# Graphical-statistical method to explore variability of hydrological time series

Charles Onyutha

## ABSTRACT

Due to increasing concern on developing measures for predictive adaptation to climate change impacts on hydrology, several studies have tended to be conducted on trends in climatic data. Conventionally, trend analysis comprises testing the null hypothesis $H_0$ (no trend) by applying the Mann–Kendall or Spearman's rho test to the entire time series. This leads to lack of information about hidden short-durational increasing or decreasing trends (hereinafter called sub-trends) in the data. Furthermore, common trend tests are purely statistical in nature and their results can be meaningless sometimes, especially when not supported by graphical exploration of changes in the data. This paper presents a graphical-statistical methodology to identify and separately analyze sub-trends for supporting attribution of hydrological changes. The method is based on cumulative sum of differences between exceedance and non-exceedance counts of data points. Through the method, it is possible to appreciate that climate variability comprises large-scale random fluctuations in terms of rising and falling hydro-climatic sub-trends which can be associated with certain attributes. Illustration on how to apply the introduced methodology was made using data over the White Nile region in Africa. Links for downloading a tool called CSD-VAT to implement the presented methodology were provided.

**Key words** | climate variability, hydrological change attribution, Mann–Kendall test, River Nile, Spearman's rho test, sub-trend analysis

**Charles Onyutha**
Department of Civil and Building Engineering,
Kyambogo University,
P.O. Box 1, Kyambogo, Kampala,
Uganda
E-mail: *conyutha@kyu.ac.ug*

## HIGHLIGHTS

- Common trend tests are purely statistical. They can yield meaningless results in some cases.
- Conventional testing of the null hypothesis (no trend) using entire data leads to lack of information about sub-trends.
- Analysis of sub-trends maximizes understanding on how changes in the data can be linked to certain driving factors.
- This paper presents a methodology for graphical-statistical analyses of sub-trends.

## INTRODUCTION

In the last decade, almost 90% of the scientific articles published on precipitation, hydrology, and extreme climatic

conditions contained the word 'trends' (Iliopoulou & Koutsoyiannis 2020). A popular practice by hydrologists in the 21st century has been fitting of trends everywhere, based on real data, and projecting them to the future (Iliopoulou & Koutsoyiannis 2020; Koutsoyiannis 2020). To do so, non-parametric or distribution-free methods including

the Mann–Kendall (MK) (Mann 1945; Kendall 1975) and Spearman's rho (SMR) (Spearman 1904; Lehmann 1975) tests are conventionally applied to the entire time series for determining whether the null hypothesis $H_0$ (no trend) is rejected. There are some crucial issues regarding the conventional approach of testing trends in climatic data:

1. Common trend detection methods tend to be applied in a purely statistical way. Results from such purely statistical trend tests can be meaningless in some cases (Kundzewicz & Robson 2000), especially if not supported by inferences from graphical exploration of changes in the data.

2. The idea of testing the $H_0$ (no trend) using the entire data record to get insight on the future climatic condition is a misleading practice. Prediction of the future using trends fitted to the entire record was recently found to be worse than simply taking the mean of the data (Iliopoulou & Koutsoyiannis 2020). By testing the $H_0$ (no trend) using an entire long-term record, a great deal of information is missed out about hidden short-durational increase or decrease (hereinafter referred to as sub-trends) in the data. In other words, application of MK or SMR test to a given long-term (for instance, 100-year) series can show that the $H_0$ (no trend) is not rejected ($p > \alpha$). This can be achieved because the effects of increasing and decreasing sub-trends within the series cancel each other and the net result of such cancellation may yield trend statistics for which the $H_0$ (no trend) cannot be rejected ($p > \alpha$). However, the $H_0$ (no trend) may be rejected ($p < \alpha$) for several sub-series from the same dataset when analyzed separately. Therefore, instead of focusing on long-term (like 100-year) trend which may be due to the effect of an external forcing on the system, analyses of sub-trends are important to characterize climate variability, a phenomenon which is crucial for consideration of designs of hydraulic structures or operation of hydrological applications. Here, there are two factors to be considered, including the relevant time scales and length of the sub-series for which the $H_0$ (no trend) is rejected ($p < \alpha$). For instance, a 15-year data period is relevant as the design life of some water supply projects or systems for an irrigation scheme.

Importantly, identification of sub-trends can reveal fascinating information on temporal variability in climatic data and this can invite further investigations to maximize our understanding on driving forces of hydrological processes over short-durational or long periods. Large-scale random fluctuations in terms of rising and falling hydrological trends could be associated with specific attributes over the relevant sub-periods. These large-scale random fluctuations may be considered deterministic, especially if they could be conditioned by a physical explanation and predictability (Koutsoyiannis 2000). If the fluctuations defy a deterministic description, a plausible option of hydrological data analysis would be to adopt a stochastic description (Montanari & Koutsoyiannis 2014; Koutsoyiannis & Montanari 2015).

The main aim of this paper was to present a graphical-statistical methodology (based on cumulative sum of differences between exceedance and non-exceedance counts of data points) for identification and analyses of sub-trends in climatic series. The innovative aspect of the presented methodology lies in the tailored partial sums of the new trend statistic to graphically separate temporal clusters of events above or below the long-term mean of the climatic dataset. The method involves testing the $H_0$ (natural randomness) in the given data using a non-overlapping window of relevant length passed from the beginning to the end of the time series. Depending on whether the $H_0$ (natural randomness) is rejected ($p < \alpha$), further investigation can be conducted to confirm if there are persistent fluctuations. This leads to testing on various other hypotheses to determine attributes of a hydrological change. Furthermore, assessment of persistence allows us to decide on which statistics to employ for further analyses of the given climatic or hydrological time series.

## THE NEW METHOD TO ANALYZE TRENDS, SUB-TRENDS AND VARIABILITY

### General

The method being introduced (which for lack of an appropriate name due to its elaborate nature for multi-purpose application can simply be referred to as Onyutha's test or method) depends on non-parametrically re-scaled series.

To go through the method in a step-wise way, let the independent and dependent variables be denoted by $X$ and $Y$, respectively. Consider that the sample size of $X$ or $Y$ is $n$. We can re-scale $X$ and $Y$ into series $d_x$ and $d_y$, respectively, using

$$d_{y,i} = n - w_{y,i} - 2t_{y,i} \text{ and } d_{x,i} = n - w_{x,i} - 2t_{x,i} \text{ for } 1 \leq i \leq n \tag{1}$$

where, $t_{y,i}$ and $t_{x,i}$ denote the number of times the $i^{th}$ observation exceeds other data points in $Y$ and $X$, respectively. Similarly, $w_{y,i}$ and $w_{x,i}$ refer to number of times the $i^{th}$ data point appears within $Y$ and $X$, respectively. For instance, a given dataset $Y$ with $n = 9$ such that $y = \{3, 4, 3, 4, 5, 2, 7, 1, 6\}$ yields $t_y = \{2, 4, 2, 4, 6, 1, 8, 0, 7\}$, $w_y = \{2, 2, 2, 2, 1, 1, 1, 1, 1\}$, and $d_y = \{3, -1, 3, -1, -4, 6, -8, 8, -6\}$. Due to re-scaling, the mean of either $d_y$ or $d_x$ is zero.

We can standardize $d_y$ and $d_x$ to obtain series $e_y$ and $e_x$, respectively, using

$$e_{y,i} = d_{y,i} \times \sqrt{(n-1) \times (c_y)^{-1}} \text{ and } e_{x,i}$$
$$= d_{x,i} \times \sqrt{(n-1) \times (c_x)^{-1}} \text{ for } 1 \leq i \leq n \tag{2}$$

where,

$$C_y = \sum_{i=1}^{n} (d_{y,i})^2 \quad \text{and} \quad C_x = \sum_{i=1}^{n} (d_{x,i})^2 \tag{3}$$

For the case when a data point $y$ (or $x$) is tied $n$ times, each value of $d_y$ will become zero, meaning that $e_{y,i} = 0$ for $1 \leq i \leq n$.

The new method's trend statistic $T$ can be given by

$$T = \sum_{j=1}^{n} \sum_{i=1}^{j} e_{y,i} \tag{4}$$

Positive and negative trends are indicated by $T > 0$ and $T < 0$, respectively. The strength of co-variation of $X$ and $Y$ can be measured using the metric $S_1$ such that

$$S_1 = \left( \sum_{j=1}^{n} \sum_{i=1}^{j} e_{x,i} \right)^{-1} \times T \tag{5}$$

Another metric $S_2$ in terms of $B = \sum_{j=1}^{n} \sum_{i=1}^{j} d_{y,i}$ for $1 \leq i \leq n$ was considered as trend statistic such that (Onyutha 2016a):

$$S_2 = \frac{6B}{n^3 - n} \tag{6}$$

If $R_x$ and $R_y$ are ranks of $X$ and $Y$, respectively, and given that the $X$'s and $Y$'s are permutations of numbers from 1 to $n$, the product moment coefficient $\rho$ of correlation between $X$ and $Y$ in terms of $F = \sum_{i=1}^{n} (R_{x,i} - R_{y,i})^2$ can be given by (Spearman 1904):

$$\rho = 1 - \frac{6F}{n^3 - n} \tag{7}$$

Based on Equations (6) and (7), we obtain

$$S_2 = (1 + \rho)\frac{B}{F} \tag{8}$$

and given that $S_1 = S_2$, the relationship between $T$ (Equation (4)) and $\rho$ (Equation (7)) can be given by

$$T = \frac{(1+\rho)(n-1)B}{2F} \sqrt{\frac{n(n+1)}{3}} \tag{9}$$

Important notes are that $\rho = S_1 = S_2$ for regular data and the mean of $\rho$ is zero. Some vital results due to Student (1921) but presented by Pearson (1907) indicated that the second moment of $\rho$ about the mean or $\mu_2(\rho)$ is given by

$$\mu_2(\rho) = \frac{1}{n-1} \tag{10}$$

Like for $\rho$, the mean of $S_1$ or $S_2$ is zero. Furthermore, as shown in Section SM1 of the Supplementary material, the variance of $S_1$ denoted as $V(S_1)$ is also $(n-1)^{-1}$ like $\mu_2(\rho)$ in Equation (10). Given that $V(S_1)$ is equal to $V(S_2)$, it follows that the variance of the term $B$ or $V(B)$ is

$$V(B) = \frac{n^2(n^2 - 1)(n+1)}{36} \tag{11}$$

Finally, the mean of $T$ (Equation (4)) is zero and for large $n$ the distribution of $T$ is approximately normal (see Figure SM1 of the Supplementary material) with the variance of $T$ or $V(T)$ (as shown in Equation (SM9) under Section SM1 in the Supplementary material) given by

$$V(T) = \frac{n(n^2 - 1)}{12} \tag{12}$$

The $V(T)$ in Equation (12) is valid for independent data. For dependent data, $V(T)$ is affected based on the form of persistence in the data. For instance, $V(T)$ is more inflated by fractional Gaussian noise (FGN) than the first-order autoregressive AR(1) model. For auto-correlated data, the corrected variance of $\rho$ or $V^A(\rho)$ is given by (Moran 1948; Hamed 2016)

$$V^A(\rho) = \frac{12^2}{2\pi n^2 (n^2 - 1)^2} \sum_{i=1}^{n} \sum_{j=1}^{n,j\neq i} \sum_{k=1}^{n} \sum_{l=1}^{n,l\neq k} (i-1)(k-1) sin^{-1}(\rho_{ijkl}) \tag{13}$$

where

$$\rho_{ijkl} = \frac{\rho_{jl} - \rho_{il} - \rho_{jk} + \rho_{ik}}{\sqrt{(2 - 2\rho_{ij})(2 - 2\rho_{kl})}} \tag{14}$$

and $\rho_{ij}$, $\rho_{kl}$, $\rho_{jl}$, and $\rho_{ik}$ denote the correlation between $y_i$ and $y_j$, $y_k$ and $y_i$, $y_j$ and $y_i$, and $y_i$ and $y_k$, respectively, for a given dataset $Y$ considering the series to be the equivalent normal variates with Gaussian dependence. Using Equation (SM7) in Section SM1 of the Supplementary material, it follows that the exact expression of the variance of $T$ or $V^C(T)$ for persistent data with multivariate Gaussian dependence can be given by

$$V^C(T) = \frac{6}{n(n+1)\pi} \sum_{i=1}^{n} \sum_{j=1}^{n,j\neq i} \sum_{k=1}^{n} \sum_{l=1}^{n,l\neq k} (i-1)(k-1) sin^{-1}(\rho_{ijkl}) \tag{15}$$

Actually, Equation (15) is so computationally arduous that it can be limited to theoretical consideration based on small samples with $n < 10$. In practice, analyses of trends

in data with persistence require long-term series. Similarly, sub-trends can occur over long periods, of say, greater than ten years. In such situations, $V(T)$ can be corrected from the influence of autocorrelation by a suitable approximation following a slight modification of the formula from Onyutha (2016b) (see Section SM4 of the Supplementary material).

For data which follow the FGN model, the corrected $V(T)$ denoted as $V^F(T)$ following 10,000 samples of FGN synthetic series generated under various circumstances of sample sizes and scaling coefficients was found to be

$$V^F(T) = V(T) \times \beta \times n^\theta \tag{16}$$

where

$$\beta = 1.4784 H_{Est}^4 + 0.5094 H_{Est}^3 - 3.9455 H_{Est}^2 + 0.8312 H_{Est} + 1.4174 \tag{17}$$

$$\theta = -0.4512 H_{Est}^4 - 0.4057 H_{Est}^3 + 1.9193 H_{Est}^2 + 0.4237 H_{Est} - 0.6144 \tag{18}$$

And $H_{Est}$ is the sample scaling coefficient obtained after removal of an apparent trend from the data. A number of approaches exist to estimate $H_{Est}$ (see Section SM6 of the Supplementary material). Distribution of $H_{Est}$ estimated from a large number of (or 50,000) synthesized white noise series was found to be approximately normal with the mean $\mu_H$ and standard deviation $\sigma_H$ given by $\mu_H = 0.5 - 2n^{-0.8}$ and $\sigma_H = (0.57 \times n^{-0.5} + n^{-1}) \times (1.0406 - 0.0001 \times n)$, respectively. Consider $Z_{\alpha/2}$ as the standard normal variate at the selected $\alpha$. If the absolute value of the standardized metric $H_z$ given by $H_z = (H_{Est} - \mu_H) \times (\sigma_H)^{-1}$ is less than $|Z_{\alpha/2}|$ the $H_0$ (the given time series is not different from a typical white noise) is not rejected ($p > \alpha$). In this case, the new method's standardized test statistic $Z$ which follows the standard normal distribution with mean (variance) of zero (one) is given by $Z = T \times (V(T))^{-0.5}$ where $V(T)$ is from Equation (12), otherwise if the $H_0$ (the dataset is like white noise) is rejected

$(p < \alpha)$, we make use of Equation (16) to compute $Z$ using

$$Z = \frac{T}{\sqrt{V^F(T)}} \tag{19}$$

The $H_0$ (no trend) is rejected $(p < \alpha)$ for $|Z| > |Z_{\alpha/2}|$; otherwise, the $H_0$ is not rejected $(p > \alpha)$. Adequacy of the variance correction approaches summarized in Equation (16) can be found in Figure SM2 of the Supplementary material. To take into account of the influence of seasonality on trend results, seasonal trend test for the new method can be found in Section SM8 of the Supplementary material.

## Graphical diagnoses of sub-trends

For graphical diagnoses of sub-trends in the data, series $a$ can be generated through sequential accumulation of $e_y$ using

$$a_j = \sum_{i=1}^{j} e_{y,i} \quad \text{for} \quad 1 \le j \le n \tag{20}$$

Temporal variation in the values of $a_j$ from Equation (20) comprises an analogy to the description of a limit Markov or memory-less process of finite-variance random walks with short-range correlation. In other words, Equation (20) employs the concept of a Brownian bridge which is based on the standard Brownian motion. Actually, a finite-variance motion converges towards the Gaussian limit distribution (Benson *et al.* 2013). This property makes the application of Brownian motion (in terms of the Brownian Bridge) so attractive in hydrology, especially in the change-point detection, see for instance, the Pettit test (Pettitt 1979) and distribution-free cumulative sum test (Pages 1954). However, this paper makes use of Equation (20) to indicate sub-periods over which the cumulative effect of temporal variation in the time series is consecutively positive or negative. To diagnose sub-trends, $a_j$ is plotted against $j$ or time of observations and this is called the sub-trend plot. To assess the significance of trend in the data, a step-wise summation is applied to $a$ of Equation (20) to obtain another series $q$ using

$$q_k = \sum_{j=1}^{k} a_j \quad \text{for} \quad 1 \le k \le n \tag{21}$$

In another step, $q_k$ is plotted against $k$ or time of observations.

## Significance of an identified sub-trend of interest

The $H_0$ (no sub-trend) is tested by adding $100(1 - \alpha)\%$ confidence interval limits (CILs) to the plot of $q_k$ versus $k$ using

$$CIL_S = \pm|Z_{\alpha/2}| \times \sqrt{V^F(T)} \tag{22}$$

If the sub-series is of reasonable length and is not different from a typical white noise, we use $V(T)$ from Equation (12) instead of $V^F(T)$. If the scatter points go outside the $100(1 - \alpha)\%$ CILs, the $H_0$ (no sub-trend) is rejected ($p < \alpha$); otherwise, the $H_0$ is not rejected ($p > \alpha$).

## Variability based on sub-trends

To test the $H_0$ (natural randomness), we select some time scale $t$ relevant to the objectives of the study. The unit of $t$ is similar to the time unit of the series being analyzed. For instance, if we are analyzing annual data, $t$ takes the unit as year. If our dataset $X$ comprises a subset $x$ from the $u^{\text{th}}$ to the $v^{\text{th}}$ value of $X$, we compute standardized trend statistic Z for each window moved in an overlapping way from the start to the end of the series. For a selected $t$, we make some term $\psi = 0.5 \times (t + 1)$ and $\psi = 0.5 \times t$ in the cases when $t$ is odd and even, respectively. To compute sub-trends we use

$$Z_j^{(t)} = f(x \subset X | x_u \le x \le x_v) \quad \text{for} \quad j = 1, 2, \ldots, n \tag{23}$$

where, $Z_j$ is the $j^{\text{th}}$ value of $Z$ (or $Z$ from the $j^{\text{th}}$ time slice), and the terms $u$ and $v$ (to define the beginning and end of each time slice) are all based on $j$ such that:

$$\left. \begin{array}{l} \text{if} \quad j < \psi, \quad v = t + j - \psi - 1 \\ \text{if} \quad j \ge \psi \quad \text{and} \quad j \le (n - \psi), \quad u = j - \psi + 1, \quad v = j + \psi \\ \text{if} \quad j > (n - \psi) \quad \text{and} \quad j \le n, \quad u = j - \psi + 1, \quad v = n \end{array} \right\}$$

$$\tag{24}$$

Normally, the number of values calculated based on a window of a particular length when moved from the

beginning to the end of the series leads to loss of information proportional to the selected window length. However, Equation (24) ensures that there is no loss of information due to the movement of the time slice or window from the start to the end of the series. In other words, the number of values computed from Equation (23) is equal to the sample size of the give dataset. To test the $H_0$ (natural randomness), thresholds on the variability can be constructed using $\pm Z_{\alpha/2}$ after plotting $Z_j$ against the corresponding $j^{\text{th}}$ time, such as, data year. If the values of Z go outside the $100(1 - \alpha)\%$ CILs, the $H_0$ (natural randomness) is rejected ($p < \alpha$); otherwise, the $H_0$ is not rejected ($p > \alpha$). If the $H_0$ (natural randomness) is not rejected ($p > \alpha$), it means temporal variation in the sub-trends occurs randomly. On the other hand, if the $H_0$ is rejected ($p < \alpha$) we suspect persistence or long-range dependence (LRD).

### Checking for persistence or LRD

To reliably assess long-term hydrological changes, natural variation in hydro-climatic variables should adequately be characterized in terms of the statistical dependence (Onyutha *et al.* 2019). Thus, instead of testing for the $H_0$ (natural randomness) using a single time scale, we look for an evidence of multiple-scale variability by quantifying LRD or the Hurst exponent $H$ (Hurst 1951).

### Attribution

To assess the attributes of temporal variation in sub-trends of river flow, we can test several hypotheses, such as (i) variability in precipitation, (ii) changes in potential evapotranspiration (PET), and (iii) land-use changes and/or human factors. Full details of the hypotheses and the procedure to test each of them can be found in Table SM1 of the Supplementary material.

## COMPARING PERFORMANCE OF THE NEW METHOD AS WELL AS MK AND SMR TESTS

Rejection rates of the new method as well as SMR and MK tests were compared under various circumstances of coefficient of variation (CV), sample size, lag-1 serial correlation

coefficient or first-order autoregressive AR(1) process, and trend slope over the ranges 0.1–0.9, 20–200, −0.9–0.9, and 0.001–0.009, respectively. Each simulation experiment was done using 5,000 synthetic series. Rejection rate was computed as the number of times the $H_0$ (no trend) was rejected divided by 5,000.

## APPLICATION OF THE NEW METHOD

### The use of synthetic data

Several synthetic series $Y_s$ with $n = 200$ or $n = 300$ were generated for the purpose of illustrating how to apply the new method for graphical-statistical diagnoses of sub-trends. To each of the series, the following steps were taken:

1. Equation (2) was applied to the full time series.
2. Sub-trend plot was made in terms of $a_j$ (Equation (20)) versus $j$ (or time of observations).
3. Sub-trends in the data were identified and separated in the form of areas enclosed between curves and the reference (or $a_j = 0$ line).
4. For each sub-trend, the corresponding sub-series (or data points over the corresponding sub-period) were separated from the full time series.
5. Equation (21) was applied to each separated sub-series.
6. A plot of $q_k$ from Step (v) against corresponding time of observations was made.
7. To the plot in Step (vi), $100(1 - \alpha)\%$ CILs were added based on Equation (22).

### Real climatic data

The new method was applied to long-term annual river flow data observed at three hydrological stations along the White Nile in Africa. Monthly river flows observed at Malakal and Mongalla (labeled Stations 1 and 2, respectively) were obtained from the Global Runoff Data Centre (GRDC), Koblenz, Germany. Monthly river flows observed at Kamdini (Station 3 from 1950 to 2000) and Jinja (Station 4 measured over the period 1900–1995) were obtained from a report submitted to Uganda Electricity board in 1997 by Kennedy & Donkin Power Ltd in Association with Sir

Alexander Gibb & Partners and Kananura Melvin Consulting Engineers. Data at Station 4 comprised Lake Victoria outflows. Monthly Lake Victoria level time series from 1900 to 1995 was also available at Station 4. Additional monthly Lake Victoria outflows at Jinja (Station 4) covering the period 1996–2005 were obtained from Ministry of Water and Environment, Uganda. An overview of the data can be seen in Table 1. Also obtained for this study were gridded hydro-climatic datasets over the upper White Nile, especially in the Equatorial region including (i) monthly precipitation and PET of the Climatic Research Unit (CRU) over the period 1901–2019 (Harris *et al.* 2020) and (ii) monthly precipitation of CenTrends v1.0 time series from 1900 to 2014 (Funk *et al.* 2015). Figure 1 shows time series plots of annual data after re-scaling or division of the series by their long-term mean (LTM).

To assess co-variation of observed flow and modeled runoff, a simplified procedure for lumped rainfall-runoff generation was adopted. Precipitation and PET were used to estimate soil moisture deficits. Runoff was generated as a function of the antecedent soil moisture when the evaporation demand was met.

**Table 1** │ Selected stations and overview of the data

| S. no. | Station name | Data period | Latitude (°) | Longitude (°) | LTM (m³/s) |
|--------|-------------|-------------|--------------|---------------|------------|
| 1 | Malakal | 1912–1982 | 9.58 | 31.62 | 939 |
| 2 | Mongalla | 1912–1982 | 5.20 | 31.77 | 1,050 |
| 3 | Kamdini | 1950–2000 | 2.24 | 32.33 | 1,075 |
| 4 | Jinja | 1900–2005 | 0.41 | 33.19 | 864 |

Temporal sub-trends in the hydro-climatic variables were computed using Equation (23). Co-variability of the hydro-climatic variables was assessed using correlation analysis. Attribution of the temporal variation in sub-trends of the river flow was done using various hypotheses and procedure presented in Table SM1 of the Supplementary material. The Hurst parameter was also estimated for each series.

## RESULTS AND DISCUSSION

### Results of the new method applied to synthetic series

Figure 2 comprises graphical illustration on diagnoses of sub-trends and their significance. In sub-trend plots, the horizontal (or $a = 0$) line is the reference indicating the case when there is completely no trend in the time series. When there is no trend (see Case 1), the reference is crossed several times with no clear area enclosed between the scatter points and the reference (Figure 2(a)). In this case, the significance of the trend in the data is determined using the entire or full time series. Using $\alpha = 0.05$, it can be noted that the data with no trend yielded values of $q_k$ which were entirely within the 95% CILs and the $H_0$ (no trend) was not rejected ($p > 0.05$) (Figure 2(j)). To statistically quantify the significance of the trend in such series, we apply Equation (19) to the entire dataset since there are no sub-trends.

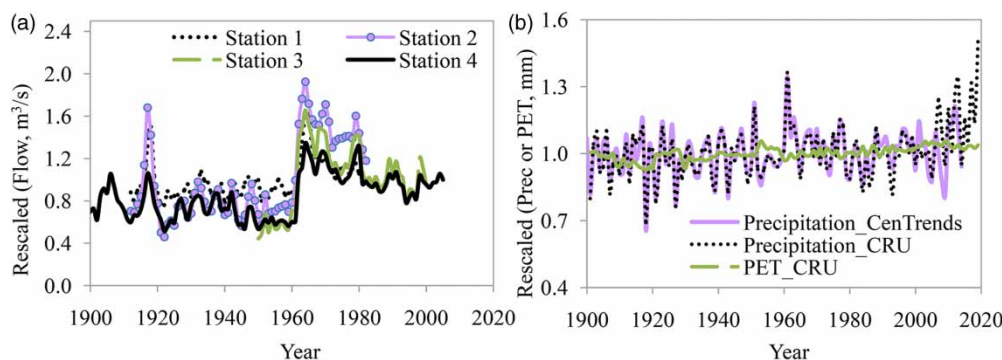Monotonic trends (Cases 2 and 3) exhibit sub-trend curves described by quadratic polynomials (Figure 2(b)



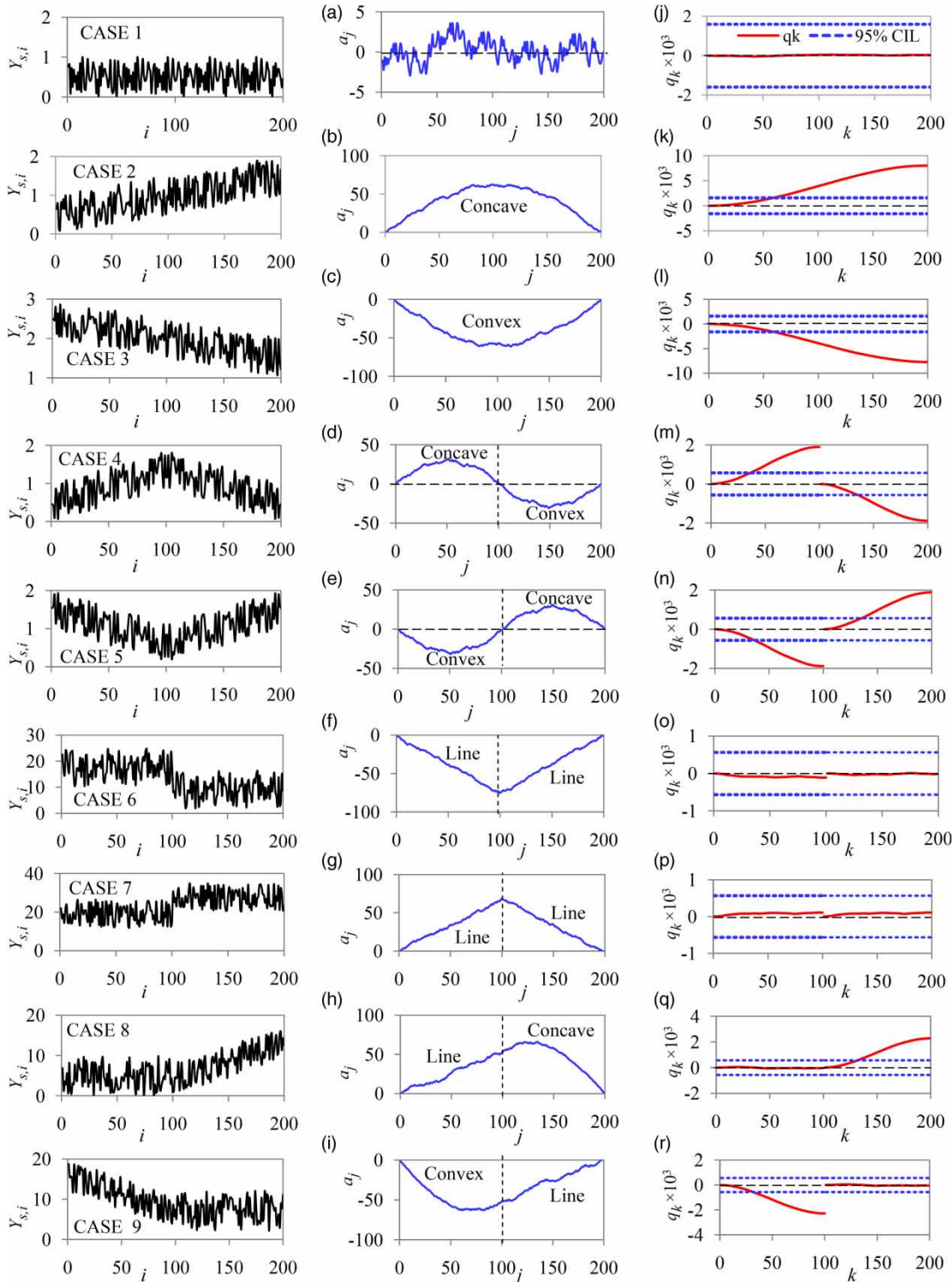**Figure 1** │ Rescaled (a) annual flow and (b) annual precipitation or PET.

**Figure 2** │ Plots for (1)–(9) synthetic series $Y_s$ of $n = 200$, with associated (a)–(i) sub-trend plots and (j–r) significance of sub-trends for the corresponding series. Legends of charts (k)–(r) are the same as that in (j).

and 2(c)). The sub-trend curve for a dataset with positive trend (Case 2) will exhibit a concave curve (Figure 2(b)). If the dataset has a negative trend (Case 3), a convex curve will be formed in the sub-trend plot (Figure 2(c)). It is worth noting that for the case with monotonic trend, the maximum absolute value of $a_j$ tends to occur around $i = n/2$ or the mid of the data period. This can easily be shown (by differentiating Equation (SM3) of the Supplementary material with respect to $i$ and equating the derivative to zero) that the point of inflection is at $i = n/2$. If the datasets are dominated by a monotonic trends as indicated by the entire sub-trend curve falling below or above the reference, the significance of the trends should be determined using the full time series. For Case 2, values of $q_k$ were above the reference and up-crossed the upper limit of the 95% confidence interval. Similarly, for Case 3, values of $q_k$ were below the reference and down-crossed the lower limit of the 95% confidence interval. Thus, for Cases 2 and 3, the $H_0$ (no trend) is rejected ($p < 0.05$) (Figure 2(k) and 2(l)). Statistically, we apply Equation (19) to the entire dataset since there are no sub-trends or the series is dominated by a monotonic increase or decrease.

When the dataset has various sub-trends, two or more curves are formed in the sub-trend plot (see Cases 4 and 5). For the part of the data with positive/negative sub-trend, the sub-trend curve will be concave/convex (Figure 2(d) and 2(e)). In Case 4 or Case 5, there are two sub-trends, one positive and the other negative. Here, the significance of each sub-trend should be determined separately using sub-data taken over the corresponding sub-period. For instance, in Case 4 or Case 5, one sub-trend occurred from $i = 1$ to $i = 100$ and the other over the sub-period $101 \leq i \leq 200$. It can be seen from Figure 2(m) and 2(n) that the $H_0$ (no trend) is rejected ($p < 0.05$) for each of these sub-trends. In other words, the negative and positive sub-trends down-crossed and up-crossed the lower and upper and 95% CILs, respectively (Figure 2(d) and 2(e)). In the conventional procedure of trend analyses without separating sub-trends, the series in Case 4 or Case 5 would yield $q_k$ at $k = n$ equal to zero (or $q_n = 0$) thereby indicating no trend, something that would be misleading. Value of $q_k$ at $k = n$ can be zero for data with two sub-trends because

the positive and negative trends cancel each other. However, in reality the $H_0$ (no trend) should be tested separately for each identified sub-trend.

For a step jump in mean (see Case 6 and Case 7), the sub-trend plot produces two lines (having opposite slopes) (Figure 2(f) and 2(g)). These lines have their point of intersection below (above) the reference in the case the data is characterized by a step downward (upward) jump in mean (Figure 2(f) and 2(g)). The significance of sub-trends in the sub-series over the two periods (one before and the other after the change point) should be determined separately. As seen from Figure 2(m) and 2(n), the $H_0$ (no trend) is not rejected ($p > 0.05$) for each of these sub-series. In other words, increase or decrease in each sub-series is minimal and the scatter points in the sub-trend plot do not cross the 95% CILs (Figure 2(o) and 2(p)).

A given dataset can have no trend over some sub-series while the other part is characterized by an increasing or decreasing sub-trend (see Cases 8 and 9 of Figure 2). In such situations, the sub-series with no trend can be indicated by a line in the sub-trend plot. However, the sub-series with increasing/decreasing sub-trend can exhibit concave/convex curve in the sub-trend plot (Figure 2(h) and 2(i)). For Cases 8 and 9, the sub-series over the periods with sub-trends should be separated from the full time series. Significance of the various sub-series should be assessed separately. For the various sub-trends, since the values of $q_k$ go beyond the 95% CILs, the $H_0$ (no trend) is rejected ($p < 0.05$) (Figure 2(q) and 2(r).

Figure 3 shows sub-trend plots for synthetic data each of $n = 300$ and having three sub-series with either trend or no trend combined. The first section is from $i = 1$ to $i = 100$ The second and third sub-series are over the sub-periods $101 \leq i \leq 200$ and $201 \leq i \leq 300$, respectively (see Cases 1–7 of Figure 3). It can be noticed that the sub-trend plots show lines for sub-trends with no trends. However, for sub-series with trends, corresponding sections of the sub-trend plot appear as curves (Figure 3(a)–3(g)). The significance of the various sub-series is determined separately (Figure 3(j)–3(r)). It is vital to note that because the sample sizes of the sub-series are the same (or $n = 100$), the widths of the 95% confidence intervals are uniform in each plot while testing the significance of the various sub-trends. In reality, sub-trends can be of different lengths of
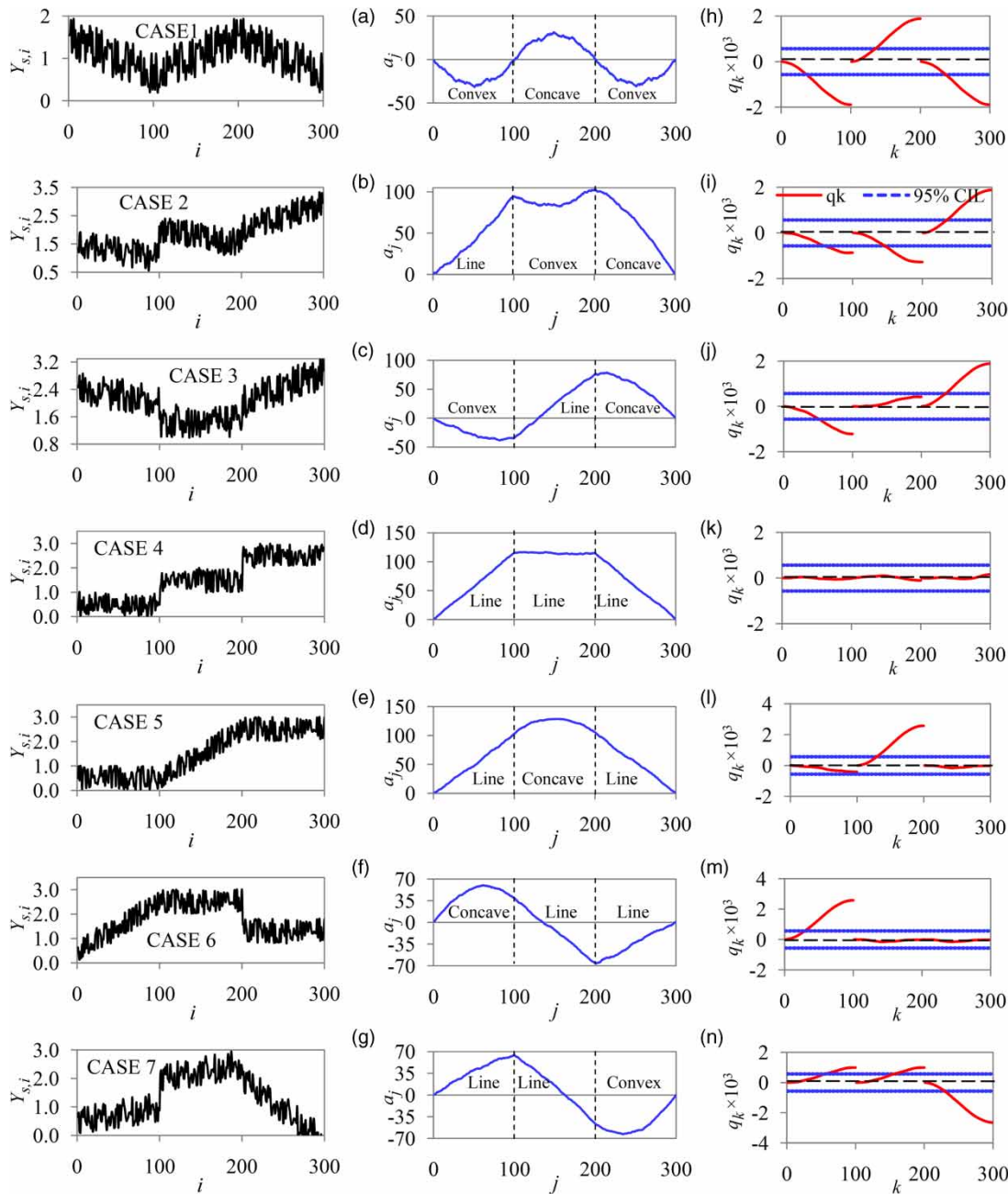
**Figure 3** │ Plots for (1)–(7) synthetic series $Y_s$ of $n = 300$ as well as (a)–(g) sub-trend plots and (h–n) significance of sub-trends for the corresponding series. Legends of charts (h)–(n) are the same as that in (i).

sub-periods and, therefore, the widths of the 95% confidence intervals can differ among sections of the sub-trend plot.

An important question to answer is: Can partial sums of MK and SMR tests be applied to produce the equivalent of sub-trend plots such as, those presented in Figures 2(a)–2(i) and 3(a)–3(g)? Application of partial sums of MK and

SMR to produce equivalent sub-trend plots as in Figures 2 and 3 has limitations. Partial sums of MK and SMR trend statistics in their current forms cannot directly be applied to separate clusters of events which occur above and below the long-term mean of the time series. This is because partial sums of the MK test statistic decrease from the

beginning towards the end of the series while fluctuating randomly regardless of whether there are sub-trends in the data. SMR test statistic comprises squared difference between ranks of $X$ and $Y$ and this makes its partial sums to always increase in magnitude even when there is an apparent negative sub-trend. Further explanations of these limitations can be found in Section SM3 of the Supplementary material.

Table 2 shows differences among methods or approaches for analyzing changes in series. It is noticeable that the conventional application of MK and SMR tests can yield trend statistic values for which the $H_0$ (no trend) is not rejected ($p > \alpha$) despite presence of apparent sub-trends. In cases where there are monotonic trends (and no sub-trends), the new methodology is consistent with conventional approaches. Here, it is important to first graphically show using sub-trend plots that the series has no sub-trends or it is dominated by a monotonic increase or decrease over the entire data record. Application of trend tests to the full time series characterized by step jump in

means can lead to rejection ($p < \alpha$) of the $H_0$ (no trend) when actually there are no apparent sub-trends. Therefore, trends in the sub-series before and after the step jump should be analyzed separately.

The value of $H$ for white noise or series with random fluctuations is 0.5. Here, the $H_0$ (natural randomness) tested using temporal variation in sub-trends was not rejected ($p > 0.05$) for the series with $H = 0.502$. Furthermore, the $H_0$ (the series is like white noise) was not rejected ($p > 0.05$) for this dataset with $H = 0.502$. For series which are not statistically different from white noise, classical statistics can be used for analyses. Application of classical statistical metrics, such as mean, coefficient of variation, and correlation follows the assumption that the sample comprises independent, and identically distributed events. Other values of $H$ in Table 2 are all above 0.5 and for the corresponding series, the $H_0$ (the given dataset is typical of white noise) was rejected ($p < 0.05$). This means that these series were characterized by persistent fluctuations. Furthermore, for each of these series, the $H_0$

**Table 2** │ Statistical trends, sub-trends, and measures of persistence in synthetic series

| S. no. | Synthetic data being analyzed | Common trends tests applied to the entire data | | The new trend test applied to sub-trends | | | Parameter $H$ of the entire data |
|--------|-------------------------------|------|------|--------|--------|--------|------|
| | | MK | SMR | Part 1 | Part 2 | Part 3 | |
| 1 | Figure 2 (Case 1) | − 0.01 | 0.04 | − 0.07 | NA | NA | 0.502 |
| 2 | Figure 2 (Case 2) | **5.46***** | **5.81***** | **4.94***** | NA | NA | 0.885 |
| 3 | Figure 2 (Case 3) | **− 5.08***** | **− 5.47***** | **− 4.68***** | NA | NA | 0.871 |
| 4 | Figure 2 (Case 4) | 0.14 | 0.13 | **3.63***** | **−3.63***** | NA | 0.856 |
| 5 | Figure 2 (Case 5) | − 0.14 | − 0.13 | **− 3.63***** | **3.63***** | NA | 0.856 |
| 6 | Figure 2 (Case 6) | **− 5.35***** | **− 5.31***** | − 0.69 | −0.12 | NA | 0.892 |
| 7 | Figure 2 (Case 7) | **5.07***** | **5.22***** | − 0.60 | −0.60 | NA | 0.876 |
| 8 | Figure 2 (Case 8) | **4.59***** | **4.94***** | − 0.28 | **3.65***** | NA | 0.888 |
| 9 | Figure 2 (Case 9) | **− 4.49***** | **− 4.80***** | **− 3.65***** | 0.28 | NA | 0.887 |
| 10 | Figure 3 (Case 1) | − 1.33 | − 1.26 | **− 3.63***** | **3.63***** | **−3.63***** | 0.864 |
| 11 | Figure 3 (Case 2) | **6.85***** | 6.87 | **− 4.93***** | **−3.72***** | **3.63***** | 0.938 |
| 12 | Figure 3 (Case 3) | **2.92***** | **3.25***** | **− 4.68***** | **2.32**** | **3.63***** | 0.939 |
| 13 | Figure 3 (Case 4) | **7.30***** | **8.37***** | − 0.15 | −0.71 | 1.00 | 0.969 |
| 14 | Figure 3 (Case 5) | **8.84***** | **9.91***** | **− 2.41**** | **3.59***** | −0.14 | 0.977 |
| 15 | Figure 3 (Case 6) | 0.14 | − 0.26 | **3.59***** | −0.14 | −0.14 | 0.948 |
| 16 | Figure 3 (Case 7) | 0.30 | 0.21 | **4.94***** | **4.94***** | **−3.53***** | 0.962 |

Bold values indicate that the $H_0$ (no trend) was rejected for $p < 0.01$ (**), and $p < 0.001$ (***). 'NA' = No other sub-trend(s).

(natural randomness) was rejected ($p < 0.05$) using temporal variation in sub-trends. Results from Table 2 show that occurrences of sub-trends in a given series can characterize statistical persistence or LRD. The case with $H > 0.5$ indicates that successive values in a given series depend on the previous values. For instance, if a given value in a series is greater than the mean, the next value is expected to exceed the sample mean with a probability greater than 0.5 (Onyutha *et al.* 2019). Generally, values of $H$ which are greater than 0.5 as shown in Table 2 indicate association with large-scale and/or multiple-scale variability or enhanced natural variability.

Enhanced natural variability means that at large horizons, we can encounter increased unpredictability. Perhaps, at short time scales, temporal clustering can be present and this can yield some potential for predictability (Dimitriadis *et al.* 2016). Generally, two frameworks exist to characterize and predict persistent processes. The first option is the Markovian framework in which the latest observations are considered to influence future probabilities. Secondly, in the Hurst–Kolomogorov (HK) (Kolmogorov 1940; Hurst 1951) framework, we consider the entire record of past realizations to condition simulations of future climatic conditions (Koutsoyiannis 2011). Application of the HK framework (which admits stationarity) can make us appreciate the coexistence of stationarity with change at all time scales. Normally, the future climatic condition tends to be predicted deterministically (such as, using linear trend). Here, the uncertainty band around the linear regression line may be narrow indicating reduced uncertainty on our prediction. However, when we apply the HK approach, the uncertainty about the future climatic conditions would be very large, and this may be realistic given the possible limited knowledge we always have especially about the (very) distant future.

## Comparing performance of the new method as well as MK and SMR tests for series

Results from various tests applied to series without considering the need to separate sub-trends were highly comparable. Details of graphical comparison can be seen in Figures SM3 and SM4 of the Supplementary material.

## Results of the new method applied to real data from the White Nile region

Figure 4 shows graphical diagnoses of sub-trends in annual hydro-climatic series. Analyses were conducted using the period over which all the hydro-climatic variables had available data records. Hereinafter, White Nile flow means the average of records from Stations 1–3 of Table 1. White Nile flows and Lake Victoria outflows exhibited an upward step jump in mean (Figure 4(a)). White Nile flows majorly come from Lake Victoria. Precipitation directly contributes to not less than 80% of the water in Lake Victoria which acts as the reservoir supplying Victoria Nile and White Nile. Less than 20% of the Lake Victoria water comes from tributaries. In other words, the Lake Victoria water level and flow from White Nile are largely influenced by precipitation over the Equatorial region. Precipitation from both sources (CenTrends and CRU) also exhibited an upward step jump in mean. However, two curves were formed by scatter points of PET over the periods 1912–1960 and 1961–2000 (Figure 4(a)). This means that PET over the Upper Nile (or White Nile) region was characterized by increasing sub-trends. However, the magnitudes of the linear increase of PET with time were different over the periods 1912–1960 and 1961–2000.

Figure 4(b) and Table 3 show the significance of identified sub-trends. Over both sub-periods 1912–1960 and 1961–2000, precipitation and river flows were characterized by decreasing trends. However, PET was characterized by increasing trends. For PET, the $H_0$ (no trend) was rejected ($p < 0.05$) for increasing trends over both sub-periods 1912–1960 and 1961–2000. Over both sub-periods, the $H_0$ (no trend) was not rejected ($p > 0.05$) for the decreasing precipitation trends. For both White Nile flow and the Lake Victoria outflow, the $H_0$ (no trend) was rejected ($p < 0.05$) for decreasing trends over the second (or 1961–2000) sub-period.

## Co-variation of sub-trends in hydro-climatic variables

Figure 5 shows co-variability of 15-year sub-trends in annual hydro-climatic variables over the White Nile region. Periods over which the new method's trend statistic $Z$ values for PET were consecutively positive (negative) were 1915–1926,
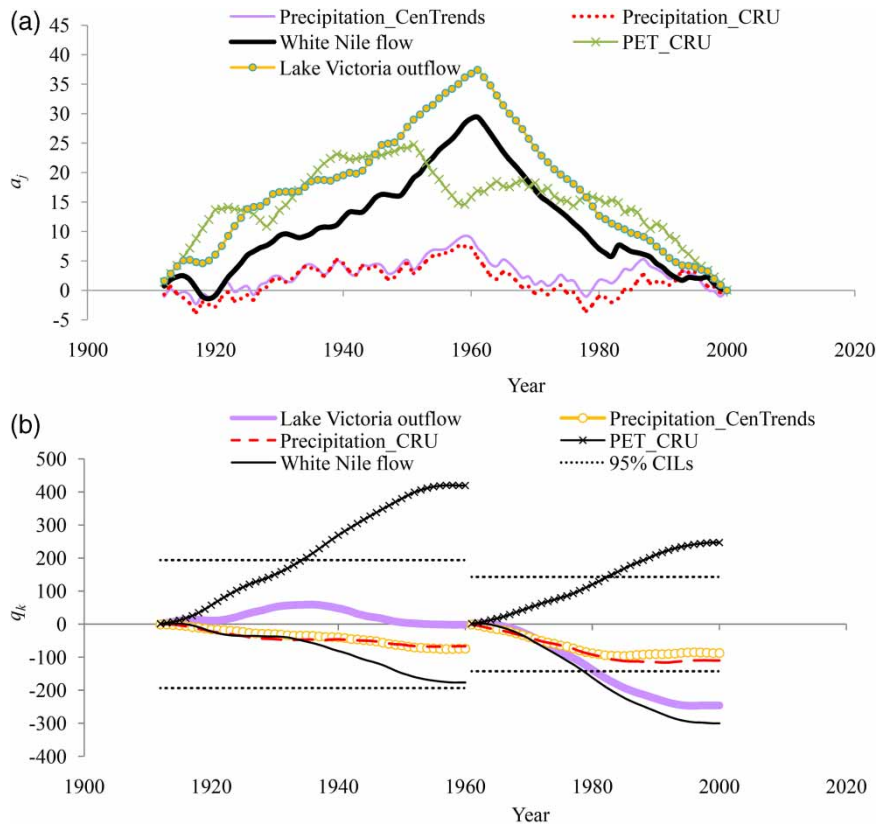
**Figure 4** │ Sub-trend plots in terms of (a) cumulative effects of temporal variation in annual hydro-climatic series and (b) significance of sub-trends.

**Table 3** │ Statistical analyses of sub-trends

| S. no. | Series name | Period 1912–1960 | | Period 1961–2000 | |
|---|---|---|---|---|---|
| | | Slope | z | Slope | z |
| 1 | Precipitation_CenTrends | − 0.80 | − 0.75 | − 1.56 | − 1.22 |
| 2 | Precipitation_CRU | − 0.71 | − 0.72 | − 2.84 | − 1.53 |
| 3 | PET_CRU | 2.16 | **4.22\*\*\*** | 1.16 | **3.38\*\*\*** |
| 4 | Lake Victoria outflow | − 0.003 | − 0.01 | − 0.76 | **−3.37\*\*\*** |
| 5 | White Nile flow | − 2.10 | − 1.78 | − 13.81 | **−4.11\*\*\*** |

A bold value with the designator \*\*\* indicates that the $H_0$ (no trend) was rejected for $p < 0.001$.

1935–1954, and 1979–2006 (1901–1914, 1927–1934, 1955–1964, and 2007 until recent 2015). The $H_0$ (natural randomness) was rejected ($p < 0.05$) over the period 1901–1914. Consecutive positive (negative) sub-trends in the Lake Victoria outflow and White Nile flow occurred over the period 1949–1967 (1967–1993). Precipitation exhibited positive (negative) sub-trends over the periods 1953–1965, 1983–

1995, and 2005 until recent 2019 (1966–1980 and 1995–2004). However, the $H_0$ (natural randomness) was not rejected ($p > 0.05$) for temporal sub-trends in precipitation. Figure 5(b) shows correlation between hydro-climatic sub-trends shown in Figure 5(a) applied to 15-year time slice or window moved in an overlapping way from the beginning towards the end of the data. Resulting correlation coefficients
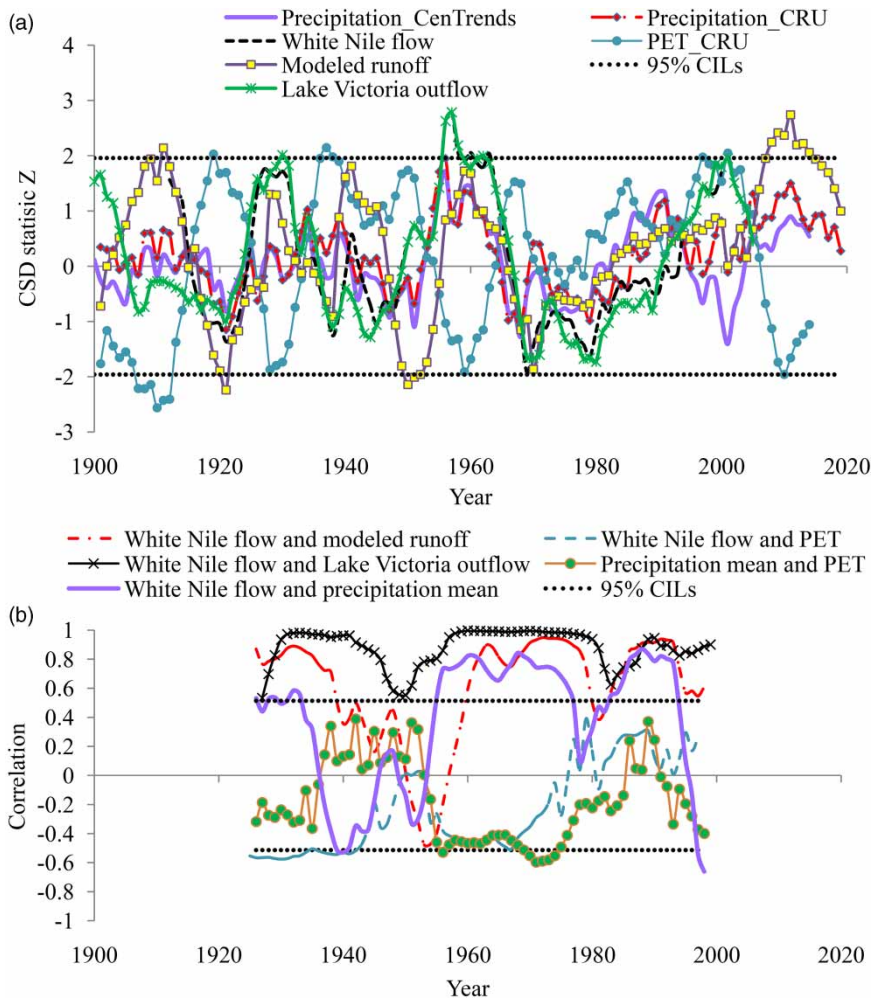
**Figure 5** | Comparison of variation in 15-year (a) sub-trends in annual data and (b) correlation among sub-trends in annual hydro-climatic data.

were plotted against the ending year of each window. Throughout the data record period, the $H_0$ (no correlation) was rejected ($p < 0.05$) for the correlation between White Nile flow and Lake Victoria outflow. The $H_0$ (no correlation) was also rejected ($p < 0.05$) for the correlation between White Nile flow and precipitation as well as PET and precipitation. Correlation between White Nile flow and precipitation was mainly positive although negative over a small epoch from the late 1930s to mid-1940s. During this period, correlation between PET and precipitation was also positive. Generally, correlation between PET and precipitation was negative except from the late 1930s to mid-1940s. The conclusive points to take from Figure 5 are that: (i) variation in the White Nile flow is significantly

($p < 0.05$) determined by changes in the Lake Victoria water level; (ii) variance in Lake Victoria water level and eventually White Nile flow can be significantly ($p < 0.05$) explained by the variability in the precipitation over the Equatorial region; (iii) variation in precipitation on land is negatively correlated with PET variability; and (iv) temporal co-variability among hydro-climatic variables depend on the period selected for analyses.

## Attribution of sub-trends in the White Nile flow

Table 4 shows results of various hypotheses put forward to explain temporal variation in sub-trends of the White Nile flows with focus on the upward step jump in mean in 1961.

**Table 4** | Hypotheses for temporal variation in sub-trends of the White Nile flows

| Hypotheses | Result | Decision/Conclusion |
|---|---|---|
| Changes in precipitation | The $H_0$ (no correlation between precipitation and river flow sub-trends) was rejected ($p < 0.05$). On a 15-year time scale, up to about 75% of the variance in the sub-trends in White Nile flow was explained by the variability in precipitation. | The upward step jump in White Nile flow mean was potentially caused by the upward step jump in precipitation over the Equatorial region. The precipitation variability in the Equatorial region could be linked to changes in the Indian Ocean Dipole and the El Nino Southern Oscillation. |
| Changes in PET | The $H_0$ (no correlation between PET and river flow sub-trends) was rejected ($p < 0.05$). Variability in PET explained up to about 30% of the changes in the White Nile flow variability. | Although the River Nile loses about 50% of its water by evaporation in the floodplains of the Sudd region in South Sudan, the PET changes over the areas upstream of the Sudd region has less influence on White Nile flow than precipitation variability. |
| Joint influence of PET, precipitation, and soil moisture on rainfall-runoff generation | On average, modeled runoff explained more variance (about 90%) in White Nile flow than when individual predictors including precipitation (75%) and PET (30%) were used. | Temporal variation in the sub-trends in White Nile flow is potentially influenced jointly by precipitation, PET and other factors. |
| Changes in the Lake Victoria water level as a reservoir supplying the White Nile | Variance in the White Nile flow explained by the changes in Lake Victoria water level was found to be more than 94%. | Temporal variability in White Nile flow sub-trends significantly ($p < 0.01$) depends on the variation in the Lake Victoria water level. In turn, about 80% of the Lake Victoria water directly comes from precipitation. |
| Changes of recording methodology and equipment | Records of Nile flows from several hydrological stations along the upper White Nile were found to temporally resonate closely with each other as well as the Lake Victoria water level. | Inconsistencies in data records were very unlikely important to explain the upward step jump in the White Nile flow in 1961. |
| Land-use and land cover (LULC) changes and/or human factors such as, urbanization, deforestation, wetland reclamation, and overgrazing | Although not investigated (for brevity), the White Nile or upper Nile sub-basin has high population growth. LULC changes follow transition in policy-institutional factors such as, shifts in land laws among the countries in the region. | LULC changes and/or human factors were unlikely important to explain the variation in the Lake Victoria water level and the White Nile flow. |
| Other human intervention such as, water abstractions, diversions, and installation or construction of hydraulic structures | Owen Falls Dam (now Nalubale Dam) through which the Lake Victoria water discharges into the Victoria Nile (where White Nile starts) operates based on a curve or level–discharge relationship agreed by Egypt and Britain. The 'agreed curve' constructed using data observed between 1939 and 1950 was to originally vary from 10.3 to 12.0 meters of water level on the Jinja gauge. The upward step jump in mean of the Lake Victoria water level in 1961 prompted water to rise above the stipulated '12-meter' mark on the Jinja gauge. This forced extension of the 'agreed curve' to operate between 10.3 and 15 meters with the main aim of ensuring that the pre-dam or original natural relationship between the Lake Victoria level and the outflow in to the Victoria Nile was retained. | Other human interventions, such as abstraction were unlikely important to explain the changes in White Nile sub-trends. Control of the Victoria Nile flow via the 'agreed curve' was unlikely to explain the upward step jump in mean of the White Nile flow. |

## Measure of long-range dependence

Table 5 shows measures of statistical dependence in hydro-climatic series over the White Nile region. The $H_0$ (the given dataset is not different from a typical white noise) was rejected ($p < 0.05$) for each series. Values of $H$ in Table 5 are all above 0.5 (and significantly different from $H = 0.5$). This means that each of the hydro-climatic series was characterized by persistent fluctuation thereby indicating association with large-scale and/or multiple-scale variability or enhanced natural variability. Results in Table 5 are consistent with the assumption that rejection of the $H_0$ (natural randomness) from Figure 5(a) was due to persistence in the hydro-climatic series. Conclusively, the temporal changes in the sub-trends of the hydro-climatic variables analyzed in this study characterized large-scale random fluctuations in the hydro-climate of the White Nile region. Since annual data were used, high values of $H$ indicate high climate variability (Koutsoyiannis 2020). Generally, temporal change is an inherent characteristic of each hydrological process and this is due to the persistent stochastic nature of the hydrological processes (Koutsoyiannis 2020).

## CONCLUSIONS

Conventional trend analysis in which Mann–Kendall and Spearman's rho tests are applied to the entire given series leads to lack of information on sub-trends (or hidden short-durational trends) in the data. Furthermore, common trend detection methods are purely statistical in nature and they can yield meaningless results in some cases if not supported by graphical exploration of changes in the data.

**Table 5** │ Values of the Hurst exponent

| S. no. | Data | H |
|--------|------|---|
| 1 | Lake Victoria water level | 0.738 |
| 2 | Lake Kyoga flow | 0.717 |
| 3 | White Nile at Malakal | 0.723 |
| 4 | White Nile at Mongalla | 0.721 |
| 5 | Precipitation over White Nile region | 0.653 |
| 6 | PET over White Nile region | 0.682 |
| 7 | Modeled runoff | 0.674 |

This paper presented a graphical-statistical methodology (which for lack of appropriate name due to its elaborate procedure, can simply be referred to as Onyutha's test (for detecting trends or sub-trends) or Onyutha's method (for other applications such as, variability analyses, and hydrological change attribution)) to identify and separately analyze sub-trends to support attribution of hydrological change. The methodology considers differences between exceedance and non-exceedance counts of data points. This method (among other things it does) can be used to (i) graphically identify sub-trend(s) over unknown period(s) of increase or decrease in the time series, (ii) statistically and graphically test for sub-trends and/or trends, (iii) analyze temporal variability based on sub-trends using time slice or window of fixed length moved from the start to the end of the dataset, and (iv) test various hypotheses for hydrological change. Furthermore, the methodology makes it possible to appreciate that climate variability comprises large-scale random fluctuations in terms of rising and falling hydrological sub-trends which could be associated with specific attributes over the relevant sub-periods. To accurately characterize large-scale random fluctuations in hydrological systems, long-term series should be considered for analyses.

To implement the CSD-based methodology presented in this paper, the reader can download a MATLAB-based tool named CSD-VAT (which stands for variability analyses tool VAT based on cumulative sum of difference CSD between exceedance and non-exceedance counts of data points) via either https://www.researchgate.net/publication/332798309 or https://sites.google.com/site/conyutha/tools-to-download

The steps for applying the presented methodology include the following:

1. Identify epochs with clusters of data points above and below long-term mean of the time series. To do so:
   (a) Select a number of time scales (for instance, 5, 10, 15, 20, 25, 30). Here, the unit of each scale is the time unit of the series. For instance, when using annual data, the unit of time scale is year.
   (b) For each time scale from Step(a), apply Equation (23) of the main paper with $Z_j$ replaced by $A_j$, where $A_j$ denotes the average of sub-series within the window running from the $u^{th}$ to the $v^{th}$ data point.

(c) To each of the series generated from Step(b), apply Equation (20).

(d) Make sub-trend plot to graphically identify sub-trends (if any) in the series of each time scale. Consult Figures 2 and 3 for a guide in graphical diagnoses of sub-trends. The choice of an epoch with a sub-trend should be based on the purpose for which analysis is being conducted. For water resources applications, sub-trend can be considered relevant for analysis if the epoch over which it occurs is equal to or greater than ten years. This means sub-trends over epochs less than ten years can be filtered and left out (unless they are relevant for intended applications). For epochs over which sub-trends exist, one needs to separate sub-series and separately assess the significance of each sub-trend. If there are no sub-trends, test the $H_0$ (no trend) using the entire data record. For testing the $H_0$ (no trend), make use of Equations (4), (12), (16), and (19) of this paper.

2. Test for $H_0$ (natural randomness) using temporal variation in sub-trends. To do so:

(a) Considering the time unit of the series, select a number of time scales (for instance, 10, 15, 20, 25, 30 years).

(b) For each time scale, apply Equation (23) as it is to the given series.

(c) Plot results from Step (b) against time of observations and add the $100(1 - \alpha)\%$ CILs.

(d) Separately test the $H_0$ (natural randomness) using results based on each time scale.

From Step 2(d), if the $H_0$ (natural randomness) is not rejected ($p > \alpha$) from results of any time scale, check whether the entire given series is statistically not different from a typical white noise. In other words, check if $H$ is statistically different from 0.5. CSD-VAT makes use of the generalized Hurst exponent $H(q^{\#})$ based on the scaling of renormalized $q^{\#}$–moment of the distribution (Di Matteo 2007). For other methods to estimate $H$, see Section SM6 of the Supplementary material. In further analyses, treat the series using classical statistics if it is not statistically different from a typical white noise.

From Step 2(d), if the $H_0$ (natural randomness) is rejected ($p < \alpha$) from results of any time scale, the given

time series can be assumed to be characterized by persistent fluctuations. Estimate the scaling exponent $H$. If the dataset is confirmed to be statistically different from a typical white noise, determine attributes of identified sub-trend(s) in the hydrological time series by testing various hypotheses. Here, make use of the procedure in Table SM1 of the Supplementary material. If we use classical statistics for series characterized by $H$ being statistically different from 0.5, only a portion of the natural uncertainty of the hydrological or hydro-climatic processes would be described stemming from an implicit assumption of a stable climate (Koutsoyiannis 2003). Therefore for further analyses, use a suitable procedure such as the Hurst–Kolmogorov (Kolmogorov 1940; Hurst 1951) framework to characterize and predict the persistent process. Importantly, the prediction should be on a short period.

## ACKNOWLEDGEMENTS

## DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

## REFERENCES

Benson, D. A., Meerschaert, M. M. & Revielle, J. 2013 Fractional calculus in hydrologic modeling: a numerical perspective. *Advances in Water Resources* **51**, 479–497.

Di Matteo, T. 2007 Multi-scaling in finance. *Quantitative Finance* **7** (1), 21–36.

Dimitriadis, P., Koutsoyiannis, D. & Tzouka, K. 2016 Predictability in dice motion: how does it differ from hydro-meteorological processes? *Hydrological Sciences Journal* **61**, 1611–1622.

Funk, C., Nicholson, S. E., Landsfeld, M., Klotter, D., Peterson, P. & Harrison, L. 2015 The Centennial Trends Greater Horn of Africa precipitation dataset. *Scientific Data* **2**, 1–17.

Hamed, K. H. 2016 The distribution of Spearman's rho trend statistic for persistent hydrologic data. *Hydrological Sciences Journal* **61** (1), 214–223.

Harris, I., Osborn, T. J., Jones, P. & Lister, D. 2020 Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data* **7**, 109.

Hurst, H. E. 1951 Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineering* **116**, 770–799.

Iliopoulou, T. & Koutsoyiannis, D. 2020 Projecting the future of rainfall extremes: better classic than trendy. *Journal of Hydrology* **558**, 125005.

Kendall, M. G. 1975 *Rank Correlation Methods*, 4th edn. Charles Griffin, London, UK.

Kolmogorov, A. N. 1940 Wiensersche Spiralen und Einige Andere Interessante Kurven in Hilbertschen Raum. *Doklady Akademii Nauk URSS* **26**, 115–118.

Koutsoyiannis, D. 2000 The Hurst phenomenon and fractional Gaussian noise made easy. *Hydrological Sciences Journal* **47** (4), 573–595.

Koutsoyiannis, D. 2003 Climate change, the Hurst phenomenon, and hydrological statistics. *Hydrological Sciences Journal* **48**, 3–24.

Koutsoyiannis, D. 2011 Hurst-Kolmogorov dynamics and uncertainty. *Journal of the American Water Resources Association* **47** (3), 481–495.

Koutsoyiannis, D. 2020 Revisiting the global hydrological cycle: is it intensifying? *Hydrology and Earth System Sciences* **24**, 3899–3932.

Koutsoyiannis, D. & Montanari, A. 2015 Negligent killing of scientific concepts: the stationarity case. *Hydrological Sciences Journal* **60**, 1174–1183. https://doi.org/10.1080/02626667.2014.959959

Kundzewicz, Z. W. & Robson, A. 2000 *Detecting Trend and Other Changes in Hydrological Data*. World Climate Program–Water, WMO/UNESCO, WCDMP-45, WMO/TD 1013, Geneva, Switzerland, p. 157.

Lehmann, E. L. 1975 *Nonparametrics, Statistical Methods Based on Ranks*. Holden-Day Inc., Oakland, CA, USA.

Mann, H. B. 1945 Nonparametric tests against trend. *Econometrica* **13** (3), 245–259.

Montanari, A. & Koutsoyiannis, D. 2014 Modeling and mitigating natural hazards: Stationarity is immortal! *Water Resources Research* **50**, 9748–9756. https://doi.org/10.1002/2014WR016092

Moran, P. A. P. 1948 Rank correlation and product-moment correlation. *Biometrika* **35**, 203–206.

Onyutha, C. 2016a Identification of sub-trends from hydro-meteorological series. *Stochastic Environmental Research Risk Assessment* **30**, 189–205.

Onyutha, C. 2016b Statistical analyses of potential evapotranspiration changes over the period 1930–2012 in the Nile River riparian countries. *Agricultural and Forest Meteorology* **226–227**, 80–95.

Onyutha, C., Rutkowska, A., Nyeko-Ogiramoi, P. & Willems, P. 2019 How well do climate models reproduce variability in observed rainfall? A case study of the Lake Victoria basin considering CMIP3, CMIP5 and CORDEX simulations. *Stochastic Environmental Research and Risk Assessment* **33** (3), 687–707.

Page, E. S. 1954 Continuous inspection schemes. *Biometrika* **41**, 100–115.

Pearson, K. 1907 *On Further Methods of Determining Correlation*. Drap. Co. Mem. Biom. Ser. iv, Dulau & Co., London, UK.

Pettitt, A. N. 1979 A non-parametric approach to the change-point problem. *Journal of the Royal Statistical Society* **28** (2), 126–135.

Spearman, C. 1904 The proof and measurement of association between two things. *The American Journal of Psychology* **15**, 72–101.

Student 1921 An experimental determination of the probable error of Dr Spearman's correlation coefficient. *Biometrika* **13** (2/3), 263–282.