



A new hybrid method to improve the ultra-short-term prediction of LOD

Sadegh Modiri^{1,2} · Santiago Belda^{3,4} · Mostafa Hoseini⁵ · Robert Heinkelmann¹ · José M. Ferrándiz⁴ · Harald Schuh^{1,2}

Received: 28 January 2019 / Accepted: 18 January 2020 / Published online: 5 February 2020
© The Author(s) 2020

Abstract

Accurate, short-term predictions of Earth orientation parameters (EOP) are needed for many real-time applications including precise tracking and navigation of interplanetary spacecraft, climate forecasting, and disaster prevention. Out of the EOP, the LOD (length of day), which represents the changes in the Earth's rotation rate, is the most challenging to predict since it is largely affected by the torques associated with changes in atmospheric circulation. In this study, the combination of Copula-based analysis and singular spectrum analysis (SSA) method is introduced to improve the accuracy of the forecasted LOD. The procedure operates as follows: First, we derive the dependence structure between LOD and the Z component of the effective angular momentum (EAM) arising from atmospheric, hydrologic, and oceanic origins (AAM + HAM + OAM). Based on the fitted theoretical Copula, we then simulate LOD from the Z component of EAM data. Next, the difference between LOD time series and its Copula-based estimation is modeled using SSA. Multiple sets of short-term LOD prediction have been done based on the IERS 05 C04 time series to assess the capability of our hybrid model. The results illustrate that the proposed method can efficiently predict LOD.

Keywords LOD · EOP · Copula-based analysis · Prediction

1 Introduction

Earth orientation parameters (EOP) are a collection of parameters that describe irregularities in the rotation of the Earth. EOP are classified into three groups: polar motion (PM) given by the x , y , parameters; diurnal rotation (e.g., ERA = Earth rotation angle, or UT1-UTC); and precession–nutation (PN) pair, which give the orientation of the conventional Celestial Intermediate Pole (CIP) in the geocentric celestial reference

frame. The EOP can be observed with modern high-precision space geodetic techniques, such as very long baseline interferometry (VLBI), satellite laser ranging (SLR), and global positioning system (GPS) (Tapley et al. 1985; Lichten et al. 1992; Schuh and Schmitz-Hübsch 2000). Real-time EOP estimation is needed for many applications including precise tracking and navigation of interplanetary spacecraft, climate forecasting, and disaster prevention. However, the complexity and time-consuming in data processing always lead to time delays. Consequently, the prediction of EOP from past observed data or combining with the geophysical phenomena is of great scientific and practical importance. In addition to the five EOP, the length of day (LOD) is used to model the variations in the Earth's rotation rate. LOD is the difference between the duration of the day measured by space geodesy and nominal day of 86,400 s duration and defined as: $LOD = -\frac{d(UT1-UTC)}{dt}$ (Freedman et al. 1994). LOD is changing due to gravitational effects from external bodies and geophysical processes occurring in different Earth layers. Consequently, its knowledge is essential for various applications related to reference frame determination and metrology, such as interplanetary navigation and

✉ Sadegh Modiri
sadegh@gfz-potsdam.de

¹ GFZ German Research Centre for Geosciences, Potsdam, Germany

² Institute for Geodesy and Geoinformation Science, Technische Universität Berlin, Berlin, Germany

³ Image Processing Laboratory (IPL) - Laboratory of Earth Observation (LEO), University of Valencia, Valencia, Spain

⁴ UAVAC, University of Alicante, Alicante, Spain

⁵ Department of Civil and Environmental Engineering, Norwegian University of Science and Technology, Trondheim, Norway

space geodesy orbitography (i.e., precise orbit determination) because of its coupling with the orbit node. However, the LOD prediction is extremely difficult due to extreme events such as El Niño which demonstrated themselves in the LOD signals (Holton and Dmowska 1989; Gross et al. 1996).

Several techniques have been developed to improve the accuracy of LOD prediction. These algorithms could be classified into two groups: first, the methods that use the information within the LOD time series, e.g., auto-covariance (AC) (Kosek et al. 1998; Kosek 2002), wavelet decomposition (Akyilmaz et al. 2011), or neural network (Schuh et al. 2002; Liao et al. 2012; Lei et al. 2015, 2017). Besides, this group includes the hybrid methods using the combination of least squares (LS) and auto-regressive (AR), auto-regressive moving average (ARMA), auto-covariance, and neural network (Kosek et al. 2005; Xu and Zhou 2015; Wu et al. 2019). In the second group, we cast the methods that take into account the axial component of effective angular momentum (EAM_z) (Freedman et al. 1994; Gross et al. 1998; Johnson et al. 2005; Niedzielski and Kosek 2008; Kosek 2012; Nastula et al. 2012; Dill et al. 2019). Freedman et al. (1994) showed that the use of atmospheric angular momentum (AAM) wind terms in the Kalman filter technique to predict LOD variations improved near-term predictions. Johnson et al. (2005) used UT1-like observations determined by AAM in the UT1-UTC combination solution to predict UT1 which showed a significant reduction in the prediction errors when compared with the previous prediction method (McCarthy and Luzum 1991). Also, Dill et al. (2019) used 6 days long predicted EAM_z values for the PM and UT1-UTC prediction using LS extrapolation and AR model. The Earth orientation parameters prediction comparison campaign (EOP PCC) took place within 2005–2009, and it was organized to assess the various prediction techniques under the same conditions and rules. One of the main results of the EOP PCC was that there is not a specific method preferred for all EOP and all prediction intervals (ultra-short term and long term). Also, as the EOP prediction accuracy benefits from AAM forecast data, EOP PCC recommended paying more attention to the analysis and prediction of atmospheric angular momentum (AAM), continental hydrology angular momentum (HAM), and ocean angular momentum (OAM) (Kalarus et al. 2010). Therefore, a new prediction method is required to fully capture the dependence structure between AAM, HAM, OAM, and EOP. Although there are approaches to quantify the dynamical relation between the geophysical fluids (the atmosphere, the ocean, and the land hydrology) and the LOD variation (Gross 2015), we ignore them in our work as we exactly want to assess this relation directly independent of theoretical implications through comparing the LOD and the effective angular momentum function z-component by the Copula method. In this paper, we introduce an algorithm to improve the LOD prediction for reaching the accuracy

goals pursued by the Global Geodetic Observing System (GGOS) of the International Association of Geodesy (IAG), i.e., 1 mm accuracy and 0.1 mm/year stability on global scales in terms of the International Terrestrial Reference Frame (ITRF) defining parameters (Plag and Pearlman 2009). We explored the combination of Copula-based analysis and singular spectrum analysis (SSA) to predict LOD. In Modiri et al. (2018), we applied the combination of SSA and Copula for the first time as a novel deterministic-stochastic tool for PM prediction. In this method, deterministic part is estimated by SSA, whereas Copula is used for modeling the stochastic part. The results indicated that the proposed approach can efficiently predict PM. Moreover, the improvement in PM prediction accuracy up to 365 days in the future is found to be 40% on average compared to the current PM prediction data from the International Earth Rotation and Reference Systems Service (IERS) Rapid Service/Prediction Center (RS/PC), hosted by the US Naval Observatory (USNO) (Petit and Luzum 2010). The Copula method contains both linear and nonlinear dependence structures between variables, and it is a powerful tool for dealing with multi-dimensional data and for modeling the relationship between parameters (Joe 1997). SSA is a subspace-based technique which makes use of empirical functions derived from the data to model the time series in a pre-specified level of details. It can be used for trend extraction and extrapolation (Alexandrov 2009; Modiri et al. 2018), periodicity detection, seasonal adjustment, smoothing, noise reduction (Golyandina et al. 2001; Ghil et al. 2002) as well as for change point detection (Moskvina and Zhigljavsky 2003; Hoseini et al. 2020). First, we derive the dependence structure between LOD and EAM_z . Based on the fitted theoretical Copula, we then simulate LOD from the EAM_z data. Next, the difference between LOD time series and its Copula-based estimation is modeled using SSA. After that, the LOD will be computed from predicted EAM_z . Finally, the difference will be predicted and will be added to LOD predicted by Copula. Multiple sets of ultra-short-term (10 days) LOD prediction have been made based on IERS 05 C04 time series to assess the capability of our hybrid model. We consider the same conditions as EOP PCC to show the effectiveness of the presented method. We compare the prediction results with those of existing techniques of EOP PCC, and the results evidence that the proposed approach can efficiently predict LOD.

2 Methods

The combination of stochastic and deterministic methods is used for LOD prediction. The Copula-based analysis technique aims to estimate the models for capturing the dependence structure between observed LOD and the EAM_z . Also,

SSA is used as a deterministic technique to obtain stochastic residuals (the difference between the observed data and the Copula generated data). Finally, let us remark that we used the IERS EOP time series available at the time of the EOP PCC so that our results could be easily compared to the former analyses. In the following section, the theoretical background of Copula theory and SSA is briefly sketched.

2.1 Copula-based analysis

The Copula approach exploits linear and nonlinear dependency between variables. Copula is a flexible tool offering an enormous improvement in capturing the real correlation pattern. This technique provides the grounds for dealing with multi-dimensional data and modeling the relation between parameters based on the marginal distribution functions of the variables (Embrechts et al. 2002).

Copula appeared in the mathematics context for the first time by Sklar (1959). Sklar's theorem indicates that a Copula function C connects a given multivariate distribution function with its univariate marginal. For bivariate distribution, there is a bivariate Copula C which models the joint cumulative probability distribution function of two variables X and Y based on the marginal cumulative distribution functions $F_X(x)$ and $F_Y(y)$.

$$P(X \leq x, Y \leq y) = F_{X,Y}(x, y) = C(F_X(x), F_Y(y)) = C(u, v) \quad (1)$$

where C describes the joint distribution function $F_{X,Y}(x, y)$. The variables u and v are transformations of X and Y to uniform distribution, respectively. The Copula is unique when the marginals are continuing functions. As the Copula is a reflection of the dependence structure itself, its construction is reduced to the study of the relationship between the variables, giving freedom for the choice of the univariate marginal distribution. Further information about Copula can be found, e.g., in Joe (1997) and Nelsen (2006). For many years, the Copula method has been used for modeling the dependence structure between random variables in different types of studies, such as Economics (Rachev and Mittnik 2000; Patton 2006, 2009), Biomedicine (Wang and Wells 2000; Escarela and Carriere 2003), Hydrology (Bárdossy and Li 2008; Bárdossy and Pegram 2009; Verhoest et al. 2015), Meteorology (Laux et al. 2011; Vogl et al. 2012), Hydro-geodesy, and Geodesy (Modiri et al. 2015, 2018). Six different bivariate Copula families are used in this research: the Archimedean 12, Archimedean 14, Clayton, Frank, Gumbel, and Joe. Further information about mathematical details on these families can be found in Appendix A, Nelsen (2006), or in Salvadori and De Michele (2007).

2.2 Singular Spectrum Analysis

SSA is a time series analysis tool which can be used to retrieve robust components of a dataset aiming to provide an easier to interpret picture of complex observations. The method diagonalizes a lag-covariance matrix concerning a basis of orthogonal eigenvectors and computes the corresponding eigenvalues (Groth and Ghil 2015). SSA is able to reveal useful information about hidden underlying processes of a time series. Within its four steps, SSA groups correlated information in a time series and offers the opportunity of reproducing new versions of the time series based on their different characteristics [see Appendix B, Golyandina et al. (2001), and Ghil et al. (2002), for more details].

2.3 Error analysis

The mean absolute error (MAE) is used in order to evaluate the prediction accuracy. The MAE is calculated for the k th day in the future as follows:

$$\text{MAE} = \frac{1}{k} \sum_{i=1}^k (|P_i - O_i|) \quad (2)$$

where P_i is the predicted value of the i th prediction day, O_i is the corresponding observed value, and k is the total prediction number.

3 Calculation and analysis

3.1 Data description

3.1.1 Length of day (LOD)

Daily time series of LOD in this contribution are from IERS EOP 05 C04 series (Gambis 2004). The LOD time series are available from (<http://hpiers.obspm.fr/eop-pc>) and span the time interval from 1996 to 2008.

3.1.2 Effective angular momentum (EAM) functions

EAM functions in both mass and motion terms explain the non-tidal geophysical excitation of the Earth's rotation due to mass redistribution in atmospheric and terrestrial hydrosphere, and in the oceans. The EAM data consist of three main components X , Y , and Z . The X and Y components are associated with the excitation of polar motion, whereas the Z component is responsible for the excitation of LOD variation (Salstein 1993; Dobsław and Thomas 2005; Dill 2008; Jungclaus et al. 2013; Dobsław 2016). The EAM functions are dimensionless with the sampling of 1 day and are

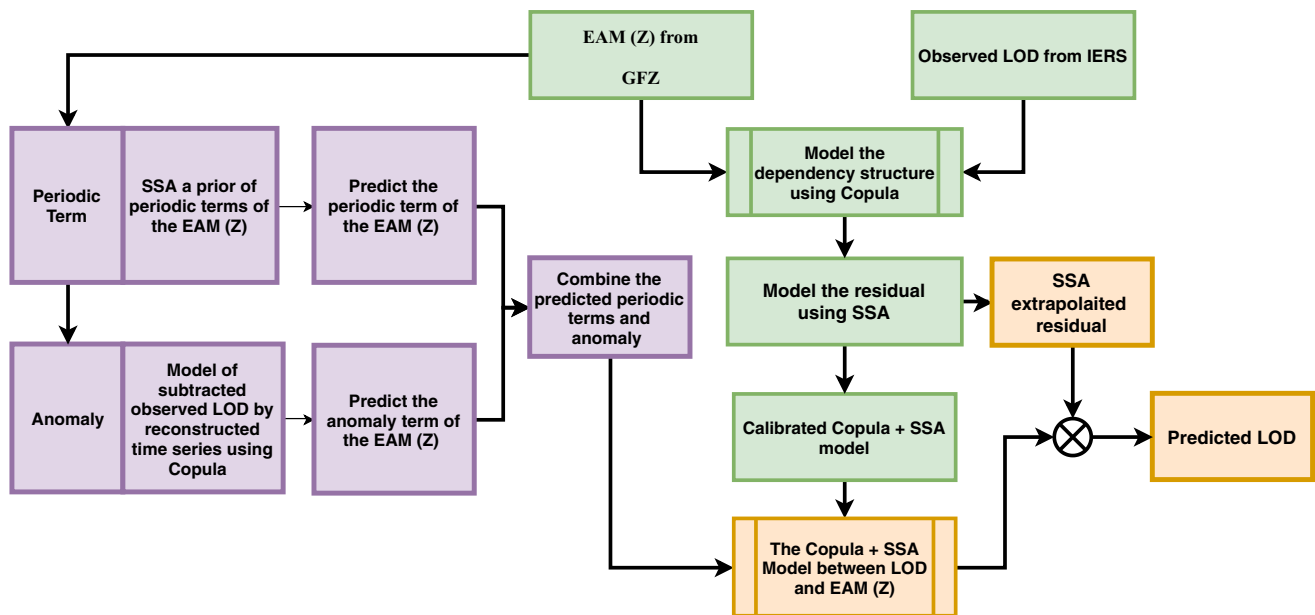


Fig. 1 Scheme of the prediction algorithm (Copula+SSA model). The Copula-based joint distribution function between LOD and EAM_Z (Calibration step) is shown in green. EAM_Z prediction is shown in purple. The prediction of LOD using the calibrated model (final step) is illustrated by orange

provided by the Earth System Modeling group at Deutsches GeoForschungsZentrum Potsdam (ESMGFZ) (Dobslaw and Dill 2018). The EAM time series are available from: (<ftp://ig2-dmz.gfz-potsdam.de/EAM/>).

3.2 Data processing and analysis

In this paper, we defined an algorithm for LOD prediction as shown in Fig. 1. It is important to note that LOD can be decomposed into several components (e.g., variations related to zonal components of solid Earth tides and ocean tides, atmospheric circulation, internal effects, and transfer of angular momentum to the Moon orbital motion). Taking into account that the accuracy of the different models may be not homogeneous, we decide to include in the modeling of this study the total variation in LOD. We are aware that it may be more challenging for testing the method performance that relying partially in previous models for certain components, but we preferred not to remove too many difficulties when testing the method as a mean to get more insight into its capabilities. Having said that, the methodology is structured as follows. We analyze the EAM_Z which is the sum of mass and motion terms of AAM, HAM, and OAM (see Fig. 2). The dependence structure between observed LOD and EAM_Z is captured and modeled by using Copula-based analysis. The difference between the observed LOD and Copula LOD estimated data is modeled using the SSA method. After that, the EAM_Z is predicted as described in detail in Modiri et al.

(2018). The prediction algorithm is demonstrated through the following steps:

1. Copula-based joint distribution function of LOD and EAM_Z
 - Model the dependency structure between LOD and EAM_Z
 - Model the periodic residual using SSA
 - Calibrate the Copula+SSA model
2. EAM_Z prediction
 - SSA periodic terms estimation
 - Copula anomaly modeling
3. LOD prediction using the calibrated Copula+SSA model
 - Sample random data from the conditional Copula;

as it is shown in Fig. 3, the training interval is between 1996 and 2003.

3.2.1 Copula-based joint distribution function of LOD and EAM_Z

In this study, the training dataset which ranges from 1996 to 2003 is used to fit a Copula-based joint distribution function; both EAM_Z and LOD are transformed to uniformly distributed values between (0, 1) interval through their empirical cumulative distribution function. Thereafter, the dependence structure between the EAM_Z and LOD is

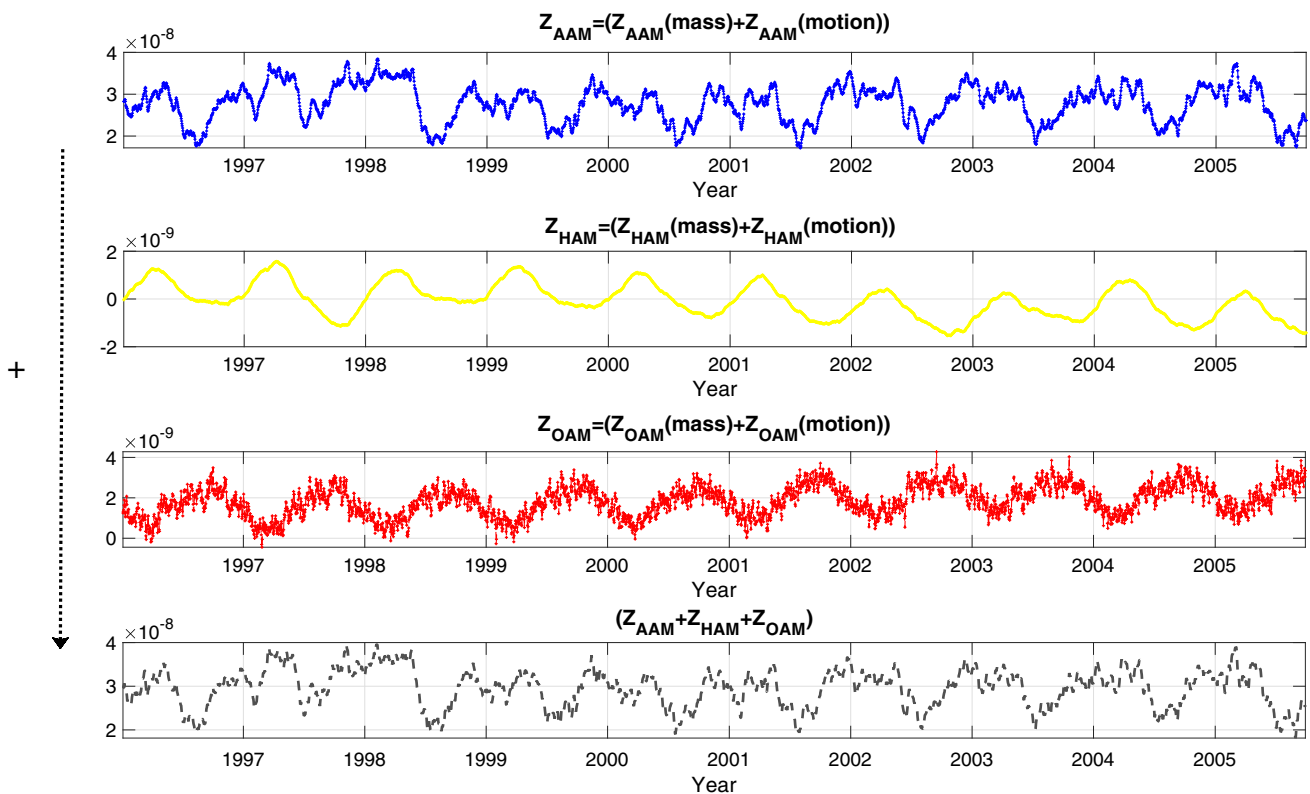
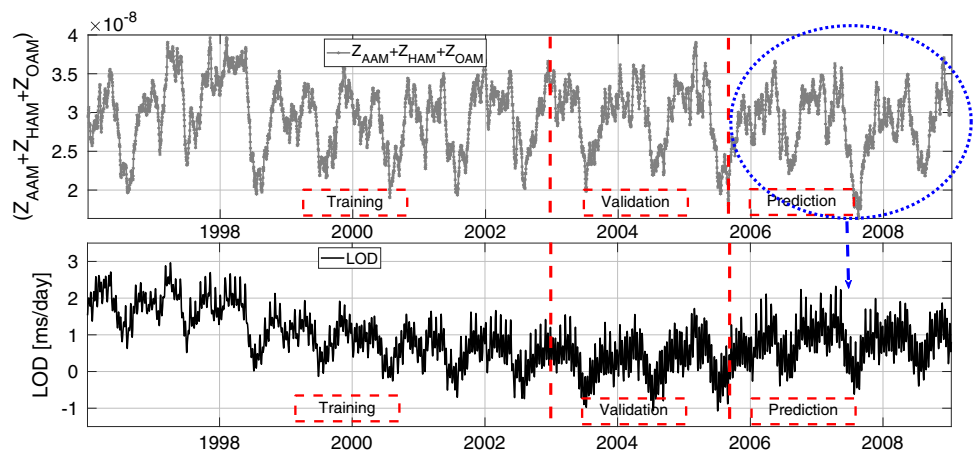


Fig. 2 The EAM_Z being the sum of mass and motion terms of AAM, HAM, and OAM ($Z_{AAM} + Z_{HAM} + Z_{OAM}$)

Fig. 3 Time series of LOD and ($Z_{AAM} + Z_{HAM} + Z_{OAM}$) between 1996 and 2008. The time series is divided into three parts: training part (1996–2003), validation (2003–2005), and prediction (2005–2008)



investigated. First, the empirical Copula is estimated using the Eq. 3. As it can be seen in Fig. 4, there is a scatter plot (upper panel) of LOD and EAM_Z , and it shows approximately a linear dependence structure with both upper and lower heavy tail which can be modeled by using the Archimedean Copula. Therefore, the theoretical bivariate Archimedean Copula functions with their estimated parameters are fitted to the estimated empirical Copula. The LOD data are sampled based on the Copula and the empirical marginal distribution of EAM_Z . Next, the residuals of the generated LOD by Copula are estimated using SSA. After

that, the ($Z_{AAM} + Z_{HAM} + Z_{OAM}$) data between 2003 and 2005 are used for the validation of LOD prediction (calibration step). Here, the predicted EAM_Z are used for the LOD prediction in the time interval between 2005 and 2008.

3.2.2 EAM_Z prediction

For this step, we defined an algorithm for EAM_Z prediction as shown in Fig. 1. The ($Z_{AAM} + Z_{HAM} + Z_{OAM}$) time series can be split up into two parts. The first part is dealing with periodic effects such as annual and semiannual

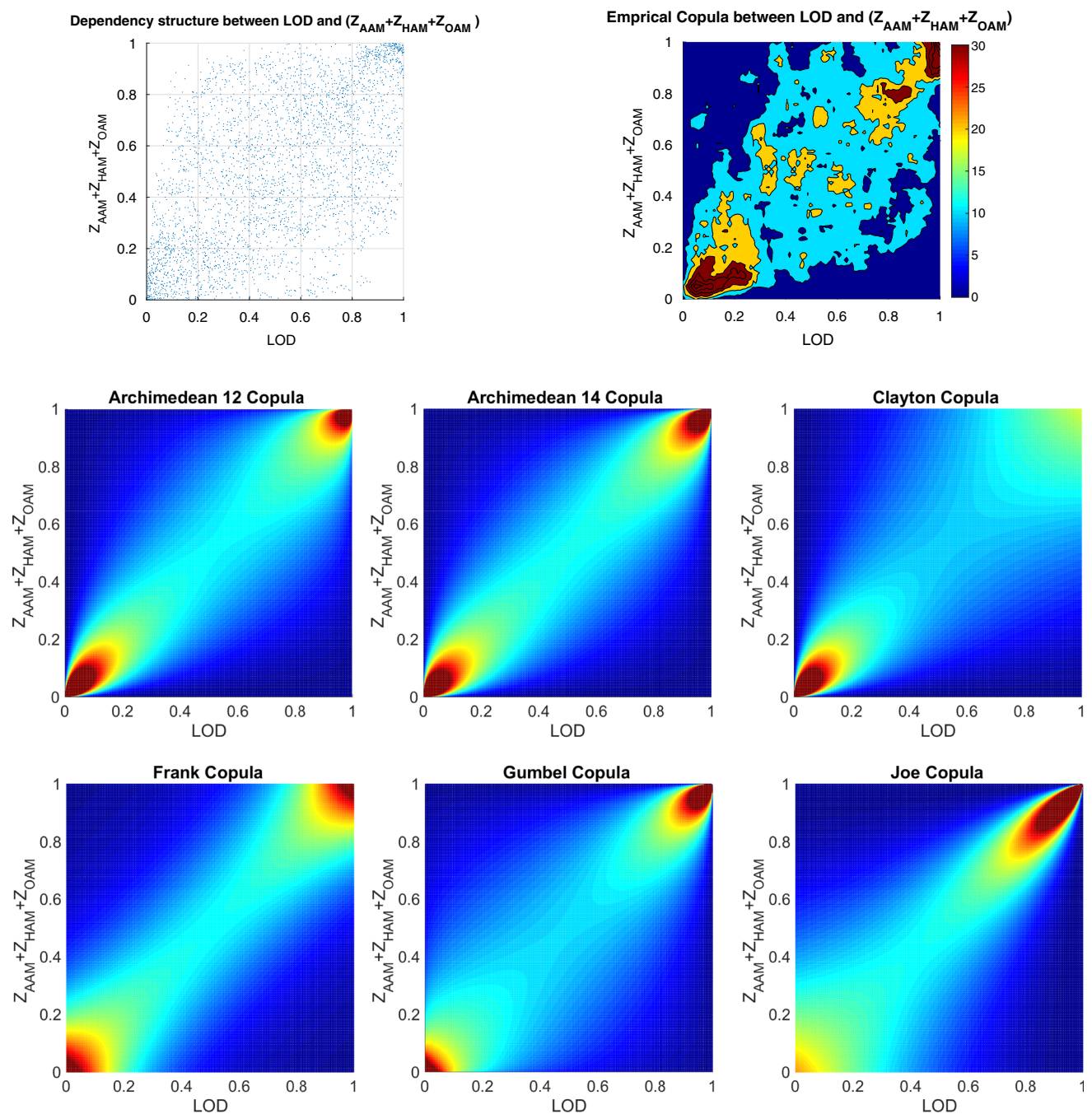


Fig. 4 Scatter plot of LOD and (EAM_Z) and its empirical Copula (upper panel). The fitted Archimedean 12 ($\theta = 1.30$), Archimedean 14 ($\theta = 1.53$), and Clayton ($\theta = 1.31$) Copula (middle panel), Frank

($\theta = 5.10$), Gumbel ($\theta = 1.69$), and Joe ($\theta = 2.92$) Copula (lower panel) between 1996 and 2003 in the rank space $[0, 1]$

variations due to the spectral analysis of EAM_Z (illustrated in Figure 5). The SSA models the periodic terms of the $(Z_{AAM} + Z_{HAM} + Z_{OAM})$ (see Fig. 6). Then, the difference between the observed EAM_Z and SSA estimated data is modeled by using the Copula-based analysis method. First, the window length ($L = 365$ days) is selected considering the main periodicity (see Figure 5). After that, the number of

singular vectors for reconstruction of the EAM_Z time series is determined. The trajectory matrix is constructed by having the window length and number of singular vectors. The cyan curve depicts the SSA-reconstructed EAM_Z time series. The periodic terms of EAM_Z are extrapolated using the SSA as a priori model. The difference between the values predicted by SSA and the predictions based on the EAM_Z is called EAM_Z

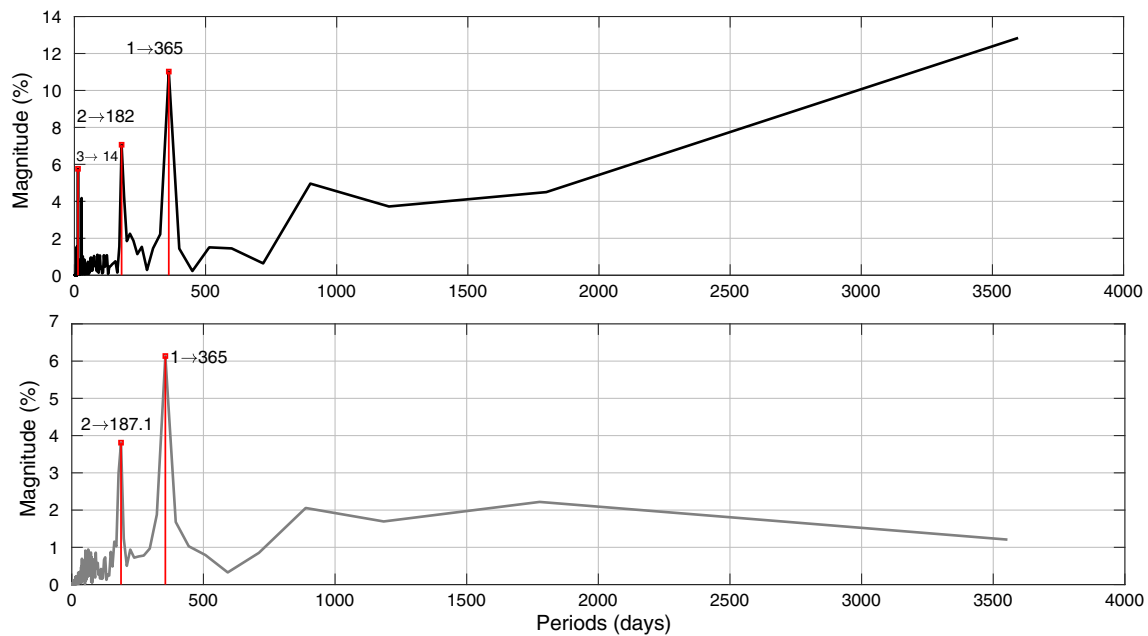
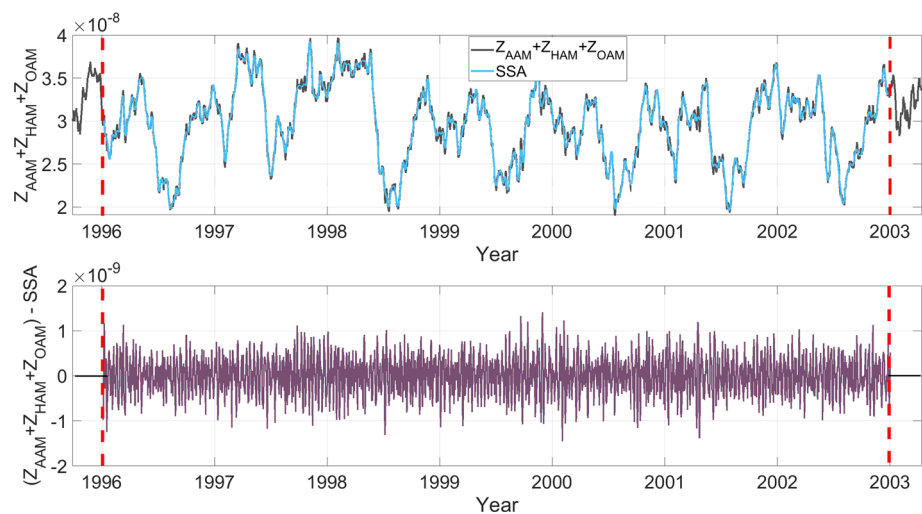


Fig. 5 Spectral analysis of the LOD (up), $Z_{AAM} + Z_{HAM} + Z_{OAM}$ (down) using fast Fourier transform (FFT)

Fig. 6 The original time series and the reconstructed time series (upper panel), and the difference between the original and reconstructed time series (lower panel) for ($Z_{AAM} + Z_{HAM} + Z_{OAM}$)



anomaly and has a stochastic nature. This anomaly part is predicted using the Copula-based model. The anomaly part is displayed in Fig. 6 (lower panel). This part is formed with the same window length L . The dependence structure between the column $_i$ and column $_{i+1}$ in the residual matrix is investigated. As can be seen in Fig. 7, the scatter plot illustrates a linear dependence between the two adjacent columns which are modeled by Archimedean Copula. Then, the empirical Copula is determined using Eq. 3. The next step is fitting a bivariate Archimedean Copula. In this study, Frank Copula is selected for predicting the EAM_Z anomaly due to its ability to capture the linear dependence structure. Finally, the Copula-based predicted anomalies are added to the pre-

viously described deterministic part. Here, we utilized seven years of EAM_Z time series from September 1998 to September 2005 for the 10 days ahead forecasting during the interval between October 2005 and 2008, i.e., the same interval as it has been used for the EOP PCC. As can be seen in Fig. 8, the MAE of the EAM_Z prediction is up to the 0.76 for the next days which is two orders of magnitude smaller than the EAM_Z magnitude. In Fig. 7, the scatter plot shows a linear relationship between column $_i$ and column $_{i+1}$. Then, the corresponding empirical Copula is estimated. Finally, the anomalies are added to the SSA-forecasted time series.

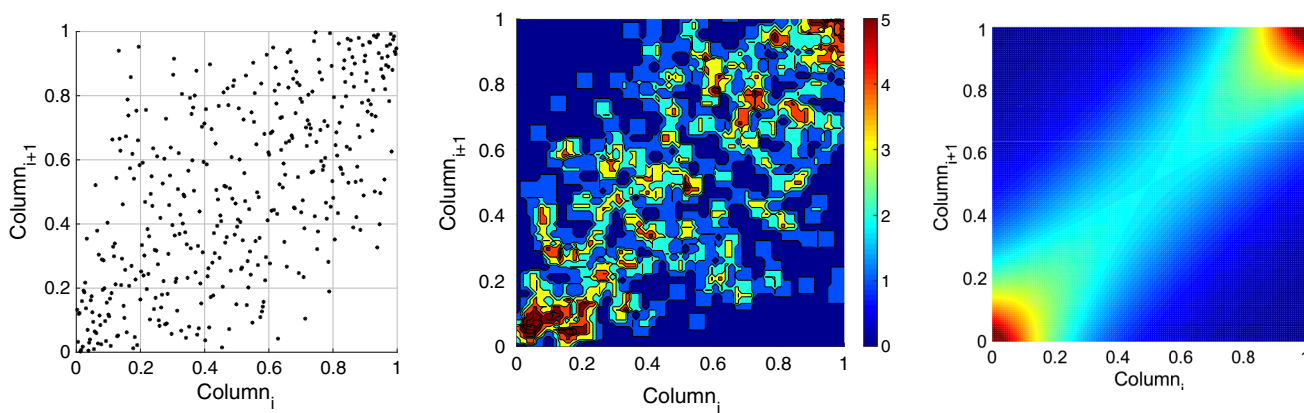


Fig. 7 Scatter plot (left) two adjacent columns in the residual matrix. The empirical Copula (middle) is estimated based on the dependency structure of two columns. The Frank Copula with $\theta = 4.79$ is fitted to the empirical Copula (right)

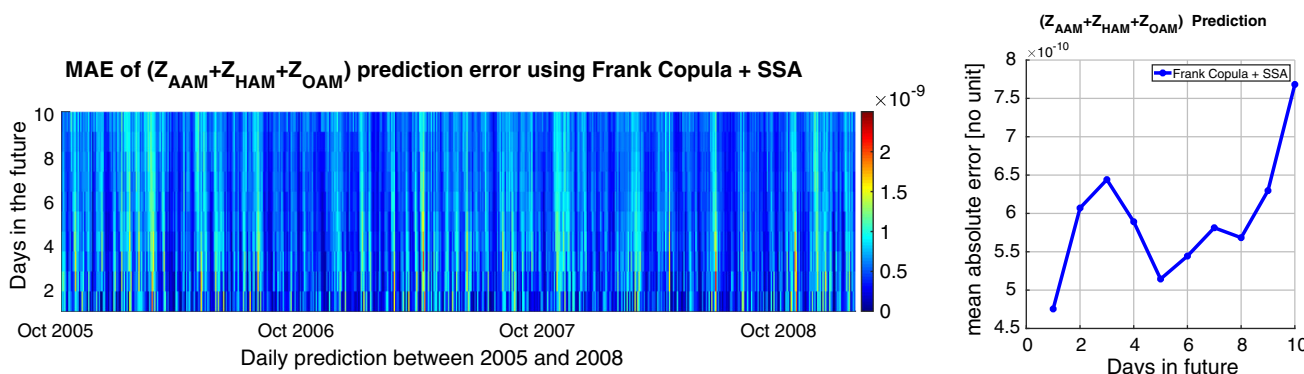


Fig. 8 MAE of $(Z_{AAM} + Z_{HAM} + Z_{OAM})$ prediction between 2005 and 2008 (left). The MAE of $(Z_{AAM} + Z_{HAM} + Z_{OAM})$ prediction between 2005 and 2008 (right)

3.3 LOD prediction from predicted EAM_Z using the calibrated Copula + SSA model

The predicted EAM_Z dataset from 2005 to 2008 is employed as input time series for the calibrated model. As can be seen in Fig. 1, the periodic terms in the residual part are predicted using SSA extrapolation as well. Finally, the Copula-based predicted data are added to the SSA-forecasted residual. To assess the proposed method, the results are compared with the EOP PCC solutions.

4 Discussion of the results

In this paper, a hybrid LOD prediction method Copula+SSA has been tested. The proposed combination method is examined based on the hind-cast experiments using the data from the past, i.e., the LOD data are predicted using the same time span (2005–2008) as the EOP PCC. Figure 9 shows the MAE of ultra-short-term prediction. The MAE of our hybrid Copula+SSA models indicates fewer errors compared to the EOP PCC solutions. However, the Kalman filter with

forecasted AAM shows a comparable performance with our proposed model with smaller MAE for the first 4 days in the future. Table 1 presents the MAE of Copula+SSA and EOP PCC results in numbers. Adoptive transform, AR, LS collocation, and NN show errors of more than 0.1 ms/day for the first day of prediction. Also, the MAE of the wavelet, LS extrapolation LS+AR, and He approaches reach more than 0.1 ms/day after 2 or 3 days in the future. From all contributions to the EOP PCC solution, Kalman filter provides the best accuracy. However, the MAE of the Kalman filter gets larger than 0.1 ms/day after 7 days. All Copula+SSA models show MAE smaller than 0.1 ms/day over 10-day prediction, except smaller Joe Copula+SSA. Figure 10 presents the MAE of 10-day-ahead prediction between 2005 and 2008 for all six hybrid models to understand better the prediction performance and its causes. Different features can be seen, and there are some common patterns such as high errors from the beginning of 2007 which might be caused by the El Niño effect or certain geomagnetic jerk events as pointed out by Shirai et al. (2005) or Malkin (2013). As NOAA National Centers for Environmental Information, State of the Climate reported, the El Niño warm event had a peak in December

Table 1 Comparison of Copula+SSA prediction and EOP PCC prediction errors (unit: ms/day)

Prediction day	1	2	3	4	5	6	7	8	9	10
Archi 12+SSA	0.047	0.060	0.063	0.063	0.059	0.064	0.066	0.061	0.072	0.083
Archi 14+SSA	0.050	0.065	0.066	0.068	0.064	0.067	0.072	0.064	0.073	0.082
Clayton+SSA	0.051	0.067	0.079	0.085	0.079	0.078	0.084	0.078	0.086	0.093
Gumbel+SSA	0.052	0.065	0.073	0.076	0.070	0.076	0.085	0.080	0.082	0.094
Frank+SSA	0.047	0.062	0.070	0.086	0.083	0.081	0.084	0.085	0.092	0.097
Joe+SSA	0.052	0.079	0.081	0.088	0.097	0.102	0.116	0.111	0.120	0.121
Kalman filter	0.042	0.051	0.057	0.062	0.071	0.084	0.094	0.107	0.119	0.128
wavelet	0.096	0.131	0.164	0.197	0.233	0.258	0.271	0.322	0.313	0.398
LSE	0.061	0.088	0.107	0.117	0.128	0.138	0.151	0.163	0.171	0.173
LS+AR EOP PC	0.070	0.097	0.118	0.133	0.142	0.143	0.154	0.171	0.179	0.188
Adaptive transform	0.165	0.158	0.162	0.159	0.160	0.160	0.154	0.179	0.528	0.593
AR	0.154	0.182	0.183	0.193	0.207	0.216	0.224	0.239	0.253	0.261
LSC	0.176	0.222	0.245	0.266	0.276	0.275	0.264	0.255	0.254	0.255
NN	0.161	0.196	0.218	0.237	0.250	0.257	0.256	0.257	0.264	0.274
HE	0.093	0.157	0.200	0.235	0.257	0.281	0.289	0.279	0.273	0.266

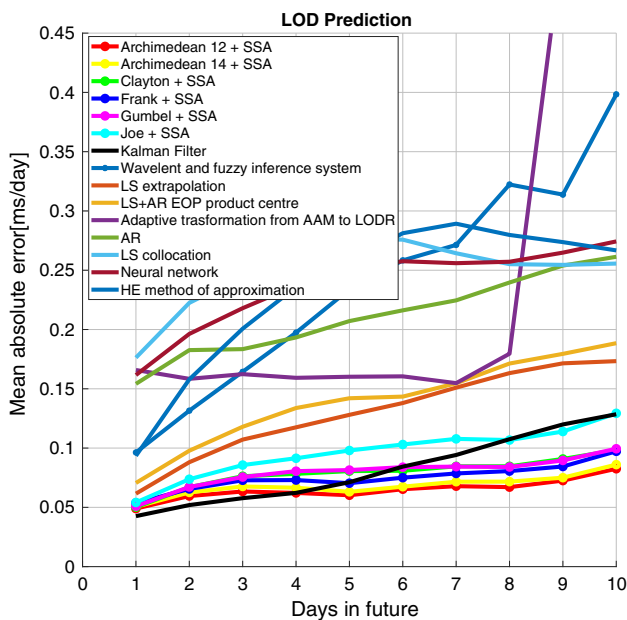


Fig. 9 Mean absolute errors of the predicted LOD using Archimedean 12+SSA, Archimedean 14+SSA, Clayton Copula+SSA, Gumbel Copula+SSA, Frank Copula+SSA, Joe Copula+SSA, and EOP PCC results

2006 and started to dissipate during January 2007. Thus, the equatorial Pacific sea-surface temperature (SST) anomalies decreased during the first two months of 2007, eventually declining to near average by the end of March 2007. Kalarus et al. (2010) suggested benefiting from the prediction of the AAM, OAM, and HAM. Therefore, our better prediction performance may be due to considering both mass and motion terms of EAM_Z for modeling the dependence structure between LOD and EAM_Z. It is important to note that the Copula+SSA Archimedean 12 and 14 Copula provide

significantly smaller errors than the other methods. On the other hand, the Joe Copula exhibits slightly larger errors than the two aforesaid models. This may have been caused by Archimedean 12 and 14 Copula’s ability to capture the upper and lower heavy tail dependence structure.

5 Summary and Conclusion

LOD represents the variation in Earth’s rotation rate which is most difficult to predict, because of the occurrence of extreme events in the LOD signal. In this paper, we introduce several approaches based on Copulas which were applied to bivariate frequency analysis. Using Copula is promising since it allows to take into account a wide range of correlation, frequently observed in time series. The presented work here is aimed at the possibility of utilizing the EAM_Z data to predict LOD data due to the existing relationship between them. In order to study this relationship, two datasets were compared: the observed LOD from IERS EOP C05 and EAM_Z derived from GFZ. The comparison with results of other methods indicates that the Copula+SSA can efficiently and precisely predict the LOD parameter at ultra-short term. All of our methods introduced here provide comparable error with the existing methods used for their evaluation in the time interval considered of up to 10 days. Besides, it is clearly demonstrated that the predicted AAM, HAM, OAM time series as additional input information can improve the LOD prediction. Among the analyzed combinations, the Archimedean 12+SSA and Archimedean 14+SSA show the most sophisticated performance with low errors. As the Kalman filter prediction provides better results within the first 3 to 4 days, we will investigate this topic further in future in order to find out how the EAM functions can deliver even better LOD predictions. The EOP PCC proved once more to be very useful.

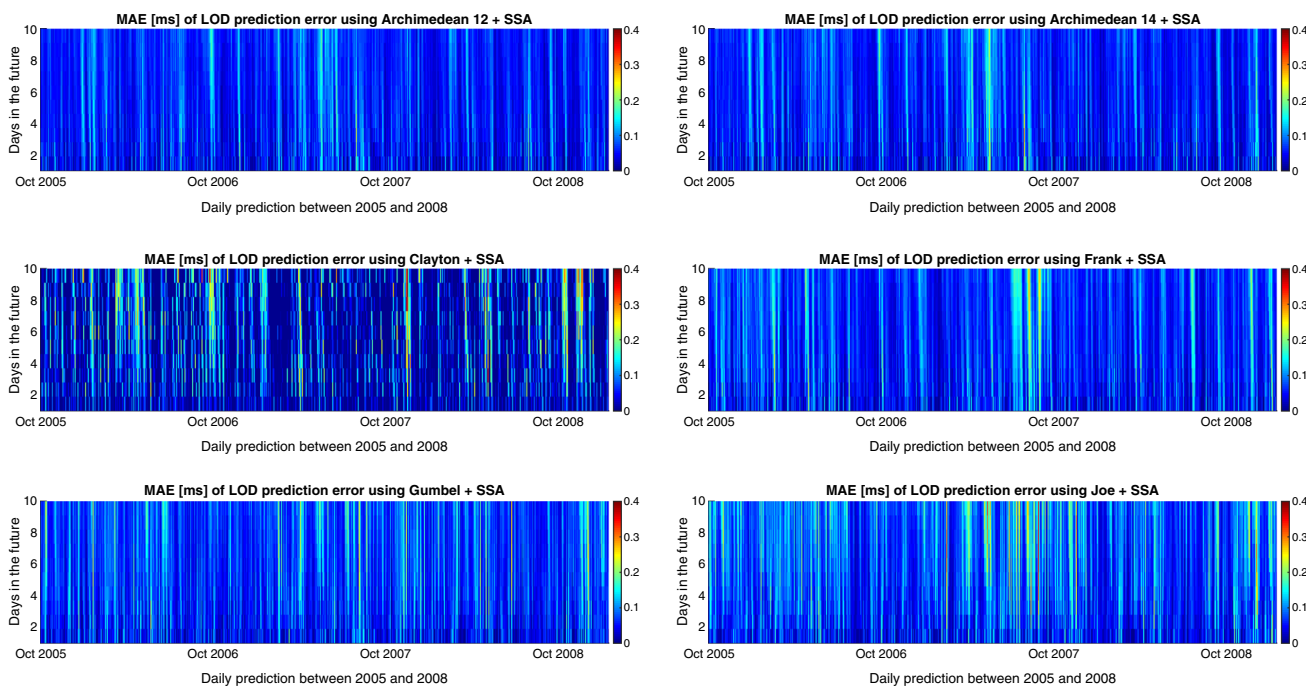


Fig. 10 Absolute errors of the predicted LOD using Archimedean 12 + SSA, Archimedean 14 + SSA, Clayton Copula + SSA, Gumbel Copula + SSA, Frank Copula + SSA, Joe Copula + SSA

As long as the data are still available for post-processing, new methods can adequately be compared in a consistent way to the methods applied in the past.

Acknowledgements Open Access funding provided by Projekt DEAL. We would like to thank the three anonymous reviewers for their insightful comments, which led to improved presentation of the results. We thank the groups from the EOP PCC (especially Dr. M. Kalarus) for sharing their results and knowledge with us. JMF and SB works were partially supported by projects AYA2016-79775-P (AEI/FEDER, UE). Also, SB was supported by the European Research Council (ERC) under the ERC-2017-STG SENTIFLEX project (Grant Agreement 755617).

Author contributions SM did most of the data analysis and writing of the manuscript. MH carried out the SSA studies and wrote a part of the manuscript. SB conceived and designed the study. RH, JMF, and HS participated in the design of the study and helped to improve the manuscript. All authors read and approved the final manuscript.

Data Availability Statement The data that support the findings of this study are available from the corresponding author upon reasonable request.

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

A Copula-based analysis

A.1 Empirical Copula

The empirical Copula approximates the unknown theoretical Copula distribution. The empirical Copula is purely based on the data, and it is defined in the rank space as follows (Genest and Rivest 1993; Genest and Favre 2007; Laux et al. 2011):

$$C_e(u, v) = \frac{1}{n} \sum_{i=1}^n \mathbf{1} \left(\frac{r_i}{n+1} \leq u, \frac{s_i}{n+1} \leq v \right) \tag{3}$$

where

$(r_1), (r_2) \dots, (r_n)$ denote the pairs of ranks of the variable $(x_1), (x_2), \dots, (x_n)$,

$(s_1), (s_2) \dots, (s_n)$ denote the pairs of ranks of the variable $(y_1), (y_2), \dots, (y_n)$,

n is the length of the data vector,

$\mathbf{1}(\dots)$ is the indicator function. If the condition is true, the indicator function is equal to 1. Otherwise, the indicator function is equal to 0.

A.2 Archimedean Copula

Some Copulas can be estimated directly with the simple form. They are named Archimedean Copulas. An Archimedean

Copula can be described in the following form:

$$C(u, v) = \phi^{-1}\{\phi(u) + \phi(v), \theta\} \tag{4}$$

where θ is the Copula parameter and the function ϕ is the generator of the Copula with the following characteristics (Nelsen 2006):

- for all $u \in (0, 1)$, $\phi(u) < 0$, ϕ is decreasing and convex (Table 2),
- $\phi(1) = 0$,

and ϕ^{-1} is defined by

$$\phi^{-1}(t) = \begin{cases} \phi^{-1}(t; \theta), & \text{if } 0 \leq t \leq \phi(0) \\ 0, & \text{if } \phi(0) \leq t \leq \infty \end{cases} \tag{5}$$

A.2.1 The relationship between Copula parameter θ and classical dependence parameter

There is a functional relationship between classical dependence parameters such as Kendall τ and Copula parameters. For one-parametric Copulas, the functional relationship between the Kendall τ Copula functions, namely:

$$\tau = 4 \int_0^1 \int_0^1 C_\theta(u, v) dC_\theta(u, v) - 1, \tag{6}$$

Equation 7 can be used to estimate the Copula parameter. Kendal τ for Archimedean Copula with the generator $\Phi(t)$ is shown:

$$\tau = 1 + 4 \int_0^1 \frac{\Phi(t)}{\Phi(t)} dt. \tag{7}$$

Thus, the relations between the Kendall τ and the Archimedean Copula parameters are illustrated in Table 3. Here D is Debye functions.

$$D_k(k) = \frac{k}{x^k} \int_{t=0}^x \frac{t}{e^t - 1} dt$$

A.3 Simulating from Copula-based conditional random data

This subsection provides the essential steps for data simulation using Copula-based conditional random data. The following steps are taken to fit the proper theoretical Copula function and simulation data (Laux et al. 2011; Vogl et al. 2012; Modiri et al. 2018).

Table 2 Six ordinary families of Archimedean Copulas (Archimedean 12, Archimedean 14, Clayton, Frank, Gumbel, and Joe Copula) and generator, parameter space, and their formula

Family	Generator	Parameter	Formula
Archimedean 12	$\phi^{Arch12}(t) = (\frac{1}{t} - 1)^\theta$	$1 \leq \theta$	$C_\theta(u, v) = (1 + [(u^{-1} - 1)^\theta + (v^{-1} - 1)^\theta]^{1/\theta})^{-1}$
Archimedean 14	$\phi^{Arch14}(t) = (t^{1/\theta} - 1)^\theta$	$1 \leq \theta$	$C_\theta(u, v) = (1 + [(u^{1/\theta} - 1)^\theta + (v^{1/\theta} - 1)^\theta]^{1/\theta})^{-\theta}$
Clayton	$\phi^{Cl}(x) = \frac{1}{\theta}(t^{-\theta} - 1)$	$-1 \leq \theta$	$C_\theta(u, v) = \max[(u^{-\theta} + v^{-\theta} - 1), 0]^{-1/\theta}$
Frank	$\phi^{Fr}(t) = -\ln\{\frac{e^{-\theta t} - 1}{e^{-\theta} - 1}\}$	$-\infty < \theta < \infty$	$C_\theta(u, v) = \frac{1}{\theta} \ln(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1})$
Gumbel	$\phi^{Gu}(t) = (-\ln t)^\theta$	$1 \leq \theta$	$C_\theta(u, v) = e^{-((-\ln(u)^\theta) + (-\ln(v)^\theta))^{1/\theta}}$
Joe	$\phi^{Joe}(t) = -\ln[1 - (1 - t)^\theta]$	$1 \leq \theta$	$C_\theta(u, v) = 1 - [(1 - u)^\theta + (1 - v)^\theta - ((1 - u)^\theta(1 - v)^\theta)]^{1/\theta}$

θ is the parameter of the Copula called the dependence parameter, which measures the dependence between the marginal (Nelsen 2006)

Table 3 The link between Archimedean Copula parameter θ and Kendall τ (Cherubini et al. 2004)

Family	τ
Archimedean 12	$\frac{\theta - \frac{2}{3}}{\theta}$
Archimedean 14	$\frac{2\theta - 1}{2\theta + 1}$
Clayton	$\frac{\theta}{\theta + 2}$
Frank	$1 - \frac{4}{\theta} [1 - D_1(\theta)]^*$
Gumbel	$\frac{\theta - 1}{\theta}$
Joe	$1 - \frac{4}{\theta} D_1(\theta)$

* $D_k(x)$ is the Debye function for any positive integer k

1. Independent identical distribution (iid)-transformation of input time series.
2. Compute the marginal distributions $F_X(x)$ and $F_Y(y)$ of the input data x and y .
3. Transform data to rank space using the estimated marginal distributions of data with u_i and v_i in rank space.
4. Compute the empirical Copula to the dependence structure of random variables using the rank-transformed data.
5. Fit a theoretical Copula function $C_\theta(u, v)$.
6. Compute the conditional Copula function.
7. Sample random data from the conditional Copula cumulative distribution function (CDF).
8. Transfer the sample back to the data space using the inverse marginal.

B Singular Spectrum Analysis

In the first step which is called embedding, we form a trajectory matrix (\mathbf{X}) by moving a window of length L over the elements of the time series (f_i), i.e.,

$$\mathbf{X} = \begin{bmatrix} f_1 & f_2 & f_3 & \dots & f_K \\ f_2 & f_3 & f_4 & \dots & f_{K+1} \\ f_3 & f_4 & f_5 & \dots & f_{K+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_L & f_{L+1} & f_{L+2} & \dots & f_N \end{bmatrix}, \begin{cases} 1 < L < K \\ K = N - L + 1 \end{cases} \quad (8)$$

The matrix \mathbf{X} is a symmetric matrix having identical elements on anti-diagonals. In the next step, we apply a singular value decomposition (SVD) to the trajectory matrix \mathbf{X} , i.e.,

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (9)$$

with the superscript T the transpose operator. The matrices \mathbf{U} and \mathbf{V} are orthonormal and are called left and right singular vectors, respectively. $\mathbf{\Sigma}$ is a diagonal matrix containing nonnegative entries, the singular values, which reflect the importance of the singular vectors.

If $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq 0$ denote diagonal entries of $\mathbf{\Sigma}$, the trajectory matrix can be written as:

$$\begin{cases} \mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 + \dots + \mathbf{X}_d \\ \mathbf{X}_i = \lambda_i \mathbf{U}_i \mathbf{V}_i^T \end{cases} \quad (10)$$

where λ_i is the corresponding singular value of \mathbf{X}_i and d is the number of nonzero singular values ($d \leq L$).

In the next step, the grouping step, we select a group from $\{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_d\}$ which will be used for reconstructing a new version of the time series. For instance, for reproducing a representative trend of the time series, we would choose the most critical part of \mathbf{X} which is retrievable using the first few singular vectors and their corresponding singular values.

The last step is dedicated to the reconstruction of the time series using the selected group. Based on the fact that the original trajectory matrix had the same entries on anti-diagonals, we reconstruct a model of the time series ($G = (g_1, g_2, \dots, g_N)$) by anti-diagonal averaging:

$$g_i = \begin{cases} \frac{1}{i} \sum_{m=1}^i \hat{x}_{m,i-m+1} & 1 \leq i < L \\ \frac{1}{L} \sum_{m=1}^L \hat{x}_{m,i-m+1} & L \leq i \leq K \\ \frac{1}{N-i+1} \sum_{m=i-K+1}^{N-K+1} \hat{x}_{m,i-m+1} & K < i \leq N \end{cases} \quad (11)$$

where $\hat{x}_{i,j}$ is an estimation of the element f_{i+j-1} of the original time series.

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

Alexandrov T (2009) A method of trend extraction using singular spectrum analysis. REVSTAT Stat J 7(1):1–22

Bárdossy A, Li J (2008) Geostatistical interpolation using copulas. Water Resour Res 44(7):W0412

Bárdossy A, Pegram G (2009) Copula based multisite model for daily precipitation simulation. Hydrol Earth Syst Sci 13(12):2299–2317

Cherubini U, Luciano E, Vecchiato W (2004) Copula methods in finance. Wiley, Hoboken

Dill R (2008) Hydrological model LSDM for operational Earth rotation and gravity field variations, vol 369. Scientific Technical Report, 08/09, Deutsches GeoForschungsZentrum, Potsdam, Germany

Dill R, Dobsław H, Thomas M (2019) Improved 90-day earth orientation predictions from angular momentum forecasts of atmosphere, ocean, and terrestrial hydrosphere. J Geodesy 93(3):287–295

Dobsław H (2016) Homogenizing surface pressure time-series from operational numerical weather prediction models for geodetic applications. J Geod Sci 6(1):61–68

Dobsław H, Dill R (2018) Predicting earth orientation changes from global forecasts of atmosphere–hydrosphere dynamics. Adv Space Res 61(4):1047–1054

- Dobslaw H, Thomas M (2005) Atmospheric induced oceanic tides from ECMWF forecasts. *Geophys Res Lett* 32(10):L10,615
- Embrechts P, McNeil A, Straumann D (2002) Correlation and dependence in risk management: properties and pitfalls. In: Dempster MAH (ed) *Risk management: value at risk and beyond*. Cambridge University Press, Cambridge, pp 176–223
- Escarela G, Carriere JF (2003) Fitting competing risks with an assumed copula. *Stat Methods Med Res* 12(4):333–349
- Freedman A, Steppe J, Dickey J, Eubanks T, Sung LY (1994) The short-term prediction of universal time and length of day using atmospheric angular momentum. *J Geophys Res Solid Earth* 99(B4):6981–6996
- Gambis D (2004) Monitoring earth orientation using space-geodetic techniques: state-of-the-art and prospective. *J Geodesy* 78(4–5):295–303
- Genest C, Favre AC (2007) Everything you always wanted to know about copula modeling but were afraid to ask. *J Hydrol Eng* 12(4):347–368
- Genest C, Rivest LP (1993) Statistical inference procedures for bivariate archimedean copulas. *J Am Stat Assoc* 88(423):1034–1043
- Ghil M, Allen M, Dettinger M, Ide K, Kondrashov D, Mann M, Robertson AW, Saunders A, Tian Y, Varadi F et al (2002) Advanced spectral methods for climatic time series. *Rev Geophys* 40(1):3–1
- Golyandina N, Nekrutkin V, Zhigljavsky AA (2001) *Analysis of time series structure: SSA and related techniques*. Chapman and Hall/CRC, Boca Raton
- Gross RS (2015) Theory of earth rotation variations. In: VIII Hotine-Marussi symposium on mathematical geodesy. Springer, pp 41–46
- Gross RS, Marcus SL, Eubanks TM, Dickey JO, Keppenne CL (1996) Detection of an enso signal in seasonal length-of-day variations. *Geophys Res Lett* 23(23):3373–3376
- Gross RS, Eubanks T, Steppe J, Freedman A, Dickey J, Runge T (1998) A Kalman-filter-based approach to combining independent earth-orientation series. *J Geodesy* 72(4):215–235
- Groth A, Ghil M (2015) Monte carlo singular spectrum analysis (SSA) revisited: detecting oscillator clusters in multivariate datasets. *J Clim* 28(19):7873–7893
- Holton JR, Dmowska R (1989) El Niño, La Niña, and the southern oscillation, vol 46. Academic press, San Diego
- Hoseini M, Alshawaf F, Nahavandchi H, Dick G, Wickert J (2020) Towards a zero-difference approach for homogenizing gnss tropospheric products. *GPS Solut* 24(1):8
- Joe H (1997) *Multivariate models and multivariate dependence concepts*. Chapman and Hall/CRC, Boca Raton
- Johnson TJ, Luzum BJ, Ray JR (2005) Improved near-term earth rotation predictions using atmospheric angular momentum analysis and forecasts. *J Geodyn* 39(3):209–221
- Jungclauss J, Fischer N, Haak H, Lohmann K, Marotzke J, Matei D, Mikolajewicz U, Notz D, Storch J (2013) Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-earth system model. *J Adv Model Earth Syst* 5(2):422–446
- Kalarus M, Schuh H, Kosek W, Akyilmaz O, Bizouard C, Gambis D, Gross R, Jovanović B, Kumakshev S, Kutterer H et al (2010) Achievements of the earth orientation parameters prediction comparison campaign. *J Geodesy* 84(10):587–596
- Kosek W (2002) Autocovariance prediction of complex-valued polar motion time series. *Adv Space Res* 30(2):375–380
- Kosek W (2012) Future improvements in EOP prediction. In: *Geodesy plant earth*. International Association Geodesy Symposia, vol 136, pp. 513–520
- Kosek W, McCarthy D, Luzum B (1998) Possible improvement of earth orientation forecast using autocovariance prediction procedures. *J Geodesy* 72(4):189–199
- Kosek W, Kalarus M, Johnson T, Wooden W, McCarthy D, Popinski W (2005) A comparison of LOD and UT1-UTC forecasts by different combined prediction techniques. *Artif Satell* 40(2):119–125
- Laux P, Vogl S, Qiu W, Knoche H, Kunstmann H (2011) Copula-based statistical refinement of precipitation in RCM simulations over complex terrain. *Hydrol Earth Syst Sci* 15(7):2401–2419
- Lei Y, Zhao D, Cai H (2015) Prediction of length-of-day using extreme learning machine. *Geodesy Geodyn* 6(2):151–159
- Lei Y, Guo M, Hu Dd, Hb Cai, Zhao Dn Hu, Zp Gao Yp (2017) Short-term prediction of UT1-UTC by combination of the grey model and neural networks. *Adv Space Res* 59(2):524–531
- Liao D, Wang Q, Zhou Y, Liao X, Huang C (2012) Long-term prediction of the earth orientation parameters by the artificial neural network technique. *J Geodyn* 62:87–92
- Lichten SM, Marcus SL, Dickey JO (1992) Sub-daily resolution of earth rotation variations with global positioning system measurements. *Geophys Res Lett* 19(6):537–540
- Malkin Z (2013) Free core nutation and geomagnetic jerks. *J Geodyn* 72:53–58
- McCarthy DD, Luzum BJ (1991) Prediction of earth orientation. *Bull Géodésique* 65(1):18–21
- Modiri S, Lorenz C, Sneeuw N, Kunstmann H (2015) Copula-based estimation of large-scale water storage changes: exploiting the dependence structure between hydrological and grace data. In: EGU general assembly conference abstracts, vol 17
- Modiri S, Belda S, Heinkelmann R, Hoseini M, Ferrándiz JM, Schuh H (2018) Polar motion prediction using the combination of ssa and copula-based analysis. *Earth Planets Space* 70(1):115
- Moskvina V, Zhigljavsky A (2003) An algorithm based on singular spectrum analysis for change-point detection. *Commun Stat Simul Comput* 32(2):319–352
- Nastula J, Gross R, Salstein DA (2012) Oceanic excitation of polar motion: identification of specific oceanic areas important for polar motion excitation. *J Geodyn* 62:16–23
- Nelsen RB (2006) *An introduction to copulas*, 2nd edn. Springer, New York
- Niedzielski T, Kosek W (2008) Prediction of UT1-UTC, LOD and AAM χ_3 by combination of least-squares and multivariate stochastic methods. *J Geodesy* 82(2):83–92
- Patton AJ (2006) Modelling asymmetric exchange rate dependence. *Int Econ Rev* 47(2):527–556
- Patton AJ (2009) Copula-based models for financial time series. In: Mikosch T, Kreiß JP, Davis R, Andersen T (eds) *Handbook of financial time series*. Springer, Berlin, pp 767–785
- Petit G, Luzum B (2010) IERS conventions (2010). Technical report, Bureau International des Poids et Mesures Sevres (France)
- Plag HP, Pearlman M (2009) *Global geodetic observing system: meeting the requirements of a global society on a changing planet in 2020*. Springer, Berlin
- Rachev S, Mittnik S (2000) *Stable Paretian models in finance*. Wiley, New York
- Salstein DA (1993) Monitoring atmospheric winds and pressures for earth orientation studies. *Adv Space Res* 13(11):175–184
- Salvadori G, De Michele C (2007) On the use of copulas in hydrology: theory and practice. *J Hydrol Eng* 12(4):369–380
- Schuh H, Schmitz-Hübsch H (2000) Short period variations in earth rotation as seen by VLBI. *Surv Geophys* 21(5–6):499–520
- Schuh H, Ulrich M, Egger D, Müller J, Schwegmann W (2002) Prediction of earth orientation parameters by artificial neural networks. *J Geodesy* 76(5):247–258
- Shirai T, Fukushima T, Malkin Z (2005) Detection of phase disturbances of free core nutation of the earth and their concurrence with geomagnetic jerks. *Earth Planets Space* 57(2):151–155
- Sklar M (1959) Fonctions de repartition an dimensions et leurs marges. *Publ Inst Stat Univ Paris* 8:229–231

- Tapley B, Schutz B, Eanes R (1985) Station coordinates, baselines, and earth rotation from Lageos laser ranging: 1976–1984. *J Geophys Res Solid Earth* 90(B11):9235–9248
- Verhoest NE, van den Berg MJ, Martens B, Lievens H, Wood EF, Pan M, Kerr YH, Al Bitar A, Tomer SK, Drusch M et al (2015) Copula-based downscaling of coarse-scale soil moisture observations with implicit bias correction. *IEEE Trans Geosci Remote Sens* 53(6):3507–3521
- Vogl S, Laux P, Qiu W, Mao G, Kunstmann H (2012) Copula-based assimilation of radar and gauge information to derive bias-corrected precipitation fields. *Hydrol Earth Syst Sci* 16(7):2311–2328
- Wang W, Wells MT (2000) Model selection and semiparametric inference for bivariate failure-time data. *J Am Stat Assoc* 95(449):62–72
- Wu F, Chang G, Deng K (2019) One-step method for predicting LOD parameters based on LS+AR model. *J Spat Sci*. <https://doi.org/10.1080/14498596.2019.1618401>
- Xu X, Zhou Y (2015) Eop prediction using least square fitting and autoregressive filter over optimized data intervals. *Adv Space Res* 56(10):2248–2253