

## Scaling of low flows in Czechia – an initial assessment

**ONDREJ LEDVINKA**

*Hydrological Database & Water Balance, Czech Hydrometeorological Institute, Na Sabatce 2050/17, 143 06 Prague  
412 – Komorany, Czechia  
ledvinka@chmi.cz*

### INTRODUCTION

Fiala *et al.* (2010) performed a study focused on the evolution of dry spells at 144 water-gauging stations covering the entire territory of Czechia. They investigated several indicators such as series of 7-day low flows, and deficit volumes and durations (both delimited based on two quantiles,  $Q_{330}$  and  $Q_{355}$ , under which the flows are commonly considered as low). Using the Mann-Kendall (MK) test and its reliable modification that accounts for short-term persistence (STP) in data, they evaluated trends in their series. However, as stated in Khaliq *et al.* (2008), low flow data may also reveal so-called long-term persistence (LTP) that may cause the MK test to detect trends falsely. Therefore, the current study aims to find out if LTP was present before Fiala *et al.* (2010) applied the MK test to the 7-day low flow series. This was done through the Hurst exponent estimation. Also, testing for a unit root was carried out, similar to Fatichi *et al.* (2009).

### DATA AND METHODS

To be consistent with the study by Fiala *et al.* (2010), the same dataset was accessed from the Czech Hydrometeorological Institute's hydrological database. There are 144 text files (one for each station), each of which comprises mean daily streamflows spanning from 1 November 1960 to 31 October 2005. Thus, the data fully represents the hydrological period of 1961–2005. No missing values occurred. From these daily data, 7-day moving averages were then acquired. Further, new hydrological years were defined, starting from 1 April, when dry spells are least likely to happen, and for these years the smallest values of 7-day moving averages were searched. Likewise, the minima corresponding to the summer and winter parts of year were selected. Note that the months from April to November represented the summer period, and December to March formed the winter period.

Because the lengths of the series obtained were rather short (44 in the case of annual and summer series, 45 in the case of winter series), many techniques aimed at the estimation of the Hurst exponent  $H$  failed. However, after some experiments, it turned out that fractionally integrated autoregressive-moving average (FARIMA) models could be fitted to the data without any serious problems. The FARIMA( $p, d, q$ ) model, in which  $p$  and  $q$  denote orders of autoregressive and moving average models, respectively, and  $d$  is the fractional differencing parameter, is defined as follows (e.g. Fatichi *et al.* 2009):

$$\Phi(B)(1-B)^d X_t = \Theta(B)\varepsilon_t \quad (1)$$

In equation (1),  $B$  is the backshift operator that has an effect such that  $BX_t = X_{t-1}$ ;  $\Phi(B)$  and  $\Theta(B)$  are autoregressive and moving-average polynomials, respectively, and  $\varepsilon_t$  is a white noise. In fact, it sufficed to estimate (e.g. via the 'forecast' R package; Hyndman and Khandakar 2008) only parameter  $d$  that is related to the Hurst exponent  $H$  according to:  $H = d + 0.5$ . The dimensionless  $H$  ranges between 0 and 1, with values of  $0 < H < 0.5$  for an antipersistent process,  $H = 0.5$  for a purely independent process and  $0.5 < H < 1$  for a persistent process (Khaliq *et al.* 2008). To assess the uncertainty in the parameter estimation, the maximum entropy bootstrap (Vinod and López-de-Lacalle 2009) was employed. First, ensembles consisting of 1000 replicates were produced and then quantiles corresponding to  $\alpha = 0.05$  were calculated.

In addition, two parametric tests were performed, which help one decide whether a random walk, a deterministic trend stationary process or a long memory process underlies the series. The

first of these tests, the Phillips-Perron (PP) test (Phillips and Perron 1988), is based on the model:

$$X_t = \alpha + \beta \cdot t + \pi \cdot X_{t-1} + \varepsilon_t \quad (2)$$

where  $\alpha$  and  $\beta$  are the parameters of a linear regression. The stationary process  $\varepsilon_t$  may be serially correlated and heteroscedastic since this issue is treated directly in the test statistic. The rejection of the null hypothesis,  $H_0 : \pi = 1$ , in favour of the alternative hypothesis,  $H_1 : \pi < 1$ , indicates that the generating process is not a random walk but the AR(1) process combined with a deterministic linear trend. The basis of the second test, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test (Kwiatkowski *et al.* 1992), is the model represented by:

$$X_t = \beta \cdot t + r_t + v_t \quad (3)$$

with the parameter of a linear trend  $\beta$ , a random walk  $r_t = r_{t-1} + \varepsilon_t$ ,  $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ , and a stationary process  $v_t$ . One may test two kinds of null hypotheses but here the author focused on the general case of trend stationarity,  $H_0 : \sigma_\varepsilon^2 = 0, \beta \neq 0$  (with  $H_1 : \sigma_\varepsilon^2 > 0$ ).

## RESULTS AND DISCUSSION

Albeit the investigated low flow series were quite short and, therefore, the uncertainty in the estimation of  $H$  was inevitably high, some remarkable insights were provided. The persistence is distinct at more than half of the stations, regardless of the season selection. Concerning the annual and summer series, the average value of  $H$  equals 0.74 and 0.7, respectively. During the winter period, the share of stations with persistence is somewhat lower ( $H = 0.55$ ). At  $\alpha = 0.05$ , the unit root tests revealed that possible non-stationarity of 38 annual and summer series and 17 winter series could not be explained by a deterministic trend. Interestingly, there were seven and three unit roots found in the annual and summer series, respectively. This is not so typical for geophysical time series. It might be caused by a lack of information in the data (thus the tests cannot distinguish), a malfunction of measuring instruments or even by a human failure.

## CONCLUSION

In recent studies, hydrologists have employed modified tests for trend that are capable of accounting for short-term persistence, such as a lag-one autoregressive process. However, long-term persistence may influence the results as well. Therefore, prior to the detection of trends, an investigation should take place to determine whether or not a long memory process generates the series. In Czechia, it also applies for the series encompassing 7-day low flows since they reveal relatively high values of the Hurst exponent (except in the winter season). In the future, other variables corresponding to streamflow drought in Czechia (deficit volumes, duration or timing) are planned to be assessed in this manner.

## REFERENCES

- Fatichi, S., *et al.* (2009) Deterministic versus stochastic trends: Detection and challenges. *Journal of Geophysical Research* 114(D18). doi:10.1029/2009JD011960
- Fiala, T., Ouarda, T. B. M. J. & Hladný, J. (2010) Evolution of low flows in the Czech Republic. *Journal of Hydrology* 393(3-4), 206–218. doi:10.1016/j.jhydrol.2010.08.018
- Hyndman, R. J. and Khandakar, Y. (2008) Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software* 27(3), 1–22.
- Khaliq, M. N., *et al.* (2008) Temporal evolution of low-flow regimes in Canadian rivers. *Water Resources Research* 44(8). doi:10.1029/2007WR006132
- Kwiatkowski, D., *et al.* (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* 54(1-3), 159–178. doi:10.1016/0304-4076(92)90104-Y
- Phillips, P. C. B. and Perron, P. (1988) Testing for a unit root in time series regression. *Biometrika* 75(2), 335–346. doi:10.1093/biomet/75.2.335
- Vinod, H. D. and López-de-Lacalle, J. (2009) Maximum entropy bootstrap for time series: The meboot R package. *Journal of Statistical Software* 29(5), 1–19.