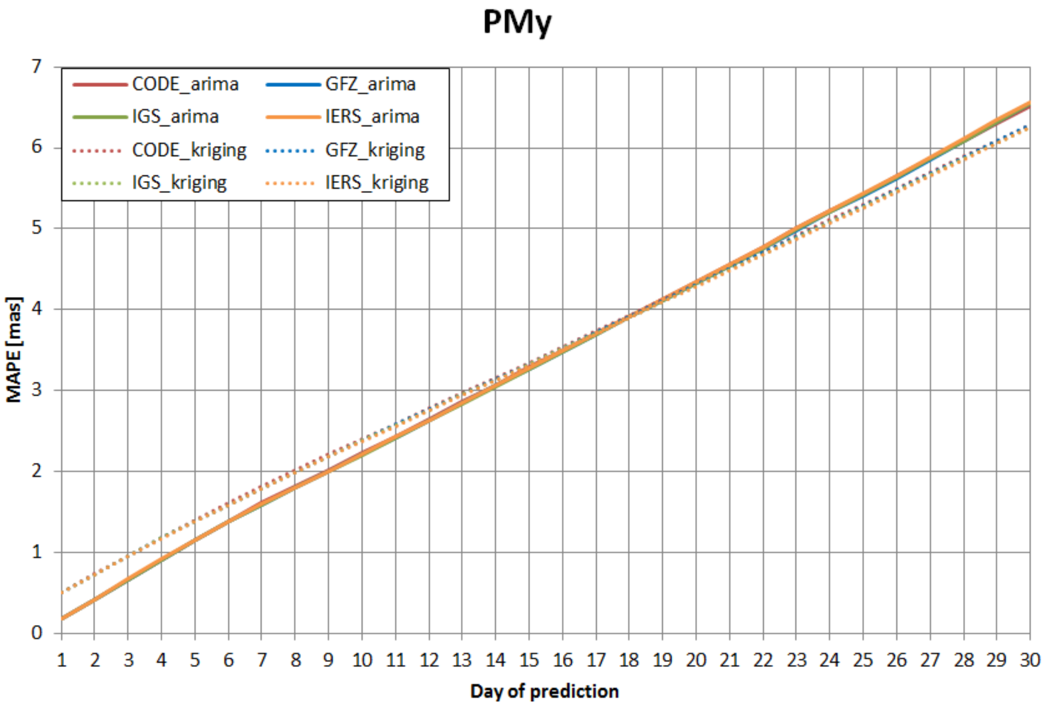


Rysunek 8 Wartości MAE (MAPE) dla 10-dniowej prognozy PMx (wykres po lewej stronie) i PMy (wykres po prawej stronie) opartej na krigingu zwyczajnym oraz z wykorzystaniem metod biorących udział w 1st EOP PCC (Michalczak i Ligas, 2021).

Rysunek 9 i rysunek 10 przedstawiają wartości MAPE dla 30-dniowej prognozy offline współrzędnych bieguna ziemskiego z wykorzystaniem krigingu zwyczajnego oraz modelu ARIMA dla szeregów czasowych pochodzących z różnych centrów obliczeniowych (Michalczak i in., 2022). W przypadku prognozy PMx można zauważyć, że metoda krigingu osiąga większe wartości MAPE do 11. dnia prognozy w porównaniu do ARIMA. Dla współrzędnej PMy ARIMA charakteryzuje się mniejszymi wartościami błędów prognozy do 18. dnia predykcji. Warto również zaznaczyć, że dokładność prognozy opartej na danych pochodzących z IERS nie różni się znacząco od dokładności uzyskanej przy wykorzystaniu szeregów czasowych z innych ośrodków obliczeniowych (CODE, GFZ i IGS).

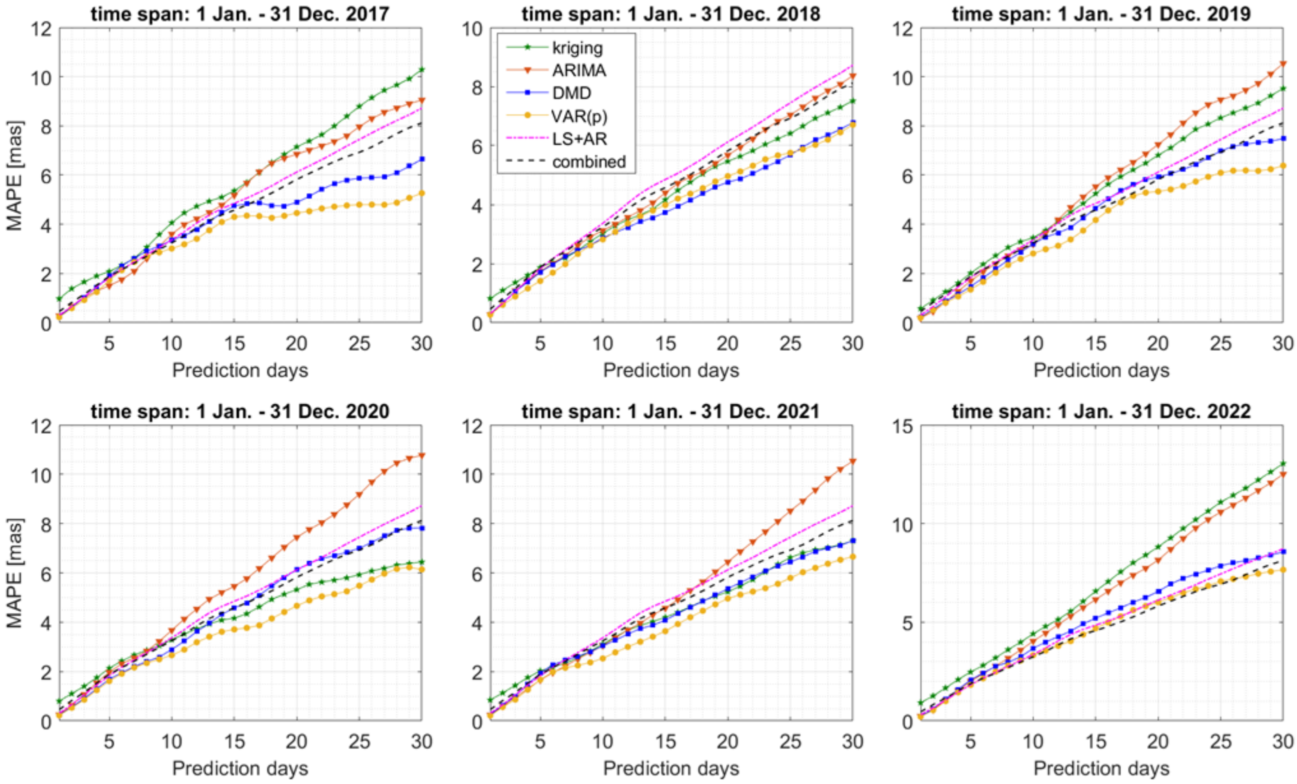


Rysunek 9 Porównanie wyników prognozy PMx z wykorzystaniem krigingu zwyczajnego i modelu ARIMA dla szeregów czasowych z różnych centrów obliczeniowych dla okresu od 1 stycznia 2017 (MJD 57754.5) do 15 kwietnia 2022 (MJD 59684.5) (Michalczak i in., 2022).



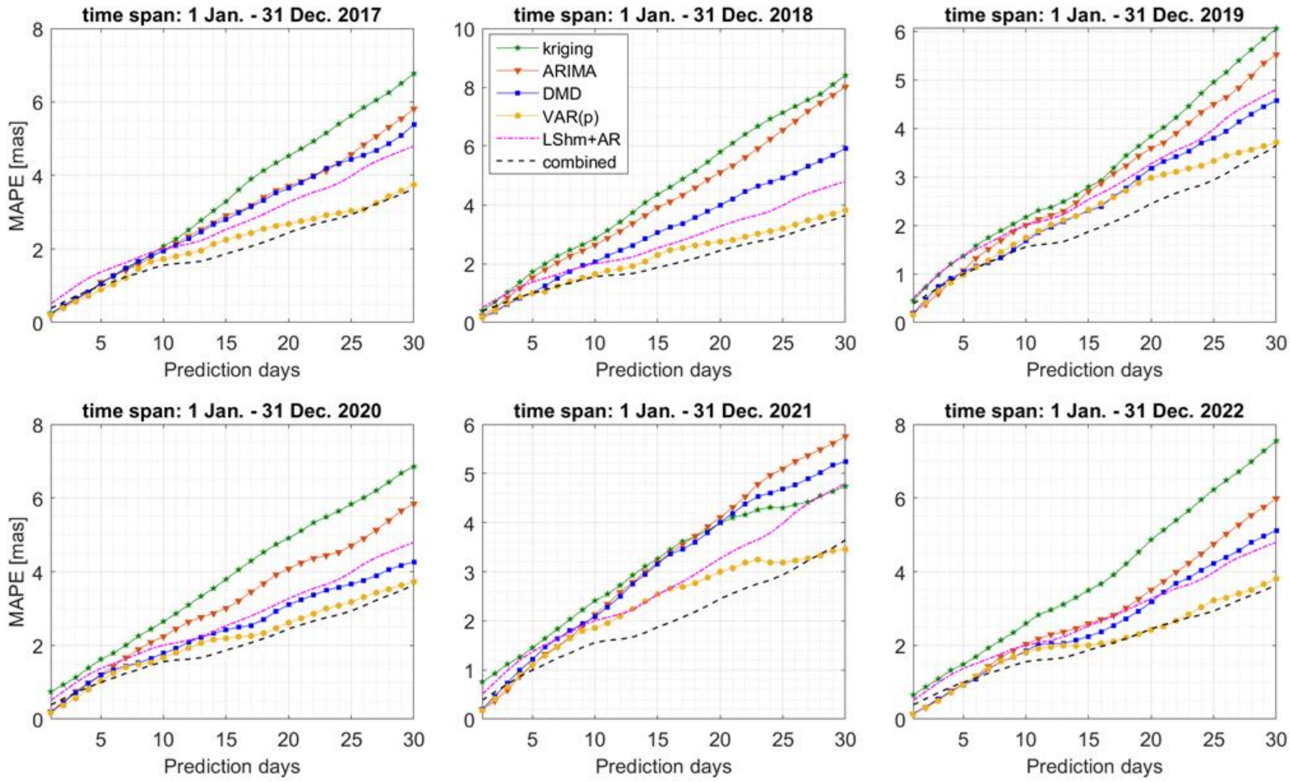
Rysunek 10 Porównanie wyników prognozy PMy z wykorzystaniem krigingu zwyczajnego i modelu ARIMA dla szeregów czasowych z różnych centrów obliczeniowych dla okresu od 1 stycznia 2017 (MJD 57754.5) do 15 kwietnia 2022 (MJD 59684.5) (Michalczak i in., 2022).

Na rysunku 11 zamieszczone zostały wyniki 30-dniowej prognozy offline PMx dla sześciu rocznych okien czasowych z wykorzystaniem krigingu zwyczajnego, modelu ARIMA, DMD i VAR(*p*), a także dwóch najdokładniejszych metod biorących udział w 1st EOP PCC: LS+AR oraz kombinacji prognoz zaproponowanej przez Kalarus i in. (2010). Wyniki pokazują, że niemalże w każdym rocznym okresie metoda VAR(*p*) charakteryzuje się najdokładniejszą prognozą. Ponadto, warto zauważyć, iż zarówno dla 10. jak i 30. dnia prognozy VAR(*p*) osiąga najmniejsze wartości MAPE. W 2022 roku można zaobserwować pogorszenie się dokładności prognozy dla niemalże wszystkich metod.



Rysunek 11 MAPE dla 30-dniowej prognozy PMx z wykorzystaniem krigingu zwyczajnego, modelu ARIMA, DMD i VAR(p), a także dwóch najdokładniejszych metod z 1st EOP PCC: LS+AR oraz kombinacji prognoz (Ligas i Michalczak, 2024).

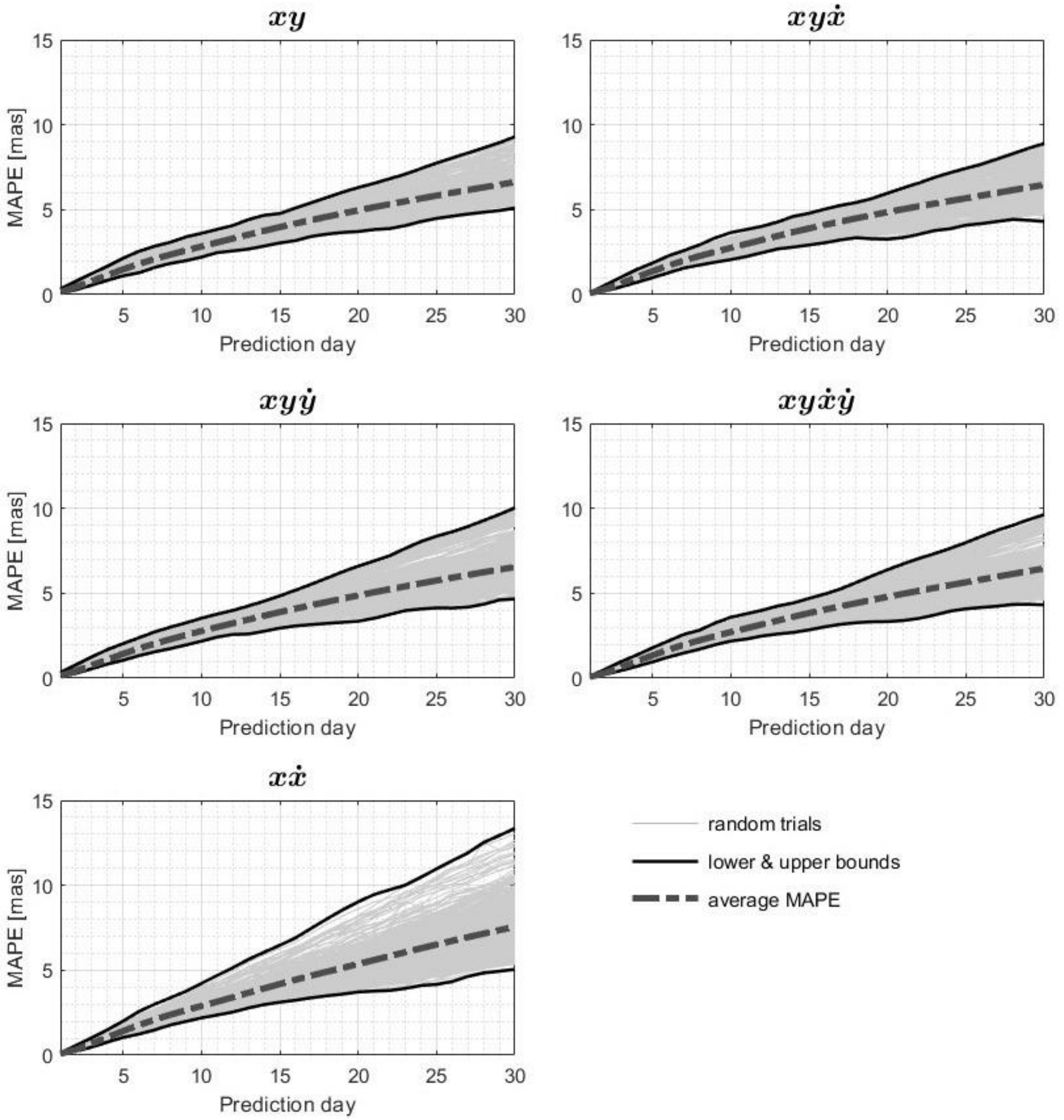
Rysunek 12 przedstawia wyniki 30-dniowej prognozy offline PMy dla sześciu rocznych okien czasowych z wykorzystaniem krigingu zwyczajnego, modelu ARIMA, DMD i VAR(p), a także dwóch najdokładniejszych metod z 1st EOP PCC: LShm+AR oraz kombinacji prognoz. Podobnie jak w przypadku prognozy PMx, najdokładniejszą metodą prognozy jest VAR(p). Dla bardzo krótkich horyzontów prognozy (do 10. dnia) DMD osiąga bardzo porównywalne wyniki do VAR(p). W odróżnieniu od prognozy PMx, w 2022 roku tylko jedna metoda, kriging zwyczajny, charakteryzuje się pogorszeniem dokładności prognozy.



Rysunek 12 MAPE dla 30-dniowej prognozy PMy z wykorzystaniem krigingu zwyczajnego, modelu ARIMA, DMD i VAR(p), a także dwóch najdokładniejszych metod z 1st EOP PCC: LShm+AR oraz kombinacji prognoz (Ligas i Michalczak, 2024).

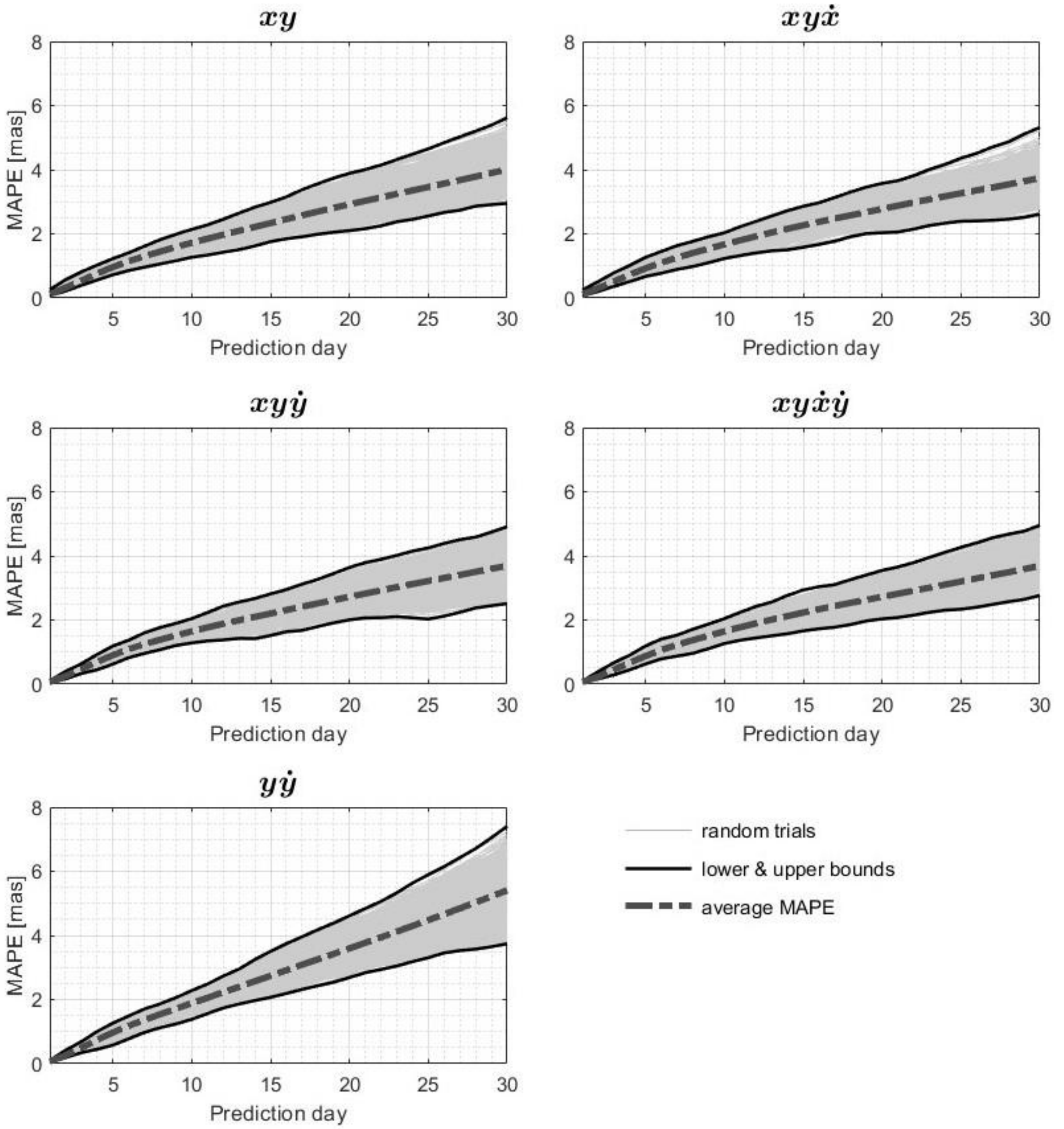
W obu przypadkach prognozy współrzędnych bieguna ziemskiego, kriging zwyczajny i ARIMA okazały się najmniej dokładnymi metodami.

Na rysunku 13 zaprezentowane zostały wartości MAPE dla różnych wariantów prognozy offline PMx z wykorzystaniem LS+VAR (Michalczak i in., 2024). Warianty te różnią się w zależności od wejściowych szeregów czasowych biorących udział w wielozmiennej prognozie. Kombinacja danych $(x, \dot{x})$, PMx i jej pochodna, która nie zawiera informacji o drugiej zmiennej (tj. PMy), charakteryzuje się najmniej dokładnymi prognozami na 30. dzień. Warto jednak zaznaczyć, że do około 7. dnia prognozy, osiąga ona lepsze wyniki niż wariant referencyjny (x, y). Spośród kombinacji, które zawierają również drugą zmienną, najmniej dokładną jest kombinacja $(x, y, \dot{y})$. Najdokładniejszym wariantem prognozy PMx, ze względu na średnią wartość MAPE, jest kombinacja zawierająca wszystkie możliwe wejściowe szeregi czasowe $(x, y, \dot{x}, \dot{y})$. Warto zaznaczyć, że kombinacja $(x, y, \dot{x})$ charakteryzuje się najmniejszymi różnicami pomiędzy najmniej dokładną i najdokładniejszą prognozą, przez co osiąga najbardziej zbliżone do siebie pojedyncze błędy prognozy.



Rysunek 13 Pojedyncze i uśrednione wartości MAPE dla PMx uzyskane z 1000 rocznych prób losowych dla różnych kombinacji danych z wykorzystaniem LS+VAR (Michalczak i in., 2024).

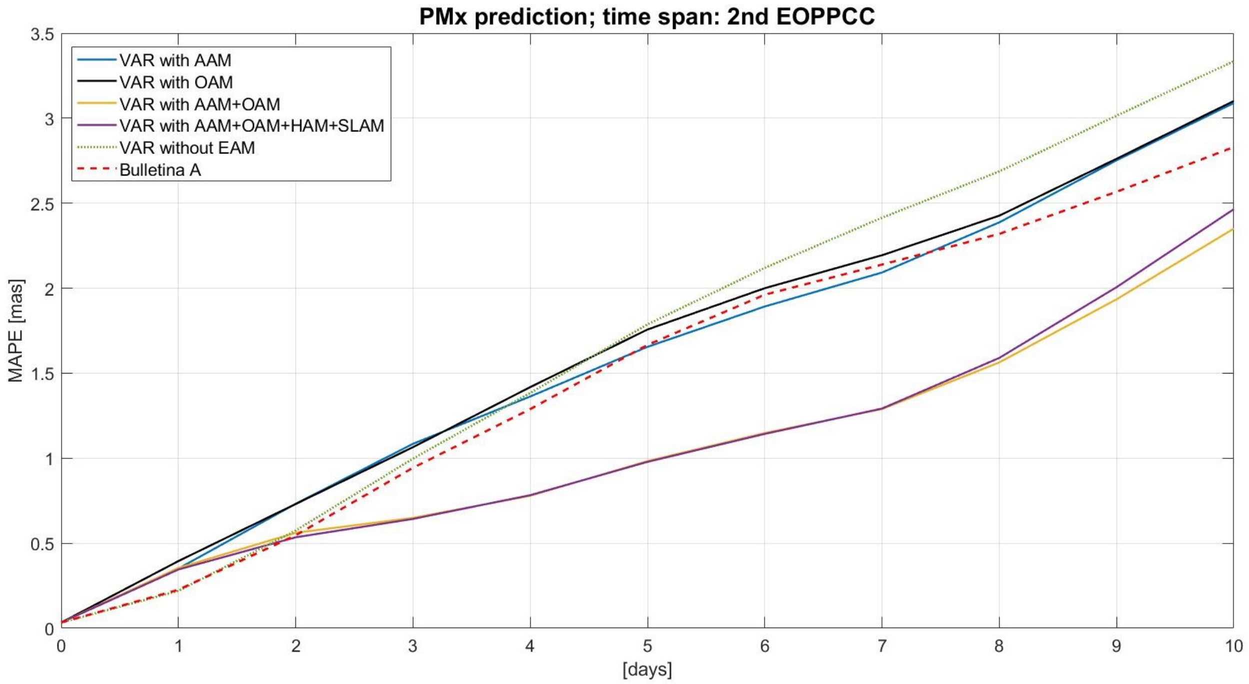
Na rysunku 14 przedstawione zostały wartości MAPE dla różnych kombinacji prognozy offline PMy metodą LS+VAR. Podobnie jak w przypadku współrzędnej PMx, kombinacja nie zawierająca drugiej zmiennej $(y, \dot{y})$ charakteryzuje się największymi wartościami MAPE. Najdokładniejszą kombinacją, pod względem średniej wartości MAPE, jest kombinacja zawierająca wszystkie szeregi czasowe $(x, y, \dot{x}, \dot{y})$. Kombinacja ta charakteryzuje się również najmniejszym rozrzutem wartości MAPE, czyli różnicą pomiędzy najdokładniejszą i najmniej dokładną pojedynczą prognozą.



Rysunek 14 Pojedyncze i uśrednione wartości MAPE dla PMy uzyskane z 1000 rocznych prób losowych dla różnych kombinacji danych z wykorzystaniem LS+VAR (Michalczak i in., 2024).

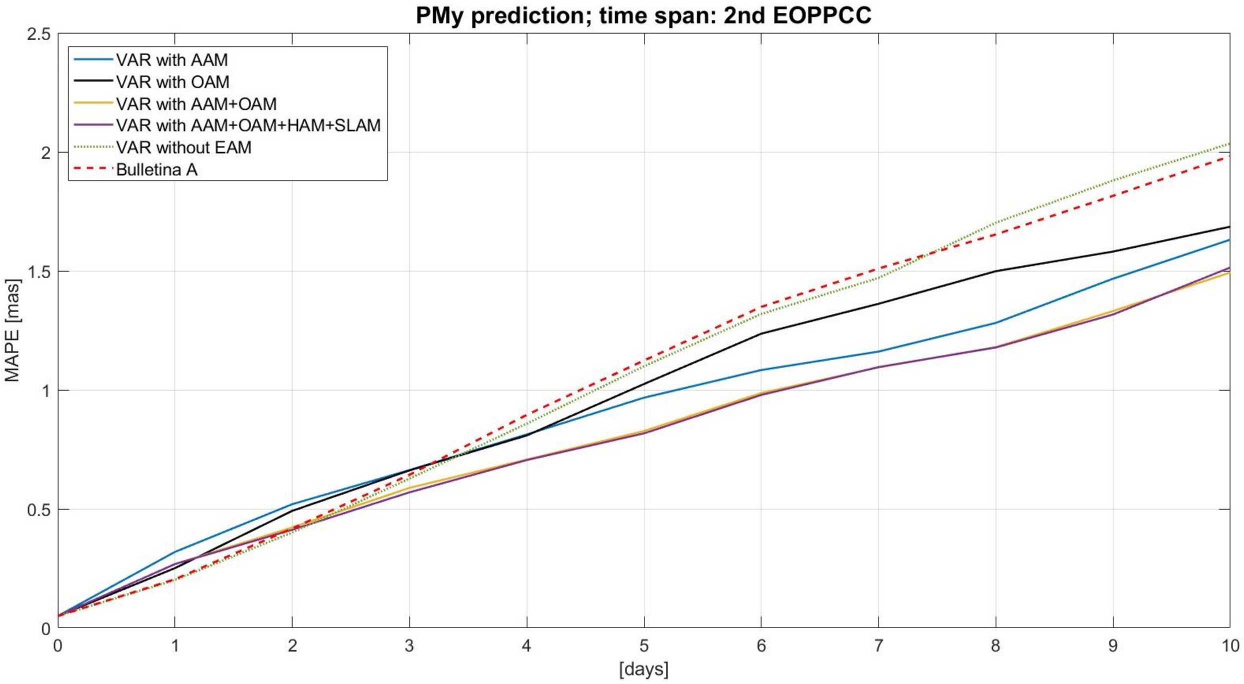
Poza wynikami prognozy offline współrzędnych bieguna ziemskiego zaprezentowanymi w niniejszej pracy, przeprowadzono również badania związane z wykorzystaniem szeregów czasowych EAM (pochodzących z ETH Zurich) w ultra-krótkiej (10-dniowej) prognozie quasi-offline parametrów PMx i PMy. Wykonano sześćdziesiąt dziewięć 10-dniowych quasi-offline prognoz z wykorzystaniem metody VAR(*p*) (zmiennymi wejściowymi są oba szeregi współrzędnych bieguna ziemskiego). W badaniach zastosowane cztery różne kombinacje komponentów EAM oraz prognozę bez uwzględnienia danych EAM. Okno czasowe opisywanych prognoz pokrywa się z okresem trwania Drugiej Kampanii Porównawczej Prognozy Parametrów Orientacji Ziemi (w dalszej części pracy określanej jako 2nd EOP PCC) – 01.09.2021 – 31.12.2022. MAPE dla poszczególnych wariantów prognozy oraz prognoz publikowanych przez IERS (Bulletin A), zaprezentowane zostały na rysunku 15 dla współrzędnej PMx i na rysunku 16 dla współrzędnej PMy.

Prognozy PMx z wykorzystaniem kombinacji AAM+OAM i AAM+OAM+HAM+SLAM charakteryzują się bardzo zbliżonymi błędami, niemalże na całej długości horyzontu predykcji. Niemniej jednak, na ostatni dzień prognozy, kombinacja AAM+OAM charakteryzuje się wartością MAPE mniejszą o około 0.1 mas w porównaniu do kombinacji AAM+OAM+HAM+SLAM. Należy zaznaczyć, że wykorzystanie jakiejkolwiek kombinacji danych EAM w procesie predykcyjnym, skutkuje zwiększeniem błędu prognozy w pierwszym dniu, w porównaniu do prognozy VAR(*p*) bez tych danych (lub Bulletin A) o około 0.1 mas. Najdokładniejsze prognozy (kombinacja AAM+OAM) cechują się MAPE dla 10. dnia równym około 2.35 mas, podczas gdy najmniej dokładne (VAR(*p*) bez danych EAM) osiągają wartość 3.35 mas. Warto zauważyć, że prognozy publikowane przez IERS charakteryzują się większymi błędami predykcji niż proponowana metoda VAR(*p*) z wykorzystaniem kombinacji AAM+OAM lub AAM+OAM+HAM+SLAM.



Rysunek 15 MAPE dla 10-dniowej prognoz PMx z wykorzystaniem VAR(*p*), czterech kombinacji danych EAM oraz bez wykorzystania szeregów EAM.

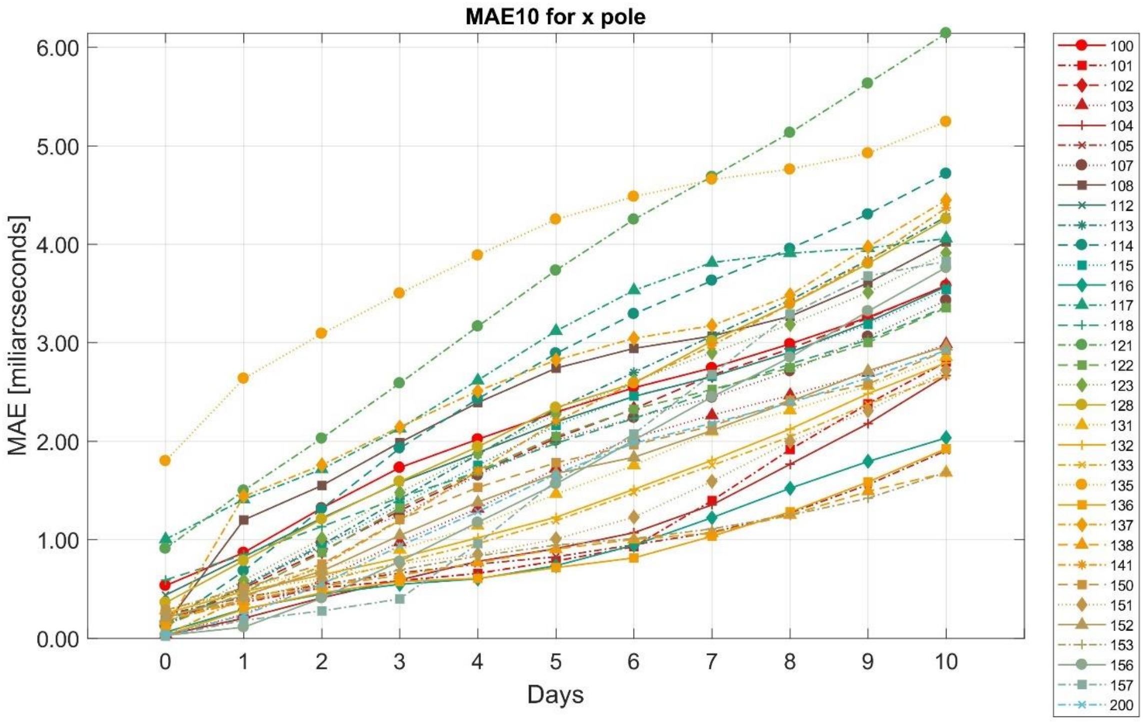
W przypadku PMy, prognozy z wykorzystaniem kombinacji danych AAM+OAM i AAM+OAM+HAM+SLAM cechują się najmniejszymi wartościami MAPE na niemalże całej 10-dniowej długości horyzontu prognozy. Wprowadzenie do procesu predykcyjnego PMy któregokolwiek z komponentów szeregu EAM zwiększa błąd prognozy na 1. dzień o około 0.1–0.2 mas. Warto podkreślić, że wprowadzenie pojedynczego szeregu EAM, tj. AAM lub OAM, do prognozy zmniejsza błędy prognozy na 10. dzień o około 0.4 mas w stosunku do prognozy bez wykorzystania danych EAM. Różnica ta zwiększa się, podczas wykorzystania kombinacji danych EAM z większą liczbą komponentów, do wartości około 0.55 mas. Należy zaznaczyć, że prognozy IERS cechują się najmniejszymi wartościami MAPE tylko dla 1. dnia prognozy. W kolejnych dniach wartości MAPE stopniowo zwiększają się, co sprawia, że Bulletin A (podobnie metoda VAR(*p*) bez wykorzystania danych EAM), staje się jedną z najmniej dokładnych spośród prezentowanych metod.



Rysunek 16 MAPE dla 10-dniowej prognoz PMy z wykorzystaniem VAR(*p*), czterech kombinacji danych EAM oraz bez wykorzystania szeregów EAM.

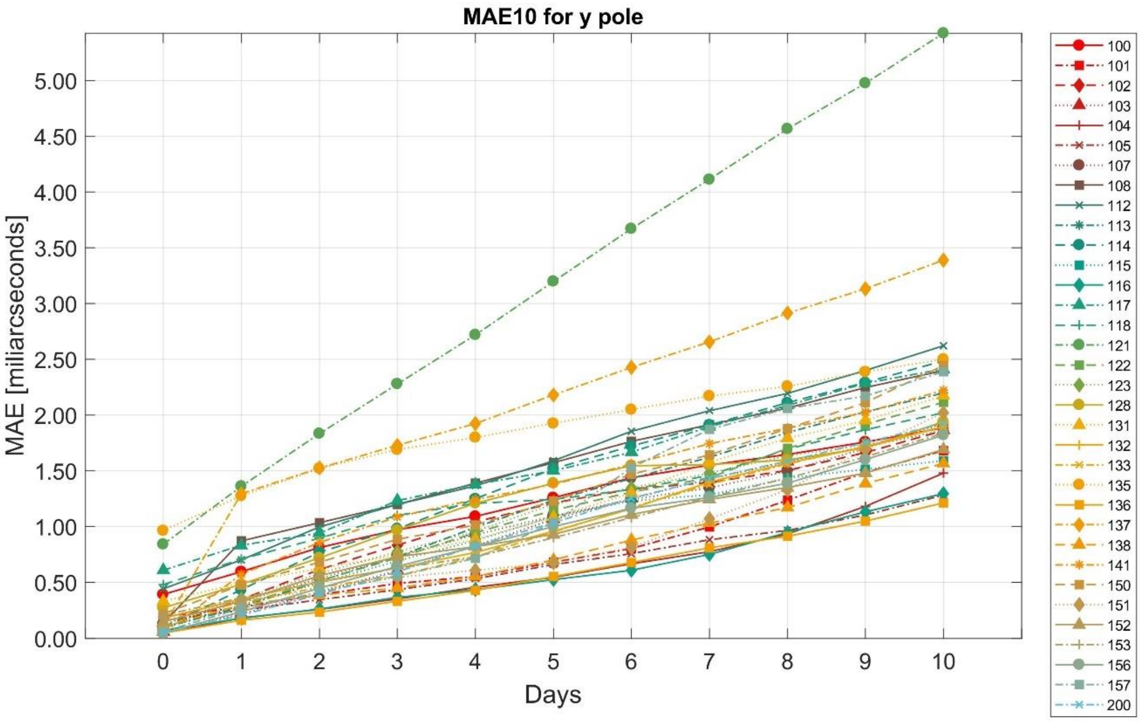
W 2nd EOP PCC zarejestrowane zostały 3 algorytmy predykcyjne (ID 141 – kriging zwyczajny, ID 156 – ARIMA i ID 157 – VAR(*p*)) do prognozy współrzędnych bieguna ziemskiego (należy zaznaczyć, że żadna z metod nie wykorzystywała danych EAM). Wyniki zaprezentowane na rysunku 17 i rysunku 18 przedstawiają wartości MAE (MAPE) dla 10-dniowej prognozy PMx i PMy uzyskanych za pomocą trzydziestu czterech metod zarejestrowanych w ramach 2nd EOP PCC.

Należy podkreślić, że dokładności prognozy PMx, uzyskane przez proponowane metody, charakteryzują się wartościami MAPE na poziomie zbliżonym do średniej w porównaniu z pozostałymi metodami. Dla pierwszych dni prognozy ID 156 oraz ID 157 charakteryzują się jednymi z najmniejszych wartości błędów, jednak w kolejnych dniach wartości te zwiększają się szybciej niż w przypadku pozostałych metod. Warto także zaznaczyć, że wartości MAPE uzyskane przez ID 141 (kriging zwyczajny) są największe spośród trzech wyżej wymienionych zarejestrowanych metod.



Rysunek 17 Wartości MAE (MAPE) dla 10-dniowej prognozy PMx uczestników 2nd EOP PCC; ID 141 – kriging zwyczajny, ID 156 – ARIMA, ID 157 – VAR(*p*). (2nd EOP PCC Workshop No. 2, 1 – 3 March 2023, Alicante, Spain & Online - Summary of the Second Earth Orientation Parameters Prediction Comparison Campaign (2nd EOP PCC)).

Podobnie jak w przypadku prognozy PMx, dokładność 10-dniowej predykcji PMy, uzyskanej za pomocą zarejestrowanych metod, oscyluje wokół średniej wartości dla wszystkich metod uczestniczących w prognozie tego parametru. Najdokładniejszą spośród zarejestrowanych (prezentowanych w tej rozprawie) metod jest model ARIMA (ID 156), natomiast najmniej dokładną jest metoda krigingu zwyczajnego (ID 141).



Rysunek 18 Wartości MAE (MAPE) dla 10-dniowej prognozy PMx uczestników 2nd EOP PCC; ID 141 – kriging zwyczajny, ID 156 – ARIMA, ID 157 – VAR(*p*). (2[nd] EOP PCC Workshop No. 2, 1 – 3 March 2023, Alicante, Spain & Online - Summary of the Second Earth Orientation Parameters Prediction Comparison Campaign (2nd EOP PCC)).

Warto zaznaczyć, że najdokładniejszymi, spośród wszystkich metod zarejestrowanych w 2nd EOP PCC, okazały się kombinacje ekstrapolacji metodą najmniejszych kwadratów (Least Squares extrapolation) oraz autoregresji (Autoregression, AR) z wykorzystaniem danych EAM.

5.2. Prognoza długości doby (LOD)

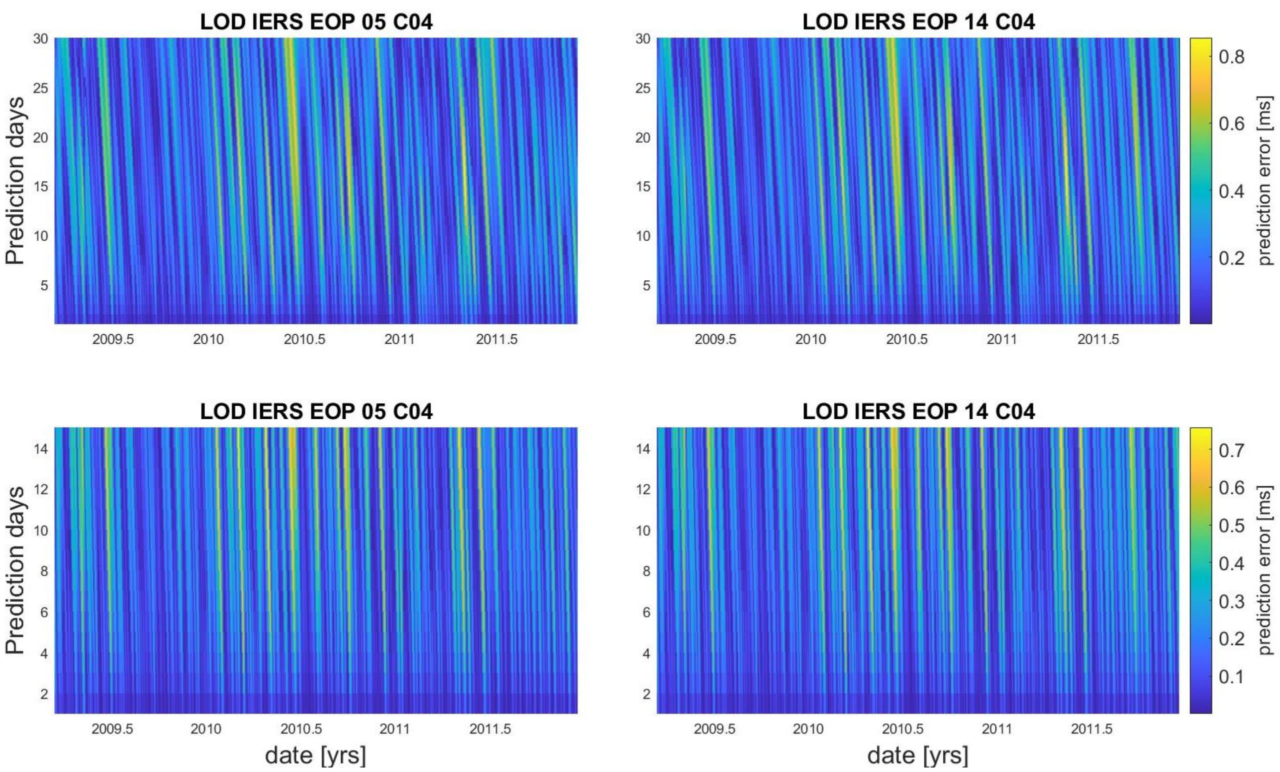
Wyniki ultrakrótkoterminowej (15 dni w przód) i krótkoterminowej (30 dni w przód) prognozy offline długości doby, uzyskane za pomocą dwóch wariantów krigingu – prostego i zwyczajnego, zostały zaprezentowane w pracy Michalczak i Ligas (2022). Wyniki zamieszczone w tabeli 4 przedstawiają wartości MAPE dla 15-dniowej prognozy LOD w zależności od metody prognozy, parametrów sterujących oraz sposobu wyboru semiwariogramu. Warto podkreślić, że wartości błędów prognozy z wykorzystaniem krigingu zwyczajnego są niższe niż dla krigingu prostego, porównując zarówno obie kombinacje parametrów sterujących jak i oba podejścia wyboru semiwariogramu. Warto zwrócić uwagę, że w 46 z 60 przypadków, adaptacyjne podejście wyboru semiwariogramu, cechuje się większymi wartościami błędu predykcji.

Tabela 4 Wartości MAPE dla 15-dniowych prognoz LOD dla krigingu prostego i zwyczajnego w podejściu adaptacyjnym i przy zastosowaniu tylko jednego semiwariogramu.

Dzień	**MAPE adaptacyjne / pojedynczy semiwariogram [ms]**			
	NN = 25 i TS = 12.2		**NN = 25 i TS = 7**	
	kriging zwyczajny	**kriging prosty**	**kriging zwyczajny**	**kriging prosty**
1	0.044 / 0.035	0.047 / 0.041	0.046 / 0.035	0.047 / 0.038
2	0.064 / 0.059	0.068 / 0.070	0.066 / 0.059	0.067 / 0.063
3	0.080 / 0.077	0.087 / 0.092	0.082 / 0.077	0.083 / 0.083

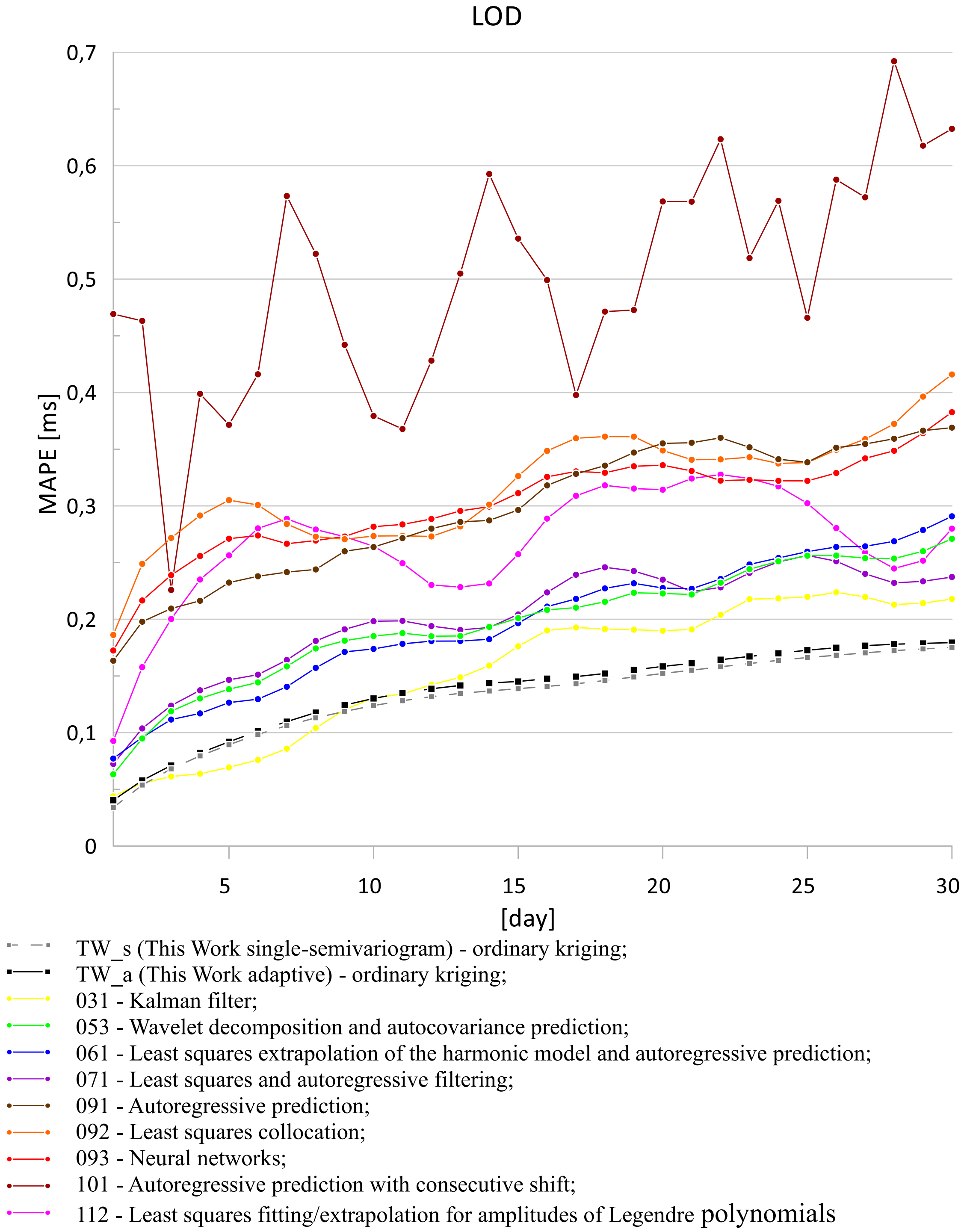
4	0.094 / 0.091	0.104 / 0.111	0.096 / 0.091	0.097 / 0.099
5	0.106 / 0.102	0.119 / 0.127	0.108 / 0.103	0.110 / 0.112
6	0.116 / 0.111	0.133 / 0.140	0.119 / 0.113	0.121 / 0.123
7	0.124 / 0.119	0.147 / 0.152	0.128 / 0.121	0.131 / 0.132
8	0.132 / 0.125	0.159 / 0.162	0.136 / 0.128	0.140 / 0.140
9	0.137 / 0.131	0.170 / 0.171	0.143 / 0.135	0.148 / 0.147
10	0.142 / 0.135	0.180 / 0.179	0.149 / 0.140	0.155 / 0.153
11	0.146 / 0.139	0.189 / 0.186	0.154 / 0.144	0.161 / 0.158
12	0.149 / 0.142	0.196 / 0.192	0.158 / 0.148	0.167 / 0.163
13	0.152 / 0.145	0.204 / 0.197	0.162 / 0.152	0.172 / 0.167
14	0.154 / 0.147	0.211 / 0.202	0.165 / 0.154	0.176 / 0.170
15	0.156 / 0.149	0.217 / 0.206	0.167 / 0.157	0.181 / 0.174

Na rysunku 19 przedstawione zostały wartości MAPE dla ultrakrótkoterminowej (15 dni w przyszłość, dolna część rysunku) i krótkoterminowej (30 dni w przyszłość, górna część rysunku) prognozy offline LOD dla każdego dnia predykcji. Największe błędy prognozy, osiągające wartości około 0.7 – 0.8 ms, występują w okolicach czerwca 2009, maja 2010, września 2010 oraz kwietnia 2011. Inspekcja dat występowania El Niño i La Niña pokazuje związek między dużymi wartościami błędów predykcji a pojawieniem się tych zjawisk (https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php).



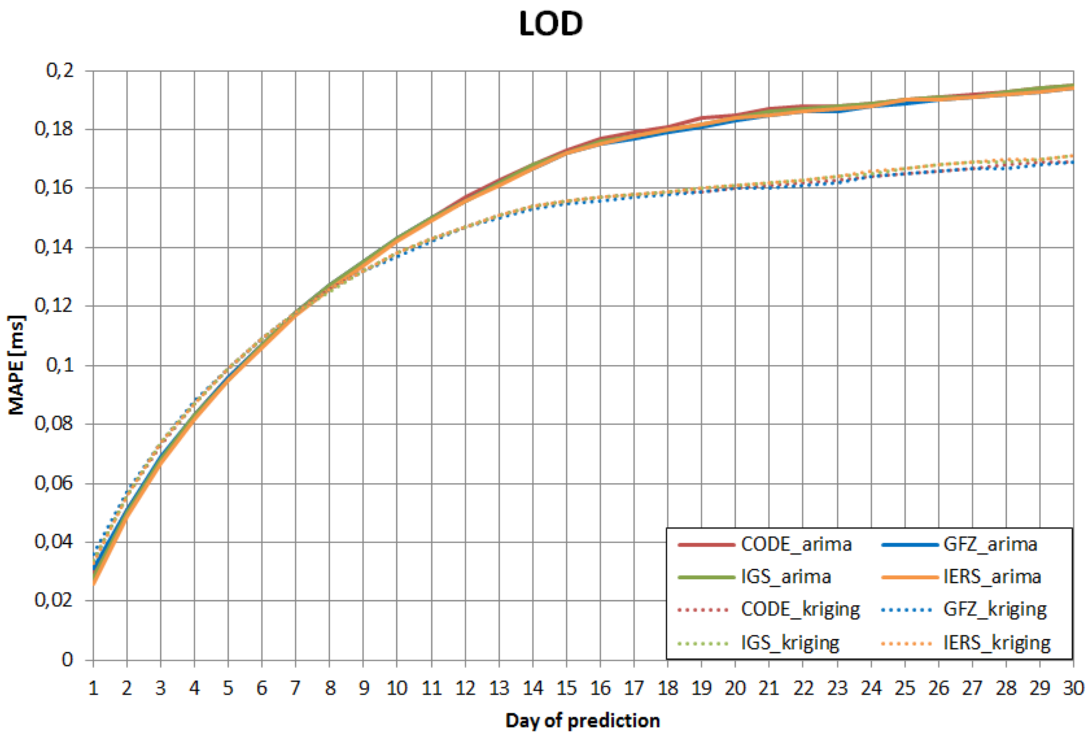
Rysunek 19 Dzienne błędy prognozy LOD dla prognoz krótkoterminowych (30 dni) i ultrakrótkoterminowych (15 dni) dla IERS EOP 05 C04 i IERS EOP 14 C04 (Michalczak i Ligas, 2022).

Rysunek 20 prezentuje wartości błędów prognozy offline LOD z wykorzystaniem krigingu zwyczajnego z zastosowaniem dwóch wariantów wyboru semiwariogramu w procesie predykcyjnym oraz metod biorących udział w 1st EOP PCC. Zestawienie to pokazuje, że metoda krigingu zwyczajnego cechuje się najmniejszymi wartościami MAPE na niemalże całej długości 30-dniowej prognozy. Jedynie dla pierwszych dni prognozy metoda filtru Kalmana charakteryzuje się większą dokładność niż proponowana metoda krigingu.



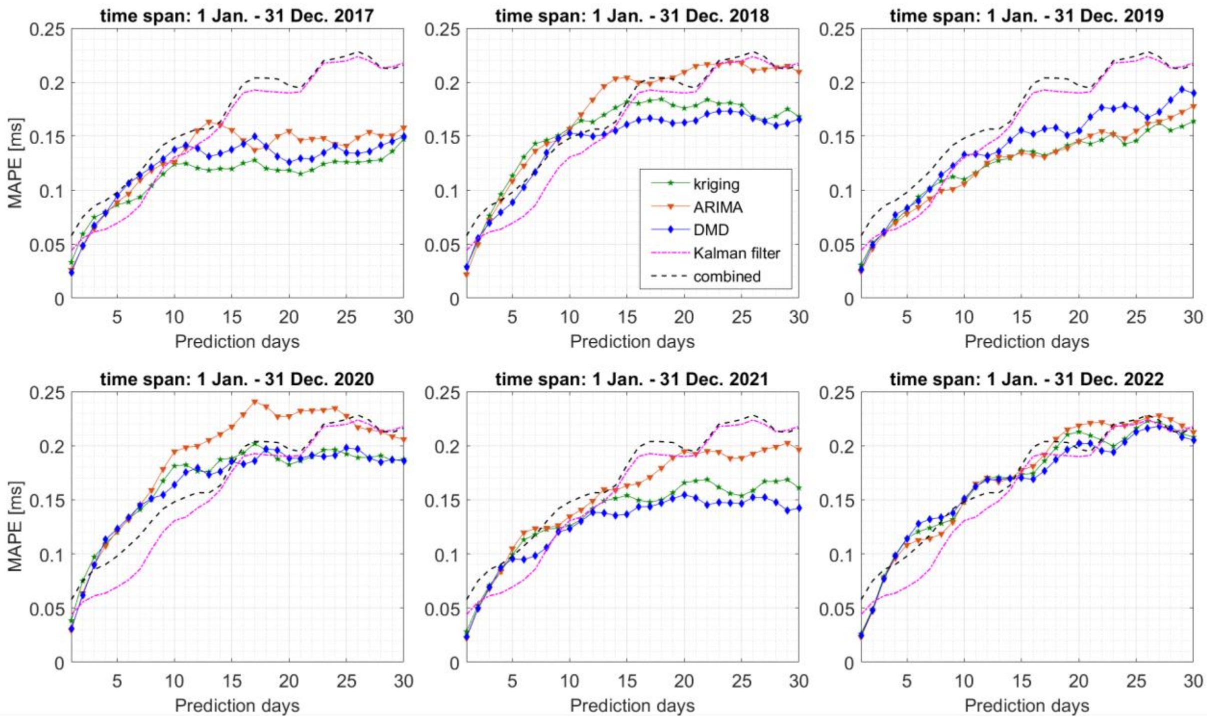
Rysunek 20 Wartości MAPE dla 30-dniowej prognozy LOD opartej na metodzie krigingu zwyczajnego i metodami zastosowanymi w 1st EOP PCC (Michalczak i Ligas, 2022).

Rysunek 21 przedstawia porównanie wartości MAPE dla 30-dniowej predykcji offline LOD metodą krigingu zwyczajnego i ARIMA. Wyniki zostały podzielone również ze względu na wykorzystane wejściowe szeregi czasowe z różnych ośrodków analitycznych (CODE, GFZ, IGS i IERS). Dla pierwszych 6 dni, prognozy z wykorzystaniem modelu ARIMA wykazują nieco wyższą dokładność. Natomiast po przekroczeniu 7. dnia prognozy, to kriging zwyczajny cechuje się mniejszymi wartościami MAPE. Maksymalne różnice pomiędzy prognozami opartymi na szeregach czasowych z różnych ośrodków sięgają wartości mniejszych niż 0.01 ms dla obu opisywanych metod.



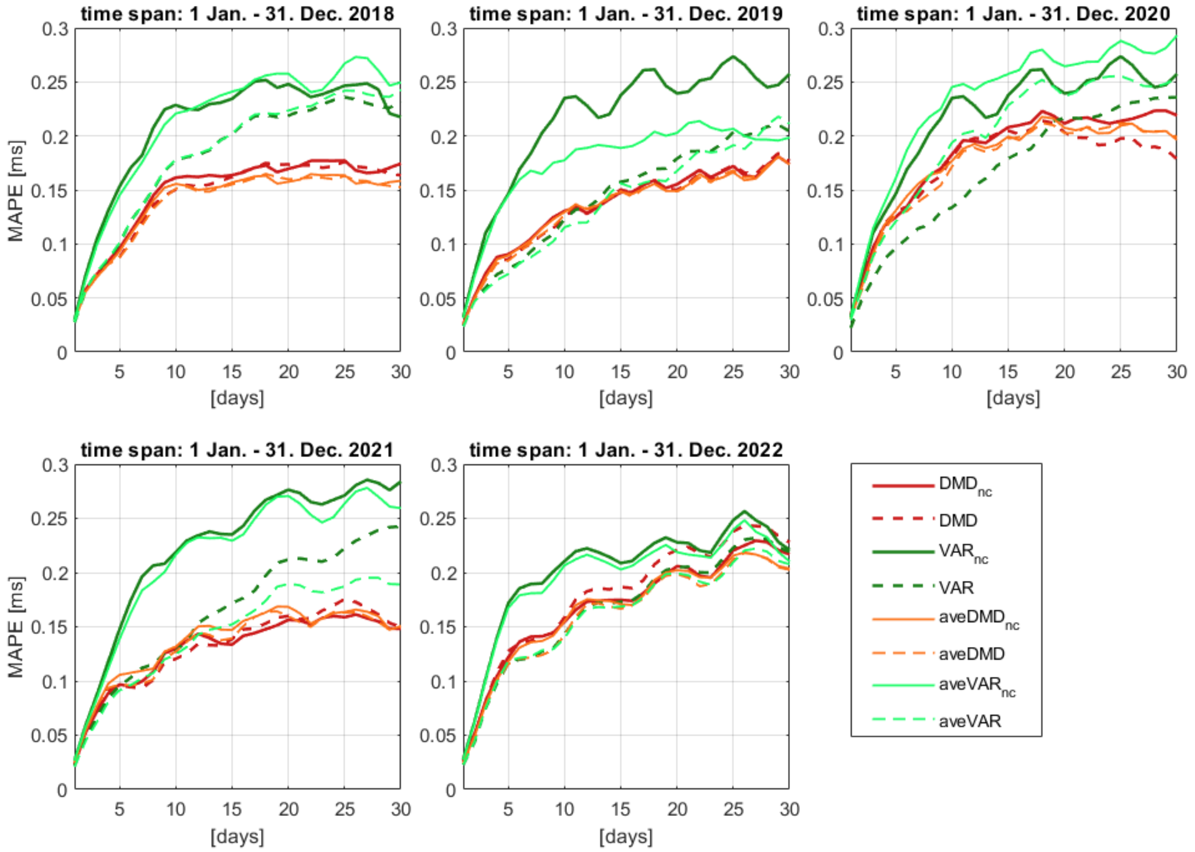
Rysunek 21 Wartości MAPE dla 30-dniowej prognozy LOD dla ARIMA i krigingu zwyczajnego dla różnych ośrodków analitycznych (Michalczak i in., 2022).

Na rysunku 22 zamieszczone zostały wyniki 30-dniowej prognozy offline LOD dla sześciu rocznych okien czasowych z wykorzystaniem krigingu zwyczajnego, modelu ARIMA i DMD, a także dwóch najdokładniejszych metod z 1st EOP PCC: filtru Kalmana oraz kombinacji prognoz (Ligas i Michalczak, 2024). Wszystkie proponowane nowe metody cechują się zbliżonymi dokładnościami aż do 10. dnia prognozy, maksymalne różnice są rzędu około 0.02 ms. Dla rocznych okresów 2018, 2020, 2021 oraz 2022 metoda DMD uzyskała najmniejsze wartości MAPE na 30. dzień predykcji. Warto zaznaczyć, że do około 10.–15. dnia metody biorące udział w 1st EOP PCC charakteryzują się mniejszymi wartościami błędu prognozy niż zaproponowane metody (kriging, ARIMA i DMD). Po około 10.-15. dniu prognozy wszystkie trzy prezentowane metody osiągają mniejsze wartości MAPE niż metody biorące udział w 1st EOP PCC.

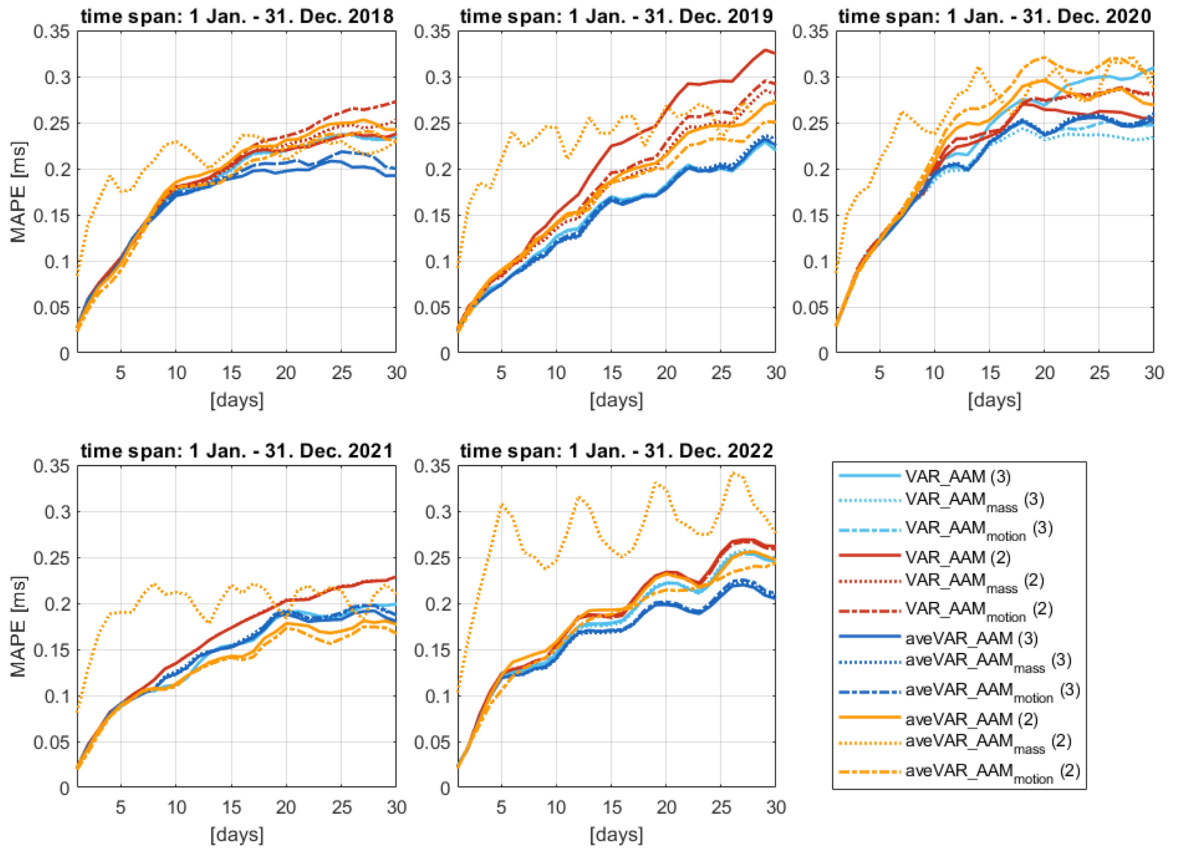


Rysunek 22 MAPE dla 30-dniowej prognozy LOD z wykorzystaniem krigingu zwyczajnego, modelu ARIMA i DMD, a także dwóch najdokładniejszych metod z 1st EOP PCC: filtr Kalmana oraz kombinacji prognoz (Ligas i Michalczak, 2024).

W pracy Michalczak i Ligas (2024) zaprezentowane zostały wyniki testów offline sprawdzające dokładność prognozy LOD z zastosowaniem dwóch metod prognozy: DMD i VAR(*p*). Ponadto badania wzbogacono o sprawdzenie wpływu stosowania korekt związanych z pływami strefowymi, kombinacji danych wejściowych w procesie predykcyjnym czy też zmiany podejścia przy wyborze zestawu parametrów sterujących (podejście adaptacyjne i uśrednione). Rysunek 23 prezentuje wartości MAPE dla 30-dniowej prognozy LOD przeprowadzonych za pomocą dwóch metod: DMD oraz VAR(*p*). Na rysunku 24 przedstawiono 30-dniową prognozę z wykorzystaniem modelu VAR(*p*) z różnymi kombinacjami danych wejściowych. Wartości błędów prognozy dla metody DMD osiągają najmniejsze wartości dla niemalże wszystkich możliwych wariantów i przedziałów czasowych. Warto zwrócić uwagę na bardzo zbliżone wartości MAPE w 2022 dla niemalże wszystkich wariantów predykcji. Ultrakrótkoterminowa (do 5-7 dni w przyszłość) prognoza LOD VAR(*p*) z dodatkowymi szeregami czasowymi AAM charakteryzuje się mniejszymi błędami niż procedura VAR(*p*) bez szeregów czasowych AAM. Średnio najdokładniejszymi metodami, spośród metod bez wykorzystywania dodatkowych danych AAM, są oba podejścia uśrednione DMD, podczas gdy z wariantów z użyciem danych AAM najdokładniejsza jest metoda VAR(*p*) z trzema zmiennymi wejściowymi i podejściem uśrednianym. Analiza dokładności prognozowania LOD z uwzględnieniem oraz bez uwzględnienia efektów pływów strefowych wykazuje, że ich eliminacja znacząco poprawia dokładność predykcji tylko w przypadku prognozy metodą VAR(*p*). Porównując wartości błędów metody VAR(*p*) przy użyciu danych AAM, można zauważyć, że modele z trzema zmiennymi wejściowymi mają mniejsze wartości MAPE niż kombinacje z dwoma zmiennymi wejściowymi. Dodatkowo, uwzględnienie obu komponentów AAM (masy i ruchu) znacząco poprawia dokładność prognoz (w porównaniu do wykorzystywania tych komponentów osobno).



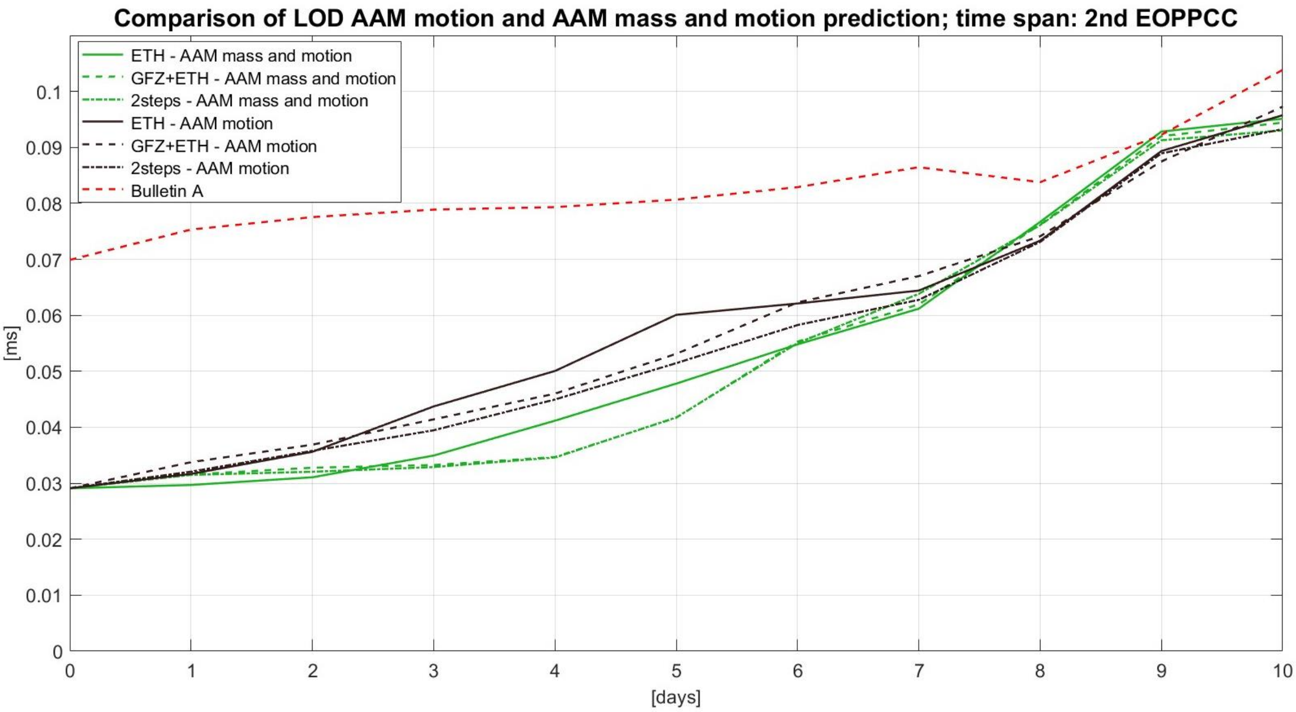
Rysunek 23 MAPE dla 30-dniowej prognozy LOD dla rocznych okiem czasowych 2018 – 2022 (Michalczak i Ligas, 2024).



Rysunek 24 MAPE dla 30-dniowej prognozy LOD z danymi AAM dla rocznych okiem czasowych 2018 – 2022 (Michalczak i Ligas, 2024).

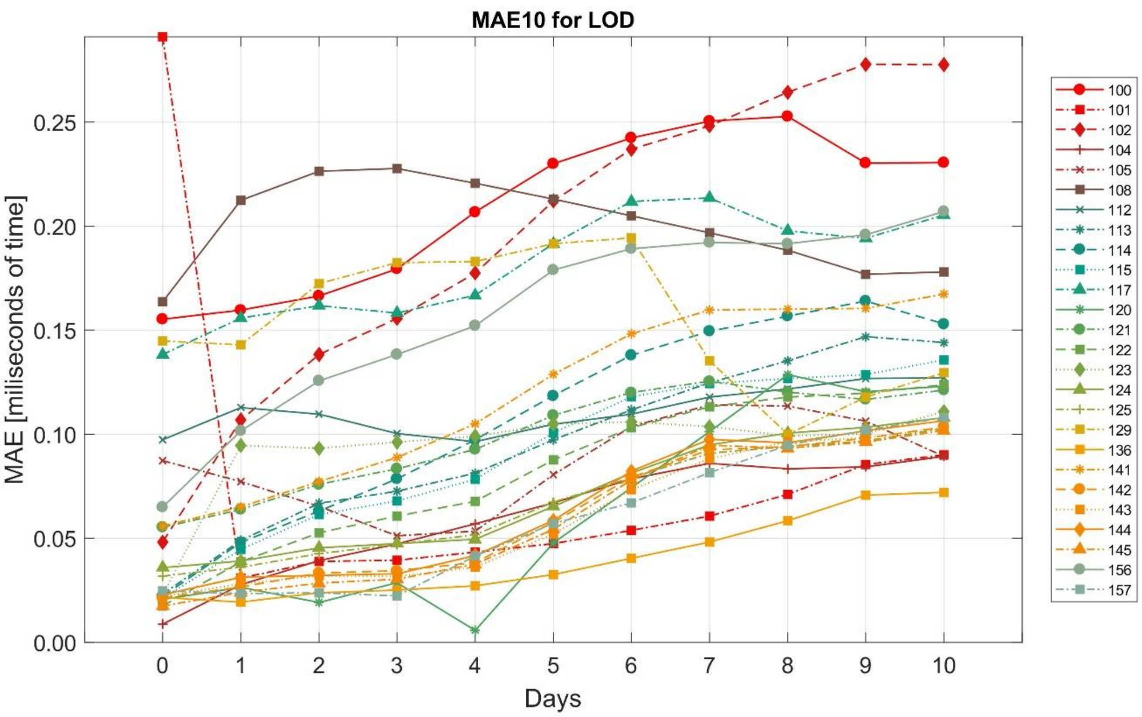
Kolejne badania dotyczyły wykorzystania danych AAM w prognozie quasi-offline LOD z zastosowaniem metody DMD. W ramach badań wykonano sześćdziesiąt dziewięć quasi-offline 10-dniowych prognoz, zgodnymi z datami rozpoczęcia i zakończenia 2nd EOP PCC. W eksperymencie przetestowano trzy warianty prognozy związane z różny danymi wejściowymi AAM. Pierwszy wariant, ETH, zakładał wykorzystanie danych AAM (10 dni) tylko z jednego centrum obliczeniowego, ETH Zurich (https://gpc.ethz.ch/products/EAM/fonode_GFZ_14d_AAM). W drugim wariancie, GFZ+ETH, dane wejściowe AAM były kombinacją danych z dwóch centrów, GFZ Potsdam (http://rz-vm115.gfz-potsdam.de:8080) (pierwsze 5 dni) i ETH Zurich (kolejne 5 dni). Trzeci wariant, 2steps, opierał się na dwuetapowej prognozie. W pierwszym etapie przeprowadzano prognozę na 5 dni z wykorzystaniem danych AAM pochodzących z GFZ Potsdam, następnie prognoza stawała się częścią szeregu wejściowego, w drugim etapie wykonywano prognozę na kolejne 5 dni bazując na danych z ETH Zurich. Każdy z opisanych wariantów można również podzielić w zależności od zastosowanych komponentów AAM – warianty z dwoma składowymi masy (z ang. „mass" – rozkład mas w obrębie Ziemi) i ruchu (z ang. „motion" – ruchy mas w obrębie Ziemi) oraz warianty w których skład wchodzi tylko komponent ruchu.

Rysunek 25 przedstawia wartości MAPE dla 10-dniowej prognozy długości doby w zależności od zastosowanej wejściowej kombinacji danych AAM. Wyniki pokazują, że przetransformowane na LOD prognozy UT1-UTC (Bulletin A) charakteryzują się najgorszą dokładnością predykcji w całym okresie trwania 2nd EOP PCC. Prognozy uwzględniające oba komponenty AAM (masy i ruchu) początkowo cechują się niższymi wartościami MAPE, jednak wraz ze wzrostem dnia prognozy stają się one zbliżone do tych które uwzględniają tylko komponent ruchu. Warianty prognozy ETH dla pierwszych 2 dni cechują się najmniejszymi wartościami MAPE. Niemniej jednak, dla kolejnych dni wyższą dokładnością charakteryzują się pozostałe 2 warianty, tj. GFZ+ETH i 2steps.



Rysunek 25 MAPE dla 10-dniowej prognozy LOD z wykorzystaniem DMD oraz danych AAM z dwóch ośrodków obliczeniowych.

W przypadku prognozy LOD podczas 2nd EOP PCC, wspomniane wcześniej ID, odpowiadały następującym metodom: ID 141 – kriging zwyczajny, ID 156 – ARIMA, zaś ID 157 – DMD (należy zaznaczyć, że żadna z metod nie wykorzystywała danych EAM). Na rysunku 26 przedstawiono wartości MAE (MAPE) dla 10-dniowej prognozy LOD uzyskanych za pomocą dwudziestu sześciu metod zarejestrowanych w ramach 2nd EOP PCC. Spośród zarejestrowanych (prezentowanych w tej rozprawie) metod najmniejsze błędy predykcji wykazała metoda DMD (ID 157), z zastrzeżeniem, że ta metoda wysłała tylko 7 prognoz. Najmniej dokładną metodą jest model ARIMA, który osiąga największe wartości MAPE. Warto zauważyć, że zarówno predykcje ID 141 jak i ID 156 znajdują się powyżej średniej wartości dla wszystkich metod uczestniczących w prognozie tego parametru. Należy zaznaczyć, że w 2nd EOP PCC najdokładniejsze prognozy LOD uzyskano dzięki połączeniu metody najmniejszych kwadratów i autoregresji z wykorzystaniem danych EAM – ID 136 i ID 101 (Śliwińska-Bronowicz i in., 2024).

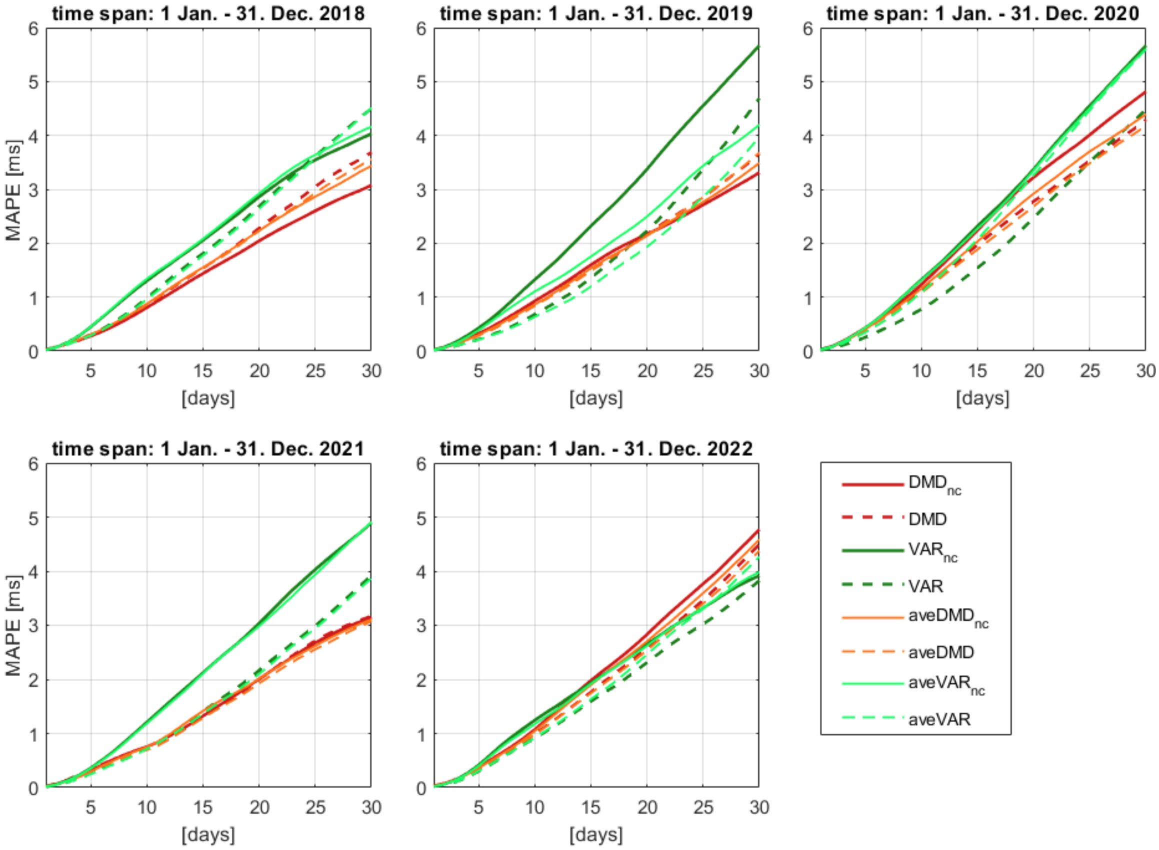


Rysunek 26 Wartości MAE (MAPE) dla 10-dniowej prognozy LOD uczestników 2nd EOP PCC; ID 141 – kriging zwyczajny, ID 156 – ARIMA, ID 157 – DMD. (2nd EOP PCC Workshop No. 2, 1 – 3 March 2023, Alicante, Spain & Online - Summary of the Second Earth Orientation Parameters Prediction Comparison Campaign (2nd EOP PCC)).

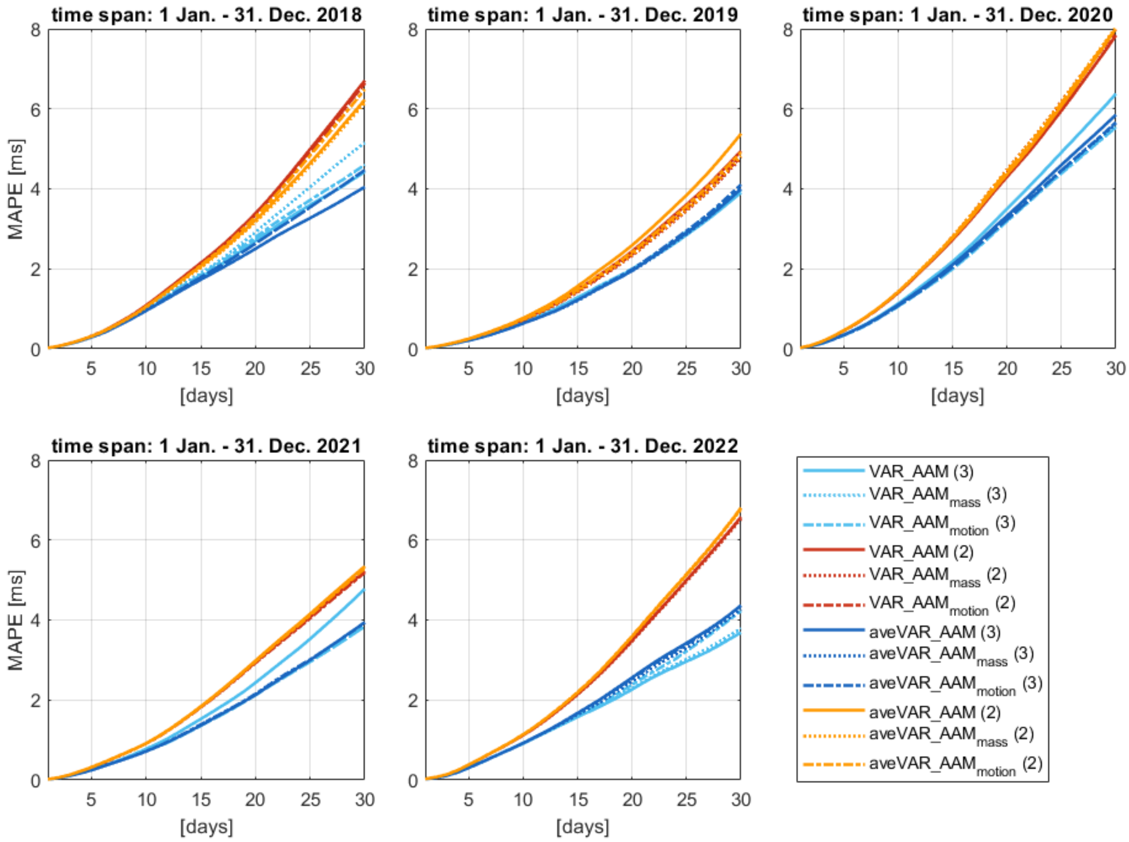
5.3. Prognoza różnicy czasów (UT1–UTC)

Badania dotyczące prognozy offline UT1-UTC (Michalczak i Ligas, 2024) obejmowały sprawdzenie dokładności prognozy w zależności od metody prognozy (DMD i VAR(*p*)), stosowania korekt związanych z pływami strefowymi, kombinacji danych wejściowych w procesie predykcyjnym czy też podejściu przy wyborze zestawu parametrów sterujących (podejście adaptacyjne lub uśrednione). Na rysunku 27 przedstawiono wartości MAPE dla 30-dniowej prognozy, uzyskane za pomocą dwóch metod: DMD oraz VAR(*p*). Rysunek 28 prezentuje prognozę z wykorzystaniem modelu VAR(*p*) z różnymi kombinacjami danych wejściowych do procesu predykcyjnego.

Wyniki ukazują, że metoda DMD cechuje się najmniejszymi wartościami MAPE na 30. dzień prognozy dla okien czasowych 2018-2021, zaś w roku 2022 prognoza metodą VAR(*p*) z 3 zmiennymi wejściowymi (UT1-UTC, LOD oraz AAM komponenty masy i ruchu) osiąga najmniejsze wartości błędów. Można zauważyć, że dla pierwszego dnia prognozy MAPE dla wszystkich metod są porównywalne i kształtują się na poziomie około 0.02 ms. Należy również odnotować, że włączenie szeregów AAM do prognozy (jako dodatkowy szereg wejściowy) VAR(*p*) zmniejsza nieznacznie wartości MAPE dla pierwszych 5-7 dni. W przypadku prognozy za pomocą metody DMD usunięcie korekt związanych z pływami strefowymi nie wpłynęło znacząco na dokładność prognozy. Natomiast w przypadku prognozy metodą VAR(*p*) zaobserwowano wyraźne zmniejszenie błędów. Na podstawie zaprezentowanych wykresów nie można jednoznacznie określić, które z podejść wyboru parametrów sterujących jest lepsze. Nieznacznie więcej niż połowa przypadków wykazuje mniejsze wartości MAPE przy zastosowaniu podejścia adaptacyjnego. Warto podkreślić, że wprowadzenie 3 szeregów czasowych do procesu prognozy metodą VAR(*p*) charakteryzuje się mniejszymi błędami predykcji, niż analogiczne podejścia wykorzystujące jedynie 2 szeregi czasowe jako dane wejściowe.



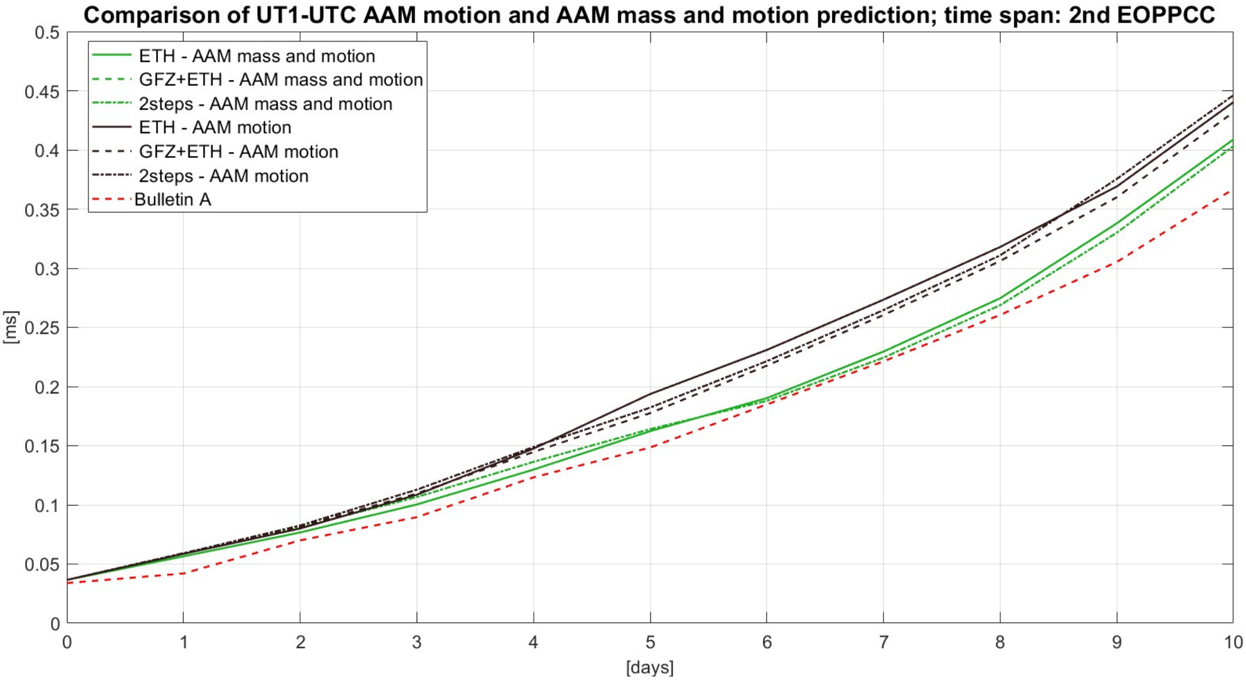
Rysunek 27 MAPE dla 30-dniowej prognozy UT1-UTC dla rocznych okiem czasowych 2018 – 2022 (Michalczak i Ligas, 2024).



Rysunek 28 MAPE dla 30-dniowej prognozy UT1-UTC z danymi AAM dla rocznych okiem czasowych 2018 – 2022 .

Przeprowadzono również badania związane z wykorzystaniem danych AAM w prognozie quasi-offline UT1-UTC metodą DMD. Zastosowano w nich podejście analogiczne do wcześniej opisanego, w którym wykorzystano dane AAM oraz metodę DMD do prognozowania LOD. Szczegóły dotyczące wariantów prognozy ETH, GFZ+ETH oraz 2steps, a także różnych kombinacji komponentów AAM (masy i ruchu), zostały przedstawione w poprzednim podrozdziale.

Wyniki zaprezentowane na rysunku 29 pokazują wartości MAPE dla 10-dniowej prognozy UT1-UTC w zależności od zastosowanej wejściowej kombinacji danych AAM. Predykcje publikowane przez IERS osiągają najmniejsze wartości błędu na przestrzeni całego horyzontu prognozy. Bulletin A osiąga wartości MAPE na 10. dzień prognozy około 0.37 ms, zaś najdokładniejszy wariant z wykorzystaniem obu komponentów AAM około 0.40 ms. W horyzoncie prognozy od 1. dnia do około 3.– 4. dnia, obie kombinacje komponentów AAM osiągają bardzo zbliżone wartości MAPE. Warianty prognozy, w których wykorzystywane są oba komponenty AAM, charakteryzują się bardzo zbliżonymi błędami prognozy. Biorąc pod uwagę ogół uzyskanych wyników można zauważyć, że najdokładniejsze prognozy, spośród zaproponowanych kombinacji, uzyskały warianty GFZ+ETH i 2steps, w których występują oba komponenty AAM.



Rysunek 29 MAPE dla 10-dniowej prognozy UT1-UTC z wykorzystaniem DMD oraz danych AAM z dwóch ośrodków obliczeniowych.

6. Podsumowanie

Znajomość Parametrów Orientacji Ziemi jest niezbędna do transformacji między niebieskim a ziemskim układem odniesienia. Są one również wykorzystywane w czasie rzeczywistym w wielu zadaniach geodezyjnych i astronomicznych, takich jak precyzyjne ustalanie pozycji zarówno w przestrzeni kosmicznej, jak i na powierzchni Ziemi, modelowanie prognozy pogody czy synchronizacja systemów czasu. Niestety, wyznaczenie Parametrów Orientacji Ziemi w czasie rzeczywistym jest niemożliwe ze względu na złożoność technik pomiarowych oraz przetwarzania danych.

W ciągu ostatnich lat opracowywano nowe metody prognozowania Parametrów Orientacji Ziemi, jednak ich dokładność wciąż pozostaje niewystarczająca. W związku z tym, konieczny jest dalszy rozwój i adaptacja coraz to nowszych metod prognozy, a także włączanie do procesy prognozy dodatkowych danych potencjalnie mogących wpłynąć na jakość prognozy.

W odpowiedzi na zidentyfikowany problem badawczy, w prezentowanej rozprawie doktorskiej podjęto próbę implementacji nowych metod prognozy Parametrów Rotacji Ziemi, w celu zwiększenia dokładność ich predykcji.

Przeprowadzone badania opierały się na zastosowaniu pięciu, w tym czterech wcześniej niewykorzystywanych do tego celu, metod (kriging prosty i kriging zwyczajny, ARIMA, DMD oraz VAR) do prognozy poszczególnych Parametrów Orientacji Ziemi. Szereg badań oraz eksperymentów pozwolił na stworzenie algorytmów działających w czasie rzeczywistym, które zostały następnie zarejestrowane w Drugiej Kampanii Porównawczej Prognozy Parametrów Orientacji Ziemi. Systematyczna archiwizacja szeregów czasowych EOP pozwoliła również na przeprowadzenie testów w trybie quasi-offline, odzwierciedlającym warunki występujące podczas trwania wspomnianej Kampanii.

Wyniki przeprowadzonych badań wskazują, że spośród zaprezentowanych metod prognozowania współrzędnych bieguna ziemskiego, bez wykorzystania danych EAM, najdokładniejsza jest metoda VAR(*p*). Kluczowym czynnikiem zwiększającym dokładność tej metody jest wykorzystanie dodatkowych szeregów czasowych w prognozie. Szeregi te mogą dostarczać potencjalnie istotnych informacji, mogących poprawić dokładność predykcji zmiennych podstawowych. W tym przypadku takimi szeregami mogą być zarówno druga współrzędna bieguna ziemskiego (dla PMx byłaby to PMy, zaś dla PMy PMx) czy szeregi zawierające informację o pochodnych obu współrzędnych. Należy zwrócić uwagę na ultrakrótkoterminową (10 dni w przód) predykcję PMx i PMy z uwzględnieniem danych o Efektywnych Momentach Pędu. Zastosowanie tych danych znacząco zwiększa dokładność prognozy, dzięki czemu metoda DMD z wykorzystaniem danych EAM może być porównywalna do obecnie najdokładniejszych metod predykcji współrzędnych bieguna ziemskiego.

Wykorzystanie szeregów czasowych jednego z komponentów EAM, to jest AAM, znacząco wpływa na dokładność ultrakrótkoterminowej predykcji długości doby. Kombinacja metody DMD i danych AAM osiąga najdokładniejsze prognozy spośród prezentowanych metod. Spośród metod niewykorzystujących danych EAM, na dłuższe horyzonty przewidywań, najmniejsze wartości błędów osiąga metoda DMD. Bardzo zbliżone do niej wartości błędów uzyskuje kombinacja modelu VAR(*p*) z trzema zmiennymi wejściowymi (LOD razem z szeregiem UT1-UTC i danymi AAM jako dodatkowymi szeregami wejściowymi). Porównując obecnie osiągalne dokładności można stwierdzić, że metoda DMD z wykorzystaniem danych AAM ma potencjał do osiągnięcia porównywalnej dokładności z obecnie najdokładniejszymi metodami prognozy, biorącymi udział w 2nd EOP PCC.

DMD charakteryzuje się najdokładniejszymi krótkoterminowymi (30-dni w przód) prognozami UT1-UTC spośród metod zaprezentowanych w niniejszej pracy. Należy zaznaczyć, że w przypadku ultrakrótkoterminowej prognozy najmniejsze wartości MAPE uzyskuje się przy zastosowaniu kombinacji metody DMD z danymi EAM. W przypadku tego parametru nie została zarejestrowana żadna metoda w 2nd EOP PCC, jednak porównując wyniki z przeprowadzonych testów quasi-offline (Rysunek 29) z wartościami MAPE opublikowanymi w Śliwińska-Bronowicz i in. (2024) można stwierdzić, że zaproponowana kombinacja DMD i danych AAM charakteryzuje się porównywalnymi wartościami błędów do najdokładniejszych metod prognozy.

Wyniki zaprezentowane w niniejszej rozprawie doktorskiej potwierdzają fakt, że zastosowanie szeregów czasowych EAM do ultrakrótkoterminowej prognozy EOP istotnie zwiększa dokładność ich predykcji (Kalarus i in., 2010).

Rezultaty uzyskane w ramach niniejszej pracy doktorskiej są potwierdzeniem postawionej tezy, że prognozowanie Parametrów Rotacji Ziemi za pomocą krigingu prostego i krigingu zwyczajnego, modelu ARIMA, VAR oraz dynamicznej dekompozycji modalnej umożliwia uzyskanie prognoz o dokładności wyższej lub porównywalnej do najdokładniejszych obecnie stosowanych metod. W szerszej perspektywie niezbędny jest ciągły rozwój zaprezentowanych algorytmów predykcyjnych, a także uwzględnianie nowych informacji, szeregów czasowych, pozwalających lepiej opisać i uwzględnić zjawiska fizyczne wpływające na ruch obrotowy Ziemi.

Literatura

Barnes, R.T.H., Hide, R., White, A.A., Wilson, C.A. (1983). Atmospheric Angular Momentum Fluctuations, Length-of-Day Changes and Polar Motion. Proceedings of The Royal Society of London, Series A: Mathematical and Physical Sciences, 387(1792), 31–73. https://doi.org/10.1098/rspa.1983.0050

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

Box, G.E.P. and Jenkins, G.M. (1976). Time Series Analysis: Forecasting and Control. Holden Day San Francisco

Cressie, N.A.C., (1993). Statistics for spatial data. John Wiley & Sons, New York. https://doi.org/10.1002/9781119115151.ch1

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

Gambis, D., Luzum, B. (2011). Earth rotation monitoring, UT1 determination and prediction. Metrologia, 48, 165-70, 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). https://doi.org/10.1007/s00190-023-01745-x

Gross, R.S., Marcus, S.L., Eubanks, T.M., Dickey, J.O., Keppenne, C.L. (1996). Detection of an ENSO Signal in Seasonal Length-of-day Variations. Geophysical Research Letters, 23(23), 3373–3376, https://doi.org/10.1029/96GL03260

Gross, R.S., Fukumori, I., Menemenlis, D. (2003). Atmospheric and oceanic excitation of the Earth’s wobbles during 1980–2000. Journal of Geophysical Research, 108(B8). https://doi.org/10.1029/2002jb002143

Gross, R.S. (2007). Earth Rotation Variations - Long Period. In: Herring TA (ed) Treatise on geophysics, geodesy, vol. 3, Elsevier, New York 2007, 239–294, https://doi.org/10.1016/B978-044452748-6.00057-2

Hatemi-J, A. (2004). Multivariate tests for autocorrelation in the stable and unstable VAR models. Economic Modelling, 21(4), 661–683. https://doi.org/https://doi.org/10.1016/j.econmod.2003.09.005

Hide R., Birch N.T., Morrison L.V., Shea D.J., White A. (1980). Atmospheric angular momentum fluctuations and changes in the length of day. Nature, 286. 114–117, https://doi.org/10.1038/286114a0

Hozakowski, W. (1989). Polar Motion Prediction by the Least -Squares Collocation Method. In: Earth Rotation and Coordinate Reference Frames, Scotland, August 10–11, 1989.

IERS Annual Report 2018 (2020). Edited by Wolfgang R. Dick and Daniela Thaller. International Earth Rotation and Reference Systems Service, Central Bureau. Frankfurt am Main: Verlag des Bundesamts für Kartographie und Geodäsie.

Kalarus, M., Schuh, H., Kosek, W., Akyilmaz, O., Bizouard, C., Gambis, D., Gross, R., Jovanović, B., Kumakshev, S., Kutterer, H., Mendes Cerveira, P. J., Pasynok, S., Zotov, L. (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

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., Belda, S., Mishra, S., Soja. B. (2024). Short-term prediction of celestial pole offsets with interpretable machine learning. Earth Planets Space 76, 18. https://doi.org/10.1186/s40623-024-01964-2

Kosek, W., Kalarus, M., Niedzielski, T. (2007). Forecasting of the Earth Orientation Parameters - Comparison of Different Algorithms. Journe´es Syste`mes de Re´fe´rence Spatio-temporels, Observatoire de Paris, 17-19 September.

Lambeck, K. (1989). The Earth’s Variable Rotation: Some Geophysical Causes. In: Kovalevsky, J., Mueller, I.I., Kolaczek, B. (eds) Reference Frames. Astrophysics and Space Science Library, vol 154. Springer, Dordrecht. https://doi.org/10.1007/978-94-009-0933-5_11

Lei, Y., Guo, M., Hu, D. dan, Cai, H. bing, Zhao, D. ning, Hu, Z. peng, Gao, Y. ping. (2017). Short-term prediction of UT1-UTC by combination of the grey model and neural networks. Advances in Space Research, 59(2), 524–531. https://doi.org/10.1016/j.asr.2016.10.030

Ligas, M., Michalczak, M. (2024). Dynamic mode decomposition and bivariate autoregressive short-term prediction of Earth rotation parameters. Journal of Applied Geodesy, vol. 18, no. 2, pp. 211-221. https://doi.org/10.1515/jag-2023-0030

McCarthy D.D., Luzum B.J. (1993). An analysis of tidal variations in the length of day. Geophysical Journal International, 114, 341–346. https://doi.org/10.1111/j.1365-246X.1993.tb03922.x

Michalczak, M., Ligas, M. (2021). Kriging-based prediction of the Earth's pole coordinates. Journal of Applied Geodesy, vol. 15, no. 3, pp. 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, vol. 70, pp. 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 centers using ARIMA and kriging – a comparison. Artificial Satellites, Proceedings of the Second Earth Orientation Parameters Prediction Comparison Campaign (2nd EOP PCC) Workshop, vol.57, no.s1, 2022, pp. 274-289, https://doi.org/10.2478/arsa-2022-0024

Michalczak, M., Ligas, M. (2024). Short-term prediction of UT1-UTC and LOD via Dynamic Mode Decomposition and combination of least-squares and vector autoregressive model. Reports on Geodesy and Geoinformatics, vol. 117, pp. 45-54. https://doi.org/10.2478/rgg-2024-0006

Michalczak, M., Ligas, M., Belda, S., Ferrándiz, J., Modiri, S. (2024). Advancing polar motion prediction with derivative information. Journal of Applied Geodesy. https://doi.org/10.1515/jag-2024-0046

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). https://doi.org/10.1007/s00190-020-01354-y

Nastula, J., Chin, T.M., Gross, R., Śliwińska J., Wińska. M. (2020). Smoothing and predicting celestial pole offsets using a Kalman filter and smoother. Journal of Geodesy, 94, 29. https://doi.org/10.1007/s00190-020-01349-9

Nelson, R.A., Mccarthy, D.D., Malys, S., Levine, J., Guinot, B., Fliegel, H.F., Beard, R.L., Bartholomew, T.R. (2001). The leap second: its history and possible future. Metrologia, 38, 509. https://doi.org/10.1088/0026-1394/38/6/6

Niedzielski, T., Kosek, W. (2008). Prediction of UT1–UTC, LOD and AAM $\chi 3$ by combination of least-squares and multivariate stochastic methods. Journal of Geodesy (2008) 82, 83–92, https://doi.org/10.1007/s00190-007-0158-9

Olea, R.A. (1999). Geostatistics for Engineers and Earth Scientists. Geostatistics for Engineers and Earth Scientists. Technometrics, 42(4), 444–445. https://doi.org/10.1080/00401706.2000.10485748

Petit, G., Luzum, B. (eds.). IERS Conventions (2010), IERS Technical Note 36, v 1.3.0. Tech. rep., International Earth Rotation and Reference Systems Service, Frankfurt am Main, Germany

Sasao, T., Wahr, J.M. (1981). An excitation mechanism for the free 'core nutation'. Geophysical Journal International 64(3), 729–746. https://doi.org/10.1111/j.1365-246X.1981.tb02692.x

Schmid, P.J. (2010). Dynamic mode decomposition of numerical and experimental data. Journal of Fluid Mechanics, 656:5–28, https://doi.org/10.1017/s0022112010001217

Schuh, H., Ulrich, M., Egger, D., Müller, J., Schwegmann, W. (2002). Prediction of Earth orientation parameters by artificial neural networks. Journal of Geodesy, 76(5), 247–258. https://doi.org/10.1007/s00190-001-0242-5

Stein, M.L. (1999). Interpolation of spatial data – some theory for kriging, Springer, New York.

Śliwińska−Bronowicz, J., Kur, T., Wińska, M., Dobslaw, H., Nastula, J., Partyka, A., Belda, S., Bizouard, C., Boggs, D., Bruni, S., Chen, L., Chin, M., Dhar, S., Dill, R., Ferrándiz, J. M., Gou, J., Gross, R., Guessoum, S., Han, S., Heinkelmann, R., Irrgang, C., Kiani Shahvandim, M., Li, L., Ligas, M., Liu, L., Lu, W., Mayer, V., Michalczak, M., Modiri, S., Otten, M., Ratcliff, T., Raut, S., Saynisch-Wagner, J., Schartner, M., Schoenemann, E., Schuh, H., Soja, B., Su, X., Thaller, D., Thomas, M., Wang, G., Wu, Y., Xu, X., Yang, X., Zhao, X., Zhou, Z. (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). https://doi.org/10.1007/s00190-024-01824-7

Tu, J.H., Rowley, W.C., Luchtenburg, D.M., Brunton, S.L., Kutz, J.N. (2014). On dynamic mode decomposition: Theory and applications. Journal of Computational Dynamics, 1(2):391–421, https://doi.org/10.3934/jcd.2014.1.391.

Xin, Z., Yuan-wei, W., Xin-yu, Y., Xu-hai, Y., Shou-gang, Z. (2022). Multi-parameter Polar Motion Prediction Based on Effective Angular Momentum Function. In Chinese Astronomy and Astrophysics (Vol. 46, Issue 4, pp. 486–499). Elsevier Ltd. https://doi.org/10.1016/j.chinastron.2022.11.009

Zotov, L.V., Xu, X.Q., Skorobogatov, A., Zhou, Y.H. (2017). Combined SAI-SHAO prediction of Earth orientation parameters since 2012 till 2017. Geodesy and Geodynamics, 9(6), pp. 485-490, KeAi, China, 2018, https://doi.org/10.1016/j.geog.2018.11.002

Załączniki

Załącznik 1.	Artykuł 1, Michalczak, M., Ligas, M., (2021). Kriging-based prediction of the Earth's pole coordinates. Journal of Applied Geodesy, vol. 15, no. 3, pp. 233-241, https://doi.org/10.1515/jag-2021-0007
Załącznik 2.	Artykuł 2, Michalczak, M., Ligas, M. (2022). The (ultra) short term prediction of length-of-day using kriging. Advances in Space Research, vol. 70, pp. 610-620, https://doi.org/10.1016/j.asr.2022.05.007
Załącznik 3.	Artykuł 3, Michalczak, M., Ligas, M., Kudrys, J. (2022). Prediction of Earth Rotation Parameters with the use of rapid products from IGS, CODE and GFZ data centers using ARIMA and kriging – a comparison. Artificial Satellites, Proceedings of the Second Earth Orientation Parameters Prediction Comparison Campaign (2nd EOP PCC) Workshop, vol.57, no.s1, 2022, pp. 274-289, https://doi.org/10.2478/arsa-2022-0024
Załącznik 4.	Artykuł 4, Michalczak, M., Ligas, M. (2024). Short-term prediction of UT1-UTC and LOD via Dynamic Mode Decomposition and combination of least-squares and vector autoregressive model. Reports on Geodesy and Geoinformatics, vol. 117, pp. 45-54. https://doi.org/10.2478/rgg-2024-0006
Załącznik 5.	Artykuł 5, Ligas, M., Michalczak, M. (2024). Dynamic mode decomposition and bivariate autoregressive short-term prediction of Earth rotation parameters. Journal of Applied Geodesy, vol. 18, no. 2, pp. 211-221. https://doi.org/10.1515/jag-2023-0030
Załącznik 6.	Artykuł 6, Michalczak, M., Ligas, M., Belda, S., Ferrándiz, J., Modiri, S. (2024). Advancing polar motion prediction with derivative information. Journal of Applied Geodesy. https://doi.org/10.1515/jag-2024-0046
Załącznik 7.	Oświadczenie określające indywidualny merytoryczny oraz procentowy wkład w powstanie rozprawy doktorskiej.

Maciej Michalczak* and Marcin Ligas

Kriging-based prediction of the Earth's pole coordinates

https://doi.org/10.1515/jag-2021-0007
Received February 16, 2021; accepted May 24, 2021

Abstract: Coordinates of the Earth's pole represent two out of five Earth orientation parameters describing Earth's rotation. They are necessary in transformation between celestial reference frame and terrestrial reference frame and what goes further in precise positioning and navigation, applications in astronomy, communication with outer space objects. Complexity of measuring techniques and data processing involved in the pole coordinates determination make it impossible to obtain them in real-time mode, hence a prediction problem of the polar motion emerges. In this study, geostatistical prediction methods, i. e., simple and ordinary kriging are applied. Millions of predictions have been performed to draw reasonable conclusions on prediction capabilities of applied kriging variants. The study is intended in ultra-short-term prediction (up to 15 days into the future) using the IERS EOP 14 C04 (IAU2000A) and IERS EOP 05 C04 (IAU2000A) series as a reference. Mean absolute prediction errors (for days 1–15) with respect to IERS 14 C04 are ranging 0.66–5.25 mas for PMx and 0.47–3.59 mas for PMy. On the other hand, for IERS 05 C04 the values are 0.60–4.95 mas and 0.44–3.29 mas for PMx and PMy; respectively. The results indicate competitiveness of the introduced methods with existing ones.

Keywords: polar motion, prediction, kriging, semivariogram, BLUP

1 Introduction

Earth Orientation Parameters (EOP) are a set of parameters that describe irregularities in the rotation of the Earth. There are five of them: coordinates of the pole (PMx, PMy), length of day (LOD) and celestial pole offsets (dX, dY). The instantaneous axis of the Earth's rotation does not stay still, that is why the coordinates of the pole change. The difference between the rotation period and the average period is called excess of length of day (LOD). Earth's rotation cannot be modeled precisely due to a complex structure [5]. The pole offsets (dX, dY) are corrections to the modeled intermediate celestial pole coordinates in the celestial reference system [9].

Transformation between the celestial and terrestrial international reference frames is required for many astronomical and geodetic purposes, including precise tracking and navigation of spacecrafts and satellites. To perform this transformation at any time, accurate EOP are needed. The process of measuring and processing data, involved in EOP determination, is complex and it is impossible to obtain them in real-time mode. The solution of this problem is EOP prediction. EOP are regularly delivered by two official IERS services: the IERS Rapid Service/Prediction Centre and the IERS EOP Product Centre and many individual analysis centres or services [8]. The official IERS EOP Product Centre publishes EOP data with thirty day latency. When necessary the latency gap may be filled by, e. g., ultra-rapid International GNSS Service (IGS) products consisting of half-observed and half-predicted parts. The accuracy of the predicted part from IGS is estimated at the level of 0.2 mas (https://www.igs.org/products/#about).

In the early 1980s polar motion was observed with space geodetic techniques such as: VLBI (Very Long Baseline Interferometry), SLR (Satellite Laser Ranging) and LLR (Lunar Laser Ranging) [19]. In the next decades, the precision of pole position measurements improved significantly due to the advent and development of GNSS (Global Navigation Satellite System) [4], [26]. EOP are calculated as a weighted combination of solutions from different space and satellite techniques [3]. Nowadays the achievable accuracy is up to about 5 µs in case of LOD and up to 30 µas or better in case of pole coordinates [3].

To forecast EOP many various prediction methods have been applied: harmonic and exponential method of approximation [11], least-squares collocation [6], neural networks [22], auto-regressive prediction, least-squares extrapolation [13], extreme learning machine and empirical mode decomposition [14], combined method of least squares and autoregressive model [24][25], combination

***Corresponding author: Maciej Michalczak,** Department of Integrated Geodesy and Cartography, Faculty of Mining Surveying and Environmental Engineering, AGH University of Science and Technology, Krakow, Poland, e-mail: mmichalc@agh.edu.pl, ORCID: https://orcid.org/0000-0001-6859-988X
Marcin Ligas, Department of Integrated Geodesy and Cartography, Faculty of Mining Surveying and Environmental Engineering, AGH University of Science and Technology, Krakow, Poland, e-mail: ligas@agh.edu.pl, ORCID: https://orcid.org/0000-0002-0970-128X

of singular spectrum analysis and Copula based analysis [17][18] or singular spectrum analysis [20].

In this contribution, geostatistical method of kriging is used to predict a residual process of the Earth pole coordinates constructed after the removal of a linear trend and periodic components. Kriging itself is a family of unbiased minimum mean square error predictors within geostatistics – a branch of spatial statistics. Kriging, originally, used in a spatial context is used here in a temporal one to predict future values of polar motion time series. Two kriging variants, i. e., simple kriging and ordinary kriging are applied. These variants differ as to the assumption on a mean value of a random process. Simple kriging assumes a mean value of a random process to be a known constant whilst in ordinary kriging it is an unknown constant. Kriging as a prediction method has settled in geodesy and positioning applications and is getting more and more attention [1], [2], [15], [16], [21].

This study is intended in ultra-short-term (up to 15 days into the future) prediction using the IERS EOP 14 C04 (IAU2000A) and IERS EOP 05 C04 (IAU2000A) series as a reference, the choice is explained in section 4. The prediction process starts on 1 January 2000 (MJD 51544) and involves millions of predictions to draw reasonable conclusions on prediction capabilities of applied kriging variants.

2 Methodology of kriging-based prediction of polar motion

2.1 Total prediction model

Total prediction model for the Earth's pole coordinates consists of two parts, i. e., deterministic and stochastic. The deterministic one is composed of an estimated linear trend and estimated periodic components. Their assumed validity is extended (extrapolated) for future time moments. In this study, only two periodic components (sine wave terms) capturing the most energetic oscillations in polar motion, i. e., Chandlerian and annual have been considered. The stochastic part is composed of predicted (kriged) residuals based on residuals obtained after the removal of trend and periodicities. Hence, the total prediction model may be descriptively expressed as:

$$\begin{aligned} \mathit{Total\ prediction} &= \mathit{linear\ trend} + \mathit{periodic\ components} \\ &\quad + \mathit{predicted\ (kriged)\ residuals} \end{aligned}$$

or

$$\hat{x}\left(t_p\right) = \hat{a}_0^x + \hat{a}_1^x t_p + \sum_{i=1}^{2} \hat{b}_i^x sin\left(\hat{c}_i^x t_p + \hat{d}_i^x\right) + \hat{\varepsilon}_{xx}\left(t_p\right) \quad (1a)$$

$$\hat{y}\left(t_p\right) = \hat{a}_0^y + \hat{a}_1^y t_p + \sum_{i=1}^{2} \hat{b}_i^y sin\left(\hat{c}_i^y t_p + \hat{d}_i^y\right) + \hat{\varepsilon}_{yy}\left(t_p\right) \quad (1b)$$

where t_p is a prediction time moment, $\hat{a}_0$, $\hat{a}_1$ are the estimated linear trend coefficients; $\hat{b}_i$, $\hat{c}_i$, $\hat{d}_i$ are the estimated amplitude, frequency and phase shift for i-th sine term; respectively and $\hat{\varepsilon}$ stands for the predicted (kriged) residual.

2.2 Stochastic part

To exploit a structure remaining in the residuals, obtained after the removal of trend and periodicities, in the prediction process the geostatistical method of kriging is applied. Kriging itself is a geostatistical term for optimal linear prediction of spatial processes [23]. In this study, on the other hand, kriging is used in a temporal context – for a stochastic process being a representation of a residual part of the polar motion prediction model. In fact, kriging stands for a broad class of unbiased minimum mean square error predictors equivalent to minimum error variance predictors due to the unbiasedness condition, i. e., BLUPs (Best Linear Unbiased Predictors). The mentioned optimality conditions may be expressed as:

$$E\left(e\right) = E\left(\hat{Z}_0 - Z_0\right) = 0 \quad \text{and} \quad E\left(ee^T\right) \rightarrow min \quad (2)$$

where e is a prediction error, $\hat{Z}_0$ is the predictor of a random variable Z_0, E is the expected value operator. In this study two variants of kriging have been used, i. e., simple and ordinary kriging. Simple kriging and ordinary kriging are optimal in the sense of BLUP if the mean value of a random function is a known constant or an unknown constant; respectively. Prediction by means of kriging requires a structure function that describes continuity and variability of a random process. The structure may be captured by a covariance function or a semivariogram, both functions are equivalent under the second order stationarity. However, in general terms, unbounded semivariograms have no covariance counterparts, e. g., a power semivariogram or its special case, a linear semivariogram. This makes the semivariogram a more general tool for describing the structure of a random process. The semivariogram has one more advantage over the covariance function, i. e., it does not require a mean value of a random process in an estimation procedure. In the following lines a model formulation, assumptions, form of the predictor, unbiasedness condition and system of equations to be solved for both kriging variants will be exposed briefly.

The following decomposition of a random function that holds for both kriging variants is considered (formulation for a single prediction):

$$\begin{cases} \mathbf{Z}(t) = \boldsymbol{\mu} + \boldsymbol{\varepsilon}(t) \\ Z(t_0) = \mu + \varepsilon(t_0) \end{cases} \tag{3}$$

or in brief by dropping an indexing time – parameter:

$$\begin{cases} \mathbf{Z} = \boldsymbol{\mu} + \boldsymbol{\varepsilon} \\ Z_0 = \mu + \varepsilon_0 \end{cases} \tag{4}$$

where t, t_0 are time moments for observables and prediction, respectively; $\mathbf{Z}(t)$, $Z(t_0)$ vector of observables and a single unobservable to be predicted; respectively (for brevity further denoted as $\mathbf{Z}$, Z_0); $\boldsymbol{\mu}$, μ mean values of a random function (constant, known for simple kriging and unknown for ordinary kriging); $\boldsymbol{\varepsilon}(t)$, $\epsilon(t_0)$ correlated disturbances (further denoted as $\boldsymbol{\varepsilon}$, ε_0).

The considered random function is equipped with a stochastic characteristic:

$$E(\boldsymbol{\varepsilon}) = \mathbf{0}, \qquad E(\varepsilon_0) = 0 \tag{5a}$$

$$E\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}\right) = Cov(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}) = Cov(\mathbf{Z}, \mathbf{Z}) = \boldsymbol{\Omega} \tag{5b}$$

$$E(\boldsymbol{\varepsilon}\varepsilon_0) = Cov(\boldsymbol{\varepsilon}, \varepsilon_0) = Cov(\mathbf{Z}, Z_0) = \boldsymbol{\omega} \tag{5c}$$

$$E\left(\varepsilon_0^2\right) = V(\varepsilon_0) = V(Z_0) = V(Z) = \sigma^2 \tag{5d}$$

where $\boldsymbol{\Omega}$ is a covariance matrix (observations – observations), $\boldsymbol{\omega}$ stands for a covariance vector (observations – prediction) and σ^2 is a variance of a random process.

Simple kriging – constant and known mean of a random function

Predictor of simple kriging takes a general affine form:

$$\hat{Z}_0 = \lambda_0 + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{Z} \tag{6}$$

Unbiasedness condition given in (2) leads to the expression for λ_0:

$$\lambda_0 = \mu - \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\mu} \tag{7}$$

and gives the following form of the predictor

$$\hat{Z}_0 = \mu + \boldsymbol{\lambda}^{\mathrm{T}}(\mathbf{Z} - \boldsymbol{\mu}) = \mu + \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\varepsilon} \tag{8}$$

System of equations resulting from optimality conditions solved for coefficients $\boldsymbol{\lambda}$ reads:

$$\boldsymbol{\lambda} = \boldsymbol{\Omega}^{-1}\boldsymbol{\omega} \tag{9}$$

Thus, the predictor takes an explicit form:

$$\hat{Z}_0 = \mu + \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}(\mathbf{Z} - \boldsymbol{\mu}) = \mu + \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\boldsymbol{\varepsilon} \tag{10}$$

or if $\mu = 0$ (this is the case in this study), then:

$$\hat{Z}_0 = \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{Z} \tag{11}$$

In both cases the prediction variance (simple kriging variance) is given by:

$$\sigma_{SK}^2 = \sigma^2 - \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\omega} = \sigma^2 - \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\boldsymbol{\omega} \tag{12}$$

Ordinary kriging – constant and unknown mean of a random function

Predictor of ordinary kriging takes a linear form:

$$\hat{Z}_0 = \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{Z} \tag{13}$$

Unbiasedness condition gives the constraints on coefficients $\boldsymbol{\lambda}$ in the form (they must sum up to unity):

$$\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{u} = 1 \tag{14}$$

where $\mathbf{u}$ is a vector of ones.

System of equations resulting from optimality conditions solved for coefficients $\boldsymbol{\lambda}$ and Lagrange multiplier κ has the following form:

$$\begin{bmatrix} \boldsymbol{\Omega} & \mathbf{u} \\ \mathbf{u}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ \kappa \end{bmatrix} = \begin{bmatrix} \boldsymbol{\omega} \\ 1 \end{bmatrix} \tag{15}$$

Symbolic solution to the system (15) gives the explicit form of the predictor:

$$\hat{Z}_0 = \hat{\mu} + \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}[\mathbf{Z} - \mathbf{u}\hat{\mu}] \tag{16}$$

where:

$$\hat{\mu} = \left(\mathbf{u}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{u}\right)^{-1}\mathbf{u}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{Z} \tag{17}$$

is a generalized least squares estimate of the expected value of a random function. The prediction variance (ordinary kriging variance) is given by:

$$\sigma_{OK}^2 = \sigma^2 - \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\omega} - \kappa = \sigma^2 - \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\boldsymbol{\omega} + \left(1 - \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{u}\right)\left(\mathbf{u}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{u}\right)^{-1}\left(1 - \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{u}\right)^{T} \tag{18}$$

3 Semivariogram model used in the study

Due to the visible parabolic behavior of estimated experimental semivariograms near the origin of coordinate system only such theoretical models have been considered. Among them the best performing one turned out to be Radon transform of exponential model of order 2 [7] (Markov 2nd order model in geodetic literature [10]), i. e.,

$$\gamma(\tau, c_0, c, a) = \begin{cases} 0 & \tau = 0 \\ c_0 + c\left[1 - \left(1 + \frac{\tau}{a}\right) e^{-\left(\frac{\tau}{a}\right)}\right] & \tau > 0 \end{cases} \tag{19}$$

and its covariance function counterpart:

$$C(\tau, c_0, c, a) = \begin{cases} c_0 + c & \tau = 0 \\ c\left(1 + \frac{\tau}{a}\right) e^{-\left(\frac{\tau}{a}\right)} & \tau > 0 \end{cases} \tag{20}$$

where: τ is a lag distance, c_0 is a “nugget effect” (or noise variance), c is a partial sill (or signal variance), a is an autocorrelation range/radius parameter.

Both structure functions are related by the formula (this holds, in general, for all second – order stationary processes):

$$\begin{aligned} \gamma(\tau, c_0, c, a) &= C(0, c_0, c, a) - C(\tau, c_0, c, a) \\ &= (c_0 + c) - C(\tau, c_0, c, a) \end{aligned} \tag{21}$$

The experimental semivariogram for 1D equidistant data has a particularly simple expression, i. e.:

$$\hat{\gamma}(\tau) = \frac{1}{2(N - \tau)} \sum_{i=1}^{N-\tau} (Z_i - Z_{i+\tau})^2 \tag{22}$$

where N stands for the number of observations and τ is a lag distance.

In this contribution, empirical semivariograms were estimated on the basis of limited number of adjacent observations preceding the first prediction time-moment. The number of these observations was a function of nearest neighbours for prediction (to construct kriging system of equation) and days covered by the forecast (15-days) as well as a constant multiplier. The number of time-lags for which the semivariogram values were estimated, was also a function of mentioned parameters. Since the entire procedure was run in an automatic mode and repeated millions of times the theoretical models of semivariograms were fit by nonlinear least squares with box constraints to assure all parameters (c_0, c, a) to be greater than or equal to zero.

4 Data description and processing

In this contribution, a daily time-series of polar motion IERS EOP 14 C04 (IAU2000A) and IERS EOP 05 C04 (IAU2000A) published by IERS have been used (https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html). The first mentioned was chosen since it is consistent with the newest ITRF solution. The latter series was used to compare results from this contribution to the results obtained from Earth Orientation Parameters Prediction Comparison Campaign (EOP PCC) [12]. Also as a by-product the quality of predictions based on two series was compared.

The processed data span the time interval from 1 January 2000 (MJD 51544) to 29 November 2020 (MJD 59182) for IERS EOP 14 C04 and from 1 January 2000 (MJD 51544) to 3 January 2012 (MJD 55929) for IERS EOP 05 C04.

The data were processed in two steps. In the first step (preliminary), the data were being split into training set (of varying length) and validation set (always 15-day) in a continuous way in the entire considered time span. This was to select the best set of steering parameters for the prediction procedure. The second step was the proper prediction based on results obtained from the previous step. The following lines describe the two steps in more detailed way.

In order to choose the best set of parameters, i. e., the number of nearest neighbours (NN) for kriging – based prediction and the time span (TS) for linear trend and periodic components estimation (Chandlerian and annual) the following sequence of instructions was repeated for each coordinate (PMx, PMy) and each method (simple kriging assuming zero mean, ordinary kriging). The time span (TS) is, in fact, the length of a moving window sliding across the analysed time interval.

– a loop was run over the number of NN. NN was varying from 10 to 30 with 1 neighbour increment.
– a loop was run over the TS. TS was varying from 6 to 9 years with 0.1 yr. increment.
– the main prediction function (dependent on NN and TS) with an internal loop running from 51544 MJD to 59182 MJD (IERS EOP 14 C04) and from 51544 MJD to 55929 MJD (IERS EOP 05 C04) with 1-day increment, i. e., for a given NN and TS a 15-day prediction was made then TS was shifted 1-day and a new 15-day prediction was made and so on.

For the series IERS EOP 14 C04, the total number of 15-day predictions for each coordinate and each method amounts to 3 180 324 (12 721 296 for all), this translates to 47 704 860 1-day predictions (190 819 440 for all). Since the ordinary kriging prediction turned out to be more accurate, only this method was applied for the series IERS EOP 05 C04, thus the total number of 15-day predictions for both coordinates amounts to 2 125 242, this translates to 31 878 630 1-day predictions. Due to the large amount of data, the results of this preliminary step are not presented in the paper but may be, in a condensed form, accessed at the website mentioned in the acknowledgements section. The described procedure enabled to select the optimal values for TS equal to 6.4 yrs. (app. the beat period of Chandlerian and annual oscillations) and NN equal to 10 immediate neighbours. The above values were selected one the basis of frequency of occurrences of minimal absolute prediction errors for given TS and NN (there were some other values for the parameters but occurred with lesser frequency).

In the proper prediction step the above-mentioned main prediction function was run with optimal values for TS and NN giving the total number of 15-day predictions equal to 5287 (IERS EOP 14 C04) and 2034 (IERS EOP 05 C04).

The results obtained from two steps served:

a) for deciding which method performs better, i. e., simple kriging or ordinary kriging,
b) for comparison to prediction methods included in EOP PCC
c) for comparison of predictability of polar motion between the two series consistent with different ITRFs, i. e., ITRF2014 and ITRF2005

In the next section, the results will be presented in a synthetic form (statistics from 5287 and 2034 15-day predictions) and the detailed results of every single 15-day prediction altogether with corresponding MJD may be accessed at the website mentioned in the acknowledgements section.

5 Results

As synthetic indicators of kriging-based prediction quality the following prediction error summary statistics were used.

Prediction error

$$e_i = O_i - P_i \tag{23}$$

where O denotes observed and P predicted polar motion coordinate.

Mean Absolute Prediction Error (MAPE)

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}|e_i| \tag{24}$$

Root Mean Square Prediction Error (RMSPE)

$$RMSPE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}e_i^2} \tag{25}$$

Median Absolute Prediction Error (MeAPE)

$$MeAPE = med\,|e_i| \tag{26}$$

The analysis of kriging-based prediction will begin with comparison of two adopted kriging variants for IERS EOP 14 C04 time series. Prediction error statistics, i. e., MAPE, MeAPE. RMSPE were computed on the basis of 5287 15-day predictions, thus constitute a reliable source of information on prediction capabilities of analysed kriging variants. Table 1 and Table 2 show a summary of prediction error statistics for PMx and PMy for each day of 15–day forecast. Results indicate a better prediction performance of ordinary kriging. MAPEs for PMx are from 5 to 16 percent higher for simple kriging (13 % on average). The least differences are visible for the first 5 days of forecast. In case of PMy, the differences vary from 2 to 9 percent (6 % on average). Similar magnitude of differences holds for the remaining summary statistics with the difference that MeAPE supresses the influence of large prediction errors (50 % of absolute errors are below this value) whilst RMSPE magnifies them. The known fact of better predictability of PMy is also confirmed. The explanation of why ordinary kriging outperforms simple kriging follows the fact that simple kriging uses the zero-mean assumption (known expected value) since the residuals after least squares detrending possess this property. But this assumption holds for the residuals from the entire time

Table 1: Prediction statistics for ordinary and simple kriging for the PMx (IERS EOP 14 C04).

	ordinary kriging			simple kriging		
Day	MAPE [mas]	MeAPE [mas]	RMSPE [mas]	MAPE [mas]	MeAPE [mas]	RMSPE [mas]
1	0.67	0.47	0.94	0.71	0.51	0.97
2	1.02	0.81	1.34	1.09	0.86	1.42
3	1.37	1.11	1.77	1.49	1.20	1.92
4	1.72	1.41	2.20	1.90	1.53	2.44
5	2.06	1.68	2.63	2.31	1.87	2.96
6	2.39	1.96	3.04	2.71	2.19	3.48
7	2.71	2.23	3.44	3.10	2.49	3.99
8	3.03	2.53	3.83	3.49	2.81	4.49
9	3.35	2.80	4.23	3.87	3.15	4.99
10	3.66	3.06	4.63	4.25	3.48	5.48
11	3.97	3.31	5.03	4.62	3.77	5.96
12	4.27	3.54	5.42	4.98	4.05	6.43
13	4.58	3.77	5.82	5.34	4.34	6.89
14	4.89	4.00	6.21	5.70	4.62	7.34
15	5.19	4.26	6.59	6.05	4.90	7.78

Table 2: Prediction statistics for ordinary and simple kriging for the PMy (IERS EOP 14 C04).

Day	ordinary kriging			simple kriging		
	MAPE [mas]	MeAPE [mas]	RMSPE [mas]	MAPE [mas]	MeAPE [mas]	RMSPE [mas]
1	0.44	0.33	0.60	0.45	0.33	0.61
2	0.69	0.57	0.88	0.70	0.58	0.90
3	0.94	0.79	1.19	0.97	0.81	1.22
4	1.19	1.02	1.49	1.23	1.05	1.54
5	1.43	1.22	1.79	1.49	1.26	1.86
6	1.66	1.40	2.08	1.74	1.46	2.18
7	1.88	1.57	2.36	1.98	1.67	2.50
8	2.09	1.75	2.63	2.23	1.84	2.81
9	2.31	1.92	2.91	2.48	2.04	3.13
10	2.53	2.12	3.19	2.73	2.26	3.45
11	2.77	2.33	3.48	2.99	2.47	3.78
12	3.00	2.53	3.78	3.26	2.69	4.12
13	3.25	2.76	4.07	3.53	2.90	4.45
14	3.49	3.02	4.38	3.80	3.14	4.79
15	3.74	3.22	4.69	4.07	3.39	5.13

Table 3: Comparison of MAPEs for ordinary kriging for PMx and PMy for IERS EOP 14 C04 and IERS EOP 05 C04.

Day	PMx			PMy		
	MAPE (14) [mas]	MAPE (05) [mas]	Difference [mas]	MAPE (14) [mas]	MAPE (05) [mas]	Difference [mas]
1	0.66	0.60	0.06	0.47	0.44	0.03
2	1.01	0.93	0.09	0.71	0.66	0.05
3	1.38	1.27	0.11	0.96	0.88	0.08
4	1.74	1.61	0.13	1.20	1.09	0.10
5	2.10	1.95	0.16	1.42	1.30	0.12
6	2.45	2.27	0.18	1.64	1.51	0.14
7	2.79	2.59	0.20	1.85	1.70	0.15
8	3.12	2.91	0.21	2.06	1.89	0.16
9	3.45	3.22	0.23	2.27	2.08	0.18
10	3.76	3.52	0.24	2.48	2.28	0.20
11	4.05	3.81	0.25	2.70	2.48	0.22
12	4.34	4.09	0.26	2.92	2.68	0.24
13	4.63	4.37	0.27	3.14	2.88	0.26
14	4.94	4.66	0.28	3.37	3.09	0.28
15	5.25	4.95	0.30	3.59	3.29	0.30

span (TS) of 6.4 yrs. but not for the limited neighbourhood as adopted in this contribution. In the limited neighbourhood, here NN = 10, the mean may significantly differ from zero, thus ordinary kriging with its estimated mean within the immediate neighbourhood gives better predictions. Hence, in further comparisons only ordinary kriging will be considered.

Figure 1 presents a comparison between the results of ultra-short prediction (10-day into the future) of EOP PCC [12] and this work (denoted as TW in the Figure 1). Denotations of prediction techniques are the same as in EOP PCC. MAPEs for ordinary kriging used in Figure 1 are listed in Table 3. The figure shows only individual solutions since the combined solution of EOP PCC is a function of individual submissions (a weighted average). It is worth noting that different data spans were used to compute the prediction errors between the EOP PCC and this contribution. In this study predictions were made in a continuous way with a one-day shift and no predictions were excluded hence, in this way, the method was tested at "good" and "bad" time moments. The results clearly indicate that the accuracy of ordinary kriging prediction of polar motion is comparable (often better) with the most accurate prediction methods that took part in EOP PCC, i. e., LS extrapolation of the harmonic model + AR prediction, spectral analysis + LS extrapolation and neural networks.

Table 3 presents a comparison between mean absolute prediction errors for polar motion time series consistent with different terrestrial frames, i. e., IERS EOP 14 C04 and IERS EOP 05 C04. The comparison incorporates the same time span 51544 MJD – 55929 MJD and consists of 2034 15-day predictions performed with the same control parameters (TS and NN). It is visible that IERS EOP 05 C04 is more predictable as measured by MAPE than IERS EOP 14 C04. This "unpredictability" of IERS EOP 14 C04 is on average on the level of 0.20 mas for PMx and 0.17 mas for PMy. So far, authors have not found a reason for these differences in predictability and the answer cannot be given whether it is method-dependent (here kriging) or a general finding.

In this place it is worth mentioning some remarks on structure functions which play a key-role in all geostatistical analyses. Figure 2 shows examples of empirical semivariograms for various 15-day forecasts for residual PMx (after removing linear trend and periodicities). Those Semivariograms were estimated on the basis of limited number of observations preceding the first day of forecast given in the Figure 2. As can be seen empirical semivariograms reveal rather negligible (or not at all) noise variance ("nugget effect" in geostatistical parlance). On the other hand, the signal variance (partial sill) changes significantly in magnitude even for 15-day predictions separated by 10 days (or less). The same holds for the range of autocorrelation which certainly is not constant.

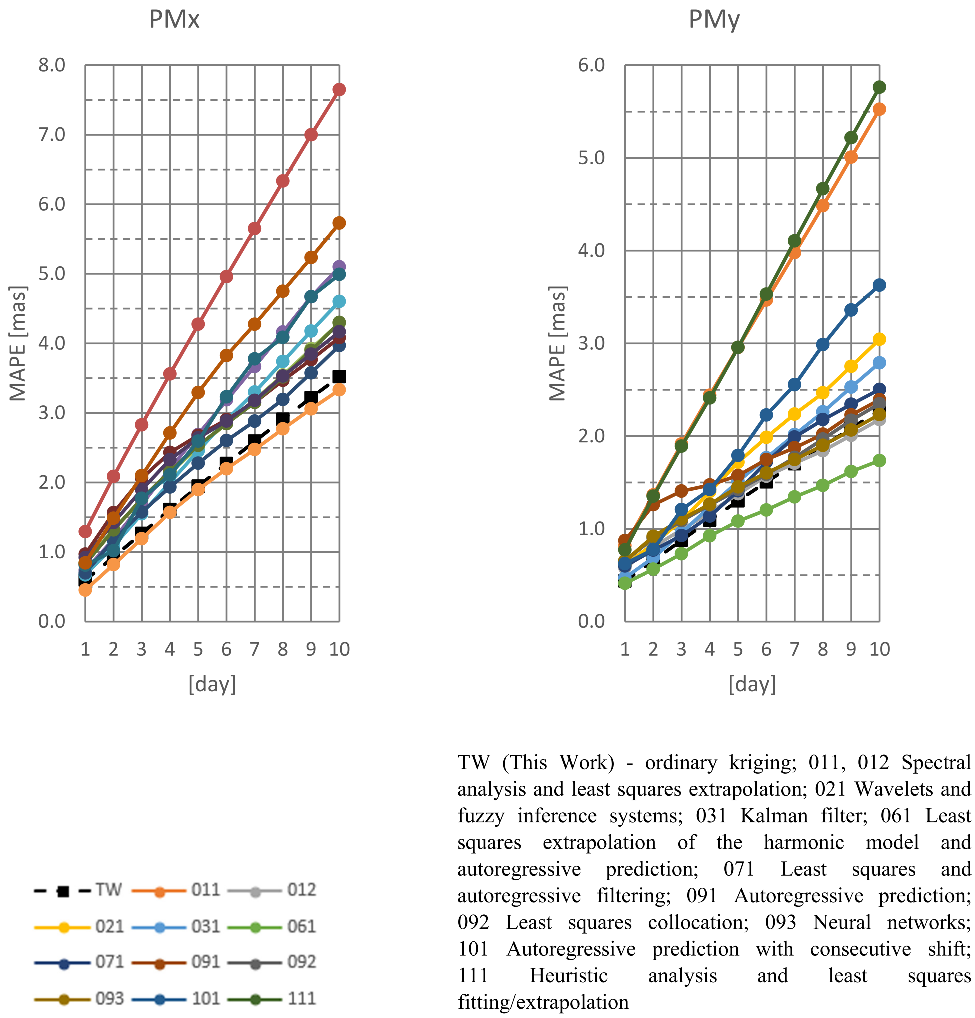


Figure 1: Comparison of ordinary kriging based ultra-short (10-day) prediction with the methods involved in EOP PCC, MAPEs for PMx (left panel) and for PMy (right panel).

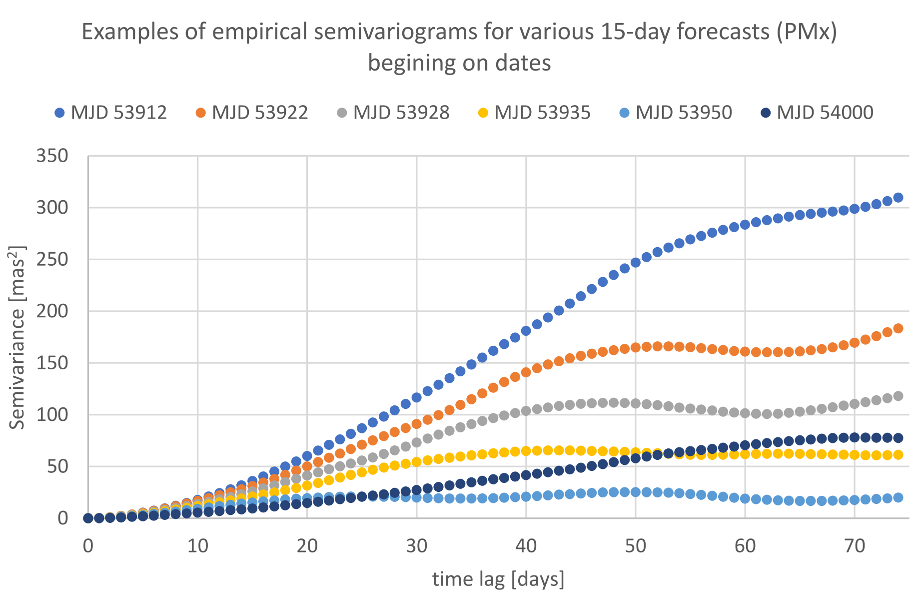


Figure 2: Examples of empirical semivariograms for various 15-day forecasts (residual PMx).

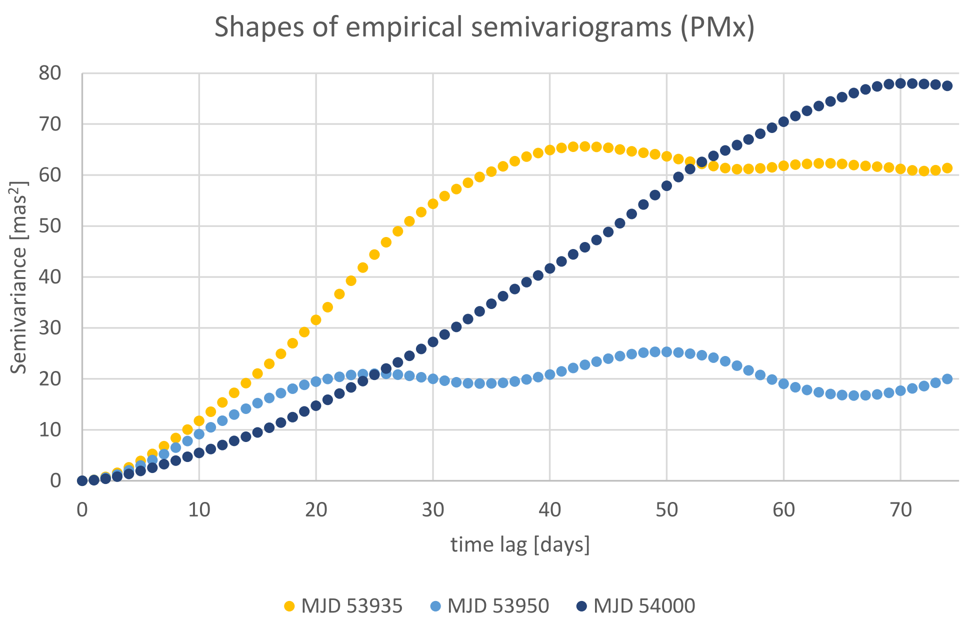


Figure 3: Shapes of empirical semivariograms for various 15-day forecasts (residual PMx).

In general, semivariograms differ in shape what perhaps is not clearly visible in Figure 2 (due to significantly different signal variances) hence in Figure 3 only semivariograms for MJDs 53935, 53950, 54000 are shown. All semivariograms reveal parabolic behavior approaching the origin of the coordinate system what proves that residual PMx is a very regular process. Some semivariograms reveal not modelled periodicities even for short time intervals like, e. g., MJD 53950. But these issues were not analyzed since only semivariogram values for the first 25 lags were necessary for prediction procedure presented herein and the semivariogram fitting procedure was so oriented to best-fit on the mentioned number of lags. The same findings are valid for semivariograms representing residual PMy.

6 Conclusions

In this contribution, ultra-short term prediction of polar motion with the use of geostatistical method of kriging has been performed on the basis of IERS EOP 14 C04 (IAU2000A) and IERS EOP 05 C04 (IAU2000A) data. Simple kriging due to the assumption of known mean, i. e., zero-mean residual process constructed after detrending is visibly a little worse than ordinary kriging where a mean value is estimated locally. The results indicate that the accuracy of ordinary kriging prediction is comparable with the most accurate prediction methods that took part in EOP PCC, i. e., LS extrapolation of the harmonic model + AR prediction, spectral analysis + LS extrapolation and neural networks. The quality of prediction is visibly coordinate-dependent, PMx is worse predictable than PMy. It is also visible that IERS EOP 05 C04 is slightly better predictable than IERS EOP 14 C04 under the same control parameters. Prediction accuracy as measured by MAPE with respect to IERS 14 C04 is ranging from 0.66 to 5.25 mas for PMx and from 0.47 to 3.59 mas for PMy (the 1st and the last day of forecast). For IERS 05 C04 the values are 0.60–4.95 mas and 0.44–3.29 mas for PMx and PMy; respectively. The great number of performed predictions gives reliable conclusions as to the applied methods. Due to the enormous number of predicted daily values only final ones in synthetic form were presented in the content of the paper but intermediate results are available at the website mentioned in the acknowledgements section.

Acknowledgment: The authors are very grateful to M. Kalarus for providing EOP PCC data.

Intermediate results are available at the website accompanied with this article home.agh.edu.pl/mmichalc.

Funding: The paper is a result of research on geospatial methods carried out within the statutory research grant no. 16.16.150.545 at the Department of Integrated Geodesy and Cartography, AGH University of Science and Technology, Krakow.

References

[1] Abdelazeem M, Çelik R N, El-Rabbany A. An accurate Kriging-based regional ionospheric model using combined GPS/BeiDou observations. Journal of Applied Geodesy, 2018 12(1), 65–76.

[2] Al-Shaery A, Lim S, Rizos C. Functional models of ordinary kriging for medium range real-time kinematic positioning based on the Virtual Reference Station technique. In: Proceedings of the ION GNSS+, Portland, OR, USA, 21–24 September 2010, pp. 2513–2521.

[3] Bizouard Ch, Lambert S, Gattano C, Becker O, Richard J.-Y. The IERS EOP 14C04 solution for Earth orientation parameters consistent with ITRF 2014. Journal of Geodesy, 2019, 93, 621–633.

[4] Dow JM, Neilan RE, Rizos C, The international GNSS Service in a changing landscape of global navigation satellite systems. J Geod, 2009, 83(3–4), 191–198.

[5] Ferrándiz, J.M.; Navarro, J.; Escapa, A.; Getino, J. The Earth Rotation: a Challenging Problem in Mathematics and Physics. Pure and Applied Geophysics, 172, pp. 57–74, 2015.

[6] Hozakowski W. Polar motion prediction by the least-squares collocation method. In: Boucher C, Wilkins GA (eds.), Earth rotation and coordinate reference frames. International Association of Geodesy Symposia, Springer, 1990, 105, 50–57.

[7] Huang Ch, Zhang H, Robeson S. On the Validity of Commonly Used Covariance and Variogram Functions on the Sphere. Mathematical Geosciences. 2011, 43. 721–733.

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

[9] IERS Conventions (2010), Petit G and Luzum B (eds.). IERS Technical Note 36, Frankfurt am Main: Verlag des Bundesamts für Kartographie und Geodäsie. 179 pp., ISBN 3-89888-989-6.

[10] Jordan S K. Self-Consistent statistical model for the gravity anomaly, vertical deflections, and undulation of the geoid. J. Geophys. Res., 1972, 77(20), 3660–3670.

[11] Jovanović B. An approximation of tabulated function. Publications De L'Institut Mathematique, 1987, 41(55), 143–148.

[12] Kalarus M, Schuh H, Kosek W, Akyilmaz O, Bizouard Ch, Gambis D, Gross, Jovanović B R, Kumakshev S, Kutterer H, Mendes Cerveira P J, Pasynok S, Zotov L. Achievements of the Earth Orientation Parameters prediction comparison campaign. Journal of Geodesy, 2010, 84, 587–596.

[13] Kosek W, Kalarus M, Niedzielski T. Forecasting Of the Earth Orientation Parameters – Comparison of Different Algorithms, Journées Systèmes de Référence Spatio-temporels 2007, Observatoire de Paris, 17–19 September 2007.

[14] Lei Y, Cai H, Zhao D. Improvement of the prediction accuracy of polar motion using empirical mode decomposition. Geodesy and Geodynamics, 2017, 8(2), 141–146.

[15] Ligas M, Kulczycki M. Kriging and moving window kriging on a sphere in geometric (GNSS/levelling) geoid modelling, Survey Review, 2018, 50(359), 155–162.

[16] Ligas M, Szombara S. Geostatistical prediction of a local geometric geoid – kriging and cokriging with the use of EGM2008 geopotential model. Studia Geophysica et Geodaetica, 2018, 62, 187–205.

[17] Modiri S, Belda S, Heinkelmann R, Hoseini M, Ferrándiz J M, Schuh H. Polar motion prediction using the combination of SSA and Copula based analysis. Earth, Planets and Space, 2018, 70, 115.

[18] Modiri S, Belda S, Heinkelmann R, Hoseini M, Ferrándiz J M, Schuh H. A new hybrid method to improve the ultra-short-term prediction of LOD, Journal of Geodesy 94(2), 2020.

[19] Moritz H, Mueller I: Earth Rotation: Theory and Observation. Ungar Pub., 1988.

[20] Okhotnikov G, Golyandina N. EOP Time Series Prediction Using Singular Spectrum Analysis, In: Corpetti T, Ienco D, Interdonato R, et al. (eds.), Proceedings of MACLEAN: MAChine learning for EArth observation workshop, RWTH Aahen University, CEUR Workshop Proceedings, 2019.

[21] Reguzzoni M, Sansò F, Venuti G. The theory of general kriging, with applications to the determination of a local geoid. Geophysical Journal International, 2005, 162, 303–314.

[22] Schuh H, Ulrich M, Egger D, Müller J, Schwegmann W. Prediction of Earth orientation parameters by artificial neural networks, Journal of Geodesy. 2002, 76, 247–258.

[23] Stein L M. Interpolation of spatial data – some theory for kriging, Springer, New York, 1999.

[24] Xu X, Zhou Y, EOP prediction using least square fitting and autoregressive filter over optimized data intervals, Advances in Space Research, 2015, 56(10), 2248–2253.

[25] Ye F, Yuan Y, Deng Z, Improved Ultra-Rapid UT1-UTC Determination and Its Preliminary Impact on GNSS Satellite Ultra-Rapid Orbit Determination, Remote Sensing, 2020, 12, 3584.

[26] Zajdel, R., Sośnica, K., Bury, G., Dach R., Prange L., System-specific systematic errors in earth rotation parameters derived from GPS, GLONASS, and Galileo. GPS Solut, 2020 24, 74.

Available online at www.sciencedirect.com

ScienceDirect

Advances in Space Research 70 (2022) 610–620

ADVANCES IN SPACE RESEARCH
(a COSPAR publication)

www.elsevier.com/locate/asr

The (ultra) short term prediction of length-of-day using kriging

Maciej Michalczak *, Marcin Ligas

Department of Integrated Geodesy and Cartography, Faculty of Geo-Data Science, Geodesy, and Environmental Engineering, AGH University of Science and Technology, Krakow, Poland

Received 10 November 2021; received in revised form 30 April 2022; accepted 5 May 2022
Available online 11 May 2022

Abstract

The contribution presents results of ultra-short-term (15 days into the future) and short-term (30 days into the future) LOD prediction using kriging. IERS EOP 14 C04 (IAU2000A) and IERS EOP 05 C04 (IAU2000A) series are used as a reference. IERS EOP 05 C04 (IAU2000A), which is no longer in force, was used for comparison with previously published results. Two strategies of semivariogram selection for prediction purposes are employed. The first one uses a single semivariogram model in the entire prediction procedure whilst the second one selects a semivariogram model, from a group of models, with the least residual sum of squares. This work's implementation of kriging offers mean absolute prediction errors in ultra-short-term forecast for days 1–15 with respect to IERS EOP 14 C04 on the level of 0.035–0.149 ms whilst for IERS EOP 05 C04 the values are 0.030–0.139 ms. For short term prediction (days 1–30) the values are 0.038–0.180 ms and 0.034–0.175 ms; respectively. The results show that LOD kriging-based prediction may compete with existing methods.

Keywords: Length-of-day; Prediction; Kriging; Semivariogram; Time series

1. Introduction

Earth Orientation Parameters (EOP) constitute a set of parameters that are crucial for the transformation between the Celestial Reference Frame (CRF) and the Terrestrial Reference Frame (TRF). The parameters, namely, coordinates of the pole (PMx, PMy), UT1-UTC and celestial pole offsets (dX, dY) are constantly changing, making this transformation time-dependent. An additional parameter that is the first negative derivative from the UT1-UTC, the length of day (LOD), is also used to model changes in the Earth's rotation rate (Modiri et al., 2020). Determination of EOP employs space geodetic techniques: VLBI (Very Long Baseline Interferometry), SLR (Satellite Laser Ranging), DORIS (Doppler Orbitography and Radiopositioning Integrated by Satellite) and GNSS (Global Navigation Satellite System) that are finally combined into a single solution (Petit and Luzum, 2010). Combination of these techniques enables to determine the polar coordinates with the accuracy of 50 µas or better and app. 10 µs in case of LOD (Dick and Thaller, 2020). International Earth Rotation and Reference System Service (IERS) publishes official EOP data via: the IERS EOP Product Centre and IERS Rapid Service/Prediction Centre (Dick and Thaller, 2020). The final IERS EOP data are published with thirty day latency; therefore, in operational settings, this gap must be filled with rapid products, e.g., Daily Rapid EOP Data published by the IERS (https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html#Rapid) or (ultra)-rapid products delivered by International GNSS Service (IGS) (https://www.igs.org/products/#about) (Dow et al., 2009).

* Corresponding author.
E-mail addresses: mmichalc@agh.edu.pl (M. Michalczak), ligas@agh.edu.pl (M. Ligas).

https://doi.org/10.1016/j.asr.2022.05.007
0273-1177/

Near real-time EOP are needed for various astronomical and geodetic purposes such as: precise satellite orbit determination, positioning and navigation, space tracking, navigation by the Deep Space Network (DSN), time-keeping and positional astronomy (Gambis and Luzum, 2011). Rapid EOP products are published daily, nevertheless it is necessary to know the EOP for the days ahead and therefore forecasting EOP is still a very relevant issue. Owing to the fact that, the pole coordinates and UT1–UTC and therefore also the LOD change rapidly over time, a prediction of these parameters is still a challenge (Niedzielski and Kosek, 2008). Variations in LOD may be of tidal, which can be modelled with high precision (Petit and Luzum, 2010) and non-tidal origin, which occur during the exchange of angular momentum between the Earth's crust and global atmosphere (Schuh et al., 2002). Length of day reveals also changes in the Earth's rotation, caused by, among other things, the tropical monsoon during El Niño events, therefore it is difficult to predict (Gross et al., 1996) (Niedzielski and Kosek, 2008).

Many prediction methods have been involved to forecast EOP parameters values in recent years. Examples of them are: least-squares collocation (Hozakowski, 1989, Zotov et al., 2018), artificial neural networks (Schuh et al., 2002; Liao et al., 2012), autocovariance prediction (Kosek, 2002), autoregression (Kosek et al., 2007), combination of least-squares and multivariate stochastic methods (Niedzielski and Kosek, 2008), wavelets and fuzzy inference systems (Akyilmaz et al., 2011), extreme learning machine (Lei et al., 2015), gaussian process regression (Lei et al., 2015), singular spectrum analysis (Okhotnikov and Golyandina, 2019), one-step forecasting method based on the LS + AR (Wu et al., 2019), combination of singular spectrum analysis and Copula-based analysis (Modiri et al., 2020) and Kalman filter (Nastula et al., 2020).

In this contribution, geostatistical method of prediction, i.e., kriging is extensively tested in LOD prediction process since it has already proven its applicability to polar motion prediction Michalczak and Ligas (2021). Kriging is the generic name of a family of univariate and multivariate linear and nonlinear predictors in the field of geostatistics. The goal of kriging is to predict in the unbiased fashion with minimization of the mean square prediction error. In case of linear methods, kriging is recognized as BLUP (Best Linear Unbiased Predictor). This study uses two variants of kriging, i.e., simple kriging and ordinary kriging. Also, two strategies of using the semivariogram function are applied. The first one uses a single-semivariogram model in the entire prediction procedure whilst the second one selects the best-fitting one chosen from a group of models.

The contribution uses IERS EOP 14 C04 (IAU2000A) and IERS EOP 05 C04 (IAU2000A) series as a reference for ultra-short-term (up to 15 days into the future) and short-term (up to 30 days into the future) LOD prediction. IERS EOP 05 C04 (IAU2000A), which is no longer updated and published, was used for comparison with previously published results. For both reference series the process of forecasting LOD starts on 1 January 2000 (MJD 51544) and ends on 18 May 2021 (MJD 59352) for IERS EOP 14 C04 and on 3 January 2012 (MJD 55929) for IERS EOP 05 C04. Due to the large number of intermediate results, they are stored on the website associated with this study and mentioned in the acknowledgement section.

2. Methodology

2.1. Data pre-processing

LOD includes well-recognized periodic effects associated with solid Earth tides with periods from 5 days up to 18.6 years. These effects may be removed from observed LOD data beforehand and restored after the prediction procedure. For this purpose, model recommended by IERS Conventions (2010) (Petit and Luzum, 2010), given in (1), was used. This model is implemented in "RG_ZONT2.f" routine and it is available on the IERS Conventions website at https://iers-conventions.obspm.fr.

$$\begin{cases} dLOD = \sum_{i=1}^{62} B_i' sin\xi_i + C_i' cos\xi_i \\ \xi_i = \sum_{j=1}^{5} a_{ij}\alpha_j \end{cases} \tag{1}$$

where values of B'_i, C'_i, a_{ij}, α_j are given in chapter 8 of the IERS Conventions (2010) (Petit and Luzum, 2010).

2.2. Prediction model

LOD prediction model consists of two parts, namely, deterministic and stochastic. The first one involves: estimated linear trend, estimated seasonal variations (sine wave terms) such as annual and semi-annual oscillations and the tidal corrections removed as mentioned in the previous section. The stochastic part consists of predicted (kriged) residuals based on residuals obtained after the removal of abovementioned components. The adopted prediction model may be given as:

$$\begin{aligned} Prediction &= linear\ trend + periodic\ components \\ &\quad + tidal corrections + predicted residuals \end{aligned}$$

or

$$\begin{aligned} LOD(t_p) &= \hat{a}_0 + \hat{a}_1 t_p + \sum_{i=1}^{3} \hat{b}_i \sin\left(\hat{c}_i t_p + \hat{d}_i\right) \\ &\quad + tidal corrections + \hat{\varepsilon}_{LOD}(t_p) \end{aligned} \tag{2}$$

where t_p is a prediction time moment, $\hat{a}_0, \hat{a}_1$ are the estimated linear trend coefficients; $\hat{b}_i, \hat{c}_i, \hat{d}_i$ are the estimated amplitude, frequency and phase shift for i-th sine wave term; respectively, tidal corrections given in (1) and $\hat{\varepsilon}$ stands for the predicted residual.

2.3. Brief on kriging

Kriging is the generic name of a family of univariate and multivariate linear and nonlinear predictors in the field of geostatistics. The name was given to honor D. G. Krige's pioneering works in evaluating mineral resources. The works were later considerably developed by G. Matheron and his followers from Ecole des Mines de Paris at Fontainebleau.

Two variants of kriging have been used in this contribution for LOD prediction, i.e., simple (SK) and ordinary kriging (OK). They are optimal in the sense of BLUP if the mean value of a random process is a known constant (SK) or an unknown constant (OK) (Cressie, 1993). Kriging is used to forecast future values of a residual part of the LOD prediction model obtained after the removal of tidal effects, trend and periodicities.

Kriging as a prediction method employs two optimality criteria, i.e., unbiasedness and minimization of the mean squared prediction error, i.e.:

$$E(e) = E\left(\hat{Z}_0 - Z_0\right) = 0 \; and \; E\left(ee^T\right) \rightarrow min \tag{3}$$

where e is a prediction error, $\widehat{Z}_0$ is the predictor of a random variable Z_0, E is the expected value operator. The carrier of information about the variability and regularity of a random process is the structure function that is essential in the prediction procedure. This structure may be captured either by a covariance function or a semivariogram.

In simple kriging the predictor of a random variable Z_0 takes an explicit form given by:

$$\hat{Z}_0 = \mu + \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}(\mathbf{Z} - \boldsymbol{\mu}) \tag{4}$$

or if $\mu = 0$, then:

$$\hat{Z}_0 = \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{Z} \tag{5}$$

with the prediction variance expressed as:

$$\sigma^2_{SK} = \sigma^2 - \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\boldsymbol{\omega} \tag{6}$$

where $\mathbf{Z}$ stands for a vector of observables; μ is the mean value of a random function, $\boldsymbol{\Omega}$ is a covariance matrix (observations – observations), $\boldsymbol{\omega}$ stands for a covariance vector (observations – prediction) and σ^2 is a variance of a random process.

Ordinary kriging predictor takes a very similar form, i.e.:

$$\hat{Z}_0 = \hat{\mu} + \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}(\mathbf{Z} - \mathbf{u}\hat{\mu}) \tag{7}$$

where a constant and known expected value of SK is replaced by its generalized least squares estimate, i.e.:

$$\hat{\mu} = \left(\mathbf{u}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{u}\right)^{-1}\mathbf{u}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{Z} \tag{8}$$

where $\mathbf{u}$ is a vector of ones.

The prediction variance of OK includes an additional positive component (the most right hand side of eq. (9)) which increases the variance due to the uncertainty introduced with the estimation of a mean value:

$$\sigma^2_{OK} = \sigma^2 - \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\boldsymbol{\omega} + \left(1 - \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{u}\right)\left(\mathbf{u}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{u}\right)^{-1}\left(1 - \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{u}\right)^T \tag{9}$$

2.4. Semivariogram models used in the study

This contribution uses two approaches of semivariogram selection for prediction purposes. The first one uses a single semivariogram model that is fitted continuously (in each window) to the empirical ones. The second one uses an adaptive strategy, i.e., selects a semivariogram model, from a group of models, with the least residual sum of squares (RSS). In this study, after many trials of fitting theoretical semivariograms, (cubic, exponential, gaussian, pentaspherical, spherical, sine hole effect, Markov 2nd order, Markov 3rd order, rational-quadratic with exponents m = 1, ½, 3/2) finally, three semivariogram models were employed for both a single-semivariogram and adaptive approach (Olea, 1999), i.e.:

a) cubic

$$\gamma(\tau, c_0, c, a) = \begin{cases} 0 & \tau = 0 \\ c_0 + c\left[7\left(\frac{\tau}{a}\right)^2 - \frac{35}{4}\left(\frac{\tau}{a}\right)^3 + \frac{7}{2}\left(\frac{\tau}{a}\right)^5 - \frac{3}{4}\left(\frac{\tau}{a}\right)^7\right] & 0 < \tau \le a \\ c_0 + c & \tau > a \end{cases} \tag{10}$$

b) exponential

$$\gamma(\tau, c_0, c, a) = \begin{cases} 0 & \tau = 0 \\ c_0 + c\left[1 - e^{-\left(\frac{\tau}{a}\right)}\right] & \tau > 0 \end{cases} \tag{11}$$

c) sine hole effect (wave effect)

$$\gamma(\tau, c_0, c, a) = \begin{cases} 0 & \tau = 0 \\ c_0 + c\left[1 - \frac{sin\left(\frac{\pi\tau}{a}\right)}{\frac{\pi\tau}{a}}\right] & \tau > 0 \end{cases} \tag{12}$$

and their covariance function counterparts; respectively:

a) cubic

$$C(\tau, c_0, c, a) = \begin{cases} c_0 + c & \tau = 0 \\ c\left[1 - 7\left(\frac{\tau}{a}\right)^2 + \frac{35}{4}\left(\frac{\tau}{a}\right)^3 - \frac{7}{2}\left(\frac{\tau}{a}\right)^5 + \frac{3}{4}\left(\frac{\tau}{a}\right)^7\right] & 0 < \tau \le a \\ 0 & \tau > a \end{cases} \tag{13}$$

b) exponential

$$C(\tau, c_0, c, a) = \begin{cases} c_0 + c & \tau = 0 \\ ce^{-\left(\frac{\tau}{a}\right)} & \tau > 0 \end{cases} \tag{14}$$

c) sine hole effect (wave effect)

$$C(\tau, c_0, c, a) = \begin{cases} c_0 + c & \tau = 0 \\ c\frac{sin\left(\frac{\pi\tau}{a}\right)}{\frac{\pi\tau}{a}} & \tau > 0 \end{cases} \tag{15}$$

where: τ is a lag distance, c_0 is a "nugget effect" (or noise variance), c is a partial sill (or signal variance), a is an autocorrelation range/radius parameter.

The experimental semivariogram for 1D equidistant data, as it is the case for LOD series, has a particularly simple expression, i.e.:

$$\widehat{\gamma}(\tau) = \frac{1}{2(N - \tau)}\sum\nolimits_{i=1}^{N-\tau}(Z_i - Z_{i+\tau})^2 \tag{16}$$

where N stands for the number of observations and τ is a lag distance.

Empirical semivariograms and kriging systems of equations were being constructed from LOD residuals located in the immediate vicinity of the prediction time span (15-days or 30-days forecast). Number of quasi-observations (residuals) adopted for empirical semivariogram estimation was a multiple of the sum of nearest neighbors for prediction and days covered by the forecast. The automatic mode of the prediction procedure forced the use of nonlinear least squares with box constraints to fit theoretical models of semivariograms in order to assure all parameters (c0, c, a) to be positive.

3. Data description and processing

Daily time series of LOD are collected from IERS EOP 14 C04 (IAU2000A) and IERS EOP 05 C04 (IAU2000A) series published by International Earth Rotation and Reference Systems Service – IERS (https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html).

The IERS EOP 05 C04 was used to compare results of this contribution to those obtained in the 1st Earth Orientation Parameters Prediction Comparison Campaign (EOP PCC) Kalarus et al. (2010). IERS EOP 14 C04 series was used since it is consistent with the latest solution of ITRF.

The processed data span the time interval from 1 January 2000 (MJD 51544) to 18 May 2021 (MJD 59352) for IERS EOP 14 C04 and from 1 January 2000 (MJD 51544) to 3 January 2012 (MJD 55929) for IERS EOP 05 C04. The last data point is the last available observation at the time of computations for IERS EOP 14 C04 and the last data point published for IERS EOP 05 C04 (no longer in force).

The raw data were processed in three steps. Firstly, tidal effects were removed according to (3). Secondly, data were split into training sets (of varying length) and validation sets (of 15-days or 30-days length) in a continuous way in the entire considered time span. This step was used to select the best set of control parameters for the final prediction procedure, i.e., the number of nearest neighbours (NN) necessary to construct kriging systems of equations and the time span (TS) necessary to estimate parameters of a linear trend and periodic components. The third step was to perform final prediction based on the best steering parameters chosen in the previous step.

To determine the best set of control parameters the following sequence of instructions was repeated for each kriging method (simple and ordinary kriging). For 15-days prediction a loop was run over:

- the number of NN, varying from 15 to 25 with 1 neighbour increment
- the TS, varying from 6 to 13 yrs. for IERS EOP 14 C04 and from 6 to 10 yrs. for IERS EOP 05 C04 with 0.1 yr. increment

After the selection of "optimal" steering parameters for 15-days prediction it was decided to limit the range of steering parameters and to use only ordinary kriging for short-term prediction (this will be clarified and confirmed by the results of the next section), therefore for 30-days prediction a loop was run over:

- the number of NN, varying from 21 to 25 with 1 neighbour increment
- the TS, varying from 11 to 13 yrs. for IERS EOP 14 C04 and from 8 to 10 yrs. for IERS EOP 05 C04 with 0.1 yr. increment

The length of the IERS EOP 05 C04 data is 4385 days, hence the TS parameter was limited to 10 yrs. (around 3650 days) to obtain a reasonable number of 15-days and 30-days predictions for further inferences on the accuracy of the kriging-based forecast. The main prediction routine (dependent on control parameters and number of sine wave terms) was run with an internal loop from 51544 MJD (1 January 2000) to 59352 MJD (18 May 2021) (IERS EOP 14 C04) and from 51544 MJD (1 January 2000) to 55929 MJD (3 January 2012) (IERS EOP 05 C04) with 1-day shift. A 15-days (and 30-days) prediction was made for a given NN and TS, then MJD was shifted 1-day and a new 15-days (and 30-days) prediction was made and so till the end of MJD.

The results obtained from both IERS EOP series allowed:

a) for a comparison of predictive capabilities of simple and ordinary kriging
b) for a comparison with prediction methods involved in 1st EOP PCC
c) for a comparison of two series consistent with different ITRFs, i.e., ITRF14 and ITRF05
d) for a comparison of strategy of selecting semivariograms, i.e., single and adaptive approach.

4. Results

Quality of kriging-based prediction was assessed by means of the mean absolute prediction error (MAPE) which is given by:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}|e_i| \tag{17}$$

where $e_i = O_i - P_i$ is the prediction error, O stands for observed and P for predicted LOD.

MAPE values presented in all subsequent tables are averaged values from all performed forecasts for a every single day (1–15 and 1–30) for a given dataset.

The idea staying behind the concept of using two approaches for semivariogram selection takes the origin from various shapes of empirical semivariograms for different time intervals. Fig. 1 (a, b, c) shows examples of behaviour of empirical semivariograms depending on the starting MJD (51565/22 January 2000, 51691/27 May 2000, 53547/26 June 2005) of the sliding window.

Each of these empirical semivariograms has a slightly different shape and it seems that one theoretical model cannot describe all of the empirical semivariograms well. This is why the authors decided to try out the adaptive approach for finding the best-fit semivariogram model. After very many trials three theoretical models that gave the best semivariogram's shape description (the lowest RSS for semivariogram fitting and MAPEs in the prediction procedure) were selected for adaptive strategy, i.e., cubic, exponential and sine hole effect. Table 1 shows MAPEs of ordinary kriging for each day of 15-day forecast in two semivariogram selection approaches for the "optimal" pair of steering parameters NN = 25, TS = 12.2 yrs. for IERS EOP 14 C04 series (this choice will be explained further in the text).

The first three columns present the values of MAPEs for a single semivariogram involved in the prediction process. The last column shows the results of all three theoretical semivariograms employed in the adaptive approach. The lowest values of MAPEs are surprisingly obtained, with a visible advantage, for a single-semivariogram approach with the exponential model. Differences between the exponential model and the adaptive approach vary from 0.003 to 0.009 ms, what favours the first mentioned. Almost the same level of accuracy as for the adaptive approach is obtained for the cubic model. The sine hole effect model gives the worst results. Although the adaptive approach seems to be more correct and should, at least theoretically, offer better results the practise of prediction showed the opposite. In fact, all the approaches (semivariograms) gave very satisfactory results in LOD prediction and may be used to control one another with the exponential model having the largest weight.

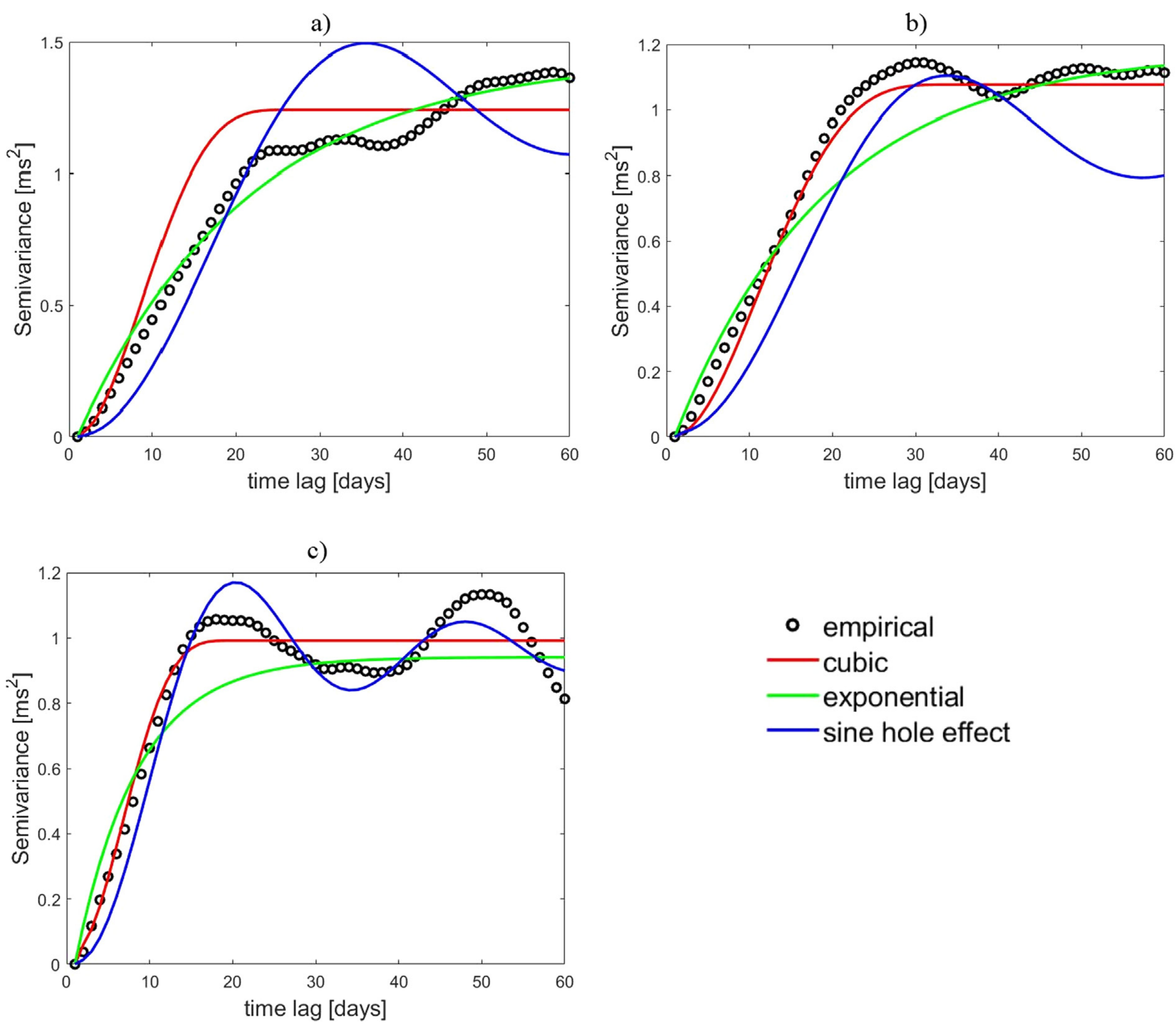


Fig. 1. Examples of empirical and fitted semivariograms for 15-day prediction, NN = 25, TS = 12.2 yrs., IERS EOP 14 C04; starting MJD a) 51565/22 January 2000, b) 51691/27 May 2000, c) 53547/26 June 2005.

Table 1
Results for single and adaptive semivariogram strategies (15-days prediction, NN = 25, TS = 12.2 yrs., IERS EOP 14 C04).

Day	MAPEs for different semivariogram selection strategies [ms]			
	Cubic	Exponential	Sine hole effect	Adaptive approach (all three)
1	0.045	0.035	0.074	0.044
2	0.064	0.059	0.088	0.064
3	0.080	0.077	0.101	0.080
4	0.094	0.091	0.113	0.094
5	0.107	0.102	0.124	0.106
6	0.117	0.111	0.133	0.116
7	0.126	0.119	0.142	0.124
8	0.132	0.125	0.150	0.132
9	0.138	0.131	0.157	0.137
10	0.142	0.135	0.162	0.142
11	0.146	0.139	0.167	0.146
12	0.149	0.142	0.170	0.149
13	0.152	0.145	0.173	0.152
14	0.155	0.147	0.175	0.154
15	0.157	0.149	0.176	0.156

The best set of control parameters, i.e., NN and TS for the 15-days LOD prediction for IERS EOP 14 C04 series was selected on the basis of minimum values of MAPEs computed from 1562 combinations of control parameters (for two variants of kriging). The intermediate results of this analysis are available on the website mentioned in the acknowledgements section. The lowest values of MAPEs in ordinary kriging were obtained for the already mentioned combination of NN = 25 and TS = 12.2 yrs. and for simple kriging NN = 25 and TS = 7 yrs. Table 2 shows MAPE values for each day of 15–day forecast in both kriging variants for two strategies of selecting semivariograms, i.e., adaptive and single-semivariogram with exponential model (separated by a slash).

As can be seen, MAPEs values for ordinary kriging based prediction are systematically lower than for simple kriging. This holds for both combinations of control parameters and both semivariograms' selection approaches.

In the first combination of parameters (NN = 25 and TS = 12.2 yrs.), the MAPEs are from 0.003 to 0.061 ms (adaptive) and from 0.003 to 0.057 ms (single-semivariogram) higher for simple kriging. In the second combination of parameters (NN = 25 and TS = 7 yrs.) results of simple kriging are higher from 0.001 to 0.014 ms (adaptive) and from 0.003 to 0.017 ms (single-semivariogram). Ordinary kriging outperforms simple kriging. This is due to no assumption of constant and known mean (zero-mean value). This would hold for the residuals from the entire time span (TS), but not for a very restricted neighbourhood (NN). Estimated mean value (constant) within the immediate neighbourhood, which is in the case of ordinary kriging, contributes to better forecasts. Hence, only ordinary kriging will be considered in further comparisons. Selected steering parameters for IERS EOP 14 C04 series (NN = 25 and TS = 12.2 yrs.) cannot be used for computations for IERS EOP 05 C04 series. This is due to the time span TS = 12.2 yrs. (app. 4453 days) is longer than

Table 2
MAPEs for ordinary and simple kriging for best steering parameters (IERS EOP 14 C04) in adaptive and single-semivariogram approach.

Day	MAPE adaptive / single-semivariogram [ms]			
	NN = 25 and TS = 12.2 yrs		NN = 25 and TS = 7 yrs	
	Ordinary kriging	Simple kriging	Ordinary kriging	Simple kriging
1	0.044 / 0.035	0.047 / 0.041	0.046 / 0.035	0.047 / 0.038
2	0.064 / 0.059	0.068 / 0.070	0.066 / 0.059	0.067 / 0.063
3	0.080 / 0.077	0.087 / 0.092	0.082 / 0.077	0.083 / 0.083
4	0.094 / 0.091	0.104 / 0.111	0.096 / 0.091	0.097 / 0.099
5	0.106 / 0.102	0.119 / 0.127	0.108 / 0.103	0.110 / 0.112
6	0.116 / 0.111	0.133 / 0.140	0.119 / 0.113	0.121 / 0.123
7	0.124 / 0.119	0.147 / 0.152	0.128 / 0.121	0.131 / 0.132
8	0.132 / 0.125	0.159 / 0.162	0.136 / 0.128	0.140 / 0.140
9	0.137 / 0.131	0.170 / 0.171	0.143 / 0.135	0.148 / 0.147
10	0.142 / 0.135	0.180 / 0.179	0.149 / 0.140	0.155 / 0.153
11	0.146 / 0.139	0.189 / 0.186	0.154 / 0.144	0.161 / 0.158
12	0.149 / 0.142	0.196 / 0.192	0.158 / 0.148	0.167 / 0.163
13	0.152 / 0.145	0.204 / 0.197	0.162 / 0.152	0.172 / 0.167
14	0.154 / 0.147	0.211 / 0.202	0.165 / 0.154	0.176 / 0.170
15	0.156 / 0.149	0.217 / 0.206	0.167 / 0.157	0.181 / 0.174

the length of the analysed data (4385 days). Therefore, the best sets of control parameters for both IERS EOP 05 C04 and truncated IERS EOP 14 C04 series were selected separately. The lowest values of MAPEs for ordinary kriging in this case were obtained for NN = 19 and TS = 9.3 yrs (IERS EOP 05 C04) and for NN = 25 and TS = 9.2 yrs (IERS EOP 14 C04).

Table 3 shows a comparison of 15-days LOD prediction results for two mentioned series. In fact, the best pair of parameters for truncated IERS EOP 14 C04 series turned out to be the second best when considered the original IERS EOP 14 C04 series. This summary is shown for two strategies of using the semivariogram, i.e., adaptive and single-semivariogram with exponential model. This LOD forecast resulted in 1012 (for TS = 9.2 yrs.) and 975 (for TS = 9.3 yrs.) 15-day predictions. As can be seen, in both cases results for IERS EOP 05 C04 are visibly better than for IERS EOP 14 C04. The same behaviour was observed for kriging based prediction of polar motion Michalczak and Ligas (2021). In the pair of parameters NN = 25 and TS = 9.2 yrs., the differences vary from 0.005 to 0.012 ms (adaptive) and from 0.004 to 0.012 ms (single-semivariogram). For the second set of parameters (NN = 19 and TS = 9.3 yrs.), the differences are on the level of 0.004–0.016 ms (adaptive) and 0.003–0.014 ms (single-semivariogram).

Table 4 presents MAPEs for each IERS EOP series for 30-days prediction for both adaptive and single-semivariogram (exponential model) strategy. The same portion of data (51544–55929 MJD, 1 January 2000 – 3 January 2012) was used for both series. In this prediction experiment the new "optimal" pair of steering parameters, for both IERS EOP series, was worked out. It turns out that both IERS EOP series have the lowest MAPEs for the same set of parameters, i.e., NN = 21 and TS = 9.2 yrs. Similarly as in 15-days forecast, IERS EOP 05 C04 is visibly better predictable than IERS EOP 14 C04. The differences vary from 0.004 to 0.010 ms (adaptive) and from 0.004 to 0.011 ms (single-semivariogram) in favour of IERS EOP 05 C04.

Fig. 2 is an extended version of the results presented in Tables 3 and 4. The figure presents absolute prediction errors, which are absolute differences between real LOD values (final IERS product) and those predicted for ultra-short-term (15 days into the future, lower panel) and short-term prediction (30 days into the future, upper panel) for every single day of the forecast. The results obtained for ordinary kriging with single-semivariogram approach, with TS = 9.2 and NN = 25. The pattern of prediction errors is noteworthy. The largest values, which are app. 0.7–0.8 ms, occur around June 2009, May 2010, September 2010 and April 2011. Inspection of dates of El Niño and La Niña occurrences shows the correspondence between the large values of the prediction errors and the appearance of these phenomena (https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php). The distribution of prediction errors presented in Fig. 2 may contribute to identifying different geophysical and meteorological phenomena that are related to the LOD prediction. Therefore, other meteorological and geophysical methods/data may be used in future studies to improve the LOD prediction.

Fig. 3 presents a comparison of MAPEs for ultra-short (10-day) ordinary kriging based LOD prediction (adaptive and single-semivariogram approach, with NN = 19 and TS = 9.3 yrs.) with the methods involved in the 1st EOP PCC. Accurate values of MAPEs for ordinary kriging are those listed in Table 3. It must be noted that different data spans were used to compute the prediction errors in the 1st EOP PCC and in this contribution. This study's predictions were made with one-day shift and no predictions were excluded hence, the results present "good" and "bad" time

Table 3
Comparison of MAPEs for ordinary kriging for two sets of steering parameters for IERS EOP 14 C04 and IERS EOP 05 C04 for 15-days prediction for adaptive and single-semivariogram approaches.

Day	**MAPE adaptive / single-semivariogram [ms]**			
	NN = 25 and TS = 9.2 yrs		**NN = 19 and TS = 9.3 yrs**	
	IERS EOP 14 C04	**IERS EOP 05 C04**	**IERS EOP 14 C04**	**IERS EOP 05 C04**
1	0.041 / 0.036	0.036 / 0.032	0.039 / 0.035	0.035 / 0.030
2	0.060 / 0.057	0.054 / 0.052	0.059 / 0.057	0.053 / 0.051
3	0.075 / 0.073	0.068 / 0.067	0.075 / 0.073	0.067 / 0.066
4	0.087 / 0.085	0.079 / 0.078	0.086 / 0.086	0.078 / 0.077
5	0.098 / 0.096	0.089 / 0.087	0.097 / 0.096	0.087 / 0.087
6	0.108 / 0.106	0.098 / 0.096	0.106 / 0.106	0.096 / 0.096
7	0.116 / 0.113	0.106 / 0.104	0.115 / 0.114	0.104 / 0.103
8	0.124 / 0.120	0.114 / 0.110	0.123 / 0.121	0.112 / 0.110
9	0.131 / 0.127	0.120 / 0.116	0.130 / 0.127	0.118 / 0.116
10	0.136 / 0.132	0.125 / 0.121	0.137 / 0.133	0.124 / 0.121
11	0.141 / 0.137	0.129 / 0.125	0.143 / 0.139	0.129 / 0.125
12	0.145 / 0.141	0.133 / 0.129	0.148 / 0.143	0.133 / 0.129
13	0.148 / 0.144	0.137 / 0.132	0.152 / 0.147	0.137 / 0.133
14	0.151 / 0.146	0.140 / 0.134	0.155 / 0.150	0.140 / 0.136
15	0.154 / 0.149	0.143 / 0.137	0.158 / 0.153	0.143 / 0.139

Table 4
Comparison of MAPEs for ordinary kriging for IERS EOP 14 C04 and IERS EOP 05 C04 for 30-days prediction for adaptive and single-semivariogram approach.

Day	MAPE adaptive/single-semivariogram [ms]		Day	MAPE adaptive/single-semivariogram [ms]	
	IERS EOP 14 C04	IERS EOP 05 C04		IERS EOP 14 C04	IERS EOP 05 C04
1	0.044 / 0.038	0.040 / 0.034	16	0.156 / 0.151	0.148 / 0.141
2	0.063 / 0.059	0.058 / 0.054	17	0.157 / 0.153	0.149 / 0.143
3	0.078 / 0.074	0.071 / 0.068	18	0.159 / 0.155	0.152 / 0.146
4	0.090 / 0.087	0.082 / 0.080	19	0.162 / 0.158	0.155 / 0.149
5	0.100 / 0.098	0.092 / 0.089	20	0.165 / 0.161	0.158 / 0.152
6	0.110 / 0.107	0.101 / 0.098	21	0.167 / 0.163	0.161 / 0.155
7	0.119 / 0.115	0.110 / 0.106	22	0.170 / 0.165	0.164 / 0.158
8	0.127 / 0.123	0.118 / 0.113	23	0.172 / 0.168	0.167 / 0.161
9	0.134 / 0.129	0.124 / 0.119	24	0.175 / 0.170	0.170 / 0.164
10	0.140 / 0.135	0.130 / 0.124	25	0.177 / 0.172	0.173 / 0.166
11	0.145 / 0.139	0.135 / 0.128	26	0.179 / 0.174	0.175 / 0.168
12	0.149 / 0.143	0.139 / 0.132	27	0.181 / 0.176	0.177 / 0.170
13	0.152 / 0.146	0.142 / 0.135	28	0.182 / 0.178	0.178 / 0.172
14	0.154 / 0.147	0.144 / 0.137	29	0.183 / 0.179	0.179 / 0.174
15	0.155 / 0.149	0.146 / 0.139	30	0.183 / 0.180	0.179 / 0.175

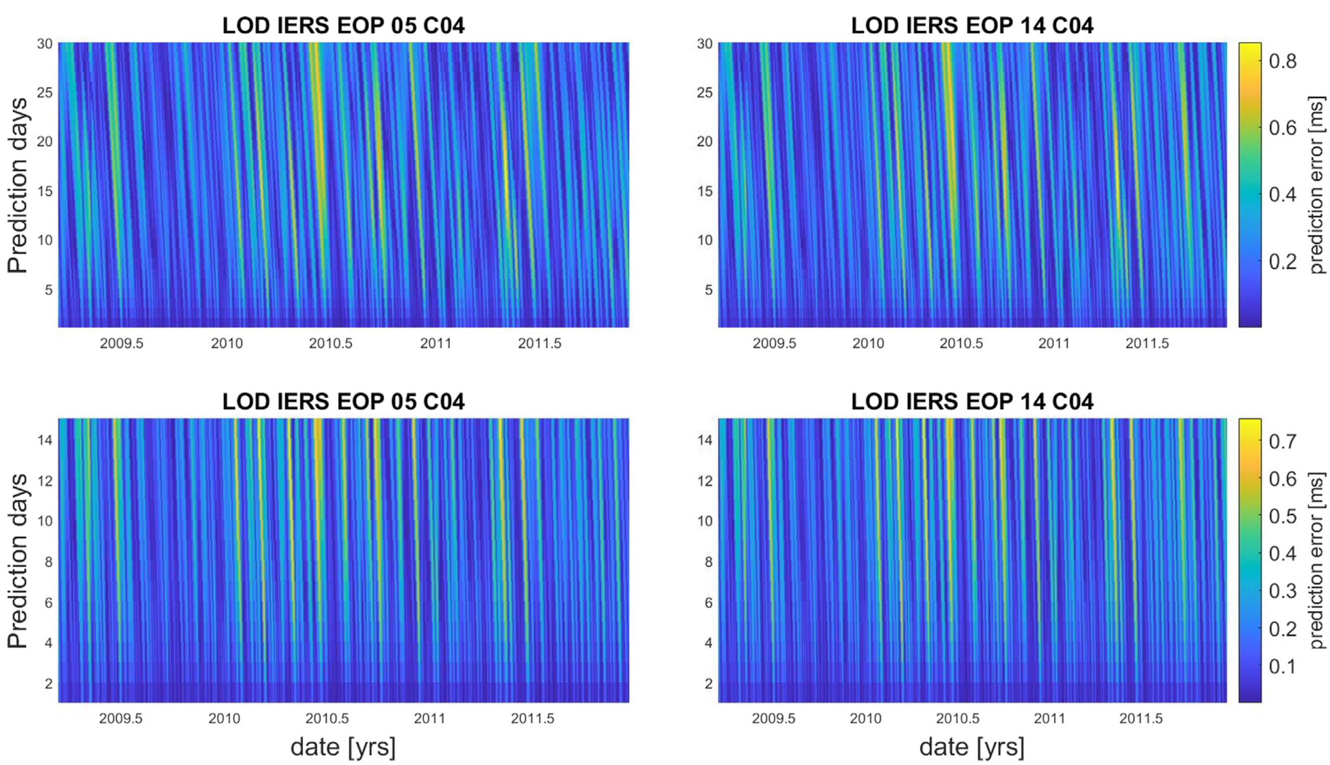


Fig. 2. Daily prediction errors for short-term (30 days) and ultra-short-term (15 days) forecast, for IERS EOP 05 C04 and IERS EOP 14 C04.

moments and may be considered as representative for the adopted method. The results indicate that ordinary kriging may compete with the most accurate prediction methods that took part in the 1st EOP PCC. The numbers show that only the Kalman filter method outperforms ordinary kriging with the largest difference between the methods less than 0.019 ms.

Fig. 4 shows MAPEs for ordinary kriging based 30-day LOD prediction (adaptive and single-semivariogram approach, with NN = 21 and TS = 9.2 yrs.) and methods involved in the 1st EOP PCC. The results show that the ordinary kriging method implemented in this contribution, in general, has the potential to outperform majority of the methods. Only Kalman filter, as it was for ultra-short term prediction, offers better accuracy up to day 11, after this day the advantage is on the side of ordinary kriging. MAPEs obtained from kriging method characterise by a small stable (linear) increase in value after about the day 10.

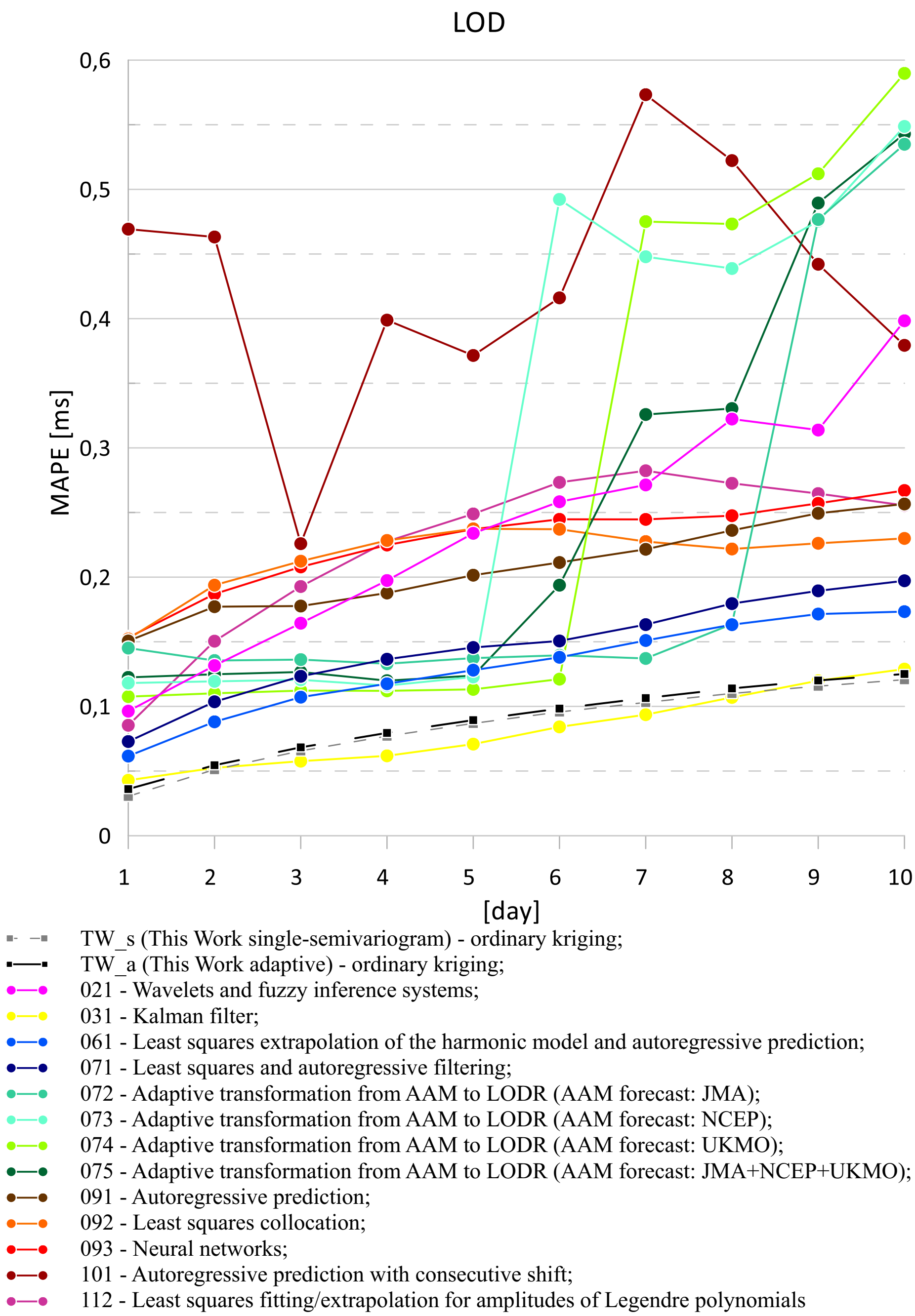


Fig. 3. The MAPEs comparison of ordinary kriging based 10-day LOD prediction with the methods involved in 1st EOP PCC.

5. Conclusion

A new kriging-based prediction method for the length-of-day (LOD), one of the Earth's rotation parameters, has been presented in this contribution. Two basic geostatistical prediction methods have been used, i.e., simple kriging and ordinary kriging. Ordinary kriging turned out to be superior to simple kriging, this is due to its immanent property of estimating a local mean value instead of a priori adopting a constant (zero-mean value here). The prediction methods were equipped with two scenarios of employing a structure function to the prediction procedure, i.e., a single-semivariogram and adaptive approach. The single-semivariogram approach uses only one semivariogram model through the entire prediction process. The adaptive approach, on the other hand, selects the best-fit semivariogram from a group of models what enables alternating among the models depending on the data behaviour. Although the adaptive approach seemed to be theoretically more sound, the computing practice showed that the single-semivariogram approach with the exponential model applied to LOD forecast gave visibly better results.

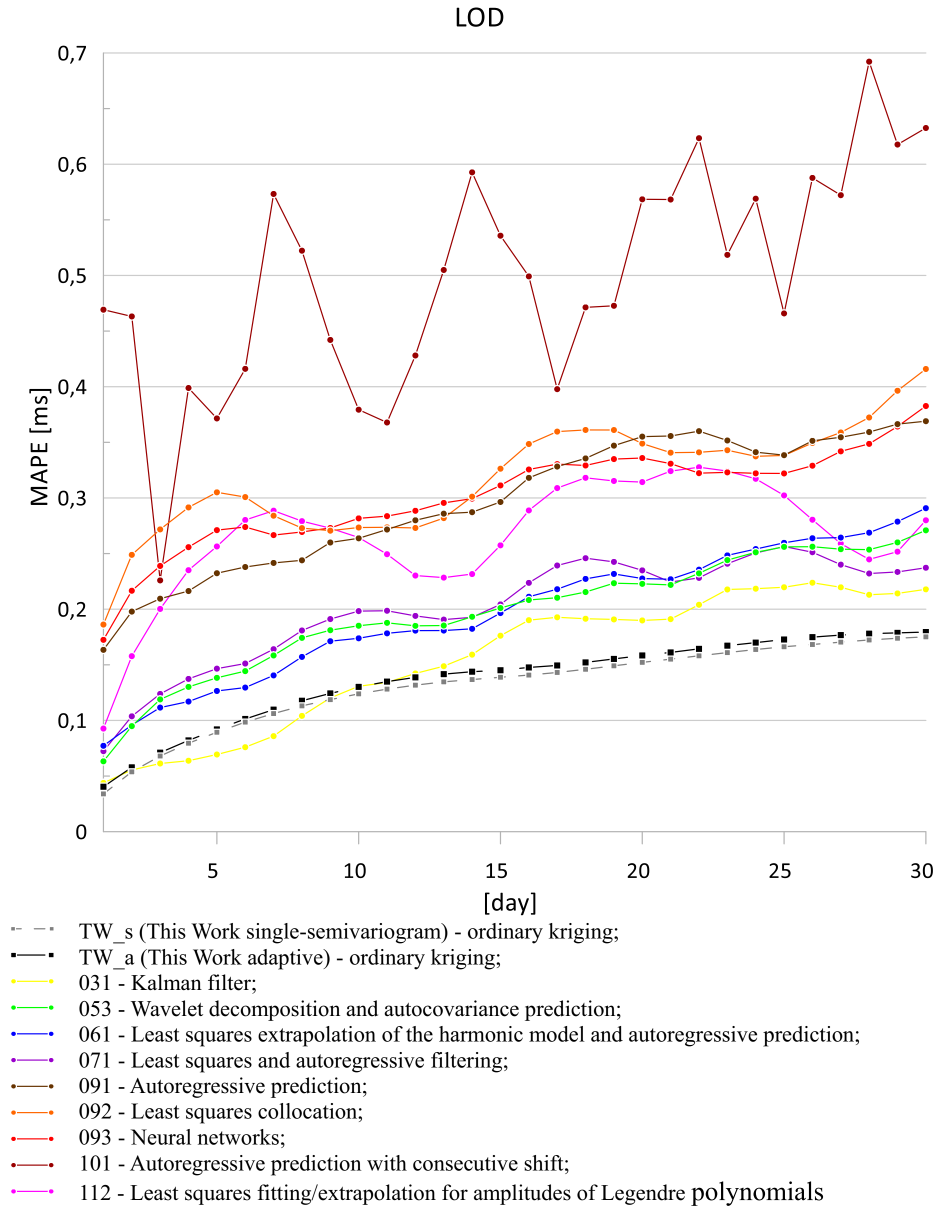


Fig. 4. The MAPEs comparison of ordinary kriging based 30-day LOD prediction with the methods involved in 1st EOP PCC.

The ultra-short (15-days) and short-term (30-days) prediction have been performed on the basis of IERS EOP 14 C04 and IERS EOP 05 C04 time series data. This contribution reports mean absolute prediction errors in ultra-short term prediction for days 1–15 with respect to IERS EOP 14 C04 on the level of 0.044 – 0.156 ms (adaptive) and 0.039 – 0.149 ms (single-semivariogram); for IERS EOP 05 C04 the values are 0.035 – 0.143 ms (adaptive) and 0.030 – 0.139 ms (single-semivariogram). For short term prediction (days 1–30) the values with respect to IERS EOP 14 C04 are on the level of 0.044 – 0.183 ms (adaptive) and 0.038 – 0.180 ms (single-semivariogram); with respect to IERS EOP 05 C04 the values are 0.040 – 0.179 ms (adaptive) and 0.034 – 0.175 ms (single-semivariogram). Comparison of prediction results for two mentioned IERS EOP series, with the same control parameters, suggest better predictability of IERS EOP 05 C04. This regularity occurs for both 15-days and 30-days prediction. The same fact has been also observed for kriging-based polar motion prediction (Michalczak and Ligas, 2021).

The pattern of LOD prediction errors given in figures of the type of Fig. 2 may be helpful in relating and identifying different meteorological and geophysical phenomena that influence LOD prediction, and may be applied in future studies.

Implementation of ordinary kriging presented in this contribution seems to be more accurate than most prediction methods that took part in the 1st EOP PCC but there is some degree of uncertainty due to different testing conditions. The most important is that forecasts presented in this contribution were made on final IERS data while the EOPPCC acted in operational settings (involvement of rapid products). The use of kriging in real-time prediction may result in a slight deterioration of the forecast accuracy. The obtained results are directly comparable with the ones of Kalman filter applied in the 1st EOP PCC (the best therein). The methods differ by no more than 0.019 ms for 10-days prediction. On the other hand, for 30-days prediction the Kalman filter method outperforms ordinary kriging in app. the first 12 days, but later on kriging takes advantage and turns more accurate. The differences in the first 12 days vary from 0.001 to 0.025 ms, later they are in the range of 0.001 to 0.051 ms in favour of this work's method. The results of this contribution are supported by a huge number of performed predictions, therefore, all intermediate results (due to their volume) are available at the website accompanied with this article and mentioned in the acknowledgements section.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors are very grateful to M. Kalarus for providing data from 1st Earth Orientation Parameters Prediction Comparison Campaign (EOP PCC).

Intermediate results of this contribution are available at the website accompanied with this article home.agh.edu.pl/mmichalc.

References

Akyilmaz, O., Kutterer, H., Shum, C., Ayan, T., 2011. Fuzzy-wavelet based prediction of earth rotation parameters. Appl. Soft Comput. 11 (1), 837–841. https://doi.org/10.1016/j.asoc.2010.01.003.

Cressie, N.A.C., 1993. Statistics for spatial data. John Wiley & Sons, New York. https://doi.org/10.1002/9781119115151.ch1.

Dick W. R. and Thaller D. (eds.) IERS Annual Report 2018. "International Earth Rotation and Reference Systems Service, Central Bureau." Frankfurt am Main: Verlagdes Bundesamts für Kartographie und Geodäsie, 2020. 207 p., ISBN 978-3-86482-136-3.

Dow, J.M., Neilan, R.E., Rizos, C., 2009. The International GNSS Service in a changing landscape of Global Navigation Satellite Systems. J. Geod. 83, 191–198. https://doi.org/10.1007/s00190-008-0300-3.

Gambis, D., Luzum, B., 2011. Earth rotation monitoring, UT1 determination and prediction. Metrologia 48, 165–170. https://doi.org/10.1088/0026-1394/48/4/S06.

Gross, R.S., Marcus, S.L., Eubanks, T.M., Dickey, J.O., Keppenne, C.L., 1996. Detection of an ENSO Signal in Seasonal Length-of-day Variations. Geophys. Res. Lett. 23 (23), 3373–3376.

Hozakowski, W., 1989. Polar Motion Prediction by the Least -Squares Collocation Method. In: Earth Rotation and Coordinate Reference Frames, Scotland, August 10–11, 1989.

Kalarus, M., Schuh, H., Kosek, W., Akyilmaz, O., Bizouard, Ch., Gambis, D., Gross, B.R., Jovanovi´c, Kumakshev, S., Kutterer, H., Mendes Cerveira, P.J., Pasynok, S., Zotov, L., 2010. Achievements of the Earth Orientation Parameters prediction comparison campaign. J. Geodesy 84, 587–596, https://doi.org/10.1007/s00190-010-0387-1.

Kosek, W., 2002. Autocovariance prediction of complex-valued polar motion time series. Advances in Space Research, Volume 30, Issue 2, Pages 375-380, ISSN 0273-1177, https://doi.org/10.1016/S0273-1177(02)00310-1.

Kosek, W., Kalarus, M., Niedzielski, T., 2007. Forecasting Of the Earth Orientation Parameters - Comparison of Different Algorithms. Journées Systèmes de Référence Spatio-temporels 2007, Observatoire de Paris, 17-19 September.

Lei, Y., Guo, M., Cai, H., Hu, D., Zhao, D., 2015. Prediction of length-of-day using Gaussian process regression. J. Navig. 68 (563–575). https://doi.org/10.1017/S0373463314000927.

Liao, D.C., Wang, Q.J., Zhou, Y.H., Liao, X.H., Huang, C.L., 2012. Long-term prediction of the Earth Orientation Parameters by the artificial neural network technique. J. Geodynamics, 62, Pages 87-92, ISSN 0264-3707, https://doi.org/10.1016/j.jog.2011.12.004.

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

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. J. Geod. 94, 23. https://doi.org/10.1007/s00190-020-01354-y.

Nastula, J., Chin, T.M., Gross, R., Śliwińska, J., Wińska, M., 2020. Smoothing and predicting celestial pole offsets using a Kalman filter and smoother. J. Geod. 94, 29. https://doi.org/10.1007/s00190-020-01349-9.

Niedzielski, T., Kosek, W., 2008. Prediction of UT1–UTC, LOD and AAM $\chi 3$ by combination of least-squares and multivariate stochastic methods. J Geod 82, 83–92. https://doi.org/10.1007/s00190-007-0158-9.

Okhotnikov, G., Golyandina, N., 2019. EOP Time Series Prediction Using Singular Spectrum Analysis, RWTH Aahen University.

Olea, R.A., 1999. Geostatistics for Engineers and Earth Scientists. Kluwer Academic Publishers. https://doi.org/10.1080/00401706.2000.10485748.

Petit, G., Luzum, B., (eds.). IERS Conventions (2010), IERS Technical Note 36, Frankfurt am Main: Verlag des Bundesamts für Kartographie und Geodäsie. 179 pp., ISBN 3-89888-989-6.

Schuh, H., Ulrich, M., Egger, D., Müller, J., Schwegmann, W., 2002. Prediction of Earth Orientation Parameters by Artificial Neural Networks. J. Geod. 76 (5), 247–258. https://doi.org/10.1007/s00190-001-0242-5.

Wu, F., Chang, G., Deng, K., One-step method for predicting LOD parameters based on LS+AR model. J. Spatial Sci., 66:2, 317-328, https://doi.org/10.1080/14498596.2019.1618401.

Zotov, L.V., Xu, X.Q., Skorobogatov, A., Zhou, Y.H., 2017. Combined SAI-SHAO prediction of Earth orientation parameters since 2012 till 2017. Geodesy and Geodynamics, 9(6), pp. 485-490, KeAi, China, 2018, https://doi.org/10.1016/j.geog.2018.11.002.

ARTIFICIAL SATELLITES, Vol. 57, No. SI1 – 2022
DOI: 10.2478/arsa-2022-0024

PREDICTION OF EARTH ROTATION PARAMETERS WITH THE USE OF RAPID PRODUCTS FROM IGS, CODE AND GFZ DATA CENTRES USING ARIMA AND KRIGING – A COMPARISON

Maciej MICHALCZAK, Marcin LIGAS, Jacek KUDRYS

Department of Integrated Geodesy and Cartography,
Faculty of Geo-Data Science, Geodesy, and Environmental Engineering,
AGH University of Science and Technology, Krakow, Poland

e-mails: mmichalc@agh.edu.pl, ligas@agh.edu.pl, jkudrys@agh.edu.pl

ABSTRACT. Real-time prediction of Earth Orientation Parameters is necessary for many advanced geodetic and astronomical tasks including positioning and navigation on Earth and in space. Earth Rotation Parameters (ERP) are a subset of EOP, consisting of coordinates of the Earth's pole (PMx, PMy) and UT1-UTC (or Length of Day – LOD). This paper presents the ultra-short-term (up to 15 days into the future) and short-term (up to 30 days into the future) ERP prediction using geostatistical method of ordinary kriging and autoregressive integrated moving average (ARIMA) model. This contribution uses rapid GNSS products EOP 14 12h from IGS, CODE and GFZ and also IERS final products – IERS EOP 14 C04 12h (IAU2000A). The results indicate that the accuracy of ARIMA prediction for each ERP is better for ultra-short prediction. The maximum differences between methods for first few days of 15-day predictions are around 0.32 mas (PMx), 0.23 mas (PMy) and 0.004 ms (LOD) in favour of ARIMA model. The maximum differences of Mean Absolute Prediction Errors (MAPEs) on the last few days of 30-day predictions are 1.91 mas (PMx), 0.30 mas (PMy) and 0.026 ms (LOD) with advantage to kriging method. For all ERPs the differences of MAPEs for time series from various analysis centres are not significant and vary up to maximum value of around 0.05 mas (PMx), 0.04 mas (PMy) and 0.005 ms (LOD).

Keywords: ARIMA, kriging, polar motion, length of day, Earth Rotation Parameters

1. INTRODUCTION

Earth Orientation Parameters (EOP), i.e., coordinates of the Earth's pole (PMx, PMy), celestial pole offsets (dX, dY), and UT1-UTC describe the irregularities in the Earth's rotation. An additional parameter that is the first negative time derivative from the UT1-UTC is the Length of Day (LOD) (Modiri et al., 2020). In this contribution, a subset of EOP, Earth Rotation Parameters (ERP), consisting of Earth's pole coordinates (PMx, PMy) and UT1-UTC (LOD) is subject to prediction.

The knowledge of EOP is essential for transformation between celestial reference frame (CRF) and terrestrial reference frame (TRF) (Gambis and Luzum, 2011), which enables precise positioning in space and on the Earth's surface, navigation, time-keeping and positional astronomy.

EOP are determined by space geodetic techniques such as VLBI (Very Long Baseline Interferometry), DORIS (Doppler Orbitography and Radiopositioning Integrated by Satellite), SLR (Satellite Laser Ranging) and GNSS (Global Navigation Satellite System), and are finally combined into a single solution (Petit and Luzum, 2010). This combination of techniques enables to determine polar coordinates with the accuracy of app. 50 μas or better and 10 μs in case of LOD (Dick and Thaller, 2020).

Complexity of measuring and data processing involved in EOP determination forces their prediction since it is impossible to obtain them in real-time mode.

The long EOP time series are published with latency by International Earth Rotation and Reference System Service (IERS) and many individual analysis centres or services (IERS Annual Report 2018). Therefore, in operational settings, this gap must be filled with rapid products, e.g., IGS (International GNSS Service), CODE (Center for Orbit Determination in Europe) or GFZ (German Research Centre for Geosciences). This may also be done by basing on rapid products published by IERS, e.g., Daily Rapid EOP Data.

A plenty of prediction methods have been applied to EOP prediction through the years; the following list is by no means exhaustive: least-squares collocation (Hozakowski, 1989), artificial neural networks (Schuh et al., 2002; Liao et al., 2012), autocovariance prediction (Kosek, 2002), autoregression (Kosek et al., 2007), combination of least-squares and multivariate stochastic methods (Niedzielski and Kosek, 2008), wavelets and fuzzy inference systems (Akyilmaz et al., 2011), extreme learning machine (Lei et al., 2015), Gaussian process regression (Lei et al., 2015), combination of least-squares extrapolation and auto-regressive modelling of EAM (Dill et al., 2019), singular spectrum analysis (Okhotnikov and Golyandina, 2019), one-step forecasting method based on the LS+AR (Wu et al., 2019), combination of singular spectrum analysis and Copula-based analysis (Modiri et al., 2020) and Kalman filter (Nastula et al., 2020), combination of Holt–Winters algorithm and angular momenta of global surficial geophysical fluids (Luo et al., 2022).

This study is intended in ultra-short-term (up to 15 days into the future) and short-term (up to 30 days into the future) ERP predictions using geostatistical method of ordinary kriging and autoregressive integrated moving average (ARIMA) model. Rapid GNSS products EOP 14 12h from IGS, CODE and GFZ and also IERS final products (IERS EOP 14 C04 12h(IAU2000A)) are used and the results are compared. For all the time series and data centres the first ERP prediction was made for 1 January 2017 (MJD 57754.5) and the last for 15 April 2022 (MJD 59684.5) with 1 day shift.

2. METHODOLOGY

2.1. Statistical agreement between final IERS and rapid CODE, IGS and GFZ products

This section attempts to statistically examine the agreement between ERP final products delivered by IERS and rapid ones delivered by CODE, IGS and GFZ. To accept or reject a hypothesis of equality, for a particular ERP in pairwise combinations of data centres, three techniques are used, i.e., matched-pair t-test, Deming regression (DR), and Passing-Bablok regression (PBR). The mentioned techniques are frequently applied in method comparison studies (e.g., in medicine or clinical chemistry) to search for biases/differences (constant/fixed or proportional) between two measurement methods (or measuring devices). In this contribution, a single pair of ERP coming from IERS (reference) and one of three (IGS, CODE, GFZ) centres on a given MJD is considered as a measurement pair. Matched-pair t-test evaluates whether the differences between variables representing two measurement methods are statistically different from a given constant (to test for equality the constant is zero); hence it is able to detect only a constant bias. The other two methods quantify and assess deviations from the line of equality (perfect agreement between measurements X and Y given by different methods or devices), i.e.,

$$Y = X \tag{1}$$

by fitting a straight line,

$$Y = aX + b \tag{2}$$

Hence, they are able to detect both fixed (intercept = b) and proportional (slope = a) biases. Deming regression is a method of fitting a straight line to scattered 2D data in an errors-in-variables framework, i.e., both the explanatory variable X and the response variable Y in (2) are considered to be erroneous. In DR, it is assumed that the variance ratio of the measurement errors λ in X and Y is constant; for $\lambda = 1$, DR yields the same results as orthogonal regression (in this study, this assumption was adopted). To assess statistical significance of slope and intercept, a pair of hypotheses is tested, i.e., H_0: $a = 1$ vs. H_1: $a \neq 1$ and H_0: mean Y - mean X = 0 vs. H_1: mean Y - mean X $\neq$ 0 to decide whether methods may be considered as equivalent (Linnet 1990).

Passing - Bablok regression is another technique of testing agreement and detecting a potential systematic bias between methods. It is a non-parametric and robust regression method where a slope a in (2) is the median of all slopes computed between any two points excluding cases of $a = 0$ or $a = \inf$ and an intercept b in (2) is calculated as $b = \text{median}(y_i - a x_i)$. For both parameters, the 1-α (usually 0.95) confidence intervals (CI) are inferred. If 0 is within the CI of b, and 1 is included in the CI of a, then according to Passing and Bablok (1983), one may infer that Y = X and the two methods are comparable. On the other hand, if 0 is not in the CI of b, there is a systematic difference (bias), and if 1 is not in the CI of a, then there is a proportional difference (bias) between the two methods.

Table 1. Matched-pair t – test between final IERS and rapid CODE, IGS and GFZ ERP products

	ERP	Mean difference [“ / s]	CI lower limit [“ / s]	CI upper limit [“ / s]	Decision
IERS-CODE	PMx	-0,00002657	-0,00002828	-0,00002486	Reject
	PMy	-0,00000364	-0,00000493	-0,00000234	Reject
	LOD	-0.00000790	-0.00000829	-0.00000751	Reject
IERS-IGS	PMx	-0,00001092	-0,00001240	-0,00000945	Reject
	PMy	-0,00000896	-0,00000996	-0,00000796	Reject
	LOD	0.00000006	-0.00000021	0.00000034	**Accept**
IERS-GFZ	PMx	-0,00000536	-0,00000719	-0,00000353	Reject
	PMy	-0,00002593	-0,00002780	-0,00002406	Reject
	LOD	-0.00001515	-0.00001561	-0.00001468	Reject

*Confidence level 1- α = 0.95 for CI

Table 2. Deming regression, final IERS vs rapid CODE, IGS and GFZ ERP products

	ERP	Slope	Intercept [" / s]	Decision
IERS-CODE	PMx	1.00003832 t = 3.04479892; p = 0.00235965 (R)	0.00002657 t = 30.52319805; p = 0.00000000 (R)	Reject
	PMy	0.99999094 t = -0.89663540; p = 0.37002469 (A)	0.00000364 t = 5.50072407; p = 0.00000004 (R)	Reject (partially satisfied)
	LOD	1.00089118 t = 3.15274029; p = 0.00164218 (R)	0.00000790 t = 39.80528652; p = 0.00000000 (R)	Reject
IERS-IGS	PMx	0.99998689 t = -1.17060095; p = 0.24190291 (A)	0.00001092 t = 14.52462344; p = 0.00000000 (R)	Reject (partially satisfied)
	PMy	0.99996457 t = -4.60782629; p = 0.00000433 (R)	0.00000896 t = 17.59538975; p = 0.00000000 (R)	Reject
	LOD	1.00020499 t = 1.16628098; p = 0.24364414 (A)	-0.00000006 t = -0.46240588; p = 0.64384211 (A)	**Accept**
IERS-GFZ	PMx	1.00008012 t = 5.82033049; p = 0.00000001 (R)	0.00000536 t = 5.74191721; p = 0.00000001 (R)	Reject
	PMy	1.00006274 t = 4.07633940; p =0.00004760 (R)	0.00002593 t = 27.24852927; p = 0.00000000 (R)	Reject
	LOD	0.99916958 t = -2.62639779; p = 0.00869725 (R)	0.00001515 t = 63.59095306; p = 0.00000000 (R)	Reject

*Significance level $\alpha = 0.025$ for hypothesis testing; (R), (A) denote Reject or Accept (no ground to reject) in separate hypothesis testing for significance of either a slope or an intercept; decision in the last column concerns the overall hypothesis of equivalence of methods; t is the value of t-statistics which take the form: (slope) t = (a – 1)/SE(a) and (intercept) t = (mean Y – mean X)/SE(mean Y – mean X) where the standard errors (SE) were calculated using the jackknife method, p is a p-value, when less than adopted $\alpha = 0.025$ then it indicates strong evidence against the null hypothesis.

Table 3. Passing-Bablok regression, final IERS vs rapid CODE, IGS and GFZ ERP products

	ERP	Slope (a / CI)	Intercept (b / CI) [“ / s]	Decision
IERS-CODE	PMx	1.0000432 (1.00002075; 1.00006572) (R)	0.0000170 (0.00001399; 0.00001988) (R)	Reject
	PMy	1.00000000 (0.99997632; 1.00001322) (A)	0.00000200 (-0.00000264; 0.00001059) (A)	**Accept**
	LOD	1.00100267 (1.00045893; 1.00155259) (R)	0.00000759 (0.00000738; 0.00000778) (R)	Reject
IERS-IGS	PMx	0.99999126 (0.99997446; 1.00000467) (A)	0.00000545 (0.00000385; 0.00000720) (R)	Reject (partially satisfied)
	PMy	0.99996812 (0.99995422; 0.99998202) (R)	0.00001770 (0.00001268; 0.00002273) (R)	Reject
	LOD	1.00029612 (1.00000000; 1.00063857) (A)	-0.00000018 (-0.00000026; -0.00000010) (R)	Reject (partially satisfied)
IERS-GFZ	PMx	1.00006766 (1.00004284; 1.00009256) (R)	-0.00000174 (-0.00000476; 0.00000115) (A)	Reject (partially satisfied)
	PMy	1.000052582 (1.00002540; 1.00008007) (R)	0.00000175 (-0.00000836; 0.00001128) (A)	Reject (partially satisfied)
	LOD	0.99948267 (0.99882881; 1.00013364) (A)	0.00001554 (0.00001524; 0.00001582) (R)	Reject (partially satisfied)

*Confidence level 1- α = 0.95 for CI; (R), (A) denote Reject or Accept (no ground to reject) in separate hypothesis testing for significance of either a slope or an intercept; decision in the last column concerns the overall hypothesis of equivalence of methods

Although the Passing-Bablok regression gives uncertain results (but very close to overall acceptance), the remaining two methods indicate statistically significant agreement (no biases) for the pair IERS-IGS in case of LOD time series, thus they may be considered as equivalent. This will be also visible when results of forecasts will be exposed. On the other hand, Passing-Bablok regression accepts hypotheses of complete agreement (no proportional and fixed biases) for the pair IERS-CODE in case of PMy; Deming regression accepts only H0 for the slope (a = 1) and rejects H0 for the intercept, hence fixed bias may be present. The value of the intercept agrees with the bias (mean difference) estimated for the matched-pair t-test. For the remaining pairs of data centres and ERPs, the methods detect either both proportional and fixed biases (rejection of two hypotheses) or either of the two. Despite the detection of differences between products belonging to IERS and considered data centres, it is worth noting that easily interpretable fixed biases (shifts) are of the order of accuracy in determining the Earth's Rotation Parameters. Thus, so far, they may be considered insignificant when confronted with the accuracy of the forecast. In addition, the results may depend on the analysed time span 1 January 2017 (MJD 57754.5) – 15 April 2022 (MJD 59684.5).

2.2. Solid Earth tides

The LOD time series includes periodic effects such as the impact of solid Earth tides with periods varying from 5 days up to 18.6 years. These tidal effects were first removed from

LOD time series by using a model recommended by IERS Conventions (2010) (Petit and Luzum, 2010) which is given by

$$\begin{cases} dLOD = \sum_{i=1}^{62} B_i' cos\xi_i + C_i' sin\xi_i \\ \xi_i = \sum_{j=1}^{5} a_{ij}\alpha_j \end{cases} \tag{3}$$

where: values of B'_i, C'_i, a_{ij}, and α_j are given in Chapter 8 of the IERS Conventions (2010) (Petit and Luzum, 2010).

2.3. Time – frequency analysis

In order to determine the number of periodic components in the prediction model, a time-frequency analysis of the length-of-day and polar motion time series based on IERS EOP 14 C04 12h (IAU2000A) was performed. Continuous Wavelet Transform with Morlet wavelets was used for this purpose with a time span of approx. 22 years (MJD 51544.5 – 59684.5). In case of LOD, the analysis resulted in a spectrogram with clearly visible lines corresponding to three periods of 0.5, 1 and 8.3 years (Figure 1). Several weaker components of periods between 3.5 and 5.5 years (5e-4 1/day, 8e-4 1/day) with increasing amplitude over time are also evident. Signals with lower frequencies visible on the spectrogram were treated as artefacts due to the length of the analysed time interval (22 years) and were not included in the prediction model.

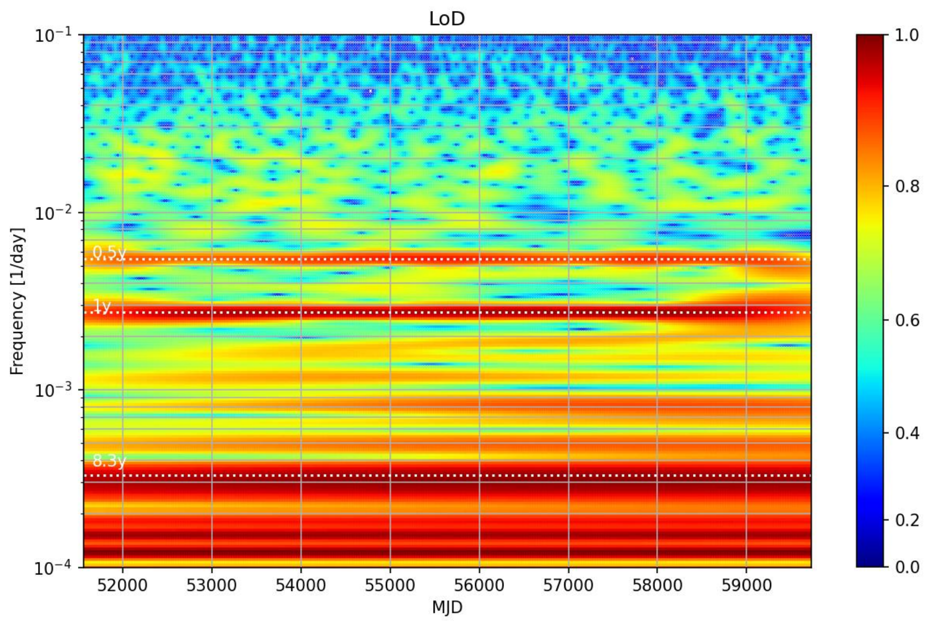


Figure 1. Spectrogram for LOD

The PMx and PMy time series were analysed in a similar manner, obtaining spectrograms with apparent domination of two components with periods of 366 and 431 (annual and Chandler oscillations). Starting at around 57000 MJD, the second component shows decreasing amplitude in both the PMx and PMy series (Figures 2 and 3). A similar effect has been observed previously around 1920-1930 and connected with the Chandler wobble phase change (Zotov and Bizouard, 2018).

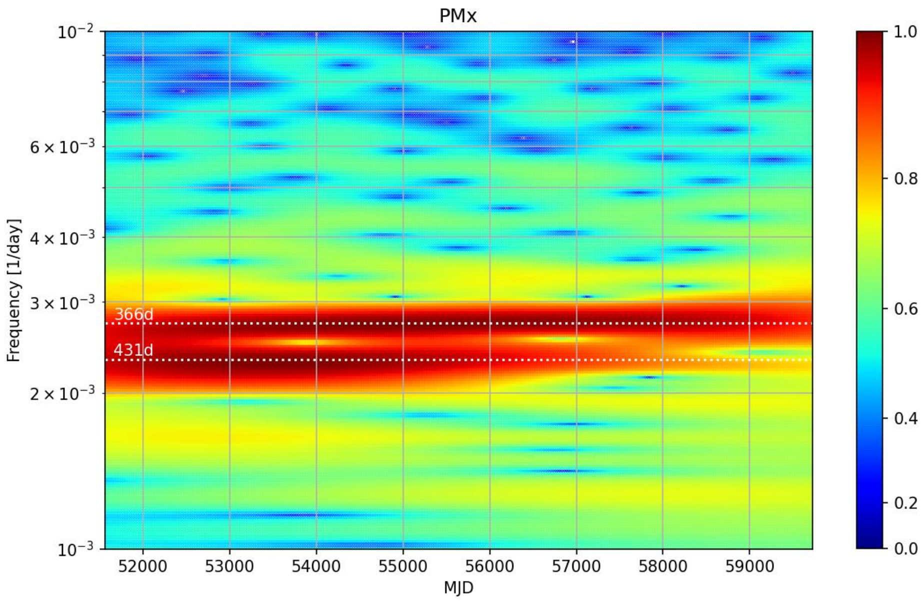


Figure 2. Spectrogram for PMx

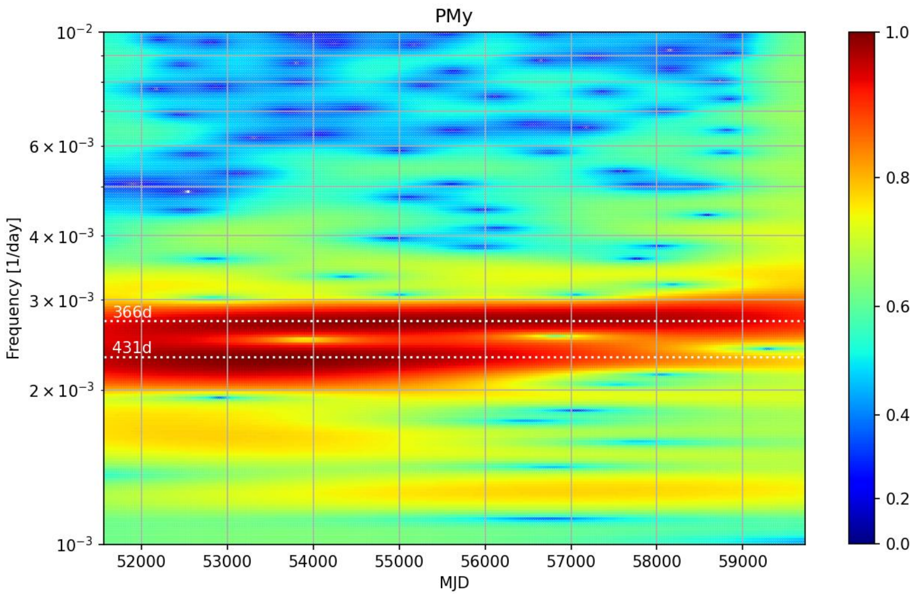


Figure 3. Spectrogram for PMy

2.4. Prediction model

Both LOD and pole coordinates are decomposed into linear and periodic components. The prediction model consists of mentioned components (extrapolated into the future) and residuals (obtained in the prediction process), and is given as follows:

$$\widehat{ERP}(t_p) = \hat{a}_0 + \hat{a}_1 t_p + \sum_{i=1}^{n} \hat{b}_i \sin(\hat{c}_i t_p + \hat{d}_i) + [dLOD] + \hat{\varepsilon}(t_p) \tag{4}$$

where $\widehat{ERP}$ is a predicted ERP; t_p is the prediction time moment; $\hat{a}_0, \hat{a}_1$ are the estimated linear trend coefficients; n is the number of terms in the series; $\hat{b}_i, \hat{c}_i, \hat{d}_i$ are the estimated amplitudes, frequencies, and phase constants for i-th sine term, respectively; *dLOD* are tidal corrections (occur only in LOD prediction) and $\hat{\varepsilon}$ stands for the predicted residual.

2.5. Ordinary kriging

Ordinary kriging, adapted to time series prediction herein, is a linear predictor of an unknown random quantity Z at a time instant t_0 (here residual PM and LOD processes), i.e.,

$$\hat{Z}(t_o) = \boldsymbol{\lambda}^T \mathbf{Z}(t) \tag{5}$$

where $\mathbf{Z}(t)$ are observed data.

It is optimal, in the sense of Best Linear Unbiased Prediction – BLUP (and Best Linear Unbiased Estimator – BLUE), if the mean value of a random process is an unknown constant (Cressie, 1993). Unknown mean value is eliminated from the prediction procedure by unbiasedness condition of summing up weights to unity, i.e., $\boldsymbol{\lambda}^T \mathbf{u} = 1$, and is internally estimated from the sample (Ligas, 2022). Optimal coefficients $\boldsymbol{\lambda}$ (kriging weights), in the sense of unbiasedness and minimization of mean squared prediction error, are computed from the system of equations:

$$\begin{bmatrix} \boldsymbol{\Omega} & \mathbf{u} \\ \mathbf{u}^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ \kappa \end{bmatrix} = \begin{bmatrix} \boldsymbol{\omega}_0 \\ 1 \end{bmatrix} \tag{6}$$

where $\boldsymbol{\Omega}$ is a covariance matrix of observations (observations – observations), $\boldsymbol{\omega}_0$ stands for a covariance vector between observations and a target location to be predicted (observations – prediction), and $\mathbf{u}$ is a vector of ones.

Additional unknown κ (Lagrange multiplier) in (6) results from the unbiasedness condition. Ordinary kriging offers a measure of prediction quality (minimized mean squared prediction error σ^2_{OK}) which may be expressed as

$$\sigma^2_{OK} = \sigma^2 - \boldsymbol{\lambda}^T \boldsymbol{\omega}_0 - \kappa \tag{7}$$

where σ^2 is a variance of observations.

Kriging is a two-step procedure. Firstly, the temporal structure of a random process is inferred (from data and a priori knowledge if available) and contained in a semivariance or covariance function. Then, the covariance structure stored in $\boldsymbol{\Omega}$ and $\boldsymbol{\omega}_0$ is employed to solve for kriging weights in (6) and to assess the quality of prediction in (7).

2.6. ARIMA

This study also uses autoregressive integrated moving average (ARIMA) model, which is the combination of autoregressive model, moving average model, and differencing process (integrated part). This model makes possible to obtain greater flexibility in fitting the model to the actual time series (Box and Jenkins, 1976). The model mentioned above can be applied not only on a stationary time series but also on series that behave in a non-stationary way. In order to check the stationarity of the time series, the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test was used (Kwiatkowski et al., 1992). ARIMA model is defined as follows (Box and Jenkins, 1976):

$$\hat{X}_t^d = c + \sum_{i=1}^{p} \varphi_i X_{t-i}^d + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \tag{8}$$

where X_t^d is the predicted value; c is a constant; p is the order of autoregressive process; $\varphi_1 \ldots \varphi_p$ are parameters of the model; d is the degree of differencing; X_{t-i}^d are the differenced past series values; q is the order of moving average process; $\theta_1 \ldots \theta_q$ are parameters of the model; ε_{t-i} are the past white noise error terms; and ε_t is the white noise.

ARIMA model is generally denoted as ARIMA (p, d, q).

3. DATA DESCRIPTION AND PROCESSING

In this contribution, rapid EOP 14 12h (IGS, CODE, and GFZ) and final EOP 14 C04 12h (IERS) daily products were used to predict ERP. The obtained predictions cover the time interval from 1 January 2017 (MJD 57754.5) to 15 April 2022 (MJD 59684.5) for all evaluated ERP time series/data centres.

Prediction procedure is based on that one described in Michalczak and Ligas (2021) and Michalczak and Ligas (2022) set of parameters such as the number of the nearest neighbors (NN) for kriging, number of observations for structure function estimation, the time span (TS) necessary to estimate parameters of a linear trend and number of periodic components. The TS parameter for PMx and PMy is equal to 6.4 yrs and for LOD 9.2 yrs; NN for PMx and PMy equals to 10 immediate neighbors and for LOD it is equal to 25.

In case of ARIMA(p, d, q) model, the best set of parameters p and q was obtained by means of corrected Akaike Information Criterion (AICc) (Akaike, 1998). The loop with AICc criterion was run over the number of p and q varying from 1 to 5; the best set of parameters was selected and then a prediction was made based on them. For each 15-day and 30-day ERP prediction, the best set of p and q was selected separately. Parameter d determines a degree of differencing to be applied in order to transform a non-stationary time series into the stationary one in the mean sense. Stationarity of each ERP time series was checked using the KPSS test. Pole coordinates time series (after removing components mentioned in 2.4.) passed the stationary test for d = 2 and for LOD time series (after removing components mentioned in 2.2. and 2.4.) d = 1.

The prediction process starts with the selection of the first day of forecast. Then, the 30 observations preceding the first day of forecast were selected (rapid products from CODE, GFZ or IGS), while further back observations were taken from the final product (IERS). Afterward, the ERP predictions were carried out with ARIMA model and kriging for 15 and 30 days ahead. Obtained results were compared with each ERP IERS final time series. The whole process is then repeated, shifting the first day of prediction to next day and repeated until the last day of the forecast is reached, i.e., 15 April 2022. The process is illustrated in Figure 4. The prediction process based on final product (IERS) did not use any rapid products (this is to compare results of prediction with and without rapid product supplementation), which means that the 30 observations preceding the first day of prediction were selected from the final product – IERS EOP 14 C04 12h.

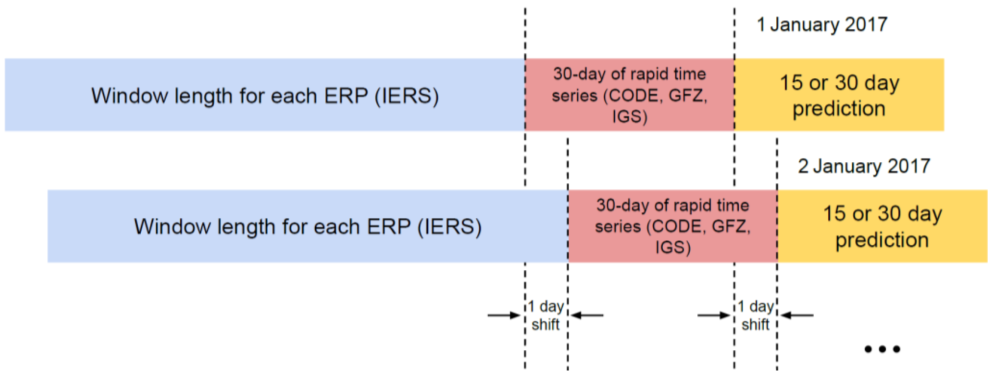


Figure 4. Diagram of the whole prediction process with rapid products

For the IGS, GFZ, and IERS, the total number of 15-day predictions for each ERP and each method amounts to 1886 and the total number of 30-day predictions is equal to 1871. Due to gaps in the CODE time series, the number of each ERP forecasts is equal to 1822 and for 30-day predictions amounts to 1807.

4. RESULTS

The ERP prediction accuracy was assessed by means of the mean absolute prediction error (MAPE) given by

$$MAPE_j = \frac{1}{n}\sum_{i=1}^{n}\left|e_{i,j}\right| \qquad j \coloneqq \text{prediction day number} \tag{9}$$

where n is a number of predictions, $e_{i,j} = O_{i,j} - P_{i,j}$ is the prediction error, O stands for observed and P for predicted ERP.

Figure 5 and 6 present the values of MAPEs for ultra-short (15-day into the future) PMx and PMy prediction using ARIMA and kriging for data time series from four analysis centres.

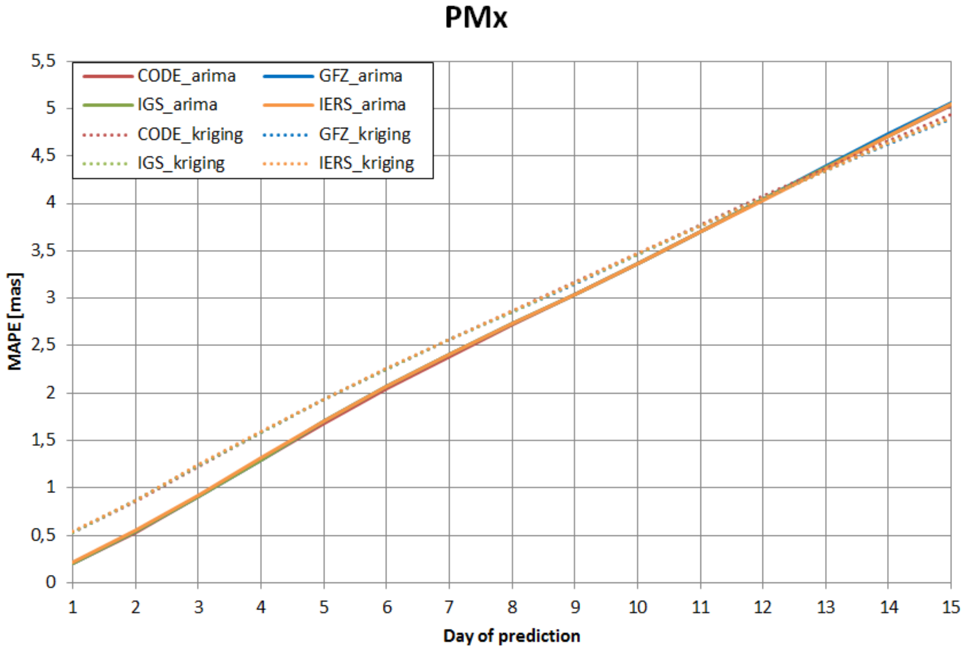


Figure 5. Comparison of MAPEs for 15-day PMx prediction for ARIMA and kriging for various analysis centres (CODE, GFZ and IGS denote rapid time series and IERS final time series)

PMx prediction with ARIMA model offers better accuracy up to day 13; after this day the advantage is on the side of kriging. For the first 5 days, the advantage is around 0.3 mas in favour of the ARIMA model. As the forecast day increases, the difference between ARIMA model and kriging decreases. For the 15th day, the difference is around 0.15 mas in favour of kriging. MAPEs for PMx for each analysis centres are very similar. The average difference between MAPEs for PMx prediction with ARIMA model is around 0.01 mas. In case of kriging, prediction performance is similar, but only for first 10 days. Then, MAPEs for CODE time series increase and differ from the other centres by about 0.04 mas.

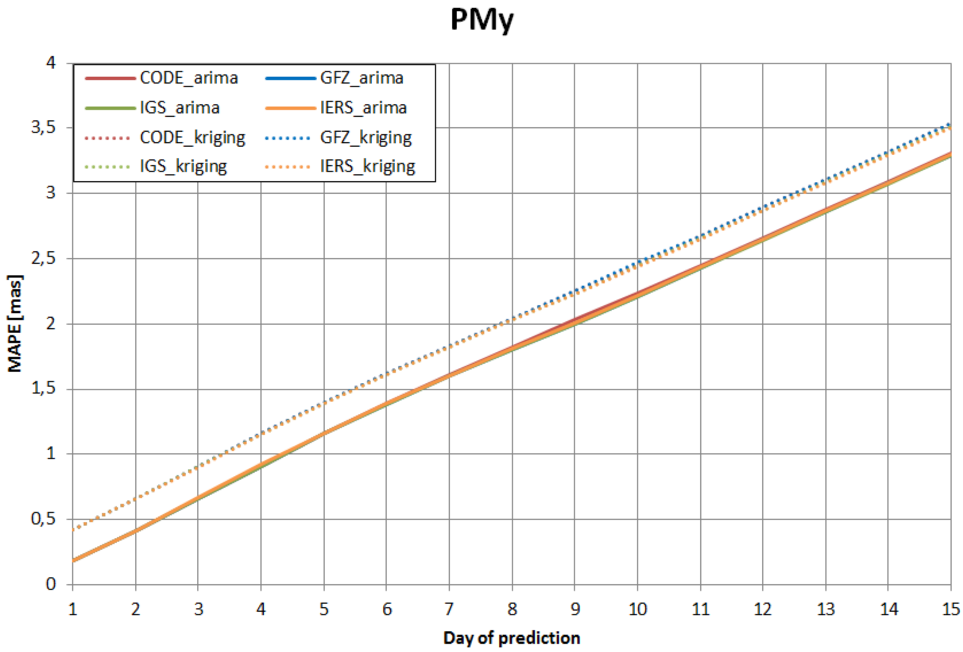


Figure 6. Comparison of MAPEs for 15-day PMy prediction for ARIMA and kriging for various analysis centres

PMy prediction with ARIMA model offers better accuracy up to the last day. The differences between ARIMA and kriging are around 0.23 mas in favour of ARIMA. Similarly to PMx results, the average difference between MAPEs for PMy prediction with ARIMA for four centres is around 0.01 mas with maximum value 0.03 mas. MAPEs for various centres for kriging prediction up to 8th day of prediction are nearly the same (with difference around 0.01 mas). On the 9th day of prediction differences between various centres increase to 0.02 mas and then on the last day, they reach values around 0.04 mas.

For both PMx and PMy, prediction accuracy based on final data (IERS) does not differ significantly from accuracy obtained with rapid products (CODE, GFZ and IGS) supplementing. This insignificant difference between the forecasts based on either final or rapid data applies also to 30-day predictions of PMx/PMy and to both LOD forecast horizons.

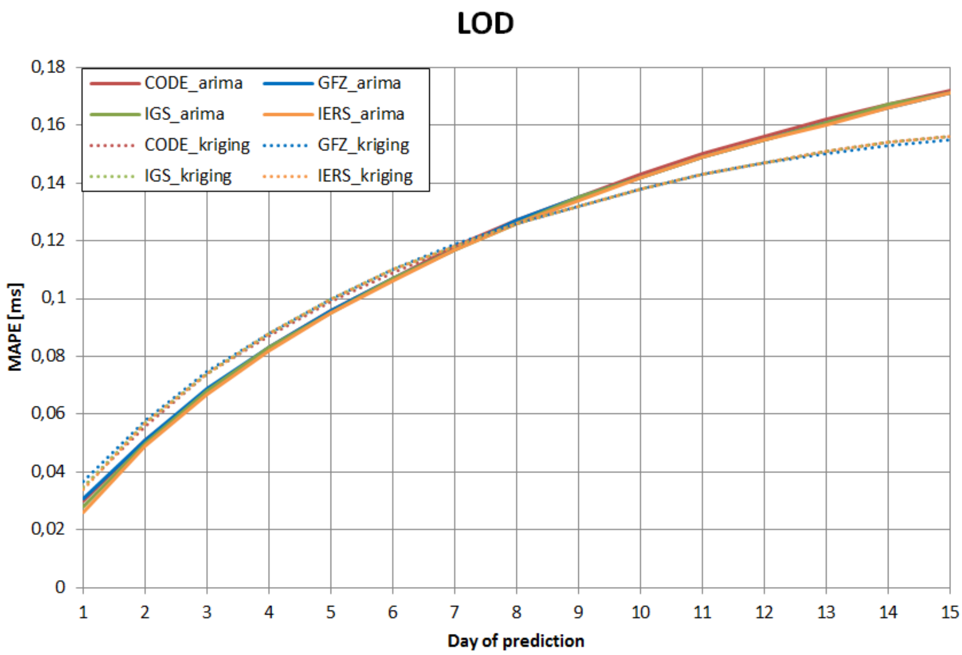


Figure 7. Comparison of MAPEs for 15-day LOD prediction for ARIMA and kriging for various analysis centres

Figure 7 presents the values of MAPEs for ultra-short (15-day into the future) LOD prediction using ARIMA and kriging for time series from four analysis centres. For the first 7 days, lower values of MAPEs are obtained, with advantage around 0.004 ms, for ARIMA. After this day, the advantage is on the side of kriging with maximum difference equal to about 0.016 ms. The accuracy of predictions within kriging and ARIMA for each time series is nearly identical.

Figures 8 and 9 present the values of MAPEs for short (30-day into the future) PMx and PMy prediction using ARIMA and kriging.

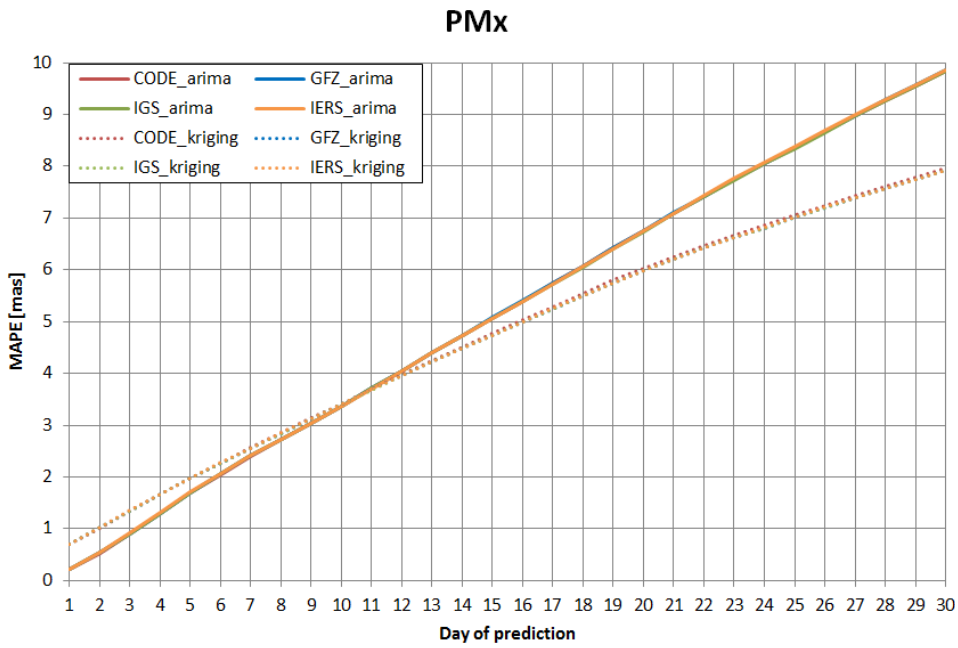


Figure 8. Comparison of MAPEs for 30-day PMx prediction for ARIMA and kriging for various analysis centres

Up to 11 day of PMx prediction, MAPEs for ARIMA are smaller than those for kriging with maximum advantage around 0.5 mas. After this day, kriging reaches better accuracy of forecast. The maximum differences between ARIMA and kriging are on the 30th day and are equal to around 1.91 mas. The accuracy of predictions based on various data centres is nearly identical; the maximum differences vary from around 0.01 mas to 0.05 mas.

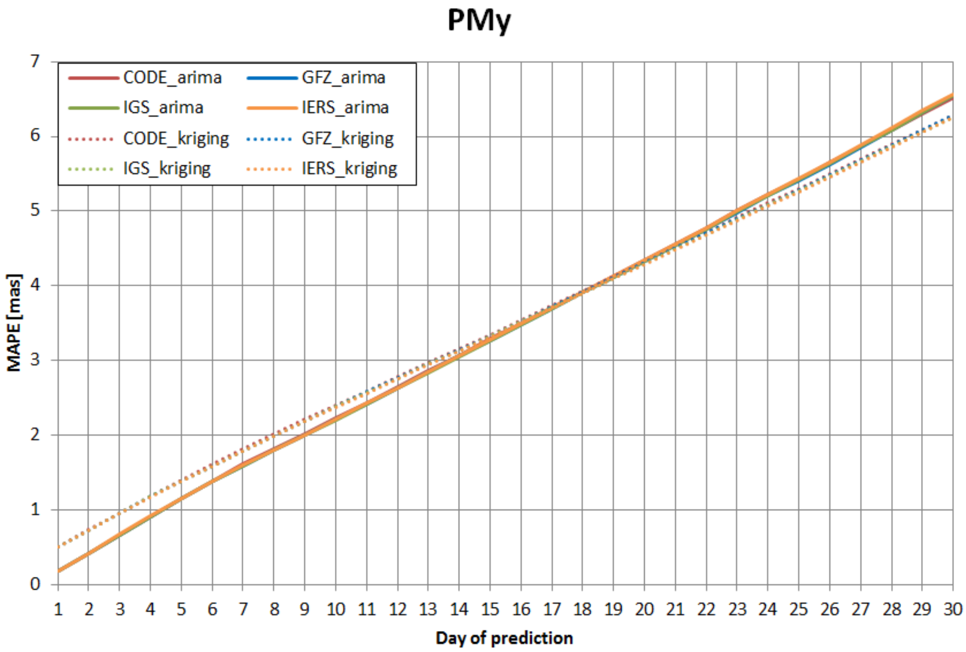


Figure 9. Comparison of MAPEs for 30-day PMy prediction for ARIMA and kriging for various analysis centres

For the first 18 days of PMy prediction, results indicate a better prediction performance of ARIMA model. From the 19th day, the advantage is in favour of kriging and reaches maximum value about 0.3 mas. The differences between predictions based on various data centres are small, around 0.01 – 0.03 mas.

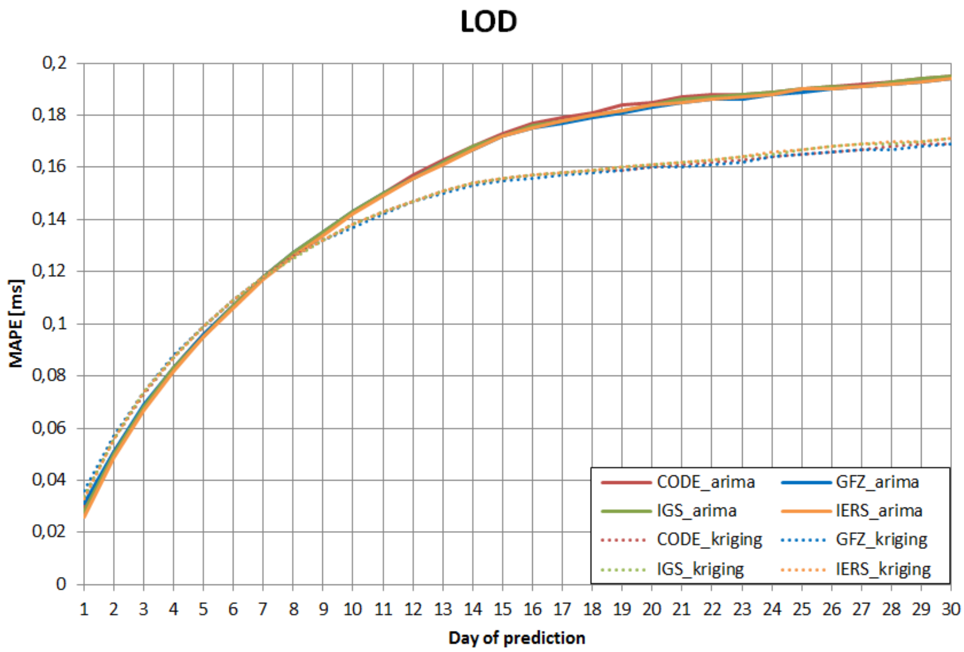


Figure 10. Comparison of MAPEs for 30-day LOD prediction for ARIMA and kriging for various analysis centres

Figure 10 shows a comparison of MAPEs for 30-day LOD prediction for ARIMA and kriging. Results indicate that for the first 6 days, prediction with ARIMA is more accurate. From the 7th day onward, kriging gives a more accurate forecast. The maximum difference between the forecasting methods is about 0.026 ms for CODE (on the last day of prediction). It is worth noting that MAPEs for IGS and IERS using kriging are most similar to each other. In majority, this also holds for ARIMA.

5. CONCLUSIONS

Earth Rotation Parameters predictions based on ordinary kriging and ARIMA model have been presented in this contribution. Predictions have been performed on the basis of rapid (CODE, IGS and GFZ) and final (IERS) ERP products. The quality of polar motion forecast is visibly coordinate-dependent, i.e., PMx is worse predictable than PMy. The results indicate that the ARIMA-based prediction is better for ultra-short prediction; for PMx it turned out to be 11 -days of forecast, for PMy 18 -days, and for LOD 8 -days. The maximum differences for first few days of 15-day predictions between methods are around 0.32 mas (PMx), 0.23 mas (PMy) and 0.004 ms (LOD) in favour of ARIMA. On the last day, the differences reach 0.15 mas (PMx) and 0.016 ms (LOD) with advantage to kriging. For the first few days of 30-day prediction, ARIMA gives lower MAPEs of approximately 0.5 mas (PMx), 0.3 mas (PMy) and 0.007 ms (LOD). The maximum differences of MAPEs on the last days of 30-day predictions are 1.91 mas (PMx), 0.3 mas (PMy) and 0.026 ms (LOD) with advantage to kriging method. For all ERP, the differences of MAPEs for time series from various analysis centres are not significant and vary up to maximum value of around 0.05 mas (PMx), 0.04 mas (PMy) and 0.005 ms (LOD). These values are on the level of accuracy of corresponding ERP determination, i.e., polar coordinates with the accuracy of app. 50 µas or better and 10 µs in case of LOD (Dick and Thaller, 2020). It is also worth noting that the accuracy of prediction on the basis of rapid products (30-day gap filling) is on the same level as on final product only. The conducted comparisons did not give a clear answer as to which method, i.e., ARIMA or kriging, is better for a given ERP and for a given forecast length, with the exception of ultra-short-term (15 days) prediction for PMy where ARIMA systematically behaved better. The results of this contribution are supported by a considerable number of performed predictions; therefore, all intermediate results are available at the website mentioned in the acknowledgements section.

Acknowledgements. The paper is a result of research on geospatial methods carried out within the statutory research grant no. 16.16.150.545 at the Department of Integrated Geodesy and Cartography, AGH University of Science and Technology, Krakow.

Intermediate results are available at the website accompanied with this article home.agh.edu.pl/mmichalc.

REFERENCES

Akaike H., (1998) Information Theory and an Extension of the Maximum Likelihood Principle, In: Parzen, E., Tanabe, K., Kitagawa, G. (eds) Selected Papers of Hirotugu Akaike. *Springer Series in Statistics, Springer*, New York, NY. https://doi.org/10.1007/978-1-4612-1694-0_15

Akyilmaz O., Kutterer H., Shum C., Ayan T., (2011) Fuzzy-wavelet based prediction of earth rotation parameters, *Applied Soft Computing,* 11(1):837–841, https://doi.org/10.1016/j.asoc.2010.01.003

Box, G.E.P. and Jenkins, G.M. (1976) Time Series Analysis: Forecasting and Control, *Holden Day* San Francisco

Cressie N.A.C. (1993) Statistics for spatial data, *John Wiley & Sons*, New York

Dick W. R. and Thaller D. (eds.) IERS Annual Report 2018 (2020) International Earth Rotation and Reference Systems Service, Central Bureau. *Frankfurt am Main: Verlagdes Bundesamts für Kartographie und Geodäsie*, 207 p., ISBN 978-3-86482-136-3

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, pp. 287—295. doi: http://doi.org/10.1007/s00190-018-1158-7

Gambis D, Luzum B., (2011) Earth rotation monitoring, UT1 determination and prediction, *Metrologia*; 48:165-70

Kalarus M., Schuh H., Kosek W., Akyilmaz O., Bizouard Ch., Gambis D., Gross, B R.. Jovanovi´c, Kumakshev S., Kutterer H., Mendes Cerveira P. J., Pasynok S., Zotov L., (2010) Achievements of the Earth Orientation Parameters prediction comparison campaign, *Journal of Geodesy*, 84:587–596

Kosek W., Kalarus M., Niedzielski T., (2007) Forecasting Of the Earth Orientation Parameters - Comparison of Different Algorithms, *Journées Systèmes de Référence Spatio-temporels*, Observatoire de Paris, 17-19 September 2007

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., Shin, Y. (1992) Testing the null hypothesis of stationarity against the alternative of a unit root, *Journal of Econometrics*. 54 (1–3): 159–178, https://doi.org/10.1016/0304-4076(92)90104-Y

Lei Y., Guo M., Cai. H, Hu D., Zhao D., (2015) Prediction of Length-of-day Using Gaussian Process Regression, *The Journal of Navigation*, 68:563–575, https://doi.org/10.1017/S0373463314000927

Liao D.C., Wang Q.J., Zhou Y.H., Liao X.H., Huang C.L., (2012) Long-term prediction of the Earth Orientation Parameters by the artificial neural network technique, *Journal of Geodynamics* Volume 62, Pages 87-92, ISSN 0264-3707, https://doi.org/10.1016/j.jog.2011.12.004

Ligas, M., (2022), Comparison of kriging and least-squares collocation – Revisited, *Journal of Applied Geodesy*, vol. 16, no. 3, 2022, pp. 217-227. https://doi.org/10.1515/jag-2021-0032

Linnet. K. (1990). Estimation of the linear relationship between the measurements of two methods with proportional errors. *Statist. Med.*. 9: 1463-1473. https://doi.org/10.1002/sim.4780091210

Luo, J., Chen, W., Ray, J., Li J., (2022), Short-Term Polar Motion Forecast Based on the Holt-Winters Algorithm and Angular Momenta of Global Surficial Geophysical Fluids. *Surv Geophys*. https://doi.org/10.1007/s10712-022-09733-0

Michalczak M., Ligas M., (2021) Kriging-based prediction of the Earth's pole coordinates, *Journal of Applied Geodesy*, vol. 15, no. 3, pp. 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*, https://doi.org/10.1016/j.asr.2022.05.007

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* 94:23

Nastula J., Chin T. M., Gross R., Śliwińska J., Wińska M., (2020) Smoothing and predicting celestial pole offsets using a Kalman filter and smoother, *Journal of Geodesy* 94:29

Niedzielski T., Kosek W., (2008) Prediction of UT1–UTC, LOD and AAM $\chi 3$ by combination of least-squares and multivariate stochastic methods, *J Geod* 82:83–92

Okhotnikov G., Golyandina N. (2019) EOP Time Series Prediction Using Singular Spectrum Analysis" *RWTH Aahen University*

Passing H. Bablok W.. (1983) A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry. Part I. *J Clin Chem Clin Biochem*. Nov;21(11):709-20. doi: 10.1515/cclm.1983.21.11.709. PMID: 6655447

Petit G and Luzum B (eds.). IERS Conventions (2010), IERS Technical Note 36, Frankfurt am Main: Verlag des Bundesamts für Kartographie und Geodäsie. 179 pp., ISBN 3-89888-989-6

Schuh H., Ulrich M., Egger D., Müller J., Schwegmann W. (2002) Prediction of Earth Orientation Parameters by Artificial Neural Networks, *Journal of Geodesy,* 76(5), 247–258, https://doi.org/10.1007/s00190-001-0242-5

Wu F., Chang G., Deng K., (2019) One-step method for predicting LOD parameters based on LS+AR model", *Journal of Spatial Science*, 66:2, 317-328, https://doi.org/10.1080/14498596

Zotov, Leonid & Bizouard, Ch. (2018). Escargot Effect and the Chandler Wobble Excitation, *Journal of Physics: Conference Series,* 955, 012033, doi: 10.1088/1742-6596/955/1/012033

Received: *2022-07-18*

Reviewed: *2022-08-19 (undisclosed name); 2022-08-22 (C. Ron)*

Accepted: *2022-10-06*

Reports on Geodesy and Geoinformatics, 2024, Vol. 117, pp. 45–54

DOI: 10.2478/rgg-2024-0006
Received: 7 November 2023 / Accepted: 20 February 2024
Published online: 6 March 2024

ORIGINAL ARTICLE

Short-term prediction of UT1-UTC and LOD via Dynamic Mode Decomposition and combination of least-squares and vector autoregressive model

Maciej Michalczak [1*] and Marcin Ligas [1]

[1]Department of Integrated Geodesy and Cartography, Faculty of Geo-Data Science, Geodesy, and Environmental Engineering, AGH University of Krakow, al. Adama Mickiewicza 30, 30-059 Kraków

*mmichalc@agh.edu.pl

Abstract

This study presents a short-term forecast of UT1-UTC and LOD using two methods, i.e. Dynamic Mode Decomposition (DMD) and combination of Least-Squares and Vector Autoregression (LS+VAR). The prediction experiments were performed separately for yearly time spans, 2018-2022. The prediction procedure started on January 1 and ended on December 31, with 7-day shifts between subsequent 30-day forecasts. Atmospheric Angular Momentum data (AAM) were used as an auxiliary time series to potentially improve the prediction accuracy of UT1-UTC and LOD in LS+VAR procedure. An experiment was also conducted with and without elimination of effect of zonal tides from UT1-UTC and LOD time series. Two approaches to using the best steering parameters for the methods were applied:. First, an adaptive approach, which observes the rule that before every single forecast, a preliminary one must be performed on the pre-selected sets of parameters, and the one with the smallest prediction error is then used for the final prediction; and second, an averaged approach, whereby several forecasts are made with different sets of parameters (the same parameters as in adaptive approach) and the final values are calculated as the averages of these predictions. Depending on the method and data combination mean absolute prediction errors (MAPE) for UT1-UTC vary from 0.63 ms to 1.43ms for the 10th day and from 3.07 ms to 8.05ms for the 30th day of the forecast. Corresponding values for LOD vary from 0.110 ms to 0.245 ms for the 10th day and from 0.148 ms to 0.325 ms for the 30th day.

Key words: UT1-UTC, length of day, dynamic mode decomposition, autoregression, prediction

1 Introduction

The Earth Orientation Parameters (EOP) describe the irregularities of the Earth's rotation and its orientation in space. They include three classes of parameters, i.e. Earth pole coordinates (PMx and PMy), the difference between Universal Time (UT1) and Universal Coordinated Time (UTC), determined by atomic clocks, i.e. UT1-UTC and Celestial Pole Offsets (dX and dY, yielded by the precession-nutation models). This set of parameters is essential to perform the conversion between Celestial Reference Frame (CRF) and Terrestrial Reference Frame (TRF) (Gambis and Luzum, 2011). Knowledge of EOP plays a significant role in several other astronomical and geodetic applications, including: space navigation, precise orbit determination, and climate forecasting and analysis.

The Length of a Day (LOD) is the first negative derivative from the UT1-UTC after removing leap seconds and is determined as the difference between the astronomically determined duration of the Earth's rotation and 86400 SI seconds. Both parameters are used to model changes in the Earth's rotation rate (Modiri et al., 2020). The following factors play an important role in the variability of UT1-UTC and LOD, e.g.: El Niño Southern Oscillation (Holton and Dmowska, 1989; Soffel, 2013), zonal wind variations in atmospheric general circulation models (Höpfner, 1998), and ice mass loss and the resulting sea level changes (Gross et al., 2004). According to Xu et al. (2022b) LOD is also strongly related to The El Nino-Southern Oscillation (ENSO) and Atmospheric Angular Momentum (AAM)

Publisher: De Gruyter

changes, across the frequencies of interannual oscillations.

Determination of EOP with a high accuracy combines such advanced geodetic techniques as Very Long Baseline Interferometry (VLBI), Global Navigation Satellite System (GNSS), Satellite Laser Ranging (SLR), Lunar Laser Ranging (LLR), and Doppler Orbitography and Radiopositioning Integrated by Satellite (DORIS). This combination enables a determination of EOP with accuracy of approximately 50 μas or higher in the case of pole coordinates and 30 μs or higher in the case of UT1-UTC (10 μs for LOD) (Dick and Thaller, 2020).

The final EOP data are usually provided with delay caused by the complexity of measuring and data processing, and published, for example, by the International Earth Rotation and Reference Systems Service (IERS) after 30 days (Dick and Thaller, 2020). This latency gap can be filled using rapid products (published by, e.g., IERS (Daily Rapid EOP Data), CODE (Center for Orbit Determination in Europe), GFZ (German Research Centre for Geosciences), or IGS (International GNSS Service)). For several astronomical and geodetic purposes, the knowledge of the near real-time EOP data is needed. The solution to this problem is EOP prediction.

EOP prediction is commonly used, as evidenced by two comparison campaigns, organised by IERS, i.e., the 1st Earth Orientation Parameters Prediction Comparison Campaign held in 2006-2008 (Kalarus et al., 2010) and the 2nd Earth Orientation Parameters Prediction Comparison Campaign held in 2021-2022 (http://eoppcc.cbk.waw.pl/, Kur et al. 2022; Śliwińska et al. 2022). Also, within a GGOS (Global Geodetic Observing System) infrastructure, a Joint Study Group 3: AI for Earth Orientation Parameter Prediction has been established (https://ggos.org/about/org/fa/ai-for-geodesy/eop-prediction/).

A number of techniques and data combinations have been used and developed to improve the accuracy of EOP prediction, e.g., neural networks (Guessoum et al., 2022; Liao et al., 2012; Schuh et al., 2002), machine learning (Kiani Shahvandi et al., 2022; Lei et al., 2017), kriging (Michalczak and Ligas, 2021, 2022), kalman filter (Gross et al., 1998; Nastula et al., 2020; Xu et al., 2012), singular spectrum analysis (Okhotnikov and Golyandina, 2019), or autoregressive models (Dill et al., 2018; Niedzielski and Kosek, 2011).

In this contribution, to predict UT1-UTC and LOD, the methods of Dynamic Mode Decomposition (DMD) and combination of least-squares and vector autoregression (LS+VAR) were used. DMD is a data-driven, equation-free technique capable of reconstructing and forecasting time series in a single numerical procedure whilst VAR is a multivariate counterpart of the Autoregressive (AR) model. DMD was applied to the separate UT1-UTC and LOD prediction whilst LS+VAR model was fed with UT1-UTC, LOD and Atmospheric Angular Momentum (AAM) series in various combinations (all used data combinations are explained in Table 1). AAM information, as an auxiliary variable, was applied to potentially strengthen the prediction of UT1-UTC and LOD in the LS+VAR model. The study also examined the difference in the accuracy of prediction with and without removing the effect of zonal tides.

This study attempts to answer the question of whether the direct incorporation of additional external data (Liouville equation not involved), such as AAM motion term, can improve LS+VAR predictions of UT1-UTC and LOD. This study also examines the impact of the removal of the effect of zonal tides from the UT1-UTC and LOD time series.

2 Prediction methods

2.1 Dynamic Mode Decomposition

DMD, based on fluid dynamics (Schmid, 2010), is a data-driven method of reconstructing and forecasting generally non-linear dynamical systems using linear techniques. It is governed by a main relation that links future states of a system to past ones through an (unknown) operator $\mathbf{A}$ that stores the dynamics of the system's evolution:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k \tag{1}$$

where $\mathbf{x}_k$ and $\mathbf{x}_{k+1}$ are the subsequent snapshots (images of a dynamical system) at discrete time instances t_k and t_{k+1}.

This equation has the form of a linear homogeneous system of difference equations with constant coefficients with some initial vector $\mathbf{x}_1$ which is known to have the solution given by:

$$\mathbf{x}_{k+1} = \mathbf{A}^k\mathbf{x}_1 = \mathbf{\Phi}\mathbf{\Lambda}^k\mathbf{\Phi}^{-1}\mathbf{x}_1 \tag{2}$$

where $\mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^{-1}$ is the eigendecomposition of $\mathbf{A}$.

DMD algorithm searches the best-fit matrix $\mathbf{A}$ in (1) (in fact its dominant eigenstructure). This is accomplished by generating a snapshot matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_m]$ that stores states of a system at discrete time instances $t = 1, 2, \dots, m$. The matrix is split into two other matrices, $\mathbf{X}_1 = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{m-1}]$ and $\mathbf{X}_2 = [\mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_m]$, that are related, according to (1), by a matrix $\mathbf{A}$, i.e.:

$$\mathbf{X}_2 = \mathbf{A}\mathbf{X}_1 \tag{3}$$

The estimate of an operator $\mathbf{A}$ is given by:

$$\mathbf{A} = \mathbf{X}_2\mathbf{X}_1^+ \tag{4}$$

where $\mathbf{X}_1^+ = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^T$ denotes the Moore-Penrose pseudo-inverse determined using the singular value decomposition (SVD). This enables for a low dimensional representation of $\mathbf{A}$ (Tu et al., 2014) by:

$$\tilde{\mathbf{A}} = \mathbf{U}^T\mathbf{A}\mathbf{U} = \mathbf{U}^T\mathbf{X}_2\mathbf{V}\mathbf{\Sigma}^{-1} \tag{5}$$

Eigenvalues of $\tilde{\mathbf{A}}$ are those of $\mathbf{A}$ and are obtained as a solution to the eigenvalue problem:

$$\tilde{\mathbf{A}}\mathbf{W} = \mathbf{W}\mathbf{\Lambda} \tag{6}$$

whilst the eigenvectors of $\mathbf{A}$ are given by:

$$\mathbf{\Phi} = \mathbf{X}_2\mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{W} \tag{7}$$

For high dimensional problems SVD of $\mathbf{X}_1$ in (4) is truncated to some prescribed rank r (low rank approximation of $\mathbf{X}_1$) yielding:

$$\mathbf{X}_1 \cong \mathbf{U}_r\mathbf{\Sigma}_r\mathbf{V}_r^T \tag{8}$$

With this approximation, the algorithm works identically as previously with components of decomposition $\mathbf{U}$, $\mathbf{\Sigma}$, $\mathbf{V}$ replaced with $\mathbf{U}_r$, $\mathbf{\Sigma}_r$, $\mathbf{V}_r$. Once the eigenstructure of $\mathbf{A}$ is determined, (2) can be used to reconstruct or predict future states of the system. For low dimensional problems (numerically tractable) operator $\mathbf{A}$ may be obtained from (4) and applied directly to (1).

In the case of time series analysis, the snapshot matrix becomes a trajectory matrix (known from, e.g., singular spectrum analysis), as described by Tirunagari et al. (2017). Hence, the input time series $\mathbf{x}$ of length T is split into L subseries $\mathbf{x}'$ of length K that are moved by one time step ahead, i.e.:

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & \cdots & x_L \\ x_2 & x_3 & \cdots & x_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_K & x_{K+1} & \cdots & x_{K+L-1} \end{bmatrix} = [\mathbf{x}'_1 \quad \mathbf{x}'_2 \quad \cdots \quad \mathbf{x}'_L] \tag{9}$$

where L is selected so that $2 \le L \le \frac{T}{2}$, K is related to L and T with

$K = T - L + 1$. The trajectory matrix is then split into two $K \times (L-1)$ matrices, as described above. The remaining steps of the DMD algorithm do not change.

2.2 Least-squares estimated linear trend and periodicities extrapolation with Vector AutoRegression on residues

This prediction procedure begins with a least-squares fit of a linear trend and periodic components (10) to the input time series (separately for each of variables V_i).

$$V_i = A_i + B_i t + \sum_{j}^{k} a_{i,j} sin\left(\omega_{i,j} t + \phi_{i,j}\right) + e_i \tag{10}$$

where V stands for a variable of interest, i.e., UT1-UTC, LOD, or AAM; A, B are the intercept and slope of a linear trend; a, ω, ϕ are the amplitude, frequency and phase of a sine wave, e stands for an error term, $[A, B, a, \omega, \phi]$ to be determined. Residuals $V_i^{Res} = \hat{e}_i$ obtained after eliminating the linear trend and periodicities in (10) are subject to a joint prediction through p − lagged vector autoregressive model LS+VAR. VAR model describes a dynamic dependence between variables, i.e., links current values of variables with their past values, and past values of other variables in the system and may be expressed as:

$$\mathbf{V}_T^{Res} = \mathbf{c} + \mathbf{A}'_p \mathbf{V}_{T-p}^{Res} + \mathbf{A}'_{p-1} \mathbf{V}_{T-p+1}^{Res} + \ldots + \mathbf{A}'_1 \mathbf{V}_{T-1}^{Res} + \mathbf{e}_T^{Res} \tag{11}$$

where a constant vector $\mathbf{c}$ ($nv \times 1$) accounts for a non-zero mean value of the vector process, each of $\mathbf{V}^{Res}$ ($nv \times 1$) stores residual values of variables at appropriate time delay, each $\mathbf{A}'$ is a $nv \times nv$ matrix of coefficients and $\mathbf{e}^{Res}$ ($nv \times 1$) is a vector of error terms, nv denotes a number of variables.

After the estimation of coefficients matrices $\mathbf{A}'$ in (11), the forecast may be performed through the successive use of the formula:

$$\hat{\mathbf{V}}_{T+h}^{Res} = [\hat{\mathbf{A}}'_1 \quad \hat{\mathbf{A}}'_2 \cdots \hat{\mathbf{A}}'_p] \begin{bmatrix} \hat{\mathbf{V}}_{T+h-1}^{Res} \\ \hat{\mathbf{V}}_{T+h-2}^{Res} \\ \vdots \\ \hat{\mathbf{V}}_{T+h-p}^{Res} \end{bmatrix} \tag{12}$$

where h denotes a time-step of forecast horizon and if $h - i \leq 0$ where $i = 1, 2, \ldots, p$ then $\hat{\mathbf{V}}_{T+h-i}^{Res} = \mathbf{V}_{T+h-i}^{Res}$. Finally, the extrapolated trend and periodic components of (10) are combined with the predicted residuals (12).

3 Data description and processing details

The short-term (30 days into the future) UT1-UTC and LOD forecast was performed using IERS EOP 14 C04 (IAU2000A) series as a reference (Bizouard et al., 2018). AAM series (Dobslaw and Dill, 2018) was also used – as an auxiliary data source in LS+VAR-based prediction procedure to potentially strengthen the prediction of UT1-UTC and LOD. The prediction experiment was completed in two variants, the first one with the removal of effect of zonal tides (recommended by IERS Conventions (2010) (Petit and Luzum, 2010)) and the other omitting this step.

The first stage in the prediction procedure involved identifying sets of best-performing steering parameters on the yearly period immediately preceding the period of the relevant forecasts (e.g., forecasts within 2018 were based on steering parameters determined for the year 2017). These sets of parameters were selected based on the condition of the least sum of squares of differences between the observed and predicted values for all 30 days of the forecasts.

The searching loop responsible for preselecting sets of parameters involved:

in the case of DMD:
- input time series length (T),
- number of snapshots (L).

in the case of LS+VAR:
- input time series length (T) for estimation of linear trend and periodic components,
- number of periodic components (PC),
- subseries of (T) for LS+VAR parameters estimation ($subT$),
- autoregression order (p).

The sets of parameters selected in this way were then applied in two kinds of prediction approaches: adaptive and averaged. In the adaptive approach, before every single proper 30-day prediction, a preliminary one was performed of the preceding 30 days. This preliminary forecast was made with pre-selected sets of parameters obtained in the aforementioned step, and the method yelding the lowest number of prediction errors was applied to the final current forecast, and the procedure was repeated until the last 30-day forecast was completed. The averaged approach differs from the adaptive one in that the forecasts are made for all sets of preselected parameters, and the actual 30-day forecast is computed as the average of these forecasts.

Table 1 presents all variants of prediction methods and data combinations used in this contribution. In the case of LS+VAR, there was a joint forecast of either two (UT1-UTC, LOD; UT1-UTC, AAM; LOD, AAM) or three variables (UT1-UTC, LOD, AAM) but the prediction procedure was optimized so that the one variable of the group was treated as the primary variable and the remaining ones as auxiliary ones. A prediction option without removing the effect of zonal tides is shown in Table 1 with the subscript “nc” (not corrected).

4 Results

To measure the quality of predictions using the examined methods and different data combinations the mean absolute prediction error (MAPE) given by (13) was used:

$$MAPE_j = \frac{1}{n} \sum_{i=1}^{n} \left| O_{i,j} - P_{i,j} \right| \tag{13}$$

where: n is a number of predictions, j is a prediction day number, O means *observed* and P - for *predicted* (here it is either UT1-UTC or LOD).

Figures 1 and 2 show MAPEs for a 30-day prediction of UT1-UTC for the yearly time spans 2018-2022 using several methods of forecast - different variations of DMD and LS+VAR. The results show that DMD, in time spans 2018-2021, is characterized by the smallest MAPEs for the 30th day of prediction. In 2022 the most accurate method is LS+VAR in both mass and motion terms using 3 input time series (UT1-UTC, LOD and AAM mass and motion terms). Clearly, for the first day of prediction, MAPEs for all methods are comparable and are at the level of around 0.02 ms. Results of predictions up to 5-7 days into the future indicate that the inclusion of AAM data in the LS+VAR prediction procedure decreases forecast errors only marginally. Interestingly, there is no significant improvement in DMD predictions of UT1-UTC when tidal effects are removed, while the improvement in the LS+VAR forecast is noticeable.

Tables 2 and 3 offers a more detailed presentation of the results of 30-day UT1-UTC predictions for all methods in the yearly time spans 2018-2022 (for selected days of forecast). For the 5th and 10th day of the forecast, the combination of LS+VAR and AAM data (using 3 input time series) achieved the smallest average errors of around 0.29 ms and 0.88 ms, respectively. The longer the range of

Table 1. Summary of variants of prediction methods and data combinations

Acronym	Description
DMD, DMD_{nc} (*)	adaptive DMD prediction of UT1-UTC or LOD
LS+VAR, $LS+VAR_{nc}$	adaptive LS+VAR prediction of UT1-UTC and LOD
aveDMD, $aveDMD_{nc}$	averaged DMD prediction of UT1-UTC or LOD
aveLS+VAR, $aveLS+VAR_{nc}$	averaged LS+VAR prediction of UT1-UTC and LOD
LS+VAR_AAM (2)	adaptive LS+VAR prediction of UT1-UTC or LOD and AAM mass and motion terms
LS+VAR_AAM (3)	adaptive LS+VAR prediction of UT1-UTC, LOD and AAM mass and motion terms
$LS+VAR\_AAM_{mass}$ (2)	adaptive LS+VAR prediction of UT1-UTC or LOD and AAM mass term
$LS+VAR\_AAM_{mass}$ (3)	adaptive LS+VAR prediction of UT1-UTC, LOD and AAM mass term
$LS+VAR\_AAM_{motion}$ (2)	adaptive LS+VAR prediction of UT1-UTC or LOD and AAM motion term
$LS+VAR\_AAM_{motion}$ (3)	adaptive LS+VAR prediction of UT1-UTC, LOD and AAM motion term
aveLS+VAR_AAM (2)	averaged LS+VAR prediction of UT1-UTC or LOD and AAM mass and motion terms
aveLS+VAR_AAM (3)	averaged LS+VAR prediction of UT1-UTC, LOD and AAM mass and motion terms
$aveLS+VAR\_AAM_{mass}$ (2)	averaged LS+VAR prediction of UT1-UTC or LOD and AAM mass term
$aveLS+VAR\_AAM_{mass}$ (3)	averaged LS+VAR prediction of UT1-UTC, LOD and AAM mass term
$aveLS+VAR\_AAM_{motion}$ (2)	averaged LS+VAR prediction of UT1-UTC or LOD and AAM motion term
$aveLS+VAR\_AAM_{motion}$ (3)	averaged LS+VAR prediction of UT1-UTC, LOD and AAM motion term

(*) subscript "nc" stands for "not corrected", (2) or (3) stands for a number of variables involved in a prediction procedure incorporating AAM data.

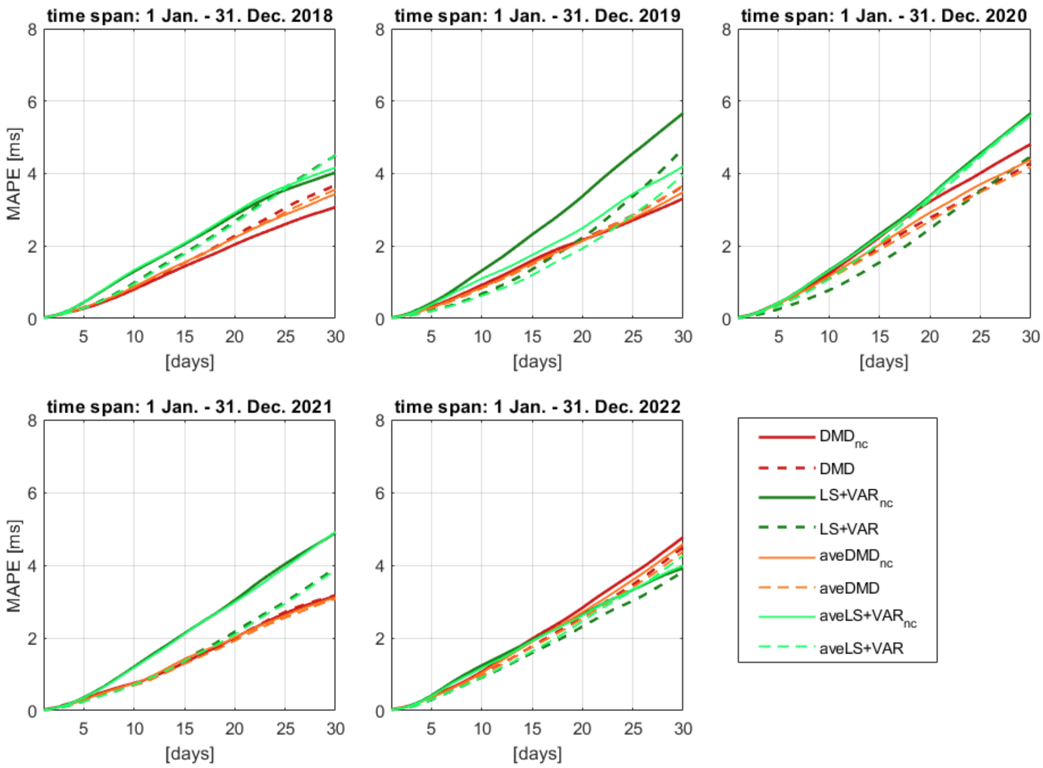


Figure 1. MAPEs of 30-day UT1-UTC prediction for the yearly time spans 2018–2022 (nc = not corrected).

the forecast, the more clearly DMD becomes the most accurate forecast method. Finally, average MAPEs on the 30th day of UT1-UTC prediction vary, depending on the method from 3.77 ms for aveDMD to 6.35 ms for aveLS+VAR AAM (2). Interestingly, the accuracy of the forecasts in 2020 dropped in almost all cases, only $LS+VAR_{nc}$ and LS+VAR maintained similar forecast error levels as in the other years. In almost half of the cases, there were smaller prediction errors using the averaged approach in UT1-UTC prediction, relative to adaptive approach. Clearly, LS+VAR AAM procedure prediction based on only 2 input time series is characterized by bigger MAPEs than the corresponding procedure with 3 input time series.

Figures 3 and 4 summarise MAPEs for 30-day predictions of LOD for the yearly time spans 2018-2022 using several methods of forecast: different variations of DMD and LS+VAR. The results indicate that for almost all time spans the most accurate method is DMD (or its variants). The prediction errors for 2022 showed a comparable accuracy of LOD prediction for almost all methods (except $LS+VAR_{nc}$, $aveLS+VAR_{nc}$ and aveLS+VAR AAM_{mass} (2)). MAPEs for the first day of forecast for all methods are very similar and are around 0.020 ms – 0.030 ms. Similarly to UT1-UTC prediction results, the ultra-short-term (up to 5-7 days into the future) LOD LS+VAR forecast with additional AAM time series is characterized by smaller errors than LS+VAR procedure without AAM time series. On average, the most accurate methods among the methods without additional AAM data for the 30the day are $aveDMD_{nc}$ and aveDMD, while among the LS+VAR variants with AAM information, it is aveLS+VAR AAM (3). A comparison of the accuracy of LOD prediction options with and without removal of the effect of zonal tides shows that removing tidal effects has significant impact on prediction accuracy using LS+VAR, while there is no significant decrease of MAPEs in DMD.

Tables 4 and 5 present the selected daily values of MAPEs for 30-day LOD prediction for all methods. For days up to the 5th day, the accuracy of prediction for almost all methods is comparable, except LS+VAR$_{nc}$, aveLS+VAR$_{nc}$ and aveLS+VAR AAM$_{mass}$ (2) , whose accuracy is significantly lower. The results indicate that LS+VAR methods with AAM data (based on 3 input time series) for 5th and 10th days for 2019, 2021 and 2022 are among the most accurate methods. Similarly to UT1-UTC prediction, introducing AAM time series into LS+VAR forecast procedure marginally increases the accuracy for 5-10 day range, relative to LS+VAR predictions without AAM data. After the 10th day the advantage of DMD-based methods is noticeable. On the 20th day of prediction, DMD-based methods reach averaged MAPEs between 0.174 ms and 0.182 ms, while other methods exceed 0.200 ms. For 2018, 2019 and 2022 aveDMD provided the most accurate LOD predictions for the 30th day, whilst in 2020 DMD$_{nc}$ and in 2021 DMD reached the smallest values of MAPEs. The results (on average) for the 30th day vary from 0.176 ms (aveDMD$_{nc}$ and aveDMD) to 0.242 ms for aveLS+VAR$_{nc}$. Removing tidal effects in LS+VAR prediction procedure decreased the averaged MAPEs for the 30th day of predictions from 0.240 ms to 0.227 ms (LS+VAR) and from 0.242 ms to 0.221 ms (aveLS+VAR). In

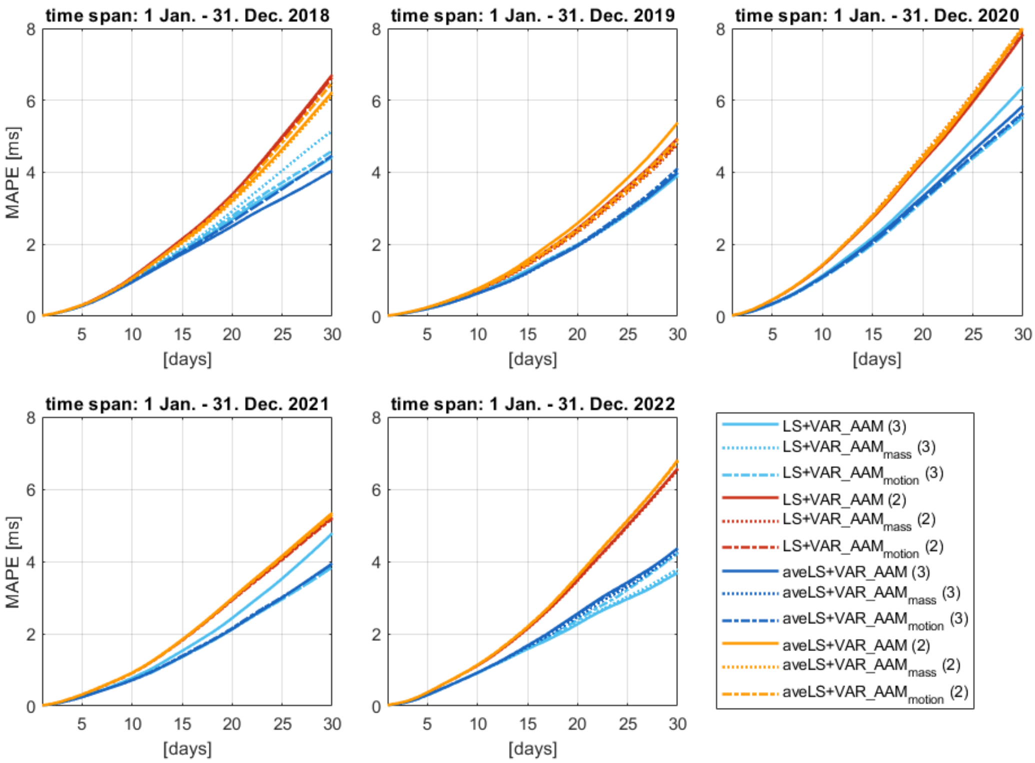


Figure 2. MAPEs of 30-day UT1-UTC prediction with AAM data for the yearly time spans 2018–2022.

Table 2. MAPEs for 30-day UT1-UTC prediction for the yearly time spans 2018-2022 [ms].

Method	Day	2018	2019	2020	2021	2022	Average
DMD$_{nc}$ / DMD	5	0.28 / 0.30	0.32 / 0.29	0.43 / 0.40	0.38 / 0.34	0.38 / 0.34	0.36 / 0.33
LS+VAR$_{nc}$ / LS+VAR		0.45 / 0.30	0.41 / 0.21	0.43 / 0.25	0.41 / 0.30	0.41 / 0.30	0.42 / 0.27
aveDMD$_{nc}$ / aveDMD		0.30 / 0.29	0.30 / 0.29	0.40 / 0.39	0.29 / 0.29	0.36 / 0.35	0.33 / 0.32
aveLS+VAR$_{nc}$ / aveLS+VAR		0.38 / 0.21	0.38 / 0.21	0.43 / 0.35	0.35 / 0.24	0.40 / 0.30	0.39 / 0.26
DMD$_{nc}$ / DMD	10	0.81 / 0.87	0.93 / 0.85	1.23 / 1.13	0.75 / 0.72	1.07 / 0.99	0.96 / 0.91
LS+VAR$_{nc}$ / LS+VAR		1.30 / 0.98	1.27 / 0.68	1.32 / 0.78	1.21 / 0.72	1.24 / 0.91	1.27 / 0.81
aveDMD$_{nc}$ / aveDMD		0.87 / 0.85	0.86 / 0.84	1.16 / 1.08	0.75 / 0.69	1.04 / 0.98	0.94 / 0.89
aveLS+VAR$_{nc}$ / aveLS+VAR		1.34 / 0.96	1.10 / 0.63	1.31 / 1.09	1.18 / 0.71	1.17 / 0.92	1.22 / 0.86
DMD$_{nc}$ / DMD	20	2.04 / 2.28	2.17 / 2.17	3.22 / 2.78	2.01 / 2,00	2.83 / 2.59	2.45 / 2.36
LS+VAR$_{nc}$ / LS+VAR		2.86 / 2.68	3.18 / 2.23	3.37 / 2.47	3.04 / 2.17	2.65 / 2.31	3.02 / 2.37
aveDMD$_{nc}$ / aveDMD		2.24 / 2.23	2.14 / 2.15	2.92 / 2.69	2.01 / 1.93	2.70 / 2.54	2.40 / 2.31
aveLS+VAR$_{nc}$ / aveLS+VAR		2.92 / 2.64	2.50 / 1.93	3.35 / 3.27	2.98 / 2.10	2.62 / 2.46	2.87 / 2.48
DMD$_{nc}$ / DMD	30	3.08 / 3.69	3.31 / 3.66	4.81 / 4.29	3.15 / 3.17	4.77 / 4.49	3.82 / 3.86
LS+VAR$_{nc}$ / LS+VAR		4.03 / 4.50	5.64 / 4.69	5.67 / 4.48	4.90 / 3.93	3.92 / 3.83	4.83 / 4.29
aveDMD$_{nc}$ / aveDMD		3.44 / 3.56	3.49 / 3.68	4.39 / 4.18	3.11 / 3.07	4.59 / 4.38	3.80 / 3.77
aveLS+VAR$_{nc}$ / aveLS+VAR		4.16 / 4.51	4.20 / 3.97	5.61 / 5.60	4.91 / 3.87	3.98 / 4.26	4.57 / 4.44

Table 3. MAPEs for 30-day UT1-UTC prediction with AAM data for the yearly time spans 2018-2022 [ms].

Method	Day	2018	2019	2020	2021	2022	Average
LS+VAR_AAM (3)/ LS+VAR_AAM (2)	5	0.29 / 0.32	0.21 / 0.25	0.35 / 0.46	0.30 / 0.32	0.30 / 0.38	0.29 / 0.34
LS+VAR_AAM$_{mass}$ (3)/ LS+VAR_AAM$_{mass}$ (2)		0.30 / 0.31	0.22 / 0.24	0.34 / 0.45	0.31 / 0.32	0.31 / 0.38	0.30 / 0.34
LS+VAR_AAM$_{motion}$ (3)/ LS+VAR_AAM$_{motion}$ (2)		0.30 / 0.31	0.21 / 0.25	0.33 / 0.45	0.31 / 0.31	0.31 / 0.38	0.29 / 0.34
aveLS+VAR_AAM (3)/ aveLS+VAR_AAM (2)		0.29 / 0.32	0.21 / 0.25	0.35 / 0.46	0.25 / 0.32	0.31 / 0.38	0.28 / 0.35
aveLS+VAR_AAM$_{mass}$ (3)/ aveLS+VAR_AAM$_{mass}$ (2)		0.29 / 0.31	0.21 / 0.25	0.34 / 0.45	0.25 / 0.32	0.30 / 0.38	0.28 / 0.34
aveLS+VAR_AAM$_{motion}$ (3)/ aveLS+VAR_AAM$_{motion}$ (2)		0.29 / 0.31	0.21 / 0.25	0.34 / 0.45	0.25 / 0.32	0.31 / 0.38	0.28 / 0.34
LS+VAR_AAM (3)/ LS+VAR_AAM (2)	10	0.97 / 1.09	0.67 / 0.74	1.13 / 1.40	0.78 / 0.92	0.91 / 1.12	0.89 / 1.05
LS+VAR_AAM$_{mass}$ (3)/ LS+VAR_AAM$_{mass}$ (2)		1.02 / 1.07	0.67 / 0.72	1.07 / 1.43	0.72 / 0.92	0.92 / 1.12	0.88 / 1.05
LS+VAR_AAM$_{motion}$ (3)/ LS+VAR_AAM$_{motion}$ (2)		1.00 / 1.07	0.67 / 0.73	1.06 / 1.40	0.72 / 0.91	0.93 / 1.13	0.88 / 1.05
aveLS+VAR_AAM (3)/ aveLS+VAR_AAM (2)		0.95 / 1.06	0.64 / 0.77	1.10 / 1.42	0.71 / 0.92	0.92 / 1.14	0.87 / 1.06
aveLS+VAR_AAM$_{mass}$ (3)/ aveLS+VAR_AAM$_{mass}$ (2)		0.97 / 1.04	0.64 / 0.73	1.08 / 1.43	0.71 / 0.92	0.92 / 1.13	0.86 / 1.05
aveLS+VAR_AAM$_{motion}$ (3)/ aveLS+VAR_AAM$_{motion}$ (2)		0.97 / 1.06	0.64 / 0.74	1.08 / 1.42	0.72 / 0.91	0.92 / 1.14	0.87 / 1.05
LS+VAR_AAM (3)/ LS+VAR_AAM (2)	20	2.70 / 3.39	1.98 / 2.44	3.50 / 4.30	2.43 / 2.95	2.27 / 3.49	2.58 / 3.31
LS+VAR_AAM$_{mass}$ (3)/ LS+VAR_AAM$_{mass}$ (2)		2.92 / 3.35	2.00 / 2.33	3.21 / 4.48	2.12 / 2.95	2.29 / 3.47	2.51 / 3.32
LS+VAR_AAM$_{motion}$ (3)/ LS+VAR_AAM$_{motion}$ (2)		2.79 / 3.36	2.00 / 2.36	3.19 / 4.41	2.12 / 2.93	2.38 / 3.49	2.50 / 3.31
aveLS+VAR_AAM (3)/ aveLS+VAR_AAM (2)		2.50 / 3.22	1.95 / 2.59	3.32 / 4.37	2.13 / 2.98	2.56 / 3.59	2.49 / 3.35
aveLS+VAR_AAM$_{mass}$ (3)/ aveLS+VAR_AAM$_{mass}$ (2)		2.63 / 3.18	1.95 / 2.35	3.24 / 4.48	2.13 / 2.98	2.46 / 3.58	2.48 / 3.31
aveLS+VAR_AAM$_{motion}$ (3)/ aveLS+VAR_AAM$_{motion}$ (2)		2.63 / 3.30	1.98 / 2.42	3.24 / 4.44	2.15 / 2.96	2.54 / 3.60	2.51 / 3.34
LS+VAR_AAM (3)/ LS+VAR_AAM (2)	30	4.43 / 6.71	3.91 / 4.94	6.38 / 7.84	4.78 / 5.22	3.68 / 6.57	4.64 / 6.26
LS+VAR_AAM$_{mass}$ (3)/ LS+VAR_AAM$_{mass}$ (2)		5.15 / 6.65	4.01 / 4.77	5.55 / 8.03	3.84 / 5.22	3.79 / 6.53	4.47 / 6.24
LS+VAR_AAM$_{motion}$ (3)/ LS+VAR_AAM$_{motion}$ (2)		4.60 / 6.64	3.96 / 4.81	5.53 / 7.96	3.84 / 5.19	4.23 / 6.57	4.43 / 6.23
aveLS+VAR_AAM (3)/ aveLS+VAR_AAM (2)		4.04 / 6.23	4.00 / 5.38	5.85 / 8.01	3.94 / 5.34	4.36 / 6.79	4.44 / 6.35
aveLS+VAR_AAM$_{mass}$ (3)/ aveLS+VAR_AAM$_{mass}$ (2)		4.47 / 6.15	3.99 / 4.82	5.63 / 8.05	3.94 / 5.32	4.26 / 6.77	4.46 / 6.22
aveLS+VAR_AAM$_{motion}$ (3)/ aveLS+VAR_AAM$_{motion}$ (2)		4.47 / 6.49	4.11 / 4.92	5.66 / 8.02	3.93 / 5.29	4.35 / 6.80	4.50 / 6.30

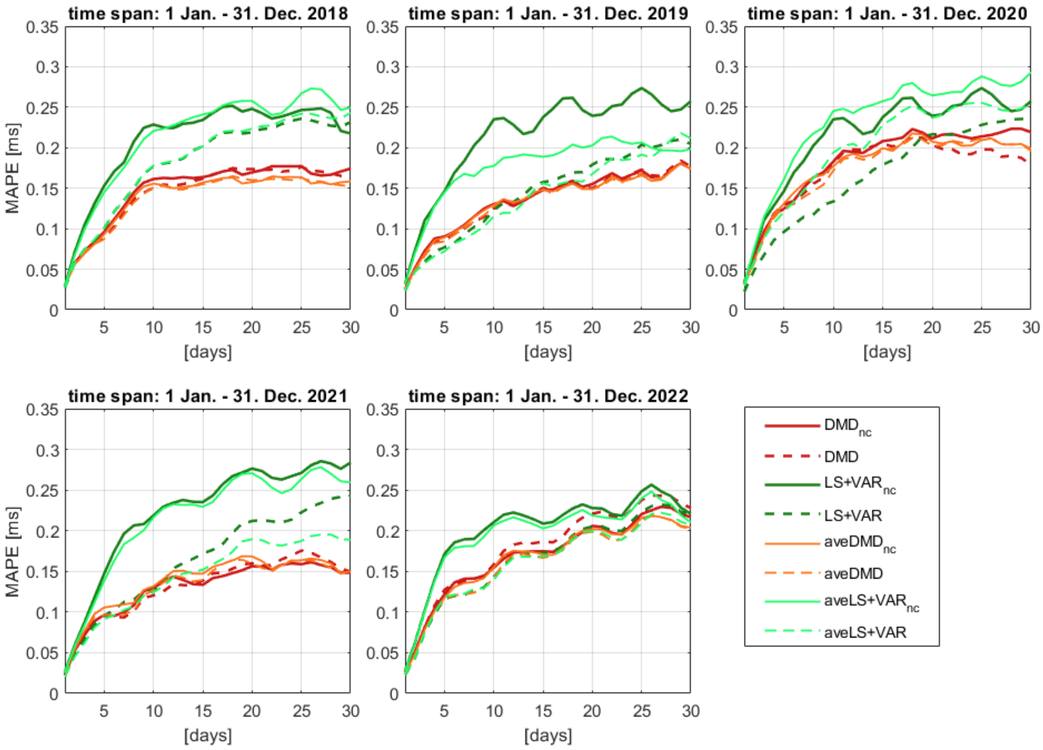


Figure 3. MAPEs of 30-day LOD prediction for the yearly time spans 2018–2022 (nc = not corrected).

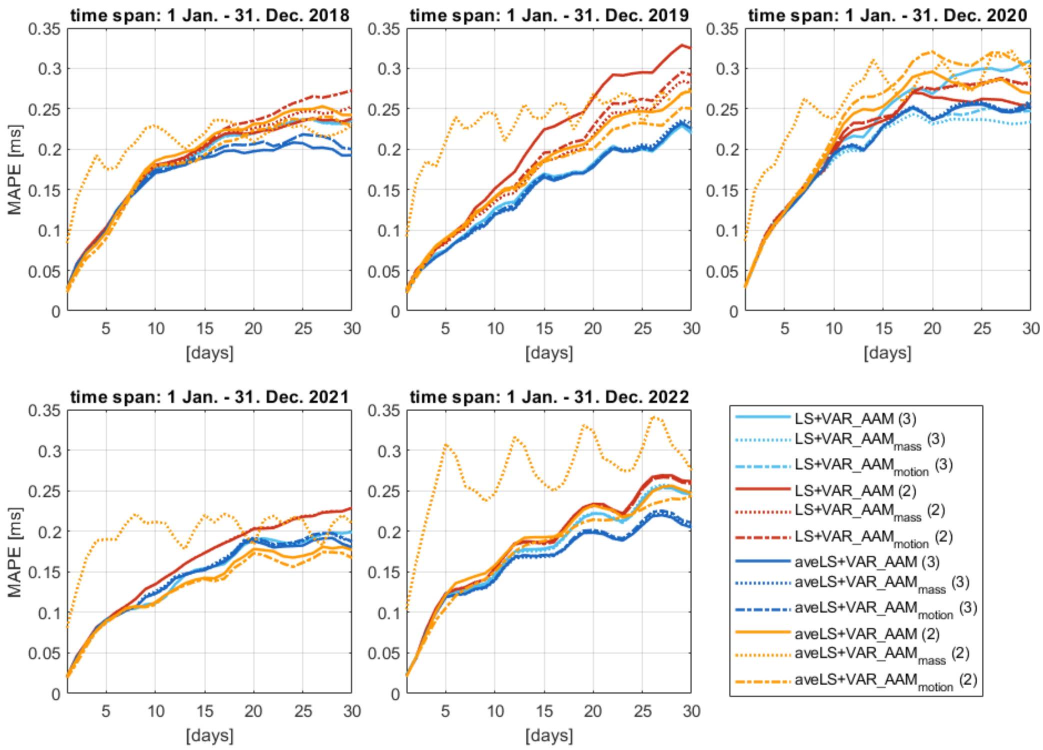


Figure 4. MAPEs of 30-day LOD prediction with AAM data for the yearly time spans 2018–2022.

the case of DMD there were no significant improvements in prediction accuracy after removing tidal effects as compared with the DMD procedure without such a removal. Results indicated a slight decrease of MAPEs in around half of the forecasts using the averaged approach in LOD prediction as compared with to the adaptive approach.

5 Conclusions

This contribution applied Dynamic Mode Decomposition and a combination of least-squares and vector autoregressive models for UT1-UTC and LOD forecast. Prediction accuracy is presented for five yearly time spans from 2018 to 2022. Prediction procedures involved adaptive and averaged modes differing by the use of sets of steering parameters identified at the preliminary stage of fore-

Table 4. MAPEs for 30-day LOD prediction for the yearly time spans 2018-2022 [ms].

Method	Day	2018	2019	2020	2021	2022	Average
DMD_{nc} / DMD	5	0.096 / 0.088	0.090 / 0.086	0.125 / 0.130	0.097 / 0.096	0.122 / 0.127	0.106 / 0.105
$LS+VAR_{nc}$ / LS+VAR		0.153 / 0.101	0.167 / 0.077	0.146 / 0.096	0.148 / 0.095	0.172 / 0.116	0.157 / 0.097
$aveDMD_{nc}$ / aveDMD		0.093 / 0.088	0.089 / 0.084	0.131 / 0.127	0.106 / 0.097	0.118 / 0.115	0.107 / 0.102
$aveLS+VAR_{nc}$ / aveLS+VAR		0.145 / 0.101	0.145 / 0.072	0.161 / 0.121	0.139 / 0.091	0.167 / 0.119	0.151 / 0.101
DMD_{nc} / DMD	10	0.161 / 0.151	0.130 / 0.125	0.183 / 0.181	0.130 / 0.120	0.155 / 0.159	0.152 / 0.147
$LS+VAR_{nc}$ / LS+VAR		0.229 / 0.178	0.197 / 0.123	0.235 / 0.134	0.219 / 0.127	0.210 / 0.143	0.218 / 0.141
$aveDMD_{nc}$ / aveDMD		0.156 / 0.151	0.129 / 0.127	0.177 / 0.172	0.132 / 0.127	0.155 / 0.142	0.150 / 0.144
$aveLS+VAR_{nc}$ / aveLS+VAR		0.221 / 0.178	0.177 / 0.115	0.245 / 0.194	0.217 / 0.125	0.207 / 0.139	0.213 / 0.150
DMD_{nc} / DMD	20	0.170 / 0.174	0.155 / 0.150	0.212 / 0.203	0.156 / 0.160	0.206 / 0.221	0.180 / 0.182
$LS+VAR_{nc}$ / LS+VAR		0.248 / 0.219	0.204 / 0.179	0.239 / 0.217	0.276 / 0.212	0.228 / 0.205	0.239 / 0.206
$aveDMD_{nc}$ / aveDMD		0.158 / 0.160	0.149 / 0.149	0.208 / 0.203	0.168 / 0.160	0.202 / 0.199	0.177 / 0.174
$aveLS+VAR_{nc}$ / aveLS+VAR		0.258 / 0.223	0.203 / 0.168	0.264 / 0.237	0.271 / 0.190	0.219 / 0.199	0.243 / 0.203
DMD_{nc} / DMD	30	0.174 / 0.164	0.174 / 0.177	0.219 / 0.178	0.148 / 0.150	0.217 / 0.228	0.186 / 0.179
$LS+VAR_{nc}$ / LS+VAR		0.217 / 0.231	0.219 / 0.204	0.257 / 0.236	0.284 / 0.243	0.221 / 0.219	0.240 / 0.227
$aveDMD_{nc}$ / aveDMD		0.158 / 0.153	0.174 / 0.174	0.196 / 0.199	0.150 / 0.151	0.204 / 0.202	0.176 / 0.176
$aveLS+VAR_{nc}$ / aveLS+VAR		0.250 / 0.243	0.198 / 0.211	0.293 / 0.253	0.259 / 0.189	0.210 / 0.208	0.242 / 0.221

cast. In addition, external data, Atmospheric Angular Momentum were included in the LS+VAR-based prediction procedure. Predictions were also performed with and without corrections due to tidal effects.

Results indicate that there is no single most accurate variant for all 30 days UT1-UTC predictions. For years 2018, 2019 and 2021 DMD_{nc} offered the lowest values of MAPEs, whilst for 2020 aveDMD and in 2022 LS+VAR AAM (3) obtained the highest accuracy. The results show that aveDMD is the most accurate method in almost all time periods for the 30th day of LOD prediction with the MAPE in the range of 0.158 – 0.204 ms.

Interestingly, the inclusion of AAM time series in the LS+VAR prediction procedure did not visibly increase the accuracy of the UT1-UTC or LOD forecasts. The results showed that incorporating AAM data into the LS+VAR prediction (comparing all variants of LS+VAR AAM and aveLS+VAR AAM with corresponding LS+VAR and aveLS+VAR) improved the UT1-UTC prediction accuracy in only 20 out of 120 cases and – in the case of LOD – in 26 out of 120 cases. The increase in prediction accuracy, if it occurs, in the case of UT1-UTC is the most evident for the 30th day of prediction and in the case of LOD – for ultra short-term forecast (up to 5 – 10 days).

In more than half of the cases, there was a slight improvement in UT1-UTC and LOD forecasts using the averaged approach comparing to the adaptive one. The results indicate that averaged approach , and not the adaptive one (for DMD and LS+VAR predictions together), improves UT1-UTC predictions (in 54 out of 80 cases) marginally more often than LOD predictions (in 47 out of 80 cases). It can be also noticed that the averaged approach, compared to the adaptive approach, improves DMD prediction 53 times (of UT1-UTC and LOD predictions together), while LS+VAR prediction – 48 times. In the case of UT1-UTC LS+VAR AAM prediction, in both approaches (using 2 and 3 input time series), there was no improvement using the averaged approach comparing to the adaptive approach. On the other hand, in LOD predictions when 2 input time series were involved in a LS+VAR prediction procedure there were reductions in forecast errors (in favour of the averaged approach over the adaptive approach) only in 18 out of 75 cases, while when using 3 input time series these reductions occurred in 11 out of 75 cases.

The results also indicate that the use of 3 variables (UT1-UTC, LOD and AAM component) in LS+VAR AAM prediction procedure is characterized by lower numbers of prediction errors than using only 2 variables (UT1-UTC/LOD and AAM component). The prediction accuracy of all variants of LS+VAR AAM (3) procedure rather than the corresponding variants of LS+VAR AAM (2), in case of UT1-UTC, increased in all cases with an average improvement of around 20.9% and in the case of LOD – in 106 out of 120 cases with improvement of around 13.8%. It can therefore be concluded that, in these cases, including an additional correlated variable in the LS+VAR prediction procedure significantly improves the accuracy of the forecast.

There was a marginal decrease of MAPEs in DMD-based predictions of UT1-UTC and LOD when tidal effects were removed comparing to the procedures without removing these effects, whilst it is much clearer in the LS+VAR-based forecast. These improvements in the 30th day of DMD and LS+VAR predictions of UT1-UTC vary from 0.04 ms to 0.52 ms and from 0.01 ms to 1.19 ms, respectively. The corresponding values for the LOD forecast are in range from 0.002 ms to 0.041 ms (in the case of DMD) and from 0.002 ms to 0.070 ms (in the case of LS+VAR).

A potential explanation for the greater impact of removing (or not removing) tidal oscillations on the LS+VAR-based forecast than on the DMD is that LS+VAR is based on the traditional approach of first removing the trend and periodic components and then applying a VAR model to the residuals. In the procedure without removing the tidal oscillations, they remain as the residuals and therefore directly affect the forecast results. DMD, on the other hand, is a data-driven method which does not rely on any prior assumptions beyond the inherent dynamics observed over time. DMD was designed to search for patterns in trends and frequencies, and their evolution in time, hence, it is very likely that it identified some of the tidal frequencies, and the remaining ones were replaced with some artificial frequencies found in the series on which the DMD model was built.

It is also worth noting that the accuracy of both UT1-UTC and LOD predictions for 2020 and 2022 decreased in almost all cases. The large values of MAPE may be caused by, among other things, such phenomena as El Niño and La Niña occurring at the same time (Xu et al. 2022a; https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php).

Table 5. MAPEs for 30-day LOD prediction with AAM data for the yearly time spans 2018-2022 [ms].

Method	Day	2018	2019	2020	2021	2022	Average
LS+VAR_AAM (3)/ LS+VAR_AAM (2)	5	0.100 / 0.104	0.075 / 0.086	0.125 / 0.120	0.088 / 0.091	0.122 / 0.123	0.102 / 0.105
LS+VAR_AAM$_{mass}$ (3)/ LS+VAR_AAM$_{mass}$ (2)		0.100 / 0.101	0.075 / 0.083	0.120 / 0.125	0.088 / 0.090	0.123 / 0.123	0.101 / 0.104
LS+VAR_AAM$_{motion}$ (3)/ LS+VAR_AAM$_{motion}$ (2)		0.100 / 0.101	0.075 / 0.085	0.120 / 0.125	0.088 / 0.091	0.124 / 0.123	0.101 / 0.105
aveLS+VAR_AAM (3)/ aveLS+VAR_AAM (2)		0.099 / 0.097	0.074 / 0.090	0.121 / 0.122	0.091 / 0.090	0.119 / 0.122	0.101 / 0.104
aveLS+VAR_AAM$_{mass}$ (3)/ aveLS+VAR_AAM$_{mass}$ (2)		0.101 / 0.175	0.074 / 0.210	0.121 / 0.209	0.092 / 0.190	0.119 / 0.308	0.101 / 0.219
aveLS+VAR_AAM$_{motion}$ (3)/ aveLS+VAR_AAM$_{motion}$ (2)		0.101 / 0.090	0.074 / 0.088	0.121 / 0.123	0.091 / 0.087	0.119 / 0.105	0.101 / 0.099
LS+VAR_AAM (3)/ LS+VAR_AAM (2)	10	0.176 / 0.181	0.126 / 0.151	0.205 / 0.196	0.110 / 0.135	0.144 / 0.154	0.152 / 0.163
LS+VAR_AAM$_{mass}$ (3)/ LS+VAR_AAM$_{mass}$ (2)		0.175 / 0.179	0.125 / 0.135	0.187 / 0.201	0.111 / 0.135	0.145 / 0.154	0.149 / 0.161
LS+VAR_AAM$_{motion}$ (3)/ LS+VAR_AAM$_{motion}$ (2)		0.176 / 0.180	0.126 / 0.140	0.191 / 0.201	0.112 / 0.135	0.147 / 0.154	0.150 / 0.162
aveLS+VAR_AAM (3)/ aveLS+VAR_AAM (2)		0.171 / 0.185	0.119 / 0.141	0.194 / 0.208	0.124 / 0.113	0.140 / 0.158	0.149 / 0.161
aveLS+VAR_AAM$_{mass}$ (3)/ aveLS+VAR_AAM$_{mass}$ (2)		0.174 / 0.229	0.119 / 0.244	0.195 / 0.238	0.126 / 0.212	0.139 / 0.247	0.151 / 0.234
aveLS+VAR_AAM$_{motion}$ (3)/ aveLS+VAR_AAM$_{motion}$ (2)		0.174 / 0.180	0.121 / 0.139	0.195 / 0.216	0.123 / 0.110	0.141 / 0.150	0.151 / 0.159
LS+VAR_AAM (3)/ LS+VAR_AAM (2)	20	0.221 / 0.220	0.181 / 0.263	0.269 / 0.264	0.190 / 0.203	0.224 / 0.234	0.217 / 0.237
LS+VAR_AAM$_{mass}$ (3)/ LS+VAR_AAM$_{mass}$ (2)		0.219 / 0.226	0.180 / 0.217	0.231 / 0.274	0.189 / 0.204	0.221 / 0.231	0.208 / 0.230
LS+VAR_AAM$_{motion}$ (3)/ LS+VAR_AAM$_{motion}$ (2)		0.220 / 0.235	0.182 / 0.228	0.237 / 0.274	0.190 / 0.203	0.222 / 0.231	0.210 / 0.234
aveLS+VAR_AAM (3)/ aveLS+VAR_AAM (2)		0.197 / 0.231	0.178 / 0.215	0.237 / 0.296	0.188 / 0.178	0.198 / 0.232	0.200 / 0.230
aveLS+VAR_AAM$_{mass}$ (3)/ aveLS+VAR_AAM$_{mass}$ (2)		0.206 / 0.209	0.179 / 0.268	0.238 / 0.295	0.192 / 0.184	0.199 / 0.323	0.203 / 0.256
aveLS+VAR_AAM$_{motion}$ (3)/ aveLS+VAR_AAM$_{motion}$ (2)		0.206 / 0.224	0.180 / 0.200	0.236 / 0.321	0.191 / 0.173	0.201 / 0.214	0.203 / 0.226
LS+VAR_AAM (3)/ LS+VAR_AAM (2)	30	0.235 / 0.238	0.220 / 0.325	0.310 / 0.255	0.199 / 0.229	0.246 / 0.261	0.242 / 0.261
LS+VAR_AAM$_{mass}$ (3)/ LS+VAR_AAM$_{mass}$ (2)		0.233 / 0.253	0.221 / 0.281	0.234 / 0.280	0.197 / 0.229	0.249 / 0.259	0.227 / 0.260
LS+VAR_AAM$_{motion}$ (3)/ LS+VAR_AAM$_{motion}$ (2)		0.234 / 0.273	0.220 / 0.292	0.247 / 0.282	0.199 / 0.229	0.246 / 0.259	0.229 / 0.267
aveLS+VAR_AAM (3)/ aveLS+VAR_AAM (2)		0.193 / 0.242	0.225 / 0.272	0.253 / 0.269	0.180 / 0.177	0.205 / 0.247	0.211 / 0.241
aveLS+VAR_AAM$_{mass}$ (3)/ aveLS+VAR_AAM$_{mass}$ (2)		0.200 / 0.230	0.232 / 0.276	0.260 / 0.287	0.188 / 0.210	0.208 / 0.276	0.218 / 0.256
aveLS+VAR_AAM$_{motion}$ (3)/ aveLS+VAR_AAM$_{motion}$ (2)		0.200 / 0.230	0.225 / 0.250	0.257 / 0.302	0.187 / 0.167	0.210 / 0.244	0.216 / 0.239

Acknowledgements

Research project supported by program „Excellence initiative – research university" for the AGH University of Krakow.

The paper is a result of research on geospatial methods carried out within the statutory research grant no. 16.16.150.545 at the Department of Integrated Geodesy and Cartography, AGH University of Krakow.

We gratefully acknowledge Poland's high-performance computing infrastructure PLGrid (HPC Centers: ACK Cyfronet AGH) for providing computer facilities and support within computational grant Nos. PLG/2022/015968 and PLG/2023/016281.

Intermediate results are available on the website accompanying this article: home.agh.edu.pl/mmichalc.

References

Bizouard, C., Lambert, S., Gattano, C., Becker, O., and Richard, J.-Y. (2018). The IERS EOP 14C04 solution for earth orientation parameters consistent with ITRF 2014. *Journal of Geodesy*, 93(5):621–633, doi:10.1007/s00190-018-1186-3.

Dick, W. R. and Thaller, D. (2020). IERS Annual Report 2018. Technical report, International Earth Rotation and Reference Systems Service, Central Bureau.

Dill, R., Dobslaw, H., and Thomas, M. (2018). Improved 90-day Earth orientation predictions from angular momentum forecasts of atmosphere, ocean, and terrestrial hydrosphere. *Journal of Geodesy*, 93(3):287–295, doi:10.1007/s00190-018-1158-7.

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

Gambis, D. and Luzum, B. (2011). Earth rotation monitoring, UT1 determination and prediction. *Metrologia*, 48(4):S165–S170, doi:10.1088/0026-1394/48/4/s06.

Gross, R. S., Eubanks, T., Steppe, J., Freedman, A., Dickey, J., and Runge, T. (1998). A Kalman-filter-based approach to combining independent Earth-orientation series. *Journal of Geodesy*, 72:215–235, doi:10.1007/s001900050162.

Gross, R. S., Fukumori, I., Menemenlis, D., and Gegout, P. (2004). Atmospheric and oceanic excitation of length-of-day variations during 1980–2000. *Journal of Geophysical Research: Solid Earth*, 109(B1), doi:10.1029/2003jb002432.

Guessoum, S., Belda, S., Ferrandiz, J. M., Modiri, S., Raut, S., Dhar, S., Heinkelmann, R., and Schuh, H. (2022). The short-term prediction of Length of Day using 1D convolutional neural networks (1D CNN). *Sensors*, 22(23):9517, doi:10.3390/s22239517.

Holton, J. R. and Dmowska, R. (1989). *El Niño, La Niña, and the southern oscillation*. Academic press, Cambridge, MA, USA.

Höpfner, J. (1998). Seasonal variations in length of day and atmospheric angular momentum. *Geophysical Journal International*, 135(2):407–437, doi:10.1046/j.1365-246X.1998.00648.x.

Kalarus, M., Schuh, H., Kosek, W., Akyilmaz, O., Bizouard, C., Gambis, D., Gross, R., Jovanović, B., Kumakshev, S., Kutterer, H., Mendes Cerveira, P. J., Pasynok, S., and Zotov, L. (2010). Achievements of the Earth orientation parameters prediction comparison campaign. *Journal of Geodesy*, 84(10):587–596, doi:10.1007/s00190-010-0387-1.

Kiani Shahvandi, M., Schartner, M., and Soja, B. (2022). Neural ODE differential learning and its application in polar motion prediction. *Journal of Geophysical Research: Solid Earth*, 127(11), doi:10.1029/2022jb024775.

Kur, T., Dobslaw, H., Śliwińska, J., Nastula, J., Wińska, M., and 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), doi:10.1186/s40623-022-01753-9.

Lei, Y., Cai, H., and Zhao, D. (2017). Improvement of the prediction accuracy of polar motion using empirical mode decomposition. *Geodesy and Geodynamics*, 8(2):141–146, doi:10.1016/j.geog.2016.09.007.

Liao, D., Wang, Q., Zhou, Y., Liao, X., and Huang, C. (2012). Long-term prediction of the Earth Orientation Parameters by the artificial neural network technique. *Journal of Geodynamics*, 62:87–92, doi:10.1016/j.jog.2011.12.004.

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

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

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

Nastula, J., Chin, T. M., Gross, R., Śliwińska, J., and Wińska, M. (2020). Smoothing and predicting celestial pole offsets using a Kalman filter and smoother. *Journal of Geodesy*, 94(3), doi:10.1007/s00190-020-01349-9.

Niedzielski, T. and Kosek, W. (2011). *Prediction Analysis of UT1-UTC Time Series by Combination of the Least-Squares and Multivariate Autoregressive Method*, pages 153–157. Springer Berlin Heidelberg, doi:10.1007/978-3-642-22078-4_23.

Okhotnikov, G. and Golyandina, N. (2019). EOP time series prediction using singular spectrum analysis. In Corpetti, T., Ienco, D., and Interdonato, R., editors, *Proceedings of MACLEAN: MAChine learning for Earth observation workshop, RWTH Aahen University, CEUR Workshop Proceedings*.

Petit, G. and Luzum, B. (2010). Iers conventions. Technical report, IERS Technical Note 36, Verlag des Bundesamts für Kartographie und Geodäsie Frankfurt am Main, Germany.

Schmid, P. J. (2010). Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656:5–28, doi:10.1017/s0022112010001217.

Schuh, H., Ulrich, M., Egger, D., Müller, J., and Schwegmann, W. (2002). Prediction of Earth orientation parameters by artificial neural networks. *Journal of Geodesy*, 76(5):247–258, doi:10.1007/s00190-001-0242-5.

Soffel, M. (2013). *Space-Time Reference Systems*. SpringerLink. Springer, Berlin. Description based upon print version of record.

Tirunagari, S., Kouchaki, S., Poh, N., Bober, M., and Windridge, D. (2017). Dynamic mode decomposition for univariate time series: analysing trends and forecasting. *hal-01463744f*.

Tu, J. H., Rowley, W. C., Luchtenburg, D. M., Brunton, S. L., and Kutz, J. N. (2014). On dynamic mode decomposition: Theory and applications. *Journal of Computational Dynamics*, 1(2):391–421, doi:10.3934/jcd.2014.1.391.

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

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

Xu, X.-Q., Zhou, Y.-H., Duan, P.-S., Fang, M., Kong, Z.-Y., Xu, C.-C., and An, X.-R. (2022b). Contributions of oceanic and continental AAM to interannual variation in ΔLOD with the detection of 2020–2021 La Nina event. *Journal of Geodesy*, 96(6), doi:10.1007/s00190-022-01632-x.

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

Marcin Ligas* and Maciej Michalczak

Dynamic mode decomposition and bivariate autoregressive short-term prediction of Earth rotation parameters

https://doi.org/10.1515/jag-2023-0030
Received April 23, 2023; accepted August 11, 2023;
published online September 5, 2023

Abstract: In this contribution two new approaches are applied to predict polar motion and length-of-day. The first one is based on Dynamic Mode Decomposition (DMD), that is purely data-driven and is capable of reconstructing and forecasting time series in one numerical procedure. The other one is based on a vector autoregression of order p – VAR(p), which is a vector counterpart of AR(p) that accounts for an evolution of variables in time and a coevolution with other variables. DMD was applied to polar motion and length-of-day whilst VAR(p) to a joint prediction of polar motion. A prediction experiment concerned 30-day forecast horizon with a 7-day shift. It was performed separately for years 2017–2022 giving 48 predictions within each year. This study uses IERS EOP 14 C04 (IAU2000) as a reference for all computations and a mean absolute prediction error (MAPE) as a measure of prediction quality. For DMD, MAPEs for x coordinate of the pole vary from 0.22–0.30 mas for the 1st day and 6.64–8.56 mas for the 30th day of prediction depending on the year whilst those values vary from 0.20–0.27 mas and 5.27–7.66 mas for VAR(p) based prediction. Corresponding values for y coordinate of the pole vary from 0.15–0.23 mas and 4.27–5.93 mas for DMD, whilst 0.13–0.21 mas and 3.46–3.82 mas for VAR(p). In case of LOD forecast, MAPEs vary from 0.023–0.031 ms for the 1st day and 0.142–0.205 ms for the 30th day depending on the year.

Keywords: polar motion; length of day; Earth rotation; dynamic mode decomposition; autoregression; prediction

***Corresponding author: Marcin Ligas**, Department of Integrated Geodesy and Cartography, Faculty of Geo-Data Science, Geodesy, and Environmental Engineering, AGH University of Krakow, Adama Mickiewicza 30, 30-059 Krakow, Poland, E-mail: ligas@agh.edu.pl. https://orcid.org/0000-0002-0970-128X
Maciej Michalczak, Department of Integrated Geodesy and Cartography, Faculty of Geo-Data Science, Geodesy, and Environmental Engineering, AGH University of Krakow, Adama Mickiewicza 30, 30-059 Krakow, Poland, E-mail: mmichalc@agh.edu.pl. https://orcid.org/0000-0001-6859-988X

1 Introduction

Earth Orientation Parameters (EOP) are the parameters necessary to perform the transformation between celestial reference frame (CRF) and terrestrial reference frame (TRF) [1]. For many space and geodetic purposes, such as real-time spacecraft navigation, satellite-based positioning on Earth, time-keeping or positional astronomy, the knowledge of real-time and future EOP data is necessary. Space geodetic techniques such as Very Long Baseline Interferometry (VLBI), Global Navigation Satellite System (GNSS), Lunar Laser Ranging (LLR), Satellite Laser Ranging (SLR), and Doppler Orbitography and Radio Positioning Integrated by Satellites (DORIS) can be used to determine EOP. Combination of these techniques enables to determine EOP with the accuracy of app. 50 µas or better in case of terrestrial pole coordinates, 30 µs or better in case of UT1-UTC and 10 µs in case of LOD [2]. However, due to the complexity of measuring techniques and data processing, EOP data, provided for example by the International Earth Rotation and Reference Systems Service (IERS), are published with 30 days latency. The solution of a problem with the availability of EOP data in real time are rapid products published by e.g., IERS (Daily Rapid EOP Data), CODE (Center for Orbit Determination in Europe), GFZ (German Research Centre for Geosciences) or IGS (International GNSS Service). To obtain future values of EOPs, a prediction based on rapid data is carried out.

The Earth Rotation Parameters (ERP) are the subset of EOP and consist of the pole coordinates (x, y) and the difference of times dUT1 = UT1 − UTC (or alternatively its first negative time derivative – length-of-day, LOD). The variability of x and y, i.e., polar motion describes fluctuations of the Celestial Intermediate Pole (CIP) against the terrestrial frame whilst the difference between rotational time and atomic time, dUT1 or LOD describe time-dependent irregularities of the Earth's rotation.

Changes of pole coordinates and LOD are caused, among other things, by the gravitational interaction between the Earth and other celestial bodies (primarily the Sun and the Moon) [3], and terrestrial processes associated with the redistribution of geophysical fluid masses near the surface [4]. Moreover, the very important role in LOD changes plays ENSO (El Nino Southern Oscillation), which makes this parameter the most challenging to predict [5, 6].

Over the years, researchers have introduced a great number of EOP prediction methods, i.e.,: auto-covariance model [7], combination of least-squares extrapolation and autoregressive models [8], Kalman filtering [9, 10], wavelet decomposition [11], fuzzy inference systems [12], copula-based analysis [13] and 1D convolutional neural networks [14]. In order to examine the accuracy of EOP prediction methods IERS has organised two comparison campaigns, one in 2006 [8] and recently in 2021 (http://eoppcc.cbk.waw.pl/).

In this contribution, to predict the ERP we used Dynamic Mode Decomposition (DMD) [15] and vector autoregression (VAR) with prior trend and periodicity removal. DMD is a purely data-driven technique with the primary goal to extract typical behavior of a dynamical system via modal analysis and once this is established the DMD is capable of reconstructing and forecasting time series in one numerical procedure. Vector autoregression of order p – VAR(p) preceded by a trend and periodic components removal (restored at the end of prediction procedure) is the second technique used in this contribution. This technique is a vector counterpart of AR(p) that accounts for an evolution of variables in time and a coevolution with other variables. Joint prediction of related variables should potentially contribute to better forecast results due to the use of additional information stored in jointly analyzed variables. In order to compare the performance of newly adapted methods for ERP prediction we have also used: ordinary kriging [16, 17], ARIMA [18] and the most accurate methods that took part in the 1st Earth Orientation Parameters Prediction Comparison Campaign [8]. The methods were applied to a short-term forecast horizon, i.e., 30 days into the future. Predictions were performed and analyzed in subsequent years 2017–2022 with 7-day shift between new predictions, giving 48 30-day predictions within each year. The results are juxtaposed with the 1st EOPPCC not for the purpose of categorical comparison but to present achievable forecast accuracy provided by the BiVAR and DMD methods. The reference to the 1st EOPPCC is only illustrative due to different conditions of research experiments (different ITRFs, time spans, operational settings). The results are presented over several consecutive years to show the relative stability (repeatability) of the solutions with the adopted parameters' sets. However, experience shows that the parameters controlling the methods need to be periodically updated to maintain a stable level of forecast accuracy.

2 Dynamic mode decomposition

Despite the fact that Dynamic Mode Decomposition [15] is a relatively young technique, it has already gained recognition and popularity in many fields of science, e.g., financial market analysis and forecasting, computer vision, neuroscience, epidemiology [19] and naturally in fluid dynamics where it was developed. In this section a Dynamic Mode Decomposition (DMD) algorithm used to predict polar motion (x, y) and length-of-day (LOD) is introduced. DMD is a data-driven, assumption free method of reconstructing and forecasting dynamical systems. The idea behind the DMD infrastructure is that the future states of a system may be inferred from past ones by means of an (unknown) operator $\mathbf{A}$ storing the dynamics of the system's evolution. This transition between states may be expressed by a simple relation:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k \tag{1}$$

This, in fact, has the form of a linear homogeneous system of difference equations with constant coefficients with some initial vector $\mathbf{x}_1$ which has the solution given by:

$$\mathbf{x}_{k+1} = \mathbf{A}^k\mathbf{x}_1 = \mathbf{\Phi}\mathbf{\Lambda}^k\mathbf{\Phi}^{-1}\mathbf{x}_1 \tag{2}$$

This may be deduced from a sequence of substitutions:

$$\begin{aligned}&\left[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_k, \mathbf{x}_{k+1}\right]\\ &\quad = \left[\mathbf{A}^0\mathbf{x}_1, \mathbf{A}^1\mathbf{x}_1, \mathbf{A}^2\mathbf{x}_1, \ldots, \mathbf{A}^{k-1}\mathbf{x}_1, \mathbf{A}^k\mathbf{x}_1\right]\end{aligned} \tag{3}$$

and an eigen-decomposition of $\mathbf{A}$:

$$\mathbf{A} = \mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^{-1} \tag{4}$$

DMD algorithm searches the best fit matrix $\mathbf{A}$ in (1) or to be more specific its dominant eigen-structure (eigenvectors/modes and eigenvalues). This is performed by constructing a snapshot matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_m]$ storing all registered (or sampled) states of a system at discrete time instances $t = 1, 2, \ldots, m$. The snapshot matrix $\mathbf{X}$ is then split into two other matrices $\mathbf{X}_1 = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{m-1}]$ and $\mathbf{X}_2 = [\mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_m]$ that are related according to (1) via matrix $\mathbf{A}$ i.e.,:

$$\mathbf{X}_2 = \mathbf{A}\mathbf{X}_1 \tag{5}$$

The estimate of an operator $\mathbf{A}$ is given by:

$$\mathbf{A} = \arg\min_{\mathbf{A}} ||\mathbf{X}_2 - \mathbf{A}\mathbf{X}_1||_F^2 = \mathbf{X}_2\mathbf{X}_1^+ \tag{6}$$

where $\mathbf{X}_1^+$ denotes the Moore–Penrose pseudo-inverse of $\mathbf{X}_1$ and is determined through the singular value decomposition (SVD).

$$\mathbf{X}_1 = \mathbf{U\Sigma V}^T \rightarrow \mathbf{X}_1^+ = \mathbf{V\Sigma}^{-1}\mathbf{U}^T \tag{7}$$

This enables for a low dimensional representation of an original operator $\mathbf{A}$ given by $\tilde{\boldsymbol{A}}$:

$$\tilde{\boldsymbol{A}} = \mathbf{U}^T\mathbf{AU} = \mathbf{U}^T\mathbf{X}_2\mathbf{V\Sigma}^{-1} \tag{8}$$

Tu et al. [20] proved that eigenvalues of $\tilde{\boldsymbol{A}}$ are those of the original matrix $\mathbf{A}$ and result from an eigenvalue problem:

$$\tilde{\boldsymbol{A}}\boldsymbol{W} = \mathbf{W\Lambda} \tag{9}$$

whilst the eigenvectors of $\mathbf{A}$ are given by:

$$\mathbf{\Phi} = \mathbf{X}_2\mathbf{V\Sigma}^{-1}\mathbf{W} \tag{10}$$

Since the DMD was developed in the fluid dynamics where the number of elements per snapshot may be in millions the SVD of $\mathbf{X}_1$ is truncated to some prescribed rank r (low rank approximation of $\mathbf{X}_1$) yielding:

$$\mathbf{X}_1 \cong \mathbf{U}_r\mathbf{\Sigma}_r\mathbf{V}_r^T \tag{11}$$

The remaining steps of the algorithm are the same as before with $\mathbf{U}$, Σ, $\mathbf{V}$ replaced with $\mathbf{U}_r$, Σ_r, $\mathbf{V}_r$. Once the eigenstructure of $\mathbf{A}$ is determined one may use (2) to reconstruct or predict future states of the system. In fact, for low dimensional problems (numerically tractable) the operator $\mathbf{A}$ may be obtained from (6) and use in either (1) or (2).

Since DMD was developed for spatio–temporal processes it requires a slight reformulation in terms of snapshot matrices to be applied to time series analysis and forecast. To do so, the input time series $\mathbf{x}$ of length T is split into L subseries $\mathbf{x}'$ of length K that are shifted by one time step ahead, this gives a trajectory matrix of the form (embedding step in singular spectrum analysis, SSA) [21].

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & \dots & x_L \\ x_2 & x_3 & \dots & x_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_K & x_{K+1} & \dots & x_{K+L-1} \end{bmatrix} = \begin{bmatrix} \mathbf{x}'_1 & \mathbf{x}'_2 & \dots & \mathbf{x}'_L \end{bmatrix} \tag{12}$$

where L is selected so that $2 \leq L \leq \frac{T}{2}$, K is related to L and T with $K = T - L + 1$.

The trajectory matrix is then split into two $K \times (L-1)$ matrices (as described above):

$$\mathbf{X}_1 = \begin{bmatrix} \mathbf{x}'_1 & \mathbf{x}'_2 & \dots & \mathbf{x}'_{L-1} \end{bmatrix} \tag{13}$$

$$\mathbf{X}_2 = \begin{bmatrix} \mathbf{x}'_2 & \mathbf{x}'_3 & \dots & \mathbf{x}'_L \end{bmatrix} \tag{14}$$

The remaining steps of the DMD algorithm do not change.

3 Bivariate autoregression with prior trend/periodicity removal

This section describes a three-step prediction procedure of the polar motion that consists of detrending, predicting and extrapolating stages. In a sense, it is similar to Remove-Compute-Restore technique in geoid modeling problem. The procedure starts with a least-squares fit of linear trend and periodic components (15) to the original time series (separately for x and y).

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} A^x + B^x t + \sum_i^k a_i^x \sin\left(\omega_i^x t + \phi_i^x\right) + \mathrm{e}^x \\ A^y + B^y t + \sum_i^k a_i^y \sin\left(\omega_i^y t + \phi_i^y\right) + \mathrm{e}^y \end{bmatrix} \tag{15}$$

where A, B are the intercept and slope of a linear trend, whereas a, ω, ϕ are the amplitude, frequency and phase of a sine wave, e^x and e^y stand for error terms, all quantities to be estimated.

The removal of mentioned components (detrending) generates least-squares residuals $x^{\mathrm{res}} = \hat{\mathrm{e}}^x$, $y^{\mathrm{res}} = \hat{\mathrm{e}}^y$ that are subject to a joint prediction process through p – lagged vector autoregressive model VAR(p). VAR(p) is a multivariate generalization of univariate AR(p) that describes dynamic dependence between variables. The model links current values of variables with their past values and past values of other variables in the system. Since, the prediction procedure uses a bivariate case of VAR(p) model the following denotation will apply to this case and will be called BiVAR(p) when necessary. The BiVAR(p) model may be expressed as:

$$\mathbf{x}'_T = \boldsymbol{c} + \mathbf{A}'_p\mathbf{x}'_{T-p} + \mathbf{A}'_{p-1}\mathbf{x}'_{T-p+1} + \cdots + \mathbf{A}'_1\mathbf{x}'_{T-1} + \mathbf{e}'_T \tag{16}$$

where a constant vector $\mathbf{c}$ (2×1) accounts for a non-zero mean value of the bivariate process, here neglected due to the detrending step, each of $\mathbf{x}'$ (2×1) is given as $\mathbf{x}' = \begin{bmatrix} x^{\mathrm{res}} = \hat{\mathrm{e}}^x \\ y^{\mathrm{res}} = \hat{\mathrm{e}}^y \end{bmatrix}$ at appropriate time delay, each $\mathbf{A}'$ is a 2×2 matrix of coefficients and $\mathbf{e}'$ (2×1) is a vector error term.

The goal now is to estimate the unknown coefficients matrices $\mathbf{A}'$ relating variables and their past states contained in delayed vectors $\mathbf{x}'$. This can be done by stacking vectors $\mathbf{x}'$ and matrices $\mathbf{A}'$ and combining them into a single system of equations of the form:

$$\begin{aligned}&\begin{bmatrix}\mathbf{x}'_{p+1} & \mathbf{x}'_{p+2} & \dots & \mathbf{x}'_{T}\end{bmatrix}\\ &\quad = \begin{bmatrix}\mathbf{A}'_1 & \mathbf{A}'_2 & \dots & \mathbf{A}'_p\end{bmatrix}\\ &\quad\quad \times \begin{bmatrix}\mathbf{x}'_{p} & \mathbf{x}'_{p+1} & \dots & \mathbf{x}'_{T-1}\\ \mathbf{x}'_{p-1} & \mathbf{x}'_{p} & \dots & \mathbf{x}'_{T-2}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{x}'_{1} & \mathbf{x}'_{2} & \dots & \mathbf{x}'_{T-p}\end{bmatrix}\end{aligned} \tag{17}$$

or simply:

$$\mathbf{X}_2 = \mathbf{A}\mathbf{X}_1 \tag{18}$$

where $\mathbf{X}_2$ is $2 \times (T-p-1)$, $\mathbf{A}$ is $2 \times (2p)$, $\mathbf{X}_1$ is $2p \times (T-p-1)$, and T is the length of bivariate vector time series, p is the order of the autoregressive process.

The best approximation, in the least-squares sense, of matrix $\mathbf{A}$ is given by:

$$\hat{\mathbf{A}} = \mathbf{X}_2\mathbf{X}_1^{+} \tag{19}$$

where $\mathbf{X}_1^{+}$ denotes the Moore–Penrose pseudo-inverse of $\mathbf{X}_1$, that minimizes the loss function $||\mathbf{X}_2 - \mathbf{A}\mathbf{X}_1||_F$ (Frobenius norm). This formulation of estimation problem is put in identical way as it is in DMD algorithm (5), (6) except for the component matrices are different.

Once the matrix $\mathbf{A}$ is estimated the prediction may be performed through successive use of the formula:

$$\hat{x}'_{T+h} = \begin{bmatrix}\hat{\mathbf{A}}'_1 & \hat{\mathbf{A}}'_2 & \dots & \hat{\mathbf{A}}'_p\end{bmatrix}\begin{bmatrix}\hat{x}'_{T+h-1}\\ \hat{x}'_{T+h-2}\\ \vdots\\ \hat{x}'_{T+h-p}\end{bmatrix} \tag{20}$$

where h denotes a time-step of forecast horizon and if $h-i \leq 0$ where $i = 1, 2, \dots, p$ then $\hat{x}'_{T+h-i} = \mathbf{x}'_{T+h-i}$.

The final step of the algorithm sums up the extrapolated trend and periodic components obtained in the detrending step (15) with predicted residuals obtained in the prediction step (20), and may be expressed by a simple formula:

$$\hat{x}'' = \hat{x} + \hat{x}' \tag{21}$$

4 Data description and processing details

This contribution uses daily sampled final IERS EOP 14 C04 time series data available on https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html as a reference for all computations. This data product is a combination of solutions obtained from four independent space

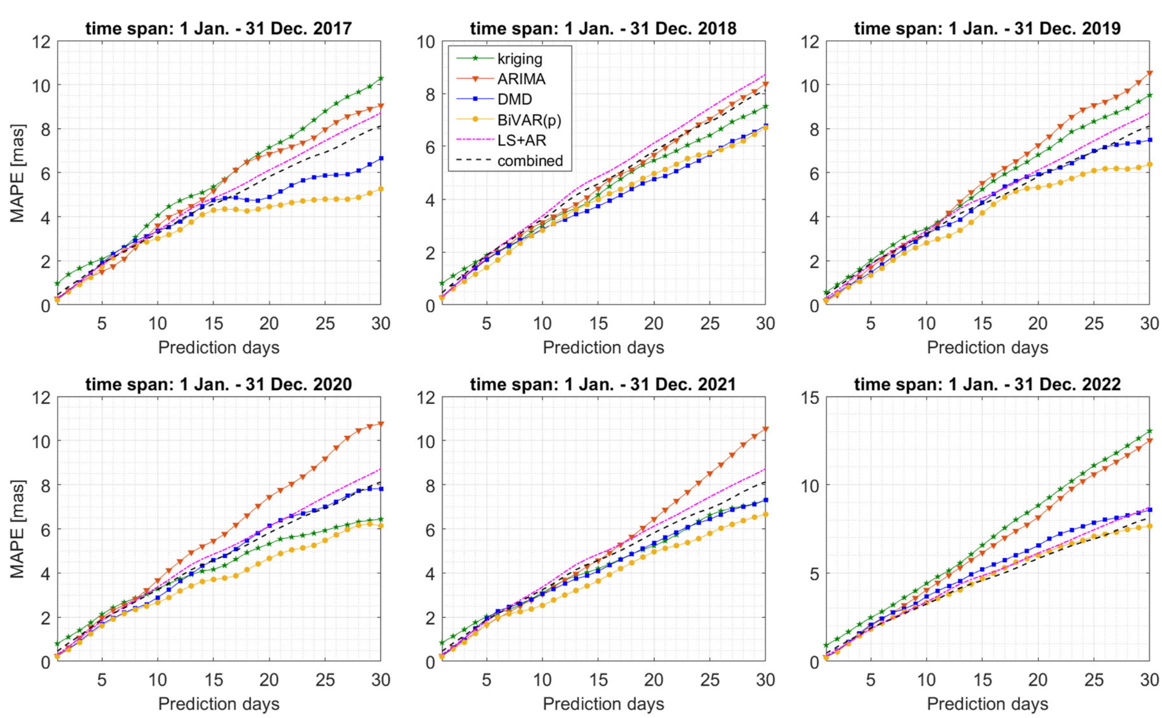


Figure 1: MAPEs of 30-day forecast for x (1st EOP PCC: LS + AR, combined).

geodetic techniques (C04; VLBI, GNSS, SLR, DORIS) and consistent with the International Terrestrial Reference Frame 2014 (14). It is published with 30-day latency. The prediction procedure was split into six subsequent years 2017–2022 giving 48 30-day predictions within each year. Forecast always started on 1 January of each year (2017-MJD 57754, 2018-MJD 58119, 2019-MJD 58484, 2020-MJD 5884, 2021-MJD 59215, 2022-MJD 59580). Within a given year, forecast was performed with 7-day shift. In fact, the prediction experiment covered the time interval from 01.01.1995 to 31.12.2022 since a forecast procedure consisted of two stages.

The first stage was necessary to select the best performing sets of parameters for the proposed methods. In case of DMD, they were: the input time series length (T) and its subdivision into a certain number of snapshots (L) (see description below Eq. (12)). In case of BiVAR(p), they were: the input time series length (T) for estimation of linear trend and periodic components, number of periodic components ($noPC$) involved, subseries of (T) for BiVAR(p) parameters estimation ($subT$) and an autoregression order (p). The first stage was completed as a result of nested loop statements for mentioned parameters. The sets of parameters were selected so that to balance the prediction accuracy for the 10th day and 30th day of forecast. The second stage was the proper prediction algorithm that for both methods was performed in an adaptive fashion. This means that every final 30-day forecast was preceded by preliminary 30-day forecasts (as many as sets of parameters) performed on 30 past observations preceding the 1st day of final forecast. The set of parameters with minimum sum of squared differences between forecast and observations was applied for

Table 1: MAPEs for 30-day x prediction for the yearly time spans 2017–2022 (units: mas).

Method	Day	2017	2018	2019	2020	2021	2022	Average
Kriging	1	0.97	0.82	0.58	0.80	0.84	0.92	0.82
ARIMA		0.28	0.27	0.19	0.25	0.23	0.24	0.24
DMD		0.25	0.30	0.22	0.24	0.25	0.23	0.25
BiVAR		0.23	0.27	0.20	0.22	0.21	0.21	0.22
LS + AR					0.31			
Combined					0.46			
Kriging	5	2.08	1.87	2.01	2.13	2.03	2.48	2.10
ARIMA		1.50	1.74	1.70	1.97	1.65	2.00	1.76
DMD		1.90	1.71	1.48	1.69	1.93	2.07	1.80
BiVAR		1.72	1.41	1.35	1.61	1.67	1.82	1.60
LS + AR					1.83			
Combined					1.89			
Kriging	10	4.06	3.03	3.45	3.29	3.11	4.41	3.56
ARIMA		3.60	3.12	3.19	3.67	3.05	4.04	3.45
DMD		3.36	2.87	3.21	2.89	3.04	3.67	3.17
BiVAR		3.01	2.82	2.81	2.66	2.53	3.33	2.86
LS + AR					3.37			
Combined					3.24			
Kriging	20	7.15	5.47	6.81	5.32	5.23	8.82	6.47
ARIMA		6.85	5.68	7.24	7.44	6.44	8.15	6.97
DMD		4.89	4.76	5.92	6.13	5.36	6.56	5.60
BiVAR		4.46	4.97	5.33	4.66	4.96	6.01	5.07
LS + AR					6.12			
Combined					5.82			
Kriging	30	10.29	7.52	9.53	6.44	7.30	13.05	9.02
ARIMA		9.05	8.37	10.54	10.77	10.54	12.51	10.30
DMD		6.64	6.78	7.48	7.82	7.31	8.56	7.43
BiVAR		5.27	6.71	6.38	6.14	6.66	7.66	6.47
LS + AR					8.73			
Combined					8.12			

the final forecast and such a procedure was repeated for the next 30-day forecast shifted by 7 days and so on.

In case of BiVAR(p), none of the single sets of parameters performed equally well for both x and y (usually one variable in a pair predicted visibly better) therefore in a joint prediction of polar motion one of the coordinates was always treated as a primary variable whilst the other as a secondary (assistant) variable.

5 Results

In order to compare the performance of newly introduced prediction methods and algorithms the results are confronted with those obtained from ordinary kriging [16, 17] and ARIMA [18] under the same conditions. In addition, the results are also presented at the background of the best performing techniques applied during the 1st Earth Orientation Parameters Prediction Comparison Campaign (1st EOP PCC) [8]. For all ERPs, a combined solution which was not an individual method of prediction but a weighted average of results provided by all methods used in 1st EOPPCC for a given EOP and the most accurate method for 30-day forecast, i.e., combination of the least-squares extrapolation and autoregressive prediction for x (W. Kosek), least-squares extrapolation of the harmonic model and autoregressive prediction for y (M. Kalarus) and Kalman filter for LOD [9, 22]. As mentioned in the Introduction section, the latter comparison is only illustrative and shows the level of accuracy provided by the new methods and its correspondence to the best performing ones of the 1st Campaign.

The quality of prediction was assessed by means of the mean absolute prediction error (MAPE) which is given by:

$$\mathrm{MAPE}_j = \frac{1}{n}\sum_{i=1}^{n}\left|O_{i,j} - P_{i,j}\right| \tag{22}$$

where: n is a number of predictions, j is a prediction day number, O stands for observed and P for predicted ERP.

Figure 1 shows MAPEs for 30-day x forecast using ordinary kriging, ARIMA, DMD, BiVAR(p) and two methods of the 1st EOP PCC, i.e., combination of the least-squares extrapolation with autoregressive prediction and a combined solution. The results show that the BiVAR(p) method, in almost all time periods, provides the smallest values of forecast error. It can also be noted that for ultra-short-term prediction (up to 8 days into the future) ARIMA is comparable to BiVAR(p) and DMD. Both newly introduced methods have the smallest values of MAPEs for time spans 2017–2021, while for the year 2022 they are very similar to the results obtained in the 1st EOP PCC.

Table 1 presents more detailed results of 30-day x forecast for all methods. Results of ultra-short-term prediction (up to 10 days into the future) indicate that the most accurate method is BiVAR(p), with MAPEs around

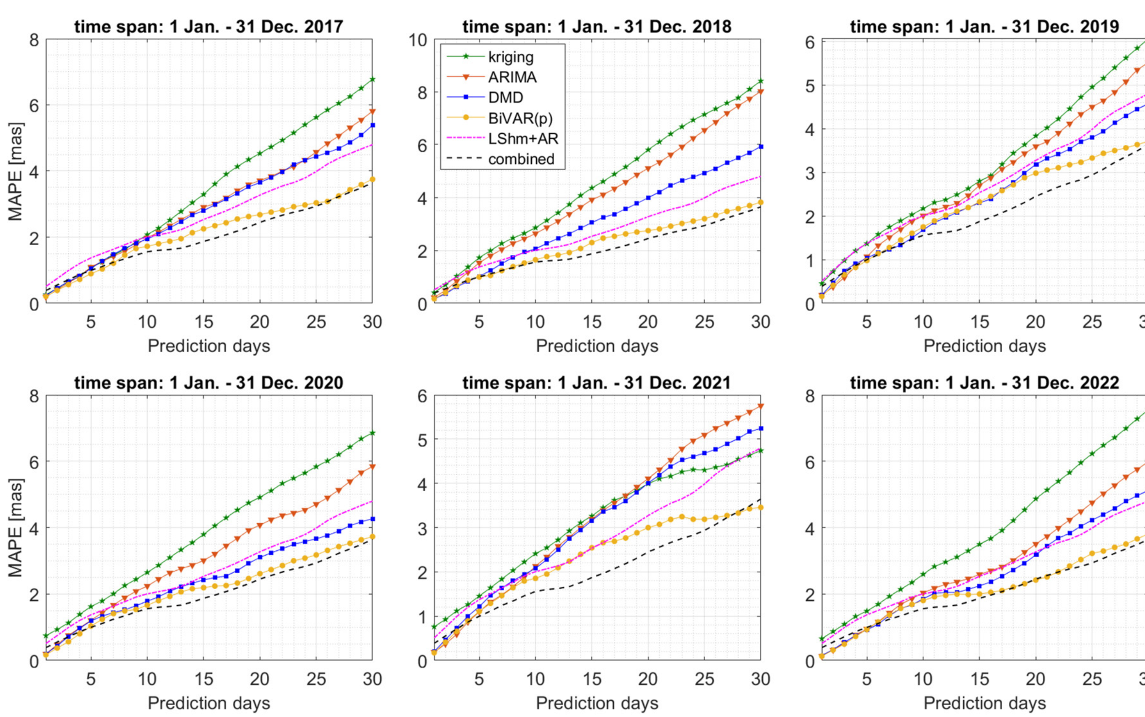


Figure 2: MAPEs of 30-day forecast for y (1st EOP PCC: LShm + AR, combined).

2.53–3.33 mas (2.86 mas on average). The MAPEs for BiVAR(p) on the 30th day of x prediction varies from 5.27 mas (2017) to 7.66 mas (2022), and are the lowest values of all methods (6.47 mas on average). It is also worth noting that the accuracy of forecast in 2022 has worsened in all cases. Inspection of periods of El Niño or La Niña occurrences reveals the correspondence between large values of the prediction errors and the incidence of these phenomena (https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php). Such correlation was also pointed out in, e.g., [23, 24] in case of polar motion and LOD prediction. However, comparing MAPEs for BiVAR(p) with those obtained during the 1st EOPPCC, it can be concluded that the overall performance with respect to the prediction accuracy for the newly introduced methods may constitute a good alternative for the existing methods or be a part of some hybrid approaches.

Figure 2 presents MAPEs for 30-day y prediction using ordinary kriging, ARIMA, DMD, BiVAR(p) and two methods involved in the 1st EOP PCC, i.e., least-squares extrapolation of the harmonic model and autoregressive prediction and a combined solution. Again, the most accurate method for 30-day y prediction among the methods used by the authors is BiVAR(p). As can be seen, in almost every yearly time span, the DMD has comparable values of MAPEs for the ultra-short forecast horizon (up to 10 days) to BiVAR(p). In the prediction range from around 10th to 25th day, the results of the combined solution (1st EOPPCC) prediction characterizes by lower forecast error values than BiVAR(p).

Table 2 shows numerical values for the selected days of the 30-day y forecast for all methods. Values of MAPEs

Table 2: MAPEs for 30-day y prediction for the yearly time spans 2017–2022 (units: mas).

Method	Day	2017	2018	2019	2020	2021	2022	Average
Kriging	1	0.26	0.40	0.45	0.75	0.76	0.65	0.55
ARIMA		0.19	0.22	0.19	0.19	0.19	0.13	0.19
DMD		0.23	0.17	0.19	0.19	0.20	0.15	0.19
BiVAR		0.21	0.17	0.16	0.17	0.17	0.13	0.17
LShm + AR					0.51			
Combined					0.39			
Kriging	5	1.07	1.72	1.37	1.62	1.45	1.49	1.45
ARIMA		1.10	1.53	1.07	1.21	1.10	0.95	1.16
DMD		1.05	1.02	1.03	1.19	1.22	0.92	1.07
BiVAR		0.89	0.99	0.98	1.03	1.11	0.93	0.99
LShm + AR					1.37			
Combined					1.00			
Kriging	10	2.08	2.85	2.17	2.65	2.41	2.59	2.46
ARIMA		1.98	2.63	2.00	2.24	2.13	2.03	2.17
DMD		1.94	2.06	1.68	1.78	2.08	1.86	1.90
BiVAR		1.72	1.64	1.75	1.66	1.85	1.80	1.74
LShm + AR					2.00			
Combined					1.56			
Kriging	20	4.53	5.80	3.83	4.91	3.99	4.87	4.66
ARIMA		3.71	5.10	3.59	4.08	4.11	3.50	4.02
DMD		3.65	3.99	3.18	3.10	4.01	3.17	3.52
BiVAR		2.68	2.75	2.98	2.61	3.00	2.42	2.74
LShm + AR					3.27			
Combined					2.45			
Kriging	30	6.78	8.41	6.06	6.85	4.74	7.55	6.73
ARIMA		5.81	8.01	5.52	5.84	5.75	5.99	6.15
DMD		5.38	5.93	4.57	4.27	5.24	5.12	5.09
BiVAR		3.75	3.82	3.71	3.73	3.46	3.81	3.71
LShm + AR					4.80			
Combined					3.64			

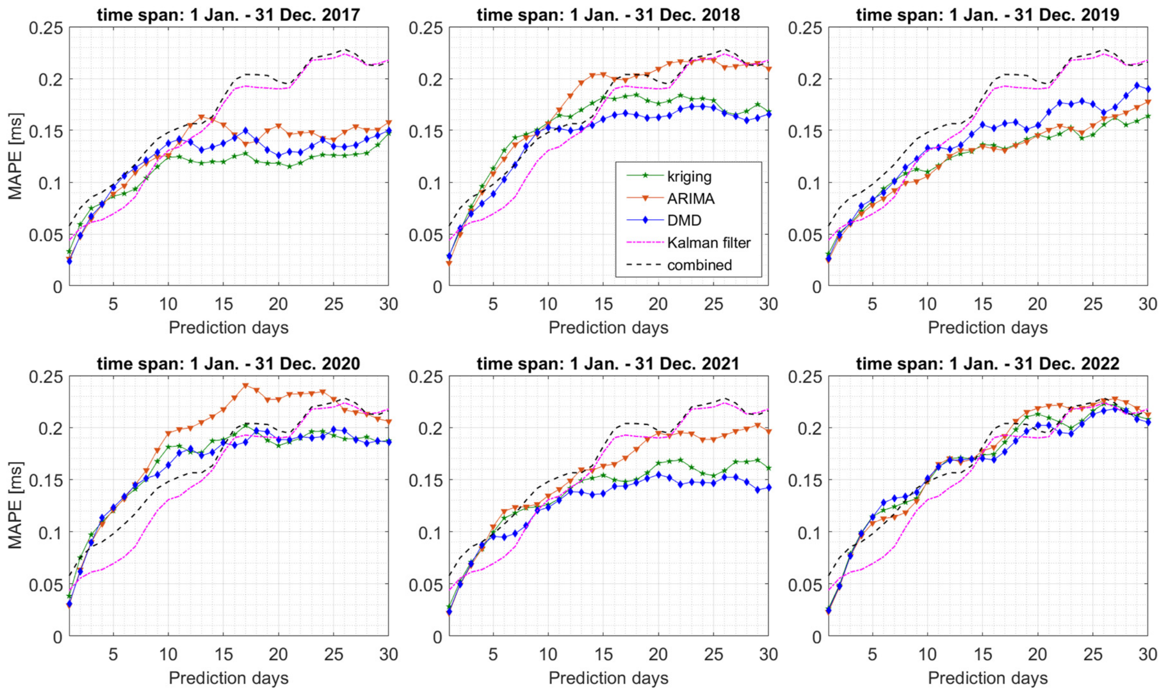


Figure 3: MAPEs of 30-day forecast for LOD (1st EOP PCC: Kalman filter, combined).

Table 3: MAPEs for 30-day LOD prediction for the yearly time spans 2017–2022 (units: ms).

Method	Day	2017	2018	2019	2020	2021	2022	Average
Kriging		0.033	0.024	0.031	0.038	0.028	0.026	0.030
ARIMA		0.026	0.022	0.025	0.029	0.022	0.023	0.025
DMD	1	0.024	0.029	0.027	0.031	0.023	0.025	0.027
Kalman filter					0.044			
Combined					0.058			
Kriging		0.086	0.113	0.082	0.120	0.099	0.114	0.102
ARIMA		0.089	0.108	0.078	0.121	0.105	0.108	0.102
DMD	5	0.095	0.088	0.083	0.123	0.095	0.114	0.100
Kalman filter					0.069			
Combined					0.098			
Kriging		0.124	0.157	0.110	0.181	0.126	0.148	0.141
ARIMA		0.126	0.157	0.106	0.195	0.135	0.149	0.145
DMD	10	0.138	0.153	0.133	0.164	0.123	0.151	0.144
Kalman filter					0.131			
Combined					0.148			
Kriging		0.118	0.176	0.146	0.213	0.166	0.213	0.172
ARIMA		0.155	0.209	0.145	0.219	0.195	0.219	0.190
DMD	20	0.126	0.163	0.155	0.202	0.155	0.202	0.167
Kalman filter					0.190			
Combined					0.197			
Kriging		0.147	0.168	0.164	0.187	0.161	0.208	0.173
ARIMA		0.158	0.209	0.178	0.206	0.196	0.212	0.193
DMD	30	0.150	0.166	0.190	0.186	0.142	0.205	0.173
Kalman filter					0.218			
Combined					0.216			

for the 10th day of prediction indicate that for ultra-short term prediction DMD and BiVAR(p) have similar mean forecast accuracy values, 1.90 mas (1.68–2.08 mas) and 1.74 mas (1.64–1.85 mas) respectively. BiVAR(p) provides a mean accuracy for the 30th day of prediction on the level of 3.71 mas, indicating similar accuracy to the combined solution of the 1st EOPPCC (3.64 mas). In contrast to x prediction, here we have seen no worsening of results in 2022.

Figure 3 presents MAPEs of 30-day LOD prediction using ordinary kriging, ARIMA, DMD and two methods of the 1st EOP PCC, i.e., Kalman filter and the combined solution. For ultra-short term (up to 10 days) prediction all newly presented methods have very similar MAPEs, for example on the 10th day MAPEs are varying between 0.141 and 0.145 ms (on average). In 2018, 2020, 2021 and 2022 DMD provides the most accurate LOD prediction for the 30th day, whilst in 2017 and 2019 ordinary kriging reaches the smallest values of MAPEs. It is visible that the results of the best methods of the 1st EOPPCC are better than those delivered by the authors' methods for ultra-short term prediction (up to 10 days). Beyond 15th day of forecast kriging, ARIMA and DMD seem to outperform the methods of the 1st EOPPCC.

Table 3 presents numerical values of MAPEs for specified days of 30-day LOD forecast using mentioned methods. Results indicate comparability of results for all the methods in ultra-short term prediction, except for Kalman filter which provided excellent result for 5th day of prediction, 0.069 ms. Beyond the 10th—15th day of the forecast the advantage of DMD and kriging can be seen. They provided roughly the same values of MAPE on the level of 17 ms on the 20th and 30th day of the forecast. Corresponding values for the 1st EOPPCC methods are slightly higher reaching app. 0.190 ms and 0.220 ms. It is worth mentioning that MAPEs for all methods have somewhat oscillatory behavior (Figure 3), in Table 3 it is visible for, e.g., DMD in 2017 (day 10 with higher MAPE than day 20) or kriging in 2018 (day 20 with higher MAPE than day 30). In case of the methods, newly introduced in this contribution, it may be partially explained by the experiment design. This means that results of only 48 30-day predictions were averaged each year. But since the methods involved in the 1st EOPPCC reveal the same character (averages from many more predictions) this may be the trace of some hidden phenomena unaccounted by the methods.

6 Conclusions

In this contribution two new methods for the Earth's Rotation Parameters prediction have been introduced, i.e., Dynamic Mode Decomposition and bivariate autoregression with prior trend/periodicity removal. The results of the 30-day prediction of ERPs were compared with the two methods previously presented by the authors (ordinary kriging and ARIMA) and results obtained during the 1st Earth Orientation Parameters Prediction Comparison Campaign. For all three ERP results are presented for six periods – yearly time spans (2017–2022), prediction always started on 1 January. The results indicate that BiVAR(p)-based prediction is the most accurate in all time periods for the 30th day of x prediction with the MAPE in the range of 5.27–7.66 mas. For ultra-short term x prediction (up to 10 days into future) mean value of MAPE for DMD and BiVAR(p) are on the level of 3.17 mas and 2.86 mas respectively. Relating these values to those obtained in the 1st EOPPCC, it can be concluded that the new methods may offer better forecast accuracy as measured by MAPE and can be an alternative to them. An additional argument is the relative stability of the results (no extreme differences) provided by these methods in all the years entering the prediction experiment. The most accurate method for the 30th day of y prediction used by the authors is BiVAR(p) with mean value of prediction error equal to 3.71 mas. As can be noticed, in almost all yearly time spans, the DMD method has comparable values of MAPEs for ultra-short term predictions (up to 10 days) to BiVAR(p), with mean MAPE values of 1.90 mas and 1.74 mas, respectively. A comparison of the results obtained in the 1st EOPPCC and presented by the authors shows that the combined solution (which is not a single method) provides, in overall, the smallest forecast error values, but BiVAR(p) (being a single method) is able to assure a comparable level of accuracy. In case of 30-day LOD prediction the most accurate method is DMD, with MAPEs on the 30th day in the range of 0.142 ms–0.205 ms (0.173 ms on average). The 10th day of LOD prediction is characterized by very similar mean MAPEs for all methods, i.e., 0.141 ms for kriging, 0.145 ms for ARIMA and 0.144 ms for DMD. Results indicate that the MAPEs of the authors methods in case of LOD prediction are compatible with those of the 1st EOPPCC. It can also be seen that newly presented methods, i.e., DMD and BiVAR(p) offer better polar motion prediction performance than previously applied by the authors methods based on kriging and ARIMA. In case of LOD prediction, only one new method was applied, i.e., DMD, which is also in most cases more accurate than mentioned kriging and ARIMA.

The results reveal that in case of BiVAR(p) additional information stored in the second variable supports the prediction of the other. DMD, on the other hand, due to low dimensionality of the prediction problem was applied with no rank truncation. This approach may be replaced with an attempt of finding some rank truncation algorithm that could replace a traditional least square fit of trend and periodic components and then residuals may be the subject of some other prediction method (a hybrid approach). The virtue of DMD is that it is a data-driven technique and no underlying equations describing a phenomenon are needed. It is equipped with only two steering parameters (this implementation), i.e., length of an input time series and its division into "snapshots" to estimate the **A** operator that projects previous state of a system (phenomenon) to the next one. It is easy to implement in an adaptive fashion (due to small number of steering parameters). In case of VAR, the advantage lies in its multivariate character, it is able to predict many variables at a time taking relations among variables into account. VAR is certainly more time consuming (in testing settings), due to more parameters to be optimized. It is harder to implement in a fully automatic mode, without a step of "human eye" inspection, without losing the prediction accuracy.

Acknowledgments: The authors are very grateful to M. Kalarus for providing data from 1st Earth Orientation Parameters Prediction Comparison Campaign (EOP PCC). We gratefully acknowledge Poland's high-performance computing infrastructure PLGrid (HPC Centers: ACK Cyfronet AGH) for providing computer facilities and support within computational grant no. PLG/2022/015968.
Research ethics: Not applicable.
Author contributions: The authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: The authors state no conflict of interest.
Research funding: The paper is a result of research on geospatial methods carried out within the statutory research no. 16.16.150.545 at the Department of Integrated Geodesy and Cartography, AGH University of Krakow.
Data availability: The data used in this contribution is publicly available on https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html.

References

1. Gambis D, Luzum B. Earth rotation monitoring, UT1 determination and prediction. Metrologia 2011;48:165–70.
2. Dick WR, Thaller D. IERS annual report 2018 international earth rotation and reference systems Service, central bureau. Frankfurt am Main: Verlagdes Bundesamts für Kartographie und Geodäsie; 2020:207 p.
3. Eubanks TM. Variations in the orientation of the earth. In: Smith DE, Turcotte DL, editors. Contributions of space geodesy in geodynamics: crustal dynamics. AGU Geodynamics Series, vol 24; 1993:1–54 pp.
4. Dill R, Dobslaw H, Hellmers H, Kehm A, Bloßfeld M, Thomas M, et al. Evaluating processing choices for the geodetic estimation of Earth orientation parameters with numerical models of global geophysical fluids. J Geophys Res Solid Earth 2020;125:e2020JB020025. https://doi.org/10.1029/2020JB020025.
5. Soffel M, Langhans R. Space-time reference systems. Berlin: Springer; 2013.
6. Raut S, Modiri S, Heinkelmann R, Balidakis K, Belda S, Kitpracha C, et al. Investigating the relationship between length of day and El-Nino using wavelet coherence method. Cham: Springer; 2022: 253–8 pp.
7. Kosek W, McCarthy D, Luzum B. Possible improvement of Earth orientation forecast using autocovariance prediction procedures. J Geodesy 1998;72:189–99.
8. Kalarus M, Schuh H, Kosek W, Akyilmaz O, Bizouard C, Gambis D, et al. Achievements of the Earth orientation parameters prediction comparison campaign. J Geodesy 2010;84:587–96.
9. Gross RS, Eubanks TM, Steppe JA, Freedman AP, Dickey JO, Runge TF. A Kalman-filter-based approach to combining independent Earth-orientation series. J Geodesy 1998;72:215–35.
10. Xu XQ, Zhou YH, Liao XH. Short-term earth orientation parameters predictions by combination of the least-squares, AR model and Kalman filter. J Geodyn 2012;62:83–6.
11. Kosek W, Kalarus M, Johnson T, Wooden W, McCarthy D, Popinski W. A comparison of LOD and UT1-UTC forecasts by different combined prediction techniques. Artif Satell 2005;40: 119–25.
12. Akyilmaz O, Kutterer H. Prediction of Earth rotation parameters by fuzzy inference systems. J Geodesy 2004;78:82–93.
13. Modiri S, Belda S, Hoseini M, Heinkelmann R, Ferrándiz JM, Schuh H. A new hybrid method to improve the ultra-short-term prediction of LOD. J Geodesy 2020;94:23.
14. Guessoum S, Belda S, Ferrandiz JM, Modiri S, Raut S, Dhar S, et al. The short-term prediction of length of day using 1D convolutional neural networks (1D CNN). Sensors 2022;22:9517.
15. Schmid PJ. Dynamic mode decomposition of numerical and experimental data. J Fluid Mech 2010;656:5–28.
16. Michalczak M, Ligas M. Kriging-based prediction of the Earth's pole coordinates. J Appl Geodesy 2021;15:233–41.
17. Michalczak M, Ligas M. The (ultra) short term prediction of length-of-day using kriging. Adv Space Res 2022;70: 610–20.
18. Michalczak M, Ligas M, Kudrys J. Prediction of earth rotation parameters with the use of rapid products from IGS, code and GFZ data centres using arima and kriging – a comparison. Artif Satell 2022;57:275–89.
19. Kutz JN, Brunton SL, Brunton BW, Proctor J. Dynamic mode decomposition: data-driven modeling of complex systems. Philadelphia SIAM; 2016.
20. Tu JH, Rowley CW, Luchtenburg DM, Brunton SL, Kutz JN. On dynamic mode decomposition: theory and applications. J Comput Dyn 2014;1:391–421.
21. Tirunagari S, Kouchaki S, Poh N, Bober M, Windridge D. Dynamic mode decomposition for univariate time series: analysing trends and forecasting, hal-01463744; 2017.
22. Freedman AP, Steppe JA, Dickey JO, Eubanks TM, Sung LY. The short-term prediction of universal time and length of day using

atmospheric angular momentum. J Geophys Res 1994;99: 6981−96.
23. Kosek W, McCarthy D, Luzum B. El Niño impact on polar motion prediction errors. Studia Geophys Geod 2001;45:347−61.
24. Lei Y, Zhao D, Guo M. Medium- and long-term prediction of length-of-day changes with the combined singular spectrum analysis and neural networks. Studia Geophys Geod 2023;58:42−55.

Advancing Polar Motion Prediction with Derivative Information

Maciej Michalczak [1], Marcin Ligas [1], Santiago Belda [2], Jose M. Ferrándiz [2], Sadegh Modiri [3]

[1] Department of Integrated Geodesy and Cartography, Faculty of Geo-Data Science, Geodesy, and Environmental Engineering, AGH University of Krakow, Krakow, Poland; mmichalc@agh.edu.pl, ligas@agh.edu.pl,
[2] UAVAC, Department of Applied Mathematics, Universidad de Alicante, Carretera San Vicente del Raspeigs/n, 03690 San Vicente del Raspeig, Alicante, Spain; santiago.belda@ua.es, jm.ferrandiz@ua.es
[3] Department Geodesy, Federal Agency for Cartography and Geodesy (BKG),60322 Frankfurt am Main, Germany; Sadegh.Modiri@bkg.bund.de

Abstract

Earth Orientation Parameters (EOP) are essential for monitoring Earth's rotational irregularities, impacting satellite navigation, space exploration, and climate forecasting. This study introduces a hybrid prediction model combining least-squares (LS) and vector autoregression (VAR) to improve Earth's Pole Coordinates (x, y) forecast accuracy. Using daily sampled IERS EOP 20 C04 data from 2013 to 2023, we conducted 1000 yearly random trials, performing 48 forecasts per year. Our method evaluates six data combinations, including primary variables (x, y) and their derivatives $(\dot{x}, \dot{y})$. Results show a systematic improvement in prediction accuracy, especially for ultra-short-term forecasts (10 days into future), with derivative information stabilizing the solutions. The best-performing combination $(x, y, \dot{x}, \dot{y})$ achieved a mean absolute prediction error (MAPE) reduction (with respect to the reference data combination) of up to 8% for the y and 7% for the x over a whole 30-day forecast horizon. These findings highlight the effectiveness of incorporating derivatives of polar motion time series into prediction procedure.

Keywords: Earth Orientation Parameters, polar motion, derivative time series, Vector Autoregression, prediction

1 Introduction

Earth Orientation Parameters (EOP) offer insights into irregularities in Earth's rotation, encompassing polar motion (PM), diurnal rotation (e.g., ERA or UT1-UTC), and precession–nutation (PN) pairs aligned with the Celestial Intermediate Pole (CIP) in geocentric reference frames. The significance of real-time EOP estimation is widely recognized, particularly in its pivotal role across diverse applications such as satellite navigation, deep space exploration, climate forecasting, and disaster prevention [1-10]. High-precision space geodetic methods, including very long baseline interferometry (VLBI), satellite laser ranging (SLR), the Global Navigation Satellite System (GNSS), and Doppler Orbitography and Radiopositioning Integrated by Satellite (DORIS), facilitate EOP observation. The complexity and time-consuming nature of data processing always result in delays. Therefore, predicting EOP from previous observations or combining them with geophysical phenomena is of great importance from a scientific and practical perspective. Despite advancements in EOP observation and estimation, accurately predicting EOP during extreme epochs remains challenging, with potential for both underestimations and over estimations [11]. This complexity arises from the intricate interplay between the atmosphere, Earth's surface, and the oceans, demanding precise modelling that considers all contributing factors [12]. Advancements in modelling, data analysis techniques, and the incorporation of superior data, higher resolution models, and refined algorithms are imperative to enhance EOP predictions. Additionally, the utilization of a novel algorithm and the integration of existing prediction models hold promise for further improving accuracy [13-15].

In this regard, the International Earth Rotation and Reference System Service (IERS) conducted two EOP prediction comparison campaigns (EOPPCC) in 2005 and 2021. These initiatives aimed to refine EOP prediction accuracy by bringing together worldwide research endeavours. Despite these efforts, the study's findings revealed that no single method consistently surpasses others across all prediction intervals.

Apart from that, in the last thirty years, researchers also dedicated substantial efforts to designed numerous high-precision EOP prediction algorithms. These include various methodologies like the least squares extrapolation (LS) utilizing harmonic models and autoregressive (AR) prediction, spectral analysis coupled with least squares extrapolation, artificial neural networks, least squares collocation (LSC), wavelet decomposition alongside auto-covariance prediction, Kalman filter forecasts, and fuzzy inference systems, among other innovative approaches.

These conventional algorithms represented meticulous design work aimed at achieving remarkable accuracy and precision in forecasting EOP values.

Nowadays, new advanced and sophisticated methods are emerging, i.e. Artificial Intelligence (AI). Machine learning (ML) and deep learning (DL) methods show significant promise in time series forecasting, particularly in their ability to autonomously grasp temporal relationships and effectively manage temporal patterns such as trends and seasonality.

However, the black-box nature of machine learning presents a significant drawback, primarily due to its lack of interpretability. While these models excel in prediction and pattern recognition, understanding how they arrive at conclusions can be challenging. This opacity poses obstacles in critical domains where explanations are crucial, such as space geodetic products. Interpreting decisions made by these models is complex, potentially leading to mistrust or incorrect assumptions about their reliability. Recently, different scientific addressed this issue trying to understand the causality (i.e. the relationship between two variables where one variable causes the other variable to change) [16].

Our previous study introduces a novel hybrid model that combines least-squares and vector autoregression [17] to enhance polar motion predictions, especially during extremum (minimum and maximum) epochs where conventional techniques often fall short. Extensive testing against observed polar motion data demonstrates the model's superior accuracy during extreme epochs compared to existing models. The model's strength lies in capturing linear trends, periodic components, and a joint prediction of the residual processes through a vector autoregressive model - VAR(p). This approach outperforms existing models in accurately depicting the (co-) dynamics of polar motion, even in the presence of outliers and seasonal effects.

In comparison to prevailing methodologies, our hybrid model represents a potential leap forward by exploring the limitations of current approaches. Many studies have been conducted that use effective angular momentum (EAM) as a technique for predicting short-term events [18–21]. Moreover, [16] suggests that in order to further improve the accuracy of prediction, we may also be able to use the uncertainty associated with polar motion in order to improve its accuracy.

Our work expands on previous efforts specializing in EOP and EAM and formal error analysis and provides a new approach to understand extremum points by leveraging historical data and polar motion rates. This contribution enhances EOP time series over a period of time. A comprehensive analysis may be achieved through the combination of polar motion rate time series with historical data in order to produce more accurate predictions due to the more accurate analysis. There is an innovative feature that accommodates changes in polar motion rate, which is particularly useful for the prediction of long-term changes in the polar movement. With the inclusion of dynamic polar motion, especially at extremum points, this pioneering method enhances prediction accuracy via the dynamic incorporation of polar motion into the model.

The organization of the paper is as follows. In Section 2 we introduce the methods. In Sections 3 we describe the data and the followed strategies. The numerical results appear in section 4. Finally, in Section 5 we state the conclusions of the study.

2 Prediction Method Description

This contribution uses extrapolation of least-squares fit of linear trend and periodic components (LS) and vector autoregression (VAR) on residuals to predict coordinates of the pole. This combination of methods will further be abbreviated to LS+VAR. The detrending step is applied to each variable separately:

$$V_i = A_i + B_i t + \sum_j^k a_{i,j} sin\left(\omega_{i,j} t + \phi_{i,j}\right) + e_i \tag{1}$$

where V stands for a variable of interest, A and B are the intercept and slope of a linear trend, whereas a, ω and ϕ are the amplitude, frequency, and phase of a sine wave, and e stands for error terms—all quantities to be estimated.

The residual series generated by subtracting the estimated trend and periodicities from the original series are subjected to a joint prediction process through a p-lagged vector autoregressive model VAR(p) given as:

$$\mathbf{V}_T^{\boldsymbol{R}} = \boldsymbol{c} + \mathbf{A}_p \mathbf{V}_{T-p}^{\boldsymbol{R}} + \mathbf{A}_{p-1} \mathbf{V}_{T-p+1}^{\boldsymbol{R}} + \cdots + \mathbf{A}_1 \mathbf{V}_{T-1}^{\boldsymbol{R}} + \mathbf{e}_T^{\boldsymbol{R}} \tag{2}$$

where a vector of constants $\mathbf{c}$ ($v\times1$) accounts for a non-zero mean of the multivariate process, entries of $\mathbf{V}^{\mathrm{R}}$ ($v\times1$) include residual values of variables at appropriate time delays and obtained in the detrending step, each $\mathbf{A}$ ($v\times v$) is a matrix of coefficients to be estimated, $\mathbf{e}^{\mathrm{R}}$ is a vector of disturbances and v stands for a number of variables in VAR(p) model.

Once the coefficient matrices $\mathbf{A}$ in (2) have been estimated, forecasting may be carried out by the successive use of equation (3).

$$\widehat{\mathbf{V}}_{T+h}^{R} = \begin{bmatrix} \widehat{\mathbf{A}}_1 & \widehat{\mathbf{A}}_2 & \cdots & \widehat{\mathbf{A}}_p \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{V}}_{T+h-1}^{R} \\ \widehat{\mathbf{V}}_{T+h-2}^{R} \\ \vdots \\ \widehat{\mathbf{V}}_{T+h-p}^{R} \end{bmatrix} \tag{3}$$

where h stands for a subsequent time-step of the forecast horizon, and if $h - i \leq 0$ where i = 1, 2, . . . , p then $\widehat{\mathbf{V}}_{T+h-i}^{R} = \mathbf{V}_{T+h-i}^{R}$.

Last step is to combine the extrapolated linear trend and periodic components of (1) with predicted residuals of (3). Details of the prediction procedure will be given in the following section.

3 Data Description and Processing Details

This contribution uses the daily sampled final IERS EOP 20 C04 time series available on https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html as a reference for all computations. The prediction experiment covers a span of 10 years, from March 1, 2013 (MJD 56352) to March 1, 2023 (MJD 60004). Within this time span, 1000 yearly random samples were generated, with the starting Modified Julian Date (MJD) of each yearly sample randomized. Within each yearly random sample, a 30-day forecast was performed every 7 days, resulting in a total of 48 forecasts in each random trial.

The experiment involves various data combinations, each of which includes derivative time series:

- $x, \dot{x}$
- $y, \dot{y}$
- $x, y, \dot{x}$
- $x, y, \dot{y}$
- $x, y, \dot{x}$
- $x, y, \dot{x}, \dot{y}$
- x, y (reference bivariate time series)

The LS+VAR prediction algorithm is a two-step process. The first step aims at selecting the best performing sets of parameters on a period of chosen length immediately preceding the forecast period (here, it is a year). This was done in a search loop responsible for preselecting sets of parameters that will later be involved in the further step of the algorithm. The searching loop involved sets of parameters listed below:

- input time series length to estimate coefficients of a linear trend and periodic terms,
- number of periodic terms,
- subseries of residual series for VAR parameters estimation that is obtained from input time series after removing trend and periodicities,
- order of autoregression.

The several sets of parameters offering the lowest mean prediction errors were identified in this step. Thus selected sets of parameters were passed to the next step where the prediction was carried out in an adaptive mode. Every proper 30-day forecast was preceded by preliminary ones performed on 30 nearest past observations before the 1st day of proper forecast. There were as many preliminary forecasts as sets of parameters identified in the first step of the algorithm. The combination of parameters assuring the lowest prediction errors on preceding 30 days was then applied to the final current forecast and this process was repeated until the last 30-day forecast in a yearly time span was reached.

Although VAR is a simultaneous procedure of predicting all variables at a time, here either two, three or four variables, only one of the main variables (x or y) was extracted since none of the single sets of parameters performed equally well for both coordinates (usually one variable from the pair predicted visibly better). Hence, in a joint LS+VAR-based prediction of the polar motion one of the coordinates was treated as a primary variable whilst others as secondary (assistant) ones and the prediction procedure was optimized with respect to the primary one.

It is worth mentioning that it was relatively easier to find a well performing set of steering parameters for combinations involving derivatives than for the reference one. This indicates some sort of stabilization of the solution with the use of derivative information.

4 Results

The basic measure of prediction accuracy in this contribution is the mean absolute prediction error (MAPE) defined as:

$$MAPE_d = \frac{1}{n}\sum_{i=1}^{n}\left|O_{i,d} - P_{i,d}\right| \tag{4}$$

where n is the number of predictions, d is the prediction day number, O stands for observed values, and P stands for predicted values.

All metrics, given below, indicating prediction accuracy improvement or decline resulting from the use of polar motion rates are based on previously defined MAPE. They are introduced in two variants, either on a specific day of forecast or for the entire forecast horizon. These measures are computed relative to the MAPEs of (x, y) data combination that does not use derivative information.

The Percentage of Improvement (PI) on a specific day d is defined as:

$$PI_d = 100\% \cdot \frac{\#differences_d < 0}{\#cases_d} \tag{5}$$

where symbol # stands for "number of", differences_d represent the difference between values of MAPE being compared and reference values of MAPE on a specific day, and $\#\text{cases}_d = 1000$ (due to 1000 random trials). Hence, this metric indicates the percentage of differences that are less than zero meaning that MAPEs being compared are smaller than the reference ones.

The Average Difference (AD) on a specific day d is calculated as:

$$AD_d = mean(differences_d) \tag{6}$$

where the mean value is taken of 1000 differences. This gives the average value on a specific day of the forecast by which MAPEs being compared differ from reference ones (in original units).

The Mean Improvement Rate (MIR) on a specific day d is defined as:

$$MIR_d = 100\% \cdot mean\left(\frac{differences_d}{reference\ values_d}\right) \tag{7}$$

where the mean value, as before, is computed from 1000 corresponding values. This gives the average percentage by which MAPEs being compared differ from reference ones.

The following measures of improvement are the overall counterparts of the ones introduced earlier, meaning that they carry general information on the improvement (or decline) within considered combination of data.

The Total Percentage of Improvement (TPI) is given by:

$$TPI = 100\% \cdot \frac{\#differences<0}{\#cases} \tag{8}$$

where #cases = 30000 (30-day forecast horizon times 1000 random trials).

The Total Average Difference (TAD) is calculated as:

$$TAD = mean(differences) \tag{9}$$

The Total Mean Improvement Rate is given by:

$$TMIR = 100\% \cdot mean\left(\frac{differences}{reference\ values}\right) \tag{10}$$

In two above metrics, the mean value is computed from 30 000 corresponding values.

Values of PI_d and TPI greater that 50 % indicate an improvement whilst values less than 50% indicate a decline. Negative values of AD_d (TAD) and MIR_d (TMIR) imply an improvement whilst positive values suggest a decline.

Figure 1 shows the prediction accuracy of the LS+VAR model for the *x*-coordinate of the Pole in all considered data combinations. It shows MAPEs for all yearly random trials generated within a 10-year time span (light grey lines), its average value (dashed black line), and highlighted lower and upper bounds (black lines) for the MAPEs. The graphs in Figure 1 are accompanied by Table 1, where detailed values of average MAPE and lower and upper bounds are given for days 1-10, 15, 20, 25, and 30. MAPEs demonstrate the standard behavior of increasing with the forecast day. Lower and upper bounds for MAPE reveal what accuracy may be expected from the method and data combinations involved on the basis of 1000 random trials. The width of the MAPE band reveals good precision (low scattering of individual MAPEs) up to approximately 10-15 days of the forecast; later on, the bounds get wider. The least precise data combination is that which does not involve the second principal variable, i.e., *y*.

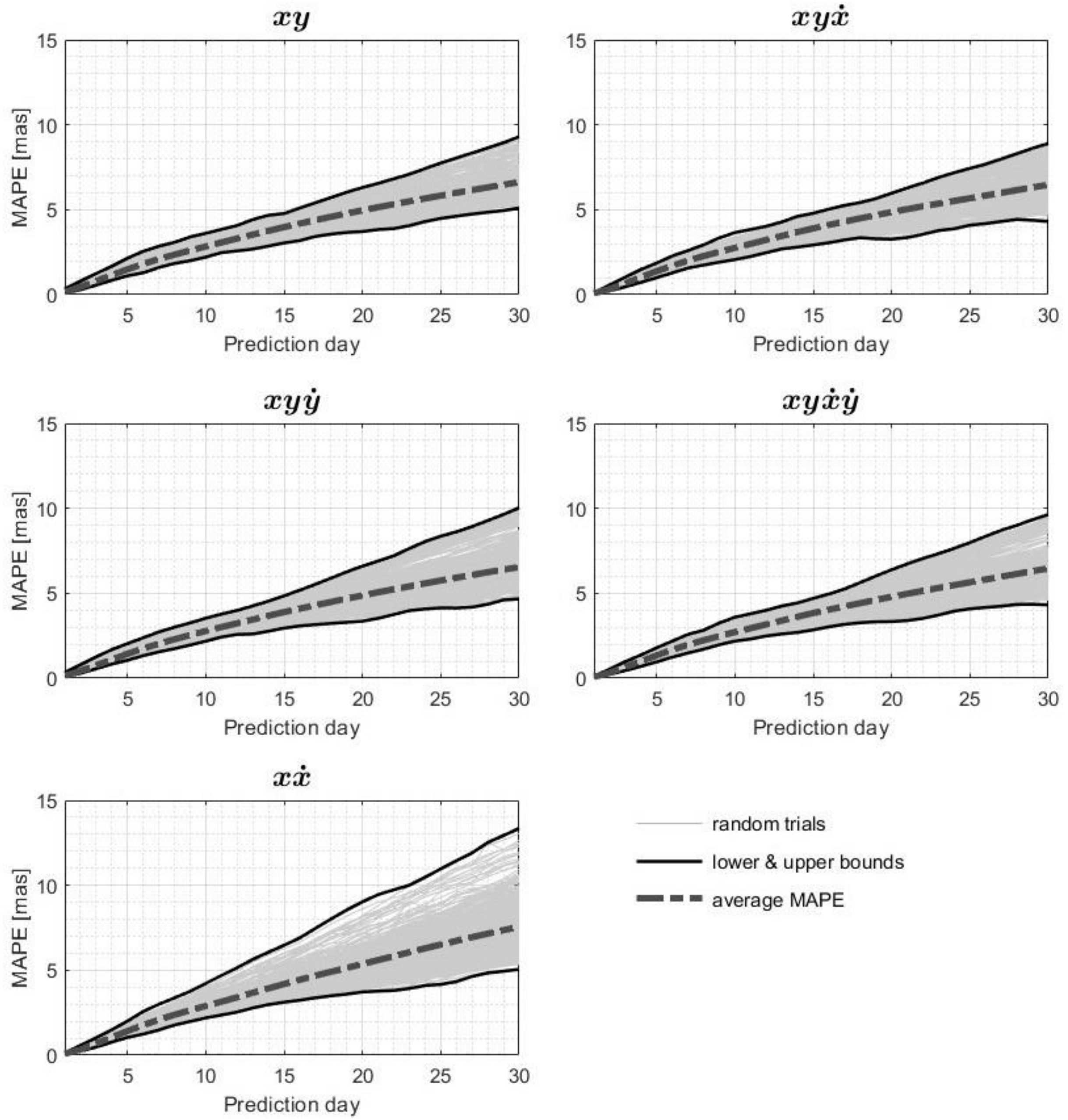


Figure 1 Individual and averaged MAPE values for polar motion *x*-coordinate derived from 1000 yearly random trials for the considered data combinations.

Table 1 Averaged, minimum and maximum MAPE values for polar motion *x*-coordinate computed from 1000 yearly random trials (units: mas).

Day	Average MAPE (minimum MAPE, maximum MAPE) (units: mas)				
	Data combination				
	x, y	$x, y, \dot{x}$	$x, y, \dot{y}$	$x, y, \dot{x}, \dot{y}$	$x, \dot{x}$
1	0.156 (0.087, 0.352)	0.079 (0.052, 0.130)	0.147 (0.080, 0.349)	0.075 (0.047, 0.118)	0.080 (0.050, 0.131)
2	0.458 (0.290, 0.796)	0.332 (0.226, 0.568)	0.431 (0.287, 0.794)	0.318 (0.222, 0.533)	0.337 (0.233, 0.583)
3	0.801	0.676	0.763	0.651	0.688

	(0.561, 1.241)	(0.460, 1.038)	(0.558, 1.256)	(0.449, 0.978)	(0.475, 1.026)
4	1.149 (0.842, 1.675)	1.031 (0.723, 1.482)	1.097 (0.830, 1.685)	0.998 (0.699, 1.365)	1.054 (0.770, 1.496)
5	1.492 (1.108, 2.143)	1.382 (0.998, 1.860)	1.431 (1.054, 2.036)	1.343 (0.958, 1.791)	1.420 (1.043, 2.008)
6	1.810 (1.286, 2.555)	1.708 (1.280, 2.277)	1.747 (1.324, 2.378)	1.671 (1.238, 2.185)	1.771 (1.241, 2.567)
7	2.096 (1.600, 2.848)	2.007 (1.563, 2.597)	2.029 (1.549, 2.722)	1.971 (1.483, 2.578)	2.090 (1.482, 2.999)
8	2.355 (1.837, 3.084)	2.274 (1.743, 2.932)	2.290 (1.738, 3.002)	2.233 (1.709, 2.828)	2.377 (1.773, 3.369)
9	2.599 (1.998, 3.400)	2.519 (1.913, 3.344)	2.537 (1.955, 3.256)	2.477 (1.969, 3.254)	2.644 (1.984, 3.760)
10	2.839 (2.208, 3.612)	2.757 (2.065, 3.657)	2.778 (2.179, 3.534)	2.716 (2.186, 3.585)	2.898 (2.195, 4.228)
15	3.977 (3.045, 4.776)	3.903 (2.923, 4.793)	3.893 (2.946, 4.851)	3.836 (2.847, 4.701)	4.190 (3.119, 6.469)
20	4.965 (3.716, 6.307)	4.861 (3.268, 5.974)	4.872 (3.357, 6.600)	4.790 (3.333, 6.387)	5.366 (3.721, 9.031)
25	5.828 (4.488, 7.753)	5.675 (4.086, 7.437)	5.751 (4.139, 8.359)	5.646 (4.083, 7.987)	6.512 (4.168, 10.968)
30	6.626 (5.099, 9.296)	6.465 (4.306, 8.895)	6.544 (4.669, 10.031)	6.462 (4.319, 9.627)	7.555 (5.050, 13.351)

Table 1 shows that all values of average MAPE in data combinations that involve two principal variables (x, y) and derivatives are lower than the ones of the reference combination. These values vary from 0.075–0.147 mas (1st day), 1.343–1.431 mas (5th day), 2.716–2.757 mas (10th day), 4.790–4.872 mas (20th day), 6.462–6.544 mas (30th day). The highest values in the above listing are for the combination $(x, y, \dot{y})$. The corresponding values for the reference combination are: 0.156 mas, 1.492 mas, 2.839 mas, 4.965 mas, 6.626 mas; respectively. The least accurate data combination $(x, \dot{x})$ does not involve the second principal variable, y, but even in this case, the improvement measured by averaged MAPE is visible for the first 7 days. A similar situation holds for lower bounds for MAPE, where 53 out of 56 values given in Table 1 are less than for the reference combination. When one considers only ultra-short-term prediction (up to 10 days), then 39 out of 40 values are less than those for reference data. Here, even the worst-performing data combination gives improved values. A slightly different situation is for upper bounds, where 36 out of 56 values are lower than those for the reference data (64%). For a 10-day forecast, 32 out of 40 values are lower, but when the worst combination $(x, \dot{x})$ is excluded, then 27 out of 30 values are better than reference data. The above shows that the use of derivatives in the prediction process stabilizes the solutions particularly for the first days of the forecast.

Figure 2 and Table 2 carry the same information but for the y component of polar motion. It is immediately apparent that the y coordinate is predicted better than the x one, a fact observed in many studies. Here, the width of the MAPE band reveals very good precision up to approximately 10 days of the forecast (~1 mas), later on, the bounds get wider. The least precise data combination is that which does not involve the second principal variable, i.e., x.

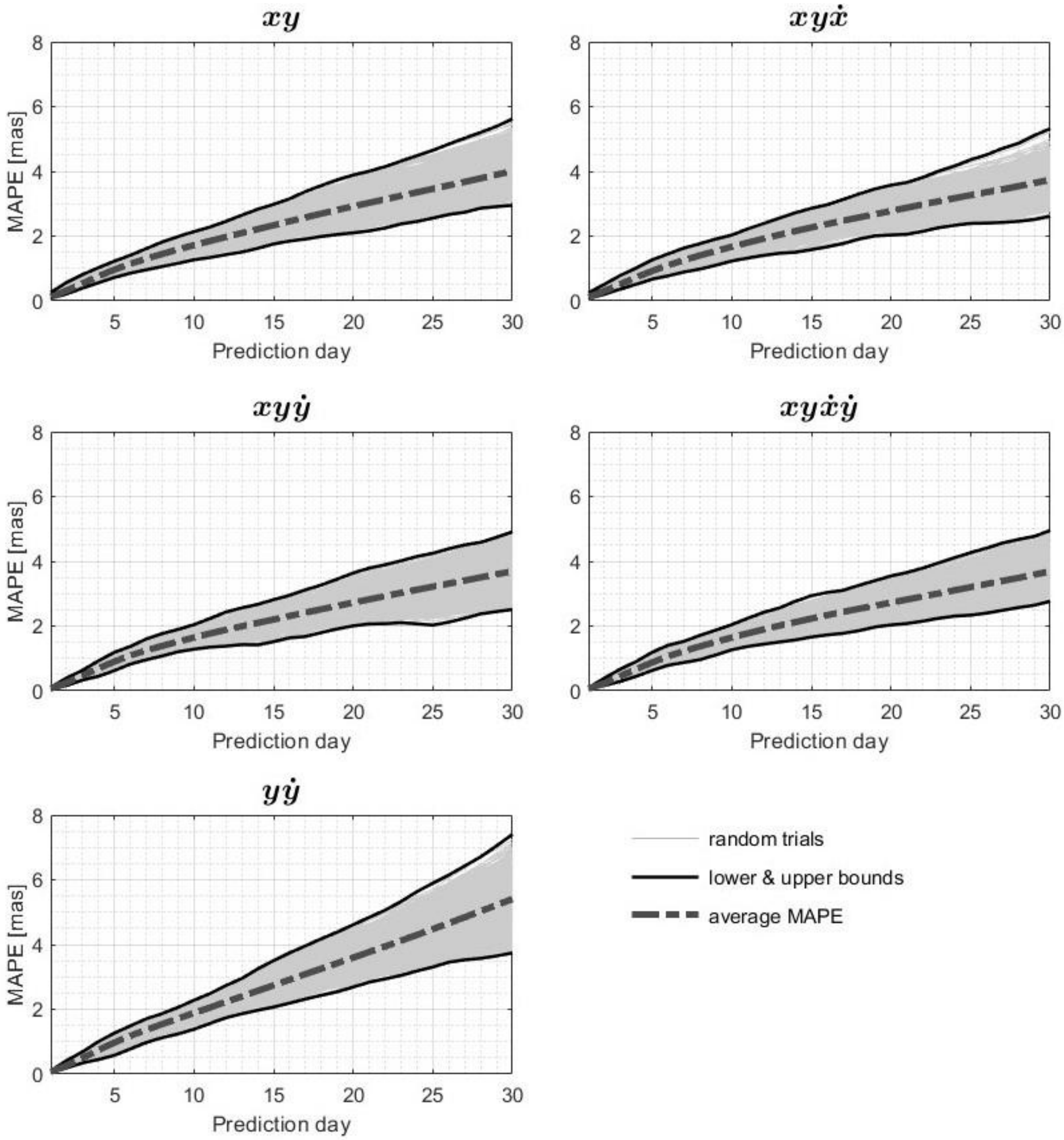


Figure 2 Individual and averaged MAPE values for polar motion *y*-coordinate derived from 1000 yearly random trials for the considered data combinations.

Table 2 Averaged, minimum and maximum MAPE values for polar motion *y*-coordinate computed from 1000 yearly random trials (units: mas).

	Average MAPE (minimum MAPE, maximum MAPE) (units: mas)				
	Data combination				
Day	x, y	$x, y, \dot{x}$	$x, y, \dot{y}$	$x, y, \dot{x}, \dot{y}$	$y, \dot{y}$
1	0.119 (0.070, 0.260)	0.115 (0.068, 0.250)	0.068 (0.033, 0.103)	0.064 (0.038, 0.096)	0.069 (0.037, 0.104)
2	0.321 (0.219, 0.570)	0.306 (0.206, 0.512)	0.255 (0.168, 0.398)	0.237 (0.164, 0.392)	0.264 (0.191, 0.419)

3	0.543 (0.393, 0.813)	0.516 (0.362, 0.786)	0.474 (0.333, 0.637)	0.454 (0.290, 0.671)	0.501 (0.346, 0.688)
4	0.763 (0.559, 1.026)	0.728 (0.512, 1.016)	0.696 (0.449, 0.926)	0.671 (0.449, 0.905)	0.742 (0.449, 1.005)
5	0.964 (0.723, 1.227)	0.924 (0.671, 1.267)	0.905 (0.625, 1.190)	0.878 (0.628, 1.191)	0.974 (0.580, 1.265)
6	1.144 (0.855, 1.410)	1.102 (0.773, 1.452)	1.090 (0.826, 1.371)	1.064 (0.789, 1.408)	1.180 (0.777, 1.489)
7	1.303 (0.961, 1.615)	1.261 (0.891, 1.633)	1.249 (0.959, 1.600)	1.226 (0.875, 1.529)	1.367 (0.976, 1.707)
8	1.449 (1.061, 1.818)	1.404 (0.982, 1.766)	1.388 (1.076, 1.766)	1.373 (0.962, 1.711)	1.543 (1.124, 1.868)
9	1.589 (1.161, 1.984)	1.540 (1.102, 1.911)	1.519 (1.206, 1.898)	1.512 (1.109, 1.877)	1.715 (1.240, 2.061)
10	1.722 (1.270, 2.141)	1.672 (1.231, 2.033)	1.646 (1.287, 2.044)	1.646 (1.266, 2.045)	1.886 (1.384, 2.282)
15	2.335 (1.758, 2.983)	2.267 (1.582, 2.859)	2.203 (1.518, 2.816)	2.231 (1.662, 2.938)	2.743 (2.071, 3.505)
20	2.921 (2.099, 3.886)	2.777 (2.035, 3.573)	2.725 (2.005, 3.638)	2.724 (2.036, 3.547)	3.597 (2.688, 4.613)
25	3.459 (2.555, 4.650)	3.265 (2.393, 4.359)	3.217 (2.029, 4.247)	3.202 (2.335, 4.268)	4.478 (3.305, 5.893)
30	4.010 (2.951, 5.622)	3.740 (2.615, 5.320)	3.690 (2.510, 4.912)	3.689 (2.763, 4.951)	5.407 (3.742, 7.404)

Table 2 shows that all values of average MAPE in data combinations that involve two principal variables (x, y) and derivatives are lower than the ones of the reference combination. These values vary from 0.064–0.115 mas (1st day), 0.878–0.924 mas (5th day), 1.646–1.672 mas (10th day), 2.724–2.777 mas (20th day), 3.689–3.740 mas (30th day). The highest values in the above listing are for the combination $(x, y, \dot{x})$. The corresponding values for the reference combination are: 0.119 mas, 0.964 mas, 1.722 mas, 2.921 mas, 4.010 mas; respectively. This time, the least accurate data combination $(y, \dot{y})$ that does not involve the second principal variable, i.e., x, improves the averaged MAPE only for the first 4 days. In this case, 45 out of 56 (80%) values of lower bounds for MAPE are less than for the reference combination. In case of 10-day prediction 33 out of 40 (82%) values are less than those for reference data. Excluding the worst data variant, it is 27 out of 30 cases with smaller values. For upper bounds, 43 out of 56 values are lower than those for the reference data (77%). For 10-day forecast 31 out of 40 (78%) values are lower, but when the worst combination $(y, \dot{y})$ is excluded then 27 out of 30 (90%) values are better than reference data. Also, in this case, the use of derivatives as additional variables improved the accuracy of the forecast.

Tables 3 and 4 present values of the above-defined improvement characteristics for x and y; respectively. All quantities are computed with respect to the data combination (x, y) that does not include rates. Metrics are given for every fifth day of the forecast horizon and also as an overall measure for the entire forecast horizon (Total). For x, day-wise percentage improvement characteristics show a general tendency of decreasing with subsequent days of the forecast. This tendency is slightly disrupted for y where the improvement has a multimodal character, i.e., high percentage improvement for the first days is gradually decreasing and around 15th – 20th day it increases again and repeats decreasing later on.

Table 3 Improvement analysis for *x*.

Data combination	Improvement measure	5th day	10th day	15th day	20th day	25th day	30th day	Total
$x, y, \dot{x}$	PI [%]	85.2	72.6	62.1	65.5	65.9	62.3	71.4
	AD [mas]	-0.110	-0.082	-0.075	-0.104	-0.153	-0.161	-0.110
	MIR [%]	-7.1	-2.8	-1.8	-2.0	-2.6	-2.3	-5.5
$x, y, \dot{y}$	PI [%]	72.6	66.9	65.6	63.8	59.1	61.7	66.5
	AD [mas]	-0.061	-0.061	-0.085	-0.093	-0.078	-0.082	-0.072
	MIR [%]	-3.9	-2.0	-2.0	-1.8	-1.3	-1.3	-2.4
$x, y, \dot{x}, \dot{y}$	PI [%]	90.6	74.5	69.1	68.4	62.0	59.3	73.6
	AD [mas]	-0.150	-0.123	-0.141	-0.175	-0.182	-0.165	-0.153
	MIR [%]	-9.7	-4.1	-3.4	-3.5	-3.2	-2.5	-7.0
$x, \dot{x}$	PI [%]	69.4	35.6	24.7	23.1	16.0	17.7	36.4
	AD [mas]	-0.073	0.059	0.212	0.401	0.683	0.929	0.299
	MIR [%]	-4.3	2.2	5.5	8.2	11.8	14.4	2.8

Table 4 Improvement analysis for *y*.

Data combination	Improvement measure	5th day	10th day	15th day	20th day	25th day	30th day	Total
$x, y, \dot{x}$	PI [%]	72.6	66.0	65.0	72.7	70.2	70.1	71.1
	AD [mas]	-0.040	-0.051	-0.069	-0.145	-0.194	-0.270	-0.109
	MIR [%]	-4.0	-2.8	-2.6	-4.5	-4.9	-5.9	-4.0
$x, y, \dot{y}$	PI [%]	76.0	72.9	77.1	74.2	70.1	69.8	75.8
	AD [mas]	-0.059	-0.076	-0.132	-0.197	-0.242	-0.320	-0.152
	MIR [%]	-5.9	-4.2	-5.5	-6.4	-6.5	-7.1	-7.4
$x, y, \dot{x}, \dot{y}$	PI [%]	89.8	77.9	74.7	81.6	78.3	78.9	81.5
	AD [mas]	-0.086	-0.076	-0.104	-0.198	-0.257	-0.321	-0.156
	MIR [%]	-8.7	-4.2	-4.2	-6.5	-7.0	-7.4	-8.3
$y, \dot{y}$	PI [%]	45.1	8.4	2.6	1.4	1.1	0.7	16.8
	AD [mas]	0.010	0.164	0.408	0.676	1.019	1.397	0.509
	MIR [%]	1.4	9.8	18.0	23.8	30.6	36.5	15.5

Overall percentage improvement quantified by cumulative measures across all days of the forecast is better visible for y (71.1% – 81.5%) than for *x* (66.5% – 73.6%). The meaning of the percentage improvement measure (PI, day-wise or overall) is that given % of MAPEs, obtained by inclusion of derivatives, out of 1000 (day-wise) or 30 000 (Total) MAPEs is of smaller value than MAPEs coming from the reference data combination. The day-wise and overall percentage metrics (last column) indicate $(x, y, \dot{x}, \dot{y})$, a combination that carries the most information, as the best performing one in LS+VAR prediction procedure. Remaining two combinations involving both coordinates of the pole and rates also reveal advantage of derivatives inclusion but to a slightly lesser degree. Here, one may observe that when predicting x inclusion of $\dot{x}$ perform better than inclusion of $\dot{y}$. The opposite may be observed for *y*, inclusion of $\dot{y}$ gives better results.

Also, the mean improvement rates (MIR) and their absolute counterparts (AD) (in original units) reveal a slight improvement by involving derivatives in the prediction process.

The relative improvement does not exceed 10% and in many cases it is not higher than 3% for x and 7% for y. On average, for the best performing data combination it is equal 7.0% (0.153 mas) for x, and 8.3% (0.156 mas) for y.

A clearly visible improvement may be observed for the first days of the forecast where percentage characteristics reach the highest values independently of the data combination (with two exceptions for y of negligible difference). Data combinations $(x, \dot{x})$ and $(y, \dot{y})$ that do not involve the second variable of the pole, either x or y reveal the worst performance however the tendency of improving the ultra-short-term forecast is also visible. The results for this data combination are presented to have a general view since they are not directly comparable to those combinations that include two coordinates of the pole. But, this proves the importance of involving two coordinates of the polar motion in a multivariate prediction process. Although the characteristics given in Tables 3 and 4 are not spectacular they show a systematic accuracy improvement (in a mean sense) by introducing derivative information.

5 Conclusions

In order to analyse and evaluate performance of LS+VAR with derivatives of polar motion as auxiliary variables, 6 data combinations have been used: $(x, \dot{x})$, $(y, \dot{y})$, $(x, y, \dot{x})$, $(x, y, \dot{y})$, $(x, y, \dot{x}, \dot{y})$ and (x, y) as the reference. The efficiency of this approach was analyzed within 10-year time span (01.03.2013 – 01.03.2023) in which 1000 yearly trials with a random beginning of the yearly period have been generated. Predictions were performed successively every 7 days giving 48 forecasts in a single yearly period. Hence, the conclusions are drawn on the basis of 48000 30-day forecasts. Results reveal a clearly visible gain in accuracy, although not substantial, it is systematic. This improvement is particularly apparent during the first few days of the forecast, which might indicate its potential use in ultra-short term prediction. Two combinations, i.e., $(x, \dot{x})$ and $(y, \dot{y})$ are not directly comparable to the reference one since they do not use the second main variable (x or y). But even in this case they turned out to be a good detector of performance of the proposed approach for ultra-short term prediction. The best performing data combination turn out to be $(x, y, \dot{x}, \dot{y})$ for both coordinates of the pole. For x coordinate it offers the accuracy, measured by average MAPE (average of 1000 MAPEs), on the level of 1.343 mas (5th day), 2.716 mas (10th day), 6.462 mas (30th day). The corresponding values for y coordinate are: 0.878 mas, 1.646 mas, 3.689 mas. The total improvement rate (across all days of the forecast horizon) with respect to the reference data combination is on the level of 7% for x and 8% for y coordinate. The total percentage of improvement (measured by number of MAPE cases lower than those for the reference combination) is equal to 74% for x and 82% for y component. Remaining two combinations $(x, y, \dot{x})$ and $(x, y, \dot{y})$ reveal also the advantage of incorporating derivatives but to a slightly lesser degree. It is worth recalling that in the latter two combinations of data prediction of x component has better performance with the inclusion of $\dot{x}$ than $\dot{y}$ whereas for y component it is the opposite. The experimental findings demonstrated in this contribution confirm that inclusion of derivatives to the prediction procedure based on LS+VAR is worth further analysis and an attempt of adapting it for operational settings (real time forecast).

Acknowledgements

We gratefully acknowledge Poland's high-performance computing infrastructure PLGrid (HPC Centers: ACK Cyfronet AGH) for providing computer facilities and support within computational grant no. PLG/2022/015968.

Research project partly supported by program „Excellence initiative – research university" for the AGH University of Krakow and it is also a result of research on geospatial methods carried out within the statutory research grant no. 16.16.150.545 at the Department of Integrated Geodesy and Cartography, AGH University of Krakow

Santiago Belda and Jose M. Ferrándiz research was partially supported by Generalitat Valenciana (PROMETEO/2021/030, SEJIGENT/2021/001), the European Union-NextGenerationEU (ZAMBRANO 21-04) and by the Project PID2020-119383GB-I00 funded by Ministerio de Ciencia e Innovación (MCIN/AEI/10.13039/501100011033/)

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Barnes, R.; Hide, R.; White, A.; Wilson, C. Atmospheric angular momentum fluctuations, length-of-day changes and polar motion. Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences 1983, 387, 31–73.

[2] Seitz, F.; Schuh, H. Earth rotation. Sciences of Geodesy-I: Advances and Future Directions 2010, pp. 185–227.

[3] Fodor, C.; Heinkelmann, R.; Schuh, H.; Varga, P. On the mutual interrelation between Earth rotation and earthquake activity. 2019.

[4] Ferrándiz, J.M.; Modiri, S.; Belda, S.; Barkin, M.; Bloßfeld, M.; Heinkelmann, R.; Schuh, H. Drift of the Earth's principal axes of inertia from GRACE and satellite laser ranging data. Remote sensing 2020, 12, 314.

[5] Modiri, S. On the improvement of earth orientation parameters estimation: using modern space geodetic techniques; Technische Universitaet Berlin (Germany), 2021.

[6] Modiri, S.; Heinkelmann, R.; Belda, S.; Malkin, Z.; Hoseini, M.; Korte, M.; Ferrándiz, J.M.; Schuh, H. Towards understanding the interconnection between celestial pole motion and Earth's magnetic field using space geodetic techniques. Sensors 2021, 21, 7555.

[7] Michalczak, M.; Ligas, M. Kriging-based prediction of the Earth's pole coordinates. Journal of Applied Geodesy 2021, 15, 233–241.

[8] Raut, S.; Modiri, S.; Heinkelmann, R.; Balidakis, K.; Belda, S.; Kitpracha, C.; Schuh, H. Investigating the relationship between Length of Day and El-Nino using wavelet coherence method. International Association of Geodesy Symposia 2022.

[9] Malkin, Z.; Belda, S.; Modiri, S. Detection of a New Large Free Core Nutation Phase Jump. Sensors 2022, 22, 5960.

[10] Shahvandi, M.K.; Dill, R.; Dobslaw, H.; Kehm, A.; Bloßfeld, M.; Schartner, M.; Mishra, S.; Soja, B. Geophysically-informed machine learning for improving rapid estimation and short-term prediction of Earth orientation parameters. Journal of Geophysical Research: Solid Earth, p. e2023JB026720.

[11] Modiri, S.; Belda, S.; Heinkelmann, R.; Hoseini, M.; Ferrándiz, J.M.; Schuh, H. Polar motion prediction using the combination of SSA and Copula-based analysis. Earth, planets and space 2018, 70, 1–18.

[12] Dobslaw, H.; Dill, R. Predicting Earth orientation changes from global forecasts of atmosphere-hydrosphere dynamics. Advances in Space Research 2018, 61, 1047–1054.

[13] Kalarus, M.; Schuh, H.; Kosek, W.; Akyilmaz, O.; Bizouard, C.; Gambis, D.; Gross, R.; Jovanović, B.; Kumakshev, S.; Kutterer, H.; et al. Achievements of the Earth orientation parameters prediction comparison campaign. Journal of Geodesy 2010, 84, 587–596.

[14] Śliwińska, J.; Kur, T.; Wińska, M.; Nastula, J.; Dobslaw, H.; Partyka, A. Second Earth Orientation Parameters Prediction Comparison Campaign (2nd EOP PCC): Overview. Artificial Satellites2022, 57, 237–253.

[15] Kur, T.; Dobslaw, H.; Śliwińska, J.; Nastula, J.; Wińska, M.; Partyka, A. 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 2022, 74, 1–9.

[16] Kiani Shahvandi, M.; Schartner, M.; Soja, B. Neural ODE differential learning and its application in polar motion prediction. Journal of Geophysical Research: Solid Earth 2022, 127, e2022JB024775.

[17] Ligas, M.; Michalczak, M. Dynamic mode decomposition and bivariate autoregressive short-term prediction of Earth rotation parameters. Journal of Applied Geodesy 2024;18(2): 211-221.

[18] Modiri, S.; Belda, S.; Hoseini, M.; Heinkelmann, R.; Ferrándiz, J.M.; Schuh, H. A new hybrid method to improve the ultra-short-term prediction of LOD. Journal of geodesy 2020, 94, 1–14.

[19] Shahvandi, M.K.; Soja, B. 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 2022, 70, 563–575.

[20] Guessoum, S.; Belda, S.; Ferrandiz, J.M.; Modiri, S.; Raut, S.; Dhar, S.; Heinkelmann, R.; Schuh, H. The Short-Term Prediction of Length of Day Using 1D Convolutional Neural Networks (1D CNN). Sensors 2022, 22, 9517.

[21] Gou, J.; Kiani Shahvandi, M.; Hohensinn, R.; Soja, B. Ultra-short-term prediction of LOD using LSTM neural networks. Journal of Geodesy 2023, 97, 52.

Akademia Górniczo-Hutnicza
im. Stanisława Staszica w Krakowie

AGH Szkoła Doktorska AGH

Michalczak Maciej **Kraków, 29.10.2024**

Nazwisko i imię Doktoranta
Full name of the Doctoral Student

Miejscowość, data (DD.MM.RRRR)
Place, Date (DD.MM.YYYY)

96032406473

Numer PESEL lub numer paszportu
PESEL or Passport number:

Inżynieria Lądowa, Geodezja i Transport

Dyscyplina naukowa lub dziedzina naukowa
Scientific discipline or field of science

OŚWIADCZENIE OKREŚLAJĄCE INDYWIDUALNY MERYTORYCZNY ORAZ PROCENTOWY WKŁAD W POWSTANIE ROZPRAWY DOKTORSKIEJ

STATEMENT DETERMINING INDIVIDUAL CONTRIBUTION TO THE DOCTORAL DISSERTATION ON A QUALITATIVE AND PERCENTAGE BASIS

1. Michalczak, M., Ligas, M. Kriging-based prediction of the Earth's pole coordinates. Journal of Applied Geodesy, 2021, vol. 15, no. 3, pp. 233-241, https://doi.org/10.1515/jag-2021-0007
   Wkład procentowy doktoranta / Doctoral Student's percentage contribution: 50%
   Wkład merytoryczny / Substantive contribution: Przygotowanie algorytmów obliczeniowych i wykonanie obliczeń, analiza wyników, opracowanie pierwszej wersji artykułu.

2. Michalczak, M., Ligas, M. The (ultra) short term prediction of length-of-day using kriging. Advances in Space Research, 2022, vol. 70, pp. 610-620, https://doi.org/10.1016/j.asr.2022.05.007
   Wkład procentowy doktoranta / Doctoral Student's percentage contribution: 50%
   Wkład merytoryczny / Substantive contribution: Propozycja koncepcji i metodologii badań, przygotowanie algorytmów obliczeniowych i wykonanie obliczeń, opracowanie pierwszej wersji artykułu.

3. Michalczak, M., Ligas, M., Kudrys, J. Prediction of Earth Rotation Parameters with the use of rapid products from IGS, CODE and GFZ data centers using ARIMA and kriging – a comparison. Artificial Satellites, Proceedings of the Second Earth Orientation Parameters Prediction Comparison Campaign (2nd EOP PCC) Workshop, 2022, vol.57, no.s1, 2022, pp. 274-289, https://doi.org/10.2478/arsa-2022-0024

Wkład procentowy doktoranta / Doctoral Student's percentage contribution: 33,3%
Wkład merytoryczny / Substantive contribution: Przygotowanie algorytmów obliczeniowych i wykonanie obliczeń, analiza wyników, opracowanie pierwszej wersji artykułu.

4. Michalczak, M., Ligas, M. Short-term prediction of UT1-UTC and LOD via Dynamic Mode Decomposition and combination of least-squares and vector autoregressive model. Reports on Geodesy and Geoinformatics, 2024, vol. 117, pp. 45-54. https://doi.org/10.2478/rgg-2024-0006
Wkład procentowy doktoranta / Doctoral Student's percentage contribution: 50%
Wkład merytoryczny / Substantive contribution: Propozycja koncepcji i metodologii badań, przygotowanie algorytmów obliczeniowych i wykonanie obliczeń, opracowanie pierwszej wersji artykułu.

5. Ligas, M., Michalczak, M. Dynamic mode decomposition and bivariate autoregressive short-term prediction of Earth rotation parameters. Journal of Applied Geodesy, 2024, vol. 18, no. 2, pp. 211-221. https://doi.org/10.1515/jag-2023-0030
Wkład procentowy doktoranta / Doctoral Student's percentage contribution: 50%
Wkład merytoryczny / Substantive contribution: Przygotowanie algorytmów obliczeniowych i wykonanie obliczeń, analiza uzyskanych wyników.

6. Michalczak, M., Ligas, M., Belda, S., Ferrándiz, J., Modiri, S. Advancing polar motion prediction with derivative information. Journal of Applied Geodesy, opublikowany online 22 sierpnia 2024, https://doi.org/10.1515/jag-2024-0046
Wkład procentowy doktoranta / Doctoral Student's percentage contribution: 20%
Wkład merytoryczny / Substantive contribution: Propozycja ogólnej koncepcji badań, analiza wyników, opracowanie pierwszej wersji artykułu.

29.10.2024 Michalczak
Data (DD.MM.RRRR) i podpis Doktoranta
Date (DD.MM.YYYY) and signature of the Doctoral Student

29.10.2024 Marcin Ligas
Data (DD.MM.RRRR) i podpis Promotora/~~Promotorów~~
Date (DD.MM.YYYY) and signature(s) of the Supervisor(s)