IS THE LENGTH-OF-DAY TIME SERIES NORMALLY DISTRIBUTED?

A.K. SEN[1], T. NIEDZIELSKI[2,3], W. KOSEK[2]
[1]Department of Mathematical Sciences, Indiana University
402 N. Blackford Street, Indianapolis, IN 46202, USA
e-mail: asen@iupui.edu
[2]Space Research Centre, Polish Academy of Sciences
Bartycka 18A, 00-716 Warsaw, Poland
e-mail: niedzielski@cbk.waw.pl, kosek@cbk.waw.pl
[3]Institute of Geography and Regional Development, University of Wrocław
pl. Uniwersytecki 1, 50-137 Wrocław, Poland

ABSTRACT. The non-tidal LOD data are analysed (data span 01.01.1962-09.01.2008) in order to provide the probabilistic characteristics of the Earth rotation rate fluctuations. The skewness of the LOD probability distribution is of -0.31 indicating that the probability distribution is asymmetrical. Moreover, the residual non-tidal LOD data is considered (after removal of the semiannual, annual, 9.3-years, and 18.6-years oscillations, and linear trend). The skewness of this residual data equals to -0.64 and indicates an increased asymmetry in the distribution. For the non-transformed LOD data the kurtosis is of 2.31 and it shows that the distribution is flattened. The kurtosis for the residuals is 5.64 indicating that the distribution is more peaked than a normal distribution. For the non-transformed LOD data, the Jonhson SB distribution provides the best fit. For the residual LOD data, the Johnson SU distribution is found to be the most appropriate model. Both the LOD and it's residual time series are appropriately modeled by probability laws that are different from a normal distribution.

1. INTRODUCTION

Many applications of time series analysis, e.g. prediction methods, require the assumption that the data are normally distributed. If the time series consists of the extreme values, the underlying probability law of data is rather different that a normal (or Gaussian) distribution. In particular, the large-scale El Niño/Southern Oscillation (ENSO) causes the extreme stochastic fluctuations in the Earth rotation rate. The Earth rotation rate is described by length-of-day (LOD, also denoted by Δ) or its integral, i.e. UT1-UTC Universal Time. The LOD increases during El Niño (due to the collapse of the tropical easterly winds), and decreases during La Niña.

The fluctuations in Δ (or equivalently UT1-UTC) are of tidal and non-tidal origins. The oscillations caused by the tides with periods from 5 days to 18.6 years are caused by gravitational interactions among the Earth, the Moon, the Sun, and - less significantly - other planets of our solar system. The tidal correction $\delta\Delta$ is determined using the model defined by the International Earth Rotation and Reference Systems Service (IERS) Conventions 2003 (McCarthy and Petit, 2004).

We consider the two time series: (1) the non-tidal length-of-day time series ($\Delta - \delta\Delta$) for the period from 01.01.1962 to 09.01.2008 computed by removing the tidal model $\delta\Delta$ from Δ time series eopc04_IAU2000.62-now (IERS, 2008) (Fig. 1(a)), and (2) the residual $\Delta - \delta\Delta$ time series denoted by $\varepsilon(\Delta - \delta\Delta)$ (after removal of the least-squares model which consists of semiannual, annual, 9.3-year, and 18.6-year oscillations, and the linear trend) (Fig. 1(b)).

Many researchers analysed $\Delta - \delta\Delta$ time series using the data processing techniques (Kosek, 1997; Kosek et al., 1998; Schuh et al., 2002; Akyilmaz and Kutterer, 2004; Niedzielski and Kosek, 2008). The predictions of $\Delta - \delta\Delta$ determined by time series methods are based upon the decomposition of $\Delta - \delta\Delta$ time series into deterministic components and stochastic residuals. Several forecasting methods assume that the residuals are normally distributed (Kosek et al., 2005; Niedzielski and Kosek, 2008).

A normal distribution is characterized by the mean and standard deviation. This distribution exhibits the zero skewness and a kurtosis value of 3. First, we discuss if a normal distribution fits both $\Delta - \delta\Delta$

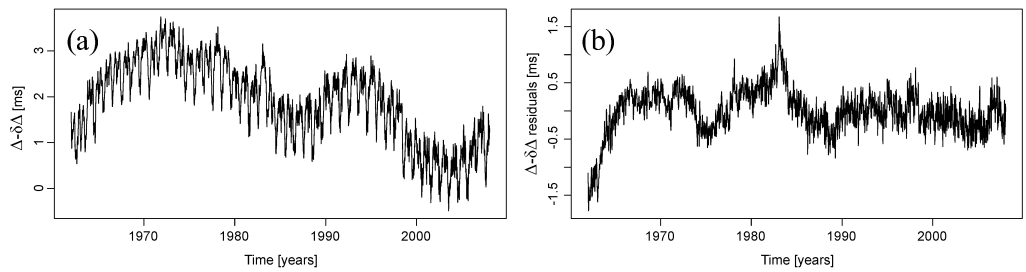

Figure 1: The $\Delta - \delta\Delta$ (a) and the $\varepsilon(\Delta - \delta\Delta)$ (b) time series.

and its residuals. Second, several non-Gaussian probability distributions are fitted to these time series.

2. FITTING A NORMAL DISTRIBUTION

Table 1 shows that the distributions of both $\Delta - \delta\Delta$ and $\varepsilon(\Delta - \delta\Delta)$ time series are skewed. The probability distribution of $\varepsilon(\Delta - \delta\Delta)$ reveals the increased asymmetry in respect to the asymmetry of $\Delta - \delta\Delta$ distribution. In both cases, the heavy-tail is present in the left-hand side of the probability density function. The analysis of kurtosis shows that the distribution of $\Delta - \delta\Delta$ is more flat than the normal distribution. In contrast, the probability density function of $\varepsilon(\Delta - \delta\Delta)$ is more peaked than the Gaussian one.

	Skewness	Kurtosis
$\Delta - \delta\Delta$	-0.31	2.31
$\varepsilon(\Delta - \delta\Delta)$	-0.64	5.64

Table 1: The skewness and kurtosis of the non-tidal LOD time series and its residuals.

The departure of $\Delta-\delta\Delta$ from the normal distribution is also detected (a) by comparing the probability density function (pdf) with the histogram of the data, (b) by comparing the cumulative distribution function (cdf) with the cdf of the data, (c) using a Probability-Probability (P-P) plot, and (d) using a Quantile-Quantile (Q-Q) plot (Fig. 2). The P-P plot is a plot of the empirical cdf values depicted against the theoretical cdf values. The Q-Q plot is a graph of the data plotted against the quantiles of the theoretical distribution. Both the P-P plot and the Q-Q plot are widely used to evaluate the goodness-of-fit (Gilchrist, 2000; Holmgren, 1995; Reiss & Thomas, 2007; Wasserman, 2002). These plots can be used to test how closely a theoretical distribution compares with the empirical distribution of the data, and to compare several fitted distributions. Figures 2(a) and 2(b) show that the empirical distribution of $\Delta - \delta\Delta$ differs from the normal distribution. The analysis of both the P-P plot (Fig. 2(c)), and the Q-Q plot (Fig. 2(d)) supports the above-mentioned conclusions.

The similar analyses confirm that the distribution of the residual time series $\varepsilon(\Delta - \delta\Delta)$ differs from the normal probability law. Indeed, Figures 3(a) and 3(b) indicate that the pdf (cdf) of the normal distribution does not fit the histogram (empirical distribution function) of $\varepsilon(\Delta - \delta\Delta)$. Similarly, the P-P plot (Fig. 3(c)) and the Q-Q plot (Fig. 3(d)) confirm the inadequacy of the normal distribution.

In the case of $\Delta - \delta\Delta$ time series, the departure from the normal distribution can be caused by the long-term oscillations driven by geological processes acting between the core and the Earth's mantle. However, the non-Gaussian distribution of $\varepsilon(\Delta - \delta\Delta)$ data is probably due to ENSO episodes which are recorded as extreme spikes in this time series.

3. FITTING NON-GAUSSIAN DISTRIBUTIONS

Following the fact, that the distributions of both $\Delta - \delta\Delta$ and $\varepsilon(\Delta - \delta\Delta)$ are different than a normal probability law, several non-Gaussian distributions are examined.

In the case of $\Delta - \delta\Delta$ time series, the 3-parameter generalized extreme value (GEV) distribution

(Kotz & Nadarajah, 2001), 4-parameter Beta (Gupta & Nadarajah, 2004), and Johnson SB distributions (Bowman & Shenton 1983) provide good fit (Fig. 2). Among them, the Johnson SB distribution appears to provide the best fit to $\Delta - \delta\Delta$ data.

The analysis of the Q-Q plot (Fig. 2(d)) shows that for small quantiles, the Johnson SB distribution provides the best fit to the empirical distribution of $\Delta - \delta\Delta$ time series. The Johnson SB distribution is rather the best in modeling the probabilities around the mode (Fig. 2(a)). However, for high quantiles both GEV and 4-parameter Beta probability distributions provide the best fit to the histogram of $\Delta - \delta\Delta$.

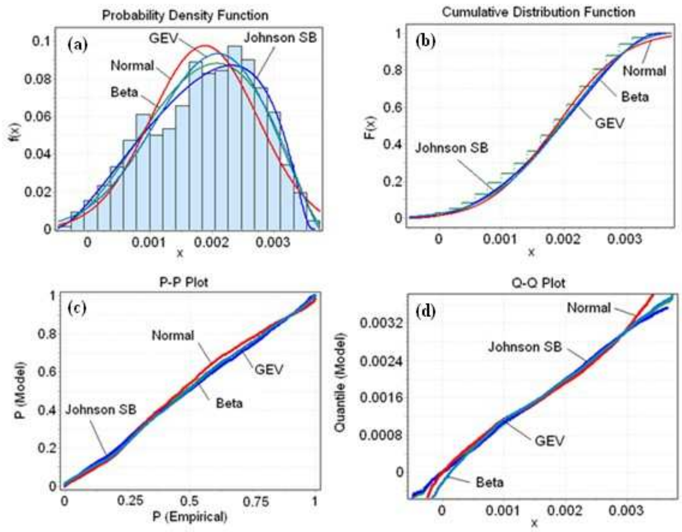

Figure 2: The analysis of $\Delta - \delta\Delta$ time series; the pdf vs. the histogram of the data (a), the cdf vs. the cdf of the data (b), the P-P plot (c), the Q-Q plot (d).

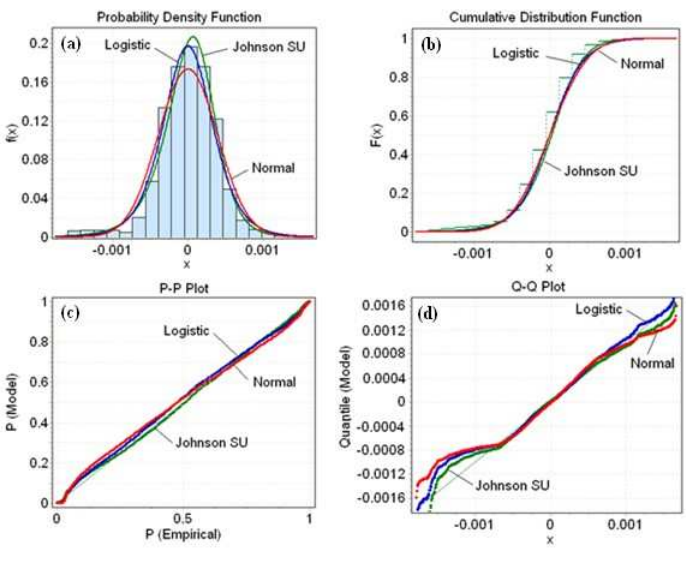

Figure 3: The analysis of $\varepsilon(\Delta - \delta\Delta)$ time series; the pdf vs. the histogram of the data (a), the cdf vs. the cdf of the data (b), the P-P plot (c), the Q-Q plot (d).

Regarding the residual $\varepsilon(\Delta - \delta\Delta)$ time series, the logistic distribution (Balakrishnan, 1991) and Johnson SU distribution (Bowman & Shenton, 1983) provide the appropriate fit (Fig. 3). In particular, the Q-Q plot (Fig. 3(d)) reveals that for high quantiles, the logistic distribution is more accurate than the Johnson SU distribution. The Johnson SU distribution outperforms the logistic distribution in the vicinity of the mode (Fig. 3(a)). For the low quantiles, however, the Johnson SU distribution provides the best fit.

4. CONCLUSIONS

The probability distribution of the non-tidal length-of-day time series cannot be accurately modeled by a normal distribution. Similarly, the distribution of the residual non-tidal length-of-day data differs from a normal probability law. We argue that these findings may be interpreted twofold. First, the departure from a normal distribution in the case of the non-tidal length-of-day time series is caused by the long-term oscillations with high amplitudes, probably due to the interactions between the core and the Earth's mantle. Second, the distribution of the residual non-tidal length-of-day data is different than a normal probability law probably due to ENSO events recorded as spikes in the residuals.

Several non-Gaussian probability distributions fit to the analysed time series. The best overall goodness-of-fit is provided by the family of the Johnson distribution. In the case of the non-tidal length-of-day time series and its residuals the Johnson SB and SU distributions, respectively, are found to be the most appropriate.

Acknowledgements. The research is supported by the Polish Ministry of Education and Science under the project No 4 T12E 039 29. Tomasz Niedzielski is supported by the Foundation for Polish Science within the Start Program (stipends for young researchers) and by the European Union within the Marie-Curie Actions. Tomasz Niedzielski and Wiesław Kosek are supported by the organizers of the Journées "Systèmes de référence spatio-temporels" 2008 and X. Lohrmann-Kolloquium, September 2008, Dresden, Germany.

5. REFERENCES

Akyilmaz, O., Kutterer, H., 2004, Prediction of Earth rotation parameters by fuzzy inference systems, J. Geodesy 78, pp. 82–93.

Balakrishnan, N., 1991, Handbook of the Logistic Distribution, Marcel Dekker, New York.

Bowman, K.O., Shenton, L.R., 1983, Johnson's system of distributions, Encyclopedia of Statistical Sciences, John Wiley & Sons, New York, 4, pp. 303–314.

Gilchrist, W., 2000, Statistical Modelling with Quantile Functions, Chapman and Hall, New York.

Gupta, A.K., Nadarajah, S., 2004, Handbook of beta functions and its applications, CRC Press, Boca Raton, Florida.

Holmgern, E.B., 1995, The P-P plot as a method of comparing treatment effects, J. Am. Stat. Assoc. 90, pp. 360–365.

IERS 2008, http://hpiers.obspm.fr/iers/eop/eopc04_05.

Kosek, W., 1997, Autocovariance prediction of short period Earth rotation parameters, Artif. Satell. 32, pp. 75–86.

Kosek, W., McCarthy, D.D., Luzum, B., 1998, Possible improvement of earth orientation forecast using autocovariance prediction procedures, J. Geodesy 72, pp. 189–199.

Kosek, W., Kalarus, M., Johnson, T.J., Wooden, W.H., McCarthy, D.D., Popiski, W., 2005, A comparison of LOD and UT1-UTC forecasts by different combined prediction techniques, Artif. Satell. 40, pp. 119–125.

Kotz, S., Nadarajah, S., 2001, Extreme Value Distributions: Theory and Applications, World Scientific, Singapore.

McCarthy, D.D., Petit, G., (Eds.), 2004, IERS Conventions, IERS Technical Note No. 32.

Niedzielski, T., Kosek, W., 2008, Prediction of UT1-UTC, LOD and AAM X3 by combination of the least-squares and multivariate stochastic methods, J. Geodesy 82, pp. 83–92.

Reiss, R.D., Thomas, M., 2007, Statistical Analysis of Extreme Values: with applications to Insurance, Finance, Hydrology and Other Fields, Birkhuser, Basel, Switzerland.

Schuh, H., Ulrich, M., Egger, D., Mueller, J., Schwegmann, W., 2002, Prediction of Earth orientation parameters by artificial neural networks, J. Geodesy 76, pp. 247–258.

Wasserman, G., 2002, Reliability Verification, Testing and Analysis in Engineering Design, CRC Press, Boca Raton, Florida.