Dependence of long-term persistence properties of precipitation on spatial and regional characteristics

Hristos Tyrallis, Panayiotis Dimitriadis, Theano Iliopoulou, Katerina Tzouka and Demetris Koutsoyiannis, National Technical University of Athens (itia.ntua.gr/1695)

1. Abstract

The long-term persistence (LTP), also known in hydrological science as the Hurst phenomenon, is a property observed in geophysical processes in which wet or dry years or days are clustered to respective long time periods. A common practice for evaluating the presence of the LTP is to test the mean annual precipitation trends in time series with the Hurst-Kolmogorov process (HKP) and estimate its Hurst parameter $H$ where high values of $H$ indicate strong LTP.

We estimate $H$ of the mean annual precipitation using instrumental data from approximately 1,500 stations which cover a large area of the earth's surface and span from 1916 to 2015. We report the $H$ estimates of all stations on their spatial and regional characteristics (i.e. their location, elevation and Köppen-Geiger climate class) using a random forest algorithm. Furthermore, we apply the Mann-Kendall test under the LTP assumption (HKP-LTP) to all time series to assess the significance of observed trends of the mean annual precipitation.

To summarize the results, the LTP seems to depend mostly on the location of the stations, while the predictive value of the fitted regression model is good. Thus, when investigating for LTP properties we recommend that they should be considered. Additionally, the application of the HKP-LTP suggests that no significant monotonic trend can characterize the global precipitation. Dominant positive significant trends are observed mostly in main climate types D (cwa), while in the other climate types the percentage of stations with positive significant trends was approximately equal to that of negative significant trends. Furthermore, 50% of all stations do exhibit significant trends at all.

2. Introduction

Long-term persistence (LTP) is an inherent property of geophysical processes in which wet or dry years or days are clustered to long time periods (Koutsoyiannis 2002).

The LTP can be modeled with the Hurst-Kolmogorov process (HKP) and characterized the magnitude of LTP (Koutsoyiannis 2003).

Estimation of $H$ is important in engineering practice (Lins and Cohn 2011).

Uncertainty increases substantially when $H$ is present in the model's parameters (Koutsoyiannis and Montanari 2007).

Significant trends under the independence assumption can be considered non-significant under the LTP assumption (Fatichi et al. 2012).

A few studies examine the LTP properties of global precipitation (Fatichi et al. 2012; Lins and Cohn 2014; Stepnowski et al. 2016) even though there is no presence in annual precipitation record in inconclusive (O’Connell et al. 2015).

There is evidence that $H$ of mean annual precipitation time series from instrumental measurements.

Investigate possible relationships between $H$ and station location features (latitude, longitude, elevation, climate type).

Examine the importance of location features in predicting $H$.

Predict $H$ using location features as predictor variables.

Estimate trends of mean annual precipitation and their significance.

Perform an exploratory analysis on the trends coupled with station location features.

3. Data and methods

Daily precipitation data from 1,535 stations (Menne et al. 2012a,b).


Earth’s surface coverage is limited to Australia, Europe, North America due to data availability.

Annual series imputation based on procedure described in Tyrallis et al. (2017).

Daily-time series are transformed to mean annual time series.


Regression of $H$ on predictor variables (longitude, latitude, spa–Cartesian coordinates, elevation, Köppen-Geiger climate class (Ktika et al. 2016)) using random forests (Breiman 2001), the closest algorithm (Strobl et al. 2007) and linear regression.

Estimation of trends and their significance using the Mann-Kendall test under the LTP assumption (MK-LTP) (Bremann et al. 2011), the closest algorithm (Breiman 2001) and linear regression.

4. Stations location and Köppen-Geiger climate types

Köppen-Geiger: climate types

1. Descriptive statistics

2. Regression output

3. Significance of annual trend estimates

4. Stations location and Köppen-Geiger climate types

Köppen-Geiger: climate types

5. $H$ estimate and climate type

6. $H$ estimate and location characteristics

7. Regression predictors and cross-validation

8. 5-fold cross-validation, variable importance

9. Significance of annual trend estimates

10. Conclusions

11. Conclusions

References

• Median is $H = 0.56$ for the dataset of 1,535 mean annual precipitation time series for the time periods 1916–2015.

• Result is consistent with Fatichi et al. (2012), Sun et al. (2014) and Iliopoulou et al. (2016).

• Location of stations is important in predicting $H$, followed by the climate type and elevation.

• However, the order of importance of the three former variables depends on the algorithm.

• The closer algorithms estimate that the climate type is the most important, while due to its simultaneous handling of continuous and categorical variables can be considered more reliable than the random forests in estimating the variable importance.

• The combinations 6 and 20 of predictor variables, which include, respectively, the Cartesian coordinates and the geographic coordinates of the stations performs well in terms of the error matrix, but most importantly, their predictions had good correlation with the tested values.

• The inclusion of the climate type and the elevation (combinations 9, 23) improved further, albeit little, the performance of the random forests. However, this marginal improvement means that the information obtained from the geographic location of the station already includes the information of the climate type.