



Research papers

The decomposition-based nonstationary flood frequency analysis



Cuauhtémoc Tonatiuh Vidrio-Sahagún, Jianxun He *

Civil Engineering, Schulich School of Engineering, University of Calgary, 2500 University Drive NW, Calgary T2N 1N4, Canada

ARTICLE INFO

ABSTRACT

Keywords:
 Nonstationarity
 Nonstationary structure
 Distribution
 Extreme events
 Uncertainty
 Fitting efficiency

The nonstationary flood frequency analysis (NS-FFA) is conducted when the assumption of stationarity in hydrologic extremes is violated. The commonly used approach is the performance-based NS-FFA, which jointly determines the distribution and the nonstationary structure based upon the model performance. However, this approach is challenged from both theoretical and practical perspectives. An alternative is the herein named decomposition-based NS-FFA, which determines the two NS-FFA model components separately and explicitly uses the available knowledge of the nonstationarity. However, it has been barely implemented in practice. This paper proposed a novel decomposition procedure that strictly follows the theoretical decomposition of nonstationary stochastic processes to advance the decomposition-based NS-FFA. The proposed decomposition procedure was compared with a previously reported method in both an analytical deduction and a simulation study. The proposed decomposition-based NS-FFA was further compared to the performance-based NS-FFA using both synthetic and real datasets from North America, which exhibit different patterns of nonstationarity. The Particle Filter was adopted for uncertainty quantification and parameter estimation. The results revealed that the proposed decomposition-based approach was advantageous in preserving the moments of underlying stochastic component, particularly the higher-order moment (i.e., skewness). In addition, the comparison of the two NS-FFA approaches demonstrated the superiority of the proposed decomposition-based approach in capturing the underlying (known) nonstationary stochastic process and being competitive with the performance-based approach from the performance perspective in real applications. The results from the simulation study and the real application also revealed several caveats of the performance-based approach, including the potential overfitting and equifinality problems as well as the selection of distinct models when adopting different performance metrics. In addition, differing from the performance-based approach, the decomposition-based NS-FFA avoided/alleviated the ergodicity violation. All these results demonstrated the advancements of the proposed decomposition-based NS-FFA and advocated its application in the NS-FFA.

1. Introduction

Assessing the recurrence of floods, which is a key task in hydrology and water resources management, is commonly conducted by flood frequency analysis (FFA). The conventional FFA predicts the flood quantiles corresponding to specific return periods and vice versa based on the core assumption that the underlying process is stationary, and thus it is often referred to as the stationary FFA (S-FFA). Nonetheless, the reliability of the S-FFA is compromised when confronted with nonstationary processes as a result of climate change and changes in land use/cover and/or water management (AghaKouchak et al., 2020; Blöschl et al., 2019; Milly et al., 2008; Milly et al., 2015). Under this circumstance, the nonstationary FFA (NS-FFA), in which the probability distribution of the floods varies as a function of a selected covariate(s), has

emerged as the theoretical solution to deal with nonstationary scenarios (François et al., 2019; López and Francés, 2013; Salas et al., 2018; Slater et al., 2021). In this context, the determination of a NS-FFA model, which is formed by coupling the probability distribution of floods and the nonstationary structure that governs its evolution over time, is a key to implementing the NS-FFA.

The conventional approach for determining a suitable NS-FFA model is through selection by ranking a set of candidate models according to the pre-selected performance metric, e.g., information-theoretic measures such as the Akaike and Bayesian Information Criteria (AIC and BIC) (e.g., Kim et al., 2017; Li et al., 2019; Ragno et al., 2019). This approach is referred to as the performance-based approach in this paper. Since the model is selected by jointly assessing the nonstationary structure and the distribution in this approach, the determination of one

* Corresponding author at: 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada.

E-mail address: jianhe@ucalgary.ca (J. He).

model component is conditioned on its performance in conjunction with the other (e.g., Agilan and Umamahesh, 2017; Mondal and Mujumdar, 2015; Strupczewski et al., 2001). Besides, this approach does not explicitly use the available knowledge of the nonstationarity beyond the time series. Such knowledge can be detailed physical derived from the causative process(es) behind the nonstationarity (e.g., through physically-based climate and hydrological models) or statistical deducted knowledge derived from underlying datasets and deductive reasoning.

In addition, the implementation of the performance-based NS-FFA has been challenged from the perspectives of both theoretical fundamental and practical application. Firstly, under nonstationarity, a time series is not a set of random independent realizations of the same distribution, and consequently, differing from a stationary time series, it does not visit most of its domain when the sample size increases (von Storch and Zwiers, 2002). As a result, the ergodicity does not hold under nonstationarity, and in turn, the statistical properties of a nonstationary process cannot be reliably inferred from a single time series only (Koutsoyiannis and Montanari, 2015; Serinaldi et al., 2018). Therefore, the model determination is questionable in this approach. Secondly, analogously to the black-box modeling, the NS-FFA model is selected for achieving the best performance while bypassing the available knowledge of the system. As a result, the optimal model may yield a spuriously high performance due to increased fitting flexibility irrespective of its correspondence to the actual nonstationary stochastic process (Luke et al., 2017; Serinaldi and Kilsby, 2015). Thirdly, the use of different performance metrics, such as for evaluating the uncertainty (e.g., the average width (AW) or coverage width index (CWI)), which has been commonly relegated to a secondary aspect, and the fitting efficiency (e.g., the AIC or BIC), often leads to selecting different optimal models (Ouarda et al., 2019; Serinaldi and Kilsby, 2015; Vidrio-Sahagún et al., 2021). Lastly, in this NS-FFA approach, there might be two or more candidate models with almost identical performance evaluated according to a pre-selected evaluation metric. Consequently, there would be uncertainty in the model selection. This issue is similar to the equifinality problem in the hydrological modeling (Beven, 2006; Beven and Freer, 2001; Khatami et al., 2019); namely, different model structures and/or parameter sets may be equally capable of reproducing the observations according to a (or a few) performance metric(s). However, such equifinal models may not necessarily yield the same results. In view of the above, inferring the NS-FFA model based on a performance metric is not necessarily reliable (Serinaldi et al., 2018).

On the other hand, nonstationary stochastic processes are commonly considered as the superposition of a stationary stochastic component and a deterministic time-dependent component (Koutsoyiannis and Montanari, 2015; Milly et al., 2015; Serinaldi et al., 2018). This notion suggests the potential of selecting the two components of the NS-FFA model separately. Namely, the nonstationary structure can be determined from the available knowledge on the deterministic time-dependent component, while the distribution can be inferred using the stationary component, in which the ergodicity holds. In this way, the ergodicity violation for statistical inference faced in the performance-based approach can be avoided when physical or statistical deducted knowledge complements the information from the time series. Despite the suggestion of incorporating physical knowledge into the NS-FFA (Gilroy and McCuen, 2012), its use is still constrained by its limitation to capture the nonstationarity of hydrological extremes (Giuntoli et al., 2021; Kundzewicz et al., 2017; Maraun et al., 2010; Maraun et al., 2017). Whereas the use of knowledge derived through deductive reasoning is considered feasible in the current practice and thus benefits the NS-FFA model determination (Serinaldi et al., 2018). It is worth highlighting that the identification of the deterministic component behind the observed changes is indispensable to justify the NS-FFA irrespective of the approach adopted (Koutsoyiannis and Montanari, 2015; Montanari and Koutsoyiannis, 2014; Serinaldi et al., 2018; Serinaldi and Kilsby, 2015). Changes produced by purely stochastic drivers

would result in doubly stochastic rather than nonstationary processes, which are preferably modeled with compound or mixed distributions (Serinaldi et al., 2018; Serinaldi and Kilsby, 2015). The lack of appropriate identification of the deterministic component can lead to unrealistic estimates and subsequently jeopardizes the reliable implementation of NS-FFA models (Serinaldi and Kilsby, 2015).

Although the NS-FFA alternative approach based on the notion of the nonstationary process decomposition (called the decomposition-based approach throughout this paper) is not new, it has not been well explored and developed yet. To present, only a few studies have applied this notion in the FFA. For instance, Cunderlik and Burn (2003) selected the distribution from detrended at-site series assuming second-order nonstationarity in the regional index-flood NS-FFA. They employed the absolute deviations from the mean of the transformed series (detrended in the mean) to remove the trend in the second statistical moment. Zhang et al. (2019) identified the overall best-fit distribution across Canada following the same detrending procedure as Cunderlik and Burn (2003) when nonstationarity presents. Sung et al. (2018) also applied this procedure to select the distribution, although considering nonstationarity in the mean only. However, such a detrending procedure does not guarantee the preservation of higher-order moments (e.g., skewness) of the stochastic component and consequently might result in poor estimates (e.g., Kalai et al., 2020). On the other hand, just a few studies have specified the nonstationary structure according to the exploratory/explanatory data analysis rather than according to the performance of the models. Namely, the identified temporally variant statistical moments of the underlying dataset have been used to determine the nonstationary structure (e.g., Cheng et al., 2014; Cunderlik and Burn, 2003; Rago et al., 2019). Despite these attempts to determine the distribution and nonstationary structure according to the trend-free datasets and the nonstationarity knowledge, the procedure to remove the deterministic component while the higher-order moments are preserved is still desired. This improvement would advance the decomposition-based approach and thus promote its practical application. Moreover, the decomposition-based approach has not been assessed, validated, or compared with the performance-based approach in the literature so far.

In view of the above, this paper proposes a novel general procedure that strictly follows the theoretical decomposition of nonstationary stochastic processes for advancing the model determination in the decomposition-based NS-FFA. The proposed procedure has the capacity to retain the statistical characteristics of the underlying stationary stochastic component after removing the deterministic component. This paper demonstrates such an advancement over the existing procedure of Cunderlik and Burn (2003) by both theoretical deduction and a simulation study in terms of preserving the statistical moments of the a priori known stationary stochastic component. In addition, the performance (namely, the fitting efficiency and uncertainty) of the proposed decomposition-based approach is further assessed and compared with the performance-based approach, which has been often adopted in the literature, using both synthetic and real datasets. These demonstrate both the advantages of the proposed decomposition-based NS-FFA and the caveats of the performance-based approach in real applications. Thirteen annual maximum series (AMS) of flow from Canada and the USA, which are perceived to have different patterns of nonstationarity, are used herein as study cases.

2. Materials and methods

Apart from the proposed decomposition-based approach, this section also introduces the performance-based approach, the distributions considered, the method used for parameter estimation and uncertainty quantification, the generation of the synthetic datasets for the simulation study, and the real datasets for the practical applications. Note that throughout this section, the random variables and their realizations are denoted by uppercase and lowercase letters, respectively, while the

vectors are denoted by boldface letters.

In practice, the temporal covariate has been commonly used in the NS-FFA as a surrogate for any time-dependent physical driver(s), especially for fitting purposes (e.g., Obeysekera and Salas, 2016; Sun et al., 2018), and/or bypassing the exploration of the physical process(es), both of which are case sensitive. This is particularly convenient for methodology-oriented rather than case-specific research objectives (e.g., Prosdocimi and Kjeldsen, 2021; Serago and Vogel, 2018). Moreover, the often low to moderate dependency of the nonstationarity on the physical driver(s) (Archfield et al., 2016; Burn and Whitfield, 2017; Ray and Goel, 2019) has hindered the use of the physical covariate(s). Thus, the temporal covariate (i.e., the time t) was adopted to illustrate the application of both the performance-based and the decomposition-based approaches in this paper.

2.1. Candidate distributions

In this paper, three probability distributions were considered in both the performance-based and the decomposition-based approaches. The candidate distributions are the Generalized Extreme Value (GEV), the Pearson type III (PE3), and the log-Pearson type III (LPE3) distributions. These three-parameter distributions have been commonly adopted in the FFA in Canada and the USA (England et al., 2019; Gado and Van Nguyen, 2016b; Stedinger et al., 1993; Zhang et al., 2020).

The cumulative GEV function of the random variable Y (F_Y) and the quantile (y_{p_e}) corresponding to an exceedance probability p_e ($p_e = 1 - F_Y(y; \theta)$) are given by:

$$F_Y(y; \theta) = \begin{cases} \exp\left(-\left[1 + \kappa\left(\frac{y - \xi}{\alpha}\right)\right]^{\frac{1}{\kappa}}\right), & \kappa \neq 0 \\ \exp\left(-\exp\left(-\left(\frac{y - \xi}{\alpha}\right)\right)\right), & \kappa = 0 \end{cases} \quad (1)$$

$$y_{p_e}(p_e; \theta) = F_Y^{-1}(p_e; \theta) = \begin{cases} \xi - \frac{\alpha}{\kappa} [1 - \{-\ln(1 - p_e)\}^{-\kappa}], & \kappa \neq 0 \\ \xi - \alpha \ln[-\ln(1 - p_e)], & \kappa = 0 \end{cases} \quad (2)$$

where the parameter vector (θ) is composed of the shape, scale, and location parameters, denoted by κ , α , and ξ , respectively.

The cumulative PE3 distribution function is given by (Hosking and Wallis, 1997):

$$F_Y(y; \theta) = \begin{cases} \frac{G\left(\alpha, \frac{y - \xi}{\beta}\right)}{\Gamma(\alpha)}, & \gamma > 0 \\ \Phi\left(\frac{y - \mu}{\sigma}\right), & \gamma = 0 \\ 1 - \frac{G\left(\alpha, \frac{\xi - y}{\beta}\right)}{\Gamma(\alpha)}, & \gamma < 0 \end{cases} \quad (3)$$

where the shape, scale, and location distribution parameters are $\alpha = 4/\gamma^2$, $\beta = \frac{1}{2}\sigma|\gamma|$, and $\xi = \mu - 2\sigma/\gamma$, respectively, if $\gamma \neq 0$; μ , σ , and γ are the mean, standard deviation, and skewness of Y , respectively; $\Gamma(\bullet)$ is the gamma function; $G(\cdot)$ is the incomplete gamma function; and $\Phi(\cdot)$ is the Normal cumulative distribution function. There is no explicit analytical form of y_{p_e} for the PE3 distribution. It has often been approximated using the frequency factor K_p through the Wilson-Hilferty transformation for $|\gamma| \leq 2$ and $0.01 \leq p \leq 0.99$ by (Stedinger et al., 1993):

$$y_{p_e} = \mu + \sigma K_p(\gamma), \quad K_p(\gamma) = \frac{2}{\gamma} \left\{ \left[1 - \left(\frac{\gamma}{6} \right)^2 + \frac{\gamma z_p}{6} \right]^3 - 1 \right\} \quad (4)$$

where $K_p(\gamma)$ is the p th quantile ($1 - p_e$) of the PE3 distribution with zero

mean, unit standard deviation and skewness γ ; and z_p is the p th quantile of the standard normal distribution.

The LPE3 distribution corresponds to the case in which the log-transformed random variable $Y^* = \log_{10}(Y)$ follows the PE3 distribution. Thus, Eqs. (3) and (4) of the PE3 distribution are applicable, while μ , σ , and γ are the mean, standard deviation, and skewness of Y^* , respectively.

In the common practice of the NS-FFA, the nonstationarity has been depicted using time-dependent distribution parameters. Thus, in this paper, a NS-FFA model consists of an aforementioned distribution whose parameters are functions of the time.

2.2. Parameter estimation and uncertainty quantification

The parameter estimation and uncertainty quantification were conducted employing the Particle Filter (PF) technique. The PF is a method that integrates the recursive Bayesian filters and the bootstrap resampling technique and is free from assumptions and reliable in uncertainty estimation (Doucet et al., 2001; Gordon et al., 1993; Särkkä, 2013). The PF method has been previously adopted in both the S-FFA and NS-FFA (Sen et al., 2020; Vidrio-Sahagún et al., 2021). In this paper, the PF method was implemented following the methodology of Vidrio-Sahagún et al. (2021).

In a nutshell, the PF estimates the hidden state vector of a dynamic system, $w_{0:i} = \{w_0, w_1, \dots, w_i\}$ (which is difficult if not impossible to be measured in the field) from the related observations, $y_{1:i} = \{y_1, y_2, \dots, y_i\}$. Here, $w \in \mathbb{R}^{N_w}$ and $y \in \mathbb{R}^{N_y}$ denote the N_w - and N_y -dimensional hidden state and observation vectors, respectively. The general idea behind the PF, as an inverse approach, is to assess the error of the model predictions with respect to the observations and inversely map the model predictions to the parameter space. In the context of the FFA, the general structure of the PF is formulated by taking the set of model parameters (θ) and the annual maximum flows (y) as the hidden state variables and the associated measurements, respectively. As the parameter estimation is conducted for a given time period in the NS-FFA, the hidden states $\theta_j = \{\theta_1, \theta_2, \dots, \theta_{N_\theta}\}$ evolve in pseudo-time, replacing time with iterations in order to incorporate all available measurements utilized as a batch at each step j , i.e., $y_j = \{y_{t_1}, y_{t_2}, \dots, y_{t_n}\} \forall j$. Here, N_θ and n are the number of NS-FFA model parameters and sample size of the AMS, respectively. This can be expressed in the Bayesian framework as the search for the joint posterior distribution of all the model parameters given the related observations. The main steps of the PF include:

1. The PF starts at $j = 1$ with a set of $m \in \mathbb{N}^{N_\theta}$ independent particles ($\Psi_j = \{\Psi_j^1, \Psi_j^2, \dots, \Psi_j^m\}$) in the parameter domain, which is initially sampled from uninformative marginal prior distributions. The initial priors were specified as uniform distributions with ranges large enough, such that encompassing all realistic values and none of the posteriors were truncated. Each particle has an identical weight $w(\Psi_j^i)$ at $j = 1$.
2. Then, the $w(\Psi_j)$ are updated over pseudo-time steps (i.e., from Ψ_j to Ψ_{j+1}) based on the particles' likelihood, and act as discrete posterior distributions. The particles' likelihoods are obtained from the error between the model outputs (i.e., the estimated quantiles computed using the Ψ_j and the mapping function $f(\cdot)$, here the quantile function) and the observations (i.e., the empirical quantiles derived using the measurements y_j and their associated empirical exceedance probability).
3. After each pseudo-time step, resampling is conducted and white noise perturbation is applied to Ψ_{j+1} to avoid particle degeneracy and maintain adequate diversity, respectively.
4. Steps 2 and 3 are repeated until the Ψ_j converges to a stable estimation with the progress of the PF over pseudo-time steps. The

Table 1

Candidate nonstationary structures and models.

Candidate model ID	Candidate nonstationary structures of distribution parameters
Model structure: $GEV(\xi(t), \alpha \text{ or } \alpha(t), \kappa)$	
$\mathcal{M}_{1,0,0}^{GEV}$	$\xi(t) = \xi_0 + \xi_1 t, \alpha = \alpha_0, \text{ and } \kappa = \kappa_0$
$\mathcal{M}_{E1,0,0}^{GEV}$	$\xi(t) = \exp(\xi_0 + \xi_1 t), \alpha = \alpha_0, \text{ and } \kappa = \kappa_0$
$\mathcal{M}_{2,0,0}^{GEV}$	$\xi(t) = \xi_0 + \xi_1 t + \rho_2 t^2, \alpha = \alpha_0, \text{ and } \kappa = \kappa_0$
$\mathcal{M}_{1,1,0}^{GEV}$	$\xi(t) = \xi_0 + \xi_1 t; \alpha(t) = \beta_0 + \beta_1 t, \text{ and } \kappa = \kappa_0$
$\mathcal{M}_{E1,1,0}^{GEV}$	$\xi(t) = \exp(\xi_0 + \xi_1 t), \alpha(t) = \beta_0 + \beta_1 t, \text{ and } \kappa = \kappa_0$
$\mathcal{M}_{1,E1,0}^{GEV}$	$\xi(t) = \xi_0 + \xi_1 t, \alpha(t) = \exp(\beta_0 + \beta_1 t), \text{ and } \kappa = \kappa_0$
$\mathcal{M}_{E1,E1,0}^{GEV}$	$\xi(t) = \exp(\xi_0 + \xi_1 t), \alpha(t) = \exp(\beta_0 + \beta_1 t), \text{ and } \kappa = \kappa_0$
Model structure: $PE3(\xi(t), \beta \text{ or } \beta(t), \alpha)$ and $LPE3(\xi(t), \beta \text{ or } \beta(t), \alpha)$	
$\mathcal{M}_{1,0,0}^{PE3}, \mathcal{M}_{1,0,0}^{LPE3}$	$\xi(t) = \xi_0 + \xi_1 t, \beta = \beta_0, \text{ and } \alpha = \alpha_0$
$\mathcal{M}_{E1,0,0}^{PE3}, \mathcal{M}_{E1,0,0}^{LPE3}$	$\xi(t) = \exp(\xi_0 + \xi_1 t), \beta = \beta_0, \text{ and } \alpha = \alpha_0$
$\mathcal{M}_{2,0,0}^{PE3}, \mathcal{M}_{2,0,0}^{LPE3}$	$\xi(t) = \xi_0 + \xi_1 t + \rho_2 t^2, \beta = \beta_0, \text{ and } \alpha = \alpha_0$
$\mathcal{M}_{1,1,0}^{PE3}, \mathcal{M}_{1,1,0}^{LPE3}$	$\xi(t) = \xi_0 + \xi_1 t, \beta(t) = \beta_0 + \beta_1 t, \text{ and } \alpha = \alpha_0$
$\mathcal{M}_{E1,1,0}^{PE3}, \mathcal{M}_{E1,1,0}^{LPE3}$	$\xi(t) = \exp(\xi_0 + \xi_1 t), \beta(t) = \beta_0 + \beta_1 t, \text{ and } \alpha = \alpha_0$
$\mathcal{M}_{1,E1,0}^{PE3}, \mathcal{M}_{1,E1,0}^{LPE3}$	$\xi(t) = \xi_0 + \xi_1 t, \beta(t) = \exp(\beta_0 + \beta_1 t), \text{ and } \alpha = \alpha_0$
$\mathcal{M}_{E1,E1,0}^{PE3}, \mathcal{M}_{E1,E1,0}^{LPE3}$	$\xi(t) = \exp(\xi_0 + \xi_1 t), \beta(t) = \exp(\beta_0 + \beta_1 t), \text{ and } \alpha = \alpha_0$

*The model ID is abbreviated in the form of $\mathcal{M}_{p_1,p_2,p_3}^D$, where the superscript D denotes the probability distribution, and the subscripts p_1 , p_2 , and p_3 denote if the nonstationarity is modeled using the time-dependent location, scale, and shape parameters, respectively. The subscript “0” indicates the temporal invariance of the corresponding distribution parameter. The subscripts “1”, “E1”, and “2” indicate that the evolution of the corresponding distribution parameter is expressed as linear, exponential, and quadratic functions of time, respectively.

parameter estimation (point estimates) can then be carried out employing the stabilized Ψ_j . The uncertainty metrics are calculated using several additional pseudo-time steps until they achieve stability as well.

In the PF, several hyperparameters need to be pre-determined. For instance, the number of particles (m) is selected in a trade-off between the precision of the estimates (converge to the exact solution when $m \rightarrow \infty$) and the computational demand. In this paper, the m was set to 5,000, while the number of pseudo-time steps to yield the point estimates was 1,000, and the number of additional pseudo-time steps to stabilize the uncertainty approximations was 100. Please refer to Sen et al. (2020) and Vidrio-Sahagún et al. (2021) for more details on the PF implementation and the setup of the hyperparameters in the S-FFA and NS-FFA.

2.3. Performance-based NS-FFA

2.3.1. Candidate nonstationary structures

Following the common practice of the performance-based approach, several candidate nonstationary structures, in which the distribution parameters (except the shape parameter) are either linear or nonlinear functions of the time, were considered and are shown in Table 1. The shape parameter was treated as constant, as its estimation is difficult due to sample size limitations and particularly unrealistic when it is allowed to vary (Coles, 2001; Katz, 2013). More complicated model structures can be taken into consideration but were not included as their parametrization could be highly uncertain (Serago and Vogel, 2018; Serinaldi and Kilsby, 2015). The optimal NS-FFA model was selected from the listed candidate models according to their performance.

2.3.2. Performance evaluation metrics

The performance was assessed in terms of fitting efficiency and uncertainty. The former has been commonly evaluated using the AIC ,

which deals with the trade-off between the goodness-of-fit offered by a model and its complexity on the basis of information loss. The latter was evaluated using the CWI (Kasiviswanathan et al., 2019). The CWI simultaneously assesses the bandwidth and the observations coverage of the uncertainty band, which conflict with each other. The AIC and CWI were calculated using the derived point estimates of the model parameters and the estimated quantiles by the full set of particles of the PF, respectively, by:

$$AIC = n \log(RMSE) + 2N_\theta \quad (5)$$

$$CWI = AW \cdot \exp\left(\left[1 - \alpha_{CWI}\right] - \frac{POC}{100}\right)^2 \quad (6)$$

where α_{CWI} is the significance level (0.05 here); and the root mean square error ($RMSE$), the AW , and percentage of coverage (POC) of the uncertainty bands are calculated by:

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (m_k - o_k)^2} \quad (7)$$

$$AW = \frac{1}{n} \sum_{k=1}^n (m_k^U - m_k^L) \quad (8)$$

$$POC = \frac{1}{n} \sum_{k=1}^n c_k, \text{ where } c_k = \begin{cases} 1 & \forall k \text{ s.t. } m_k^L \leq o_k \leq m_k^U \\ 0 & \text{elsewhere} \end{cases} \quad (9)$$

where o_k are the empirical quantiles according to the plotting position formula $p_{1:n} = (r - 0.35)/n$ (where r is the rank of the k^{th} observation); m_k are the corresponding modeled quantiles; and m_k^U and m_k^L are the upper and lower uncertainty bounds of the estimates of the k^{th} observation, respectively. The empirical plotting position formula was selected due to its better performance for estimating the empirical quantiles compared to other formulas for extreme value distributions (Hosking, 1990). A model yielding a smaller AIC is more efficient in fitting, while a model with a lower CWI is less uncertain.

2.4. Decomposition-based NS-FFA

2.4.1. Proposed decomposition procedure

The decomposition-based NS-FFA relies on the widely adopted statistical representation of nonstationary processes (Y_t) in the hydrometeorological literature (e.g., Koutsoyiannis and Montanari, 2015; Milly et al., 2015; Serinaldi et al., 2018), namely:

$$G[Y_t] = G[X_t] + Z_t \quad (10)$$

where X_t and Z_t are the stationary stochastic and the time-dependent deterministic components, respectively; and $G[\cdot]$ is a generic operation, such as $E[\cdot]$ or $Var[\cdot]$. The stochastic processes Y_t and X_t are the families of random variables that describe the evolution of the underlying time-dependent and time-independent physical processes over the period of interest (i.e., for $t = t_1, t_2, \dots, t_n$), respectively. In this general decomposition, Z_t governs the changes of $G[Y_t]$ over time.

In the decomposition-based approach, X_t is estimated by removing Z_t from Y_t . A novel procedure that strictly follows the theoretical decomposition of nonstationary stochastic processes was proposed herein to remove Z_t in the first two statistical moments according to the definition of weak/second-order stationarity (Lindgren et al., 2013; von Storch and Zwiers, 2002). Differing from the previous methodology of Cunderlik and Burn (2003), the proposed procedure mathematically guarantees the removal of the deterministic component in the first two moments without altering higher-order moments and thus preserves the stationary stochastic component. More details on the procedure of the benchmark method and its issues in preserving the stochastic component and its statistical moments according to the theoretical deduction are provided in Appendix A. In this paper, μ_{X_t} , σ_{X_t} , and γ_{X_t} denote the

unconditional (stationary) mean, standard deviation, and skewness of X_t , while μ_{Y_t} , σ_{Y_t} , and γ_{Y_t} are the conditional (nonstationary) mean, standard deviation, and skewness of Y_t . The conditional statistics refer to the time-dependent statistical properties of the underlying nonstationary stochastic process. Based upon the second-order stationarity, three general classes of nonstationary scenarios were considered herein; namely, Class 1 (C1) – presence of Z_t in the μ_{Y_t} only, Class 2 (C2) – presence of Z_t in the σ_{Y_t} only, and Class 3 (C3) – presence of Z_t in both the μ_{Y_t} and σ_{Y_t} .

For datasets in C1 (i.e., whose process decomposition consists of $E[Y_t] = E[X_t] + Z_t$), the stationary dataset $x_t = \{x_{t_1}, x_{t_2}, \dots, x_{t_n}\}$ was obtained following an additive scheme by simply subtracting Z_t in the form of $f(t)$ from the AMS ($y_t = \{y_{t_1}, y_{t_2}, \dots, y_{t_n}\}$):

$$x_t = y_t - f(t) \quad (11)$$

When $f(t) = 0 \forall t$, it is the special case of the absence of Z_t in the mean, and thus Y_t is stationary. Equation (11) ensures that only the mean of the dataset is perturbed after Z_t is removed, while both the conditional standard deviation and skewness are preserved:

$$\mu_{X_t} \equiv E[X_t] = E[Y_t - Z_t] = E[Y_t] - Z_t = \mu_{Y_t} - f(t) \quad (12)$$

$$\begin{aligned} \sigma_{X_t}^2 &\equiv E[(X_t - \mu_{X_t})^2] = E[(\{Y_t - Z_t\} - \{\mu_{Y_t} - Z_t\})^2] = E[(Y_t - \mu_{Y_t})^2] \\ &= \sigma_{Y_t}^2 \end{aligned} \quad (13)$$

$$\begin{aligned} \gamma_{X_t} &\equiv \frac{E[(X_t - \mu_{X_t})^3]}{\left\{E[(X_t - \mu_{X_t})^2]\right\}^{3/2}} = \frac{E[(\{Y_t - Z_t\} - \{\mu_{Y_t} - Z_t\})^3]}{\left\{E[(\{Y_t - Z_t\} - \{\mu_{Y_t} - Z_t\})^2]\right\}^{3/2}} \\ &= \frac{E[(Y_t - \mu_{Y_t})^3]}{\left\{E[(Y_t - \mu_{Y_t})^2]\right\}^{3/2}} = \gamma_{Y_t} \end{aligned} \quad (14)$$

For datasets in C2 (i.e., whose process decomposition consists of $\text{Var}[Y_t]^{1/2} = \text{Var}[X_t]^{1/2} + Z_t$), the x_t was obtained following a multiplicative scheme by removing Z_t in the form of $g(t)$:

$$x_t = \hat{\mu}_{Y_t} + (y_t - \hat{\mu}_{Y_t}) \cdot h(t); \text{ where } h(t) = \frac{\hat{\sigma}_{X_t}}{\hat{\sigma}_{X_t} + g(t)} \quad (15)$$

where $\hat{\mu}_{Y_t}$ is the estimated mean of Y_t , which is time-independent for datasets in C2; $h(t)$ is a normalized multiplicative term depicting Z_t ; $g(t)$ is the estimated Z_t in the σ_{Y_t} ; and $\hat{\sigma}_{X_t}$ is the estimated standard deviation of X_t , using the beginning of the observation period as the reference. When $g(t) = 0 \forall t$ (and thus $h(t) = 1$), it is the special case of the absence of Z_t in the standard deviation, and thus Y_t is stationary. Similar to illustrated in Eqs. (12) – (14), Eq. (15) also ensures that only the standard deviation is perturbed (as $\sigma_{X_t}^2 = h(t)^2 \cdot \sigma_{Y_t}^2$) after removing Z_t from the dataset, while both the conditional mean and skewness are preserved, i.e., $\mu_{X_t} = \mu_{Y_t}$ and $\gamma_{X_t} = \gamma_{Y_t}$ (refer to Appendix B). Note that since Z_t is deterministic and/or independent, its covariances with X_t and Y_t are zero and thus are not included in the formulation.

For datasets in C3 (i.e., whose process decomposition consists of $E[Y_t] = E[X_t] + Z_{1,t}; \text{Var}[Y_t]^{1/2} = \text{Var}[X_t]^{1/2} + Z_{2,t}$), the x was obtained by combining Eqs. (11) and (15):

$$x = \hat{\mu}_{Y_t-f(t)} + (\{y_t - f(t)\} - \hat{\mu}_{Y_t-f(t)}) \cdot h(t); \text{ where } h(t) = \frac{\hat{\sigma}_{X_t}}{\hat{\sigma}_{X_t} + g(t)} \quad (16)$$

where $\hat{\mu}_{Y_t-f(t)}$ is the estimated mean of the partially decomposed dataset after $Z_{1,t}$ ($f(t)$) is removed from μ_{Y_t} but $Z_{2,t}$ ($g(t)$) is still present in σ_{Y_t} . It can be also demonstrated that Eq. (16) ensures that only the mean and standard deviation are perturbed (as $\mu_{X_t} = \mu_{Y_t} - f(t)$ and $\sigma_{X_t}^2 = h(t)^2 \cdot \sigma_{Y_t}^2$) after removing Z_t from both the mean and standard deviation, while the

conditional skewness is preserved, i.e., $\gamma_{X_t} = \gamma_{Y_t}$ (refer to Appendix B). The above Eqs. (10) – (16) are applicable for any form of Z_t in the mean ($f(t)$) and/or the standard deviation ($g(t)$).

In the proposed decomposition procedure, the deterministic component Z_t is a function of time that captures the available knowledge of the nonstationarity. The Z_t is often expressed in a parsimonious mathematical form determined by exploratory statistical analysis. Hence, in this paper, the Z_t was determined as a linear function of time and parameterized using the non-parametric Sen's slope estimator for estimating the trend in the μ_{Y_t} and the Sen's slope estimator coupled with the moving window technique for estimating the trend in the σ_{Y_t} . The hyperparameters of the moving window method, namely the window length and step, were set to 10 and 5 years, as they were found optimal in terms of bias among the ranges of window length (from 10 to 50 years) and window step (from 1 year to the window length) for the synthetic nonstationary GEV and PE3 datasets (Appendix C). Note that in the proposed decomposition procedure, the specification of Z_t is not limited to the methods mentioned above but can be further tailored whenever a more elaborate characterization of the nonstationarity (e.g., using a nonlinear function) is available.

2.4.2. Determination of the NS-FFA model

The distribution was selected based on the stationary component, i.e., x_t . Specifically, the distribution was selected from the pool of candidate distributions using the L-moment ratio diagram (Hosking, 1990), which has been commonly used and acknowledged to be robust in the S-FFA (e.g., Nguyen et al., 2017; Ouarda and Charron, 2019; Papalexiou and Koutsoyiannis, 2013). The difference between the L-kurtosis of the sample and the candidate theoretical distributions, which is referred to as L-kurtosis discrepancy, was used to select the optimal distribution (Hosking and Wallis, 1997).

The nonstationary structure of the NS-FFA model was identified according to the distribution moment equations (e.g., see Stedinger et al., 1993 or Gado and Van Nguyen, 2016a) based on Z_t . For instance, a trend in the mean would be properly captured by a model with a time-dependent location parameter, while a trend in the standard deviation would be expressed by both the time-dependent location and scale parameters for the three candidate distributions. As linear trend(s) were considered in this paper, the nonstationary structures considered are $\mathcal{M}_{1,0,0}^D$ or $\mathcal{M}_{1,1,0}^D$ (where D is the distribution) in Table 1.

2.5. Simulation study

A simulation study was conducted to demonstrate the advancements of the proposed decomposition method over the method employed by Cunderlik and Burn (2003) (referred to as the benchmark method herein). The comparison between these methods relied on the relative bias of the statistical moments of the derived stationary stochastic component with respect to the true statistical moments (known a priori). The synthetic nonstationary datasets (AMSSs) have stationary stochastic components following the GEV and PE3 distributions and are called the nonstationary GEV and PE3 datasets, respectively. Three classes of nonstationary datasets were generated, namely C1 with $\mu_{Y_t} = \mu_{X_t} + at$, where $\mu_{X_t} = 100$, and $a = 0.4$; C2 with $\sigma_{Y_t} = \sigma_{X_t} + bt$, where $\sigma_{X_t} = 45$ and $b = 0.2$; and C3 with $\mu_{Y_t} = \mu_{X_t} + at$ and $\sigma_{Y_t} = \sigma_{X_t} + bt$ with the same a and b as C1 and C2, respectively. The skewness of all datasets, γ_{X_t} , is time-independent and is equal to 3. All these selected values are within their commonly reported ranges in several studies (Agilan and Umapath, 2017; Ouarda et al., 2018; Papalexiou and Koutsoyiannis, 2013; Sun et al., 2015; Villarini and Smith, 2010). The nonstationary GEV datasets for each class were generated by randomly sampling Eq. (2) with the time-dependent distribution parameters derived from the moment equations (Stedinger et al., 1993; Stedinger, 2017). The nonstationary PE3 datasets were generated employing Eq. (4) and the designated time-dependent moment(s) for each class. The sample size of

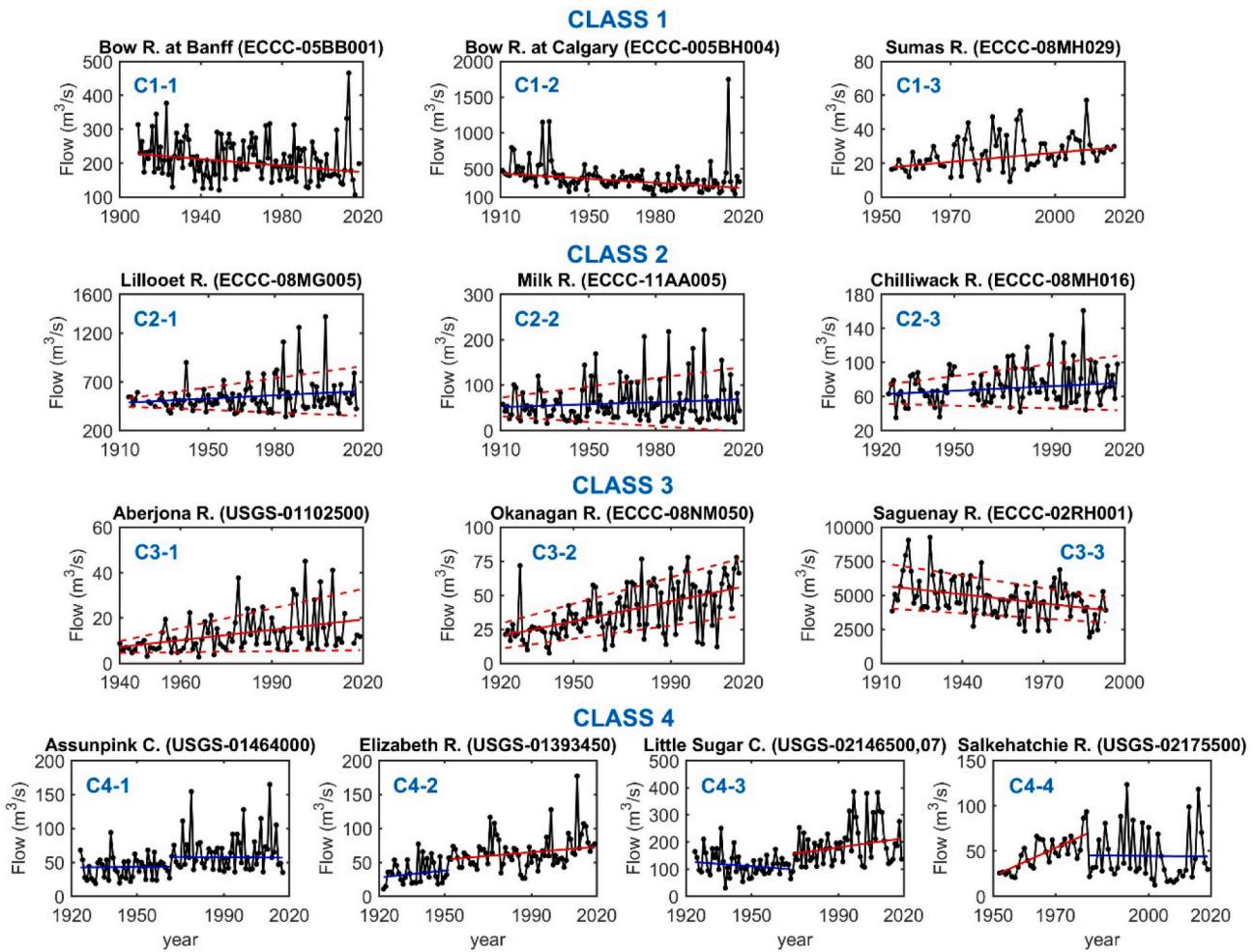


Fig. 1. Classification of the annual maximum flow datasets according to the perceived patterns of nonstationarity, namely C1, C2, C3, and C4 (i.e., exhibiting significant temporal trends in the mean, standard deviation, both mean and standard deviation, and a significant change point, respectively). The significant trend in the mean is indicated by the solid red line; the significant trend in the standard deviation around the mean is indicated by the dashed red lines; and the insignificant trend in the mean is indicated by the blue solid line (in C2 and C4), which indicates the time-invariant mean.

all the synthetic datasets was fixed to 100, which is often considered long in practice and is similar to the sample size of the real datasets utilized in this paper. For each distribution and class, 5,000 datasets were generated, and thus a total of 30,000 datasets were employed in this simulation study.

To further investigate the proposed decomposition-based approach in the NS-FFA, its performance was assessed and compared with the performance-based approach. The fitting efficiency and uncertainty of the models determined by both approaches were examined in terms of AIC and CWI, respectively. These metrics were estimated in two different ways, namely, with respect to the empirical and the true quantiles, which are derived from the sample datasets and the a priori known true nonstationary distribution, respectively. These two assessment settings are named the *theoretical* and *empirical* assessments herein. The use of different references to calculate the performance metrics in the two assessment settings led to the differences in the AIC and CWI. The AIC was employed as the selection criterion in the performance-based approach for the comparison, as it has been commonly used in the literature. The number of nonstationary datasets considered for the simulation study was determined such that the estimates for all the assessments were stabilized.

2.6. Application on real datasets

Flow AMSSs of thirteen stations retrieved from the Environment and

Climate Change Canada (ECCC) and the United States Geological Survey (USGS) were employed. The exploratory analyses were performed to detect the presence of temporal trends and change points in these datasets, which are indicative of the presence of nonstationarity. The non-parametric Mann-Kendall (MK) test was employed to detect monotonic trends in the mean. The MK test was coupled with the moving window method with the window length of 10 and the window step of 5 to investigate the temporal trends in the standard deviation. Higher-order statistical moments were not investigated due to their high sensitivity to extreme events, limited sample sizes, and consequently the difficulty in their estimation (Griffis and Stedinger, 2009; Papalexiou and Koutsoyiannis, 2013). The non-parametric Mann-Kendall-Sneyers (MKS) and Mann-Whitney-Pettitt (MWP) tests were employed to detect the presence of change point(s) in the mean. The detection of change points in the standard deviation was not conducted due to the limited sample sizes. A significance level of 0.05 was used in all these tests.

The perceived nonstationarity in these datasets exhibits different patterns in the form of significant trends and/or change points. Based upon the presence or absence of significant trends in the mean, standard deviation, and/or change points, the AMSSs were grouped into the C1, C2, and C3, plus an additional class (Class 4 (C4)) – more complex patterns containing a significant change point. No change points were detected in the datasets of C1, C2, and C3. The classification of the datasets, as well as their corresponding IDs, are presented in Fig. 1. Among the thirteen datasets, several of them (including C1-1, C3-1, C4-

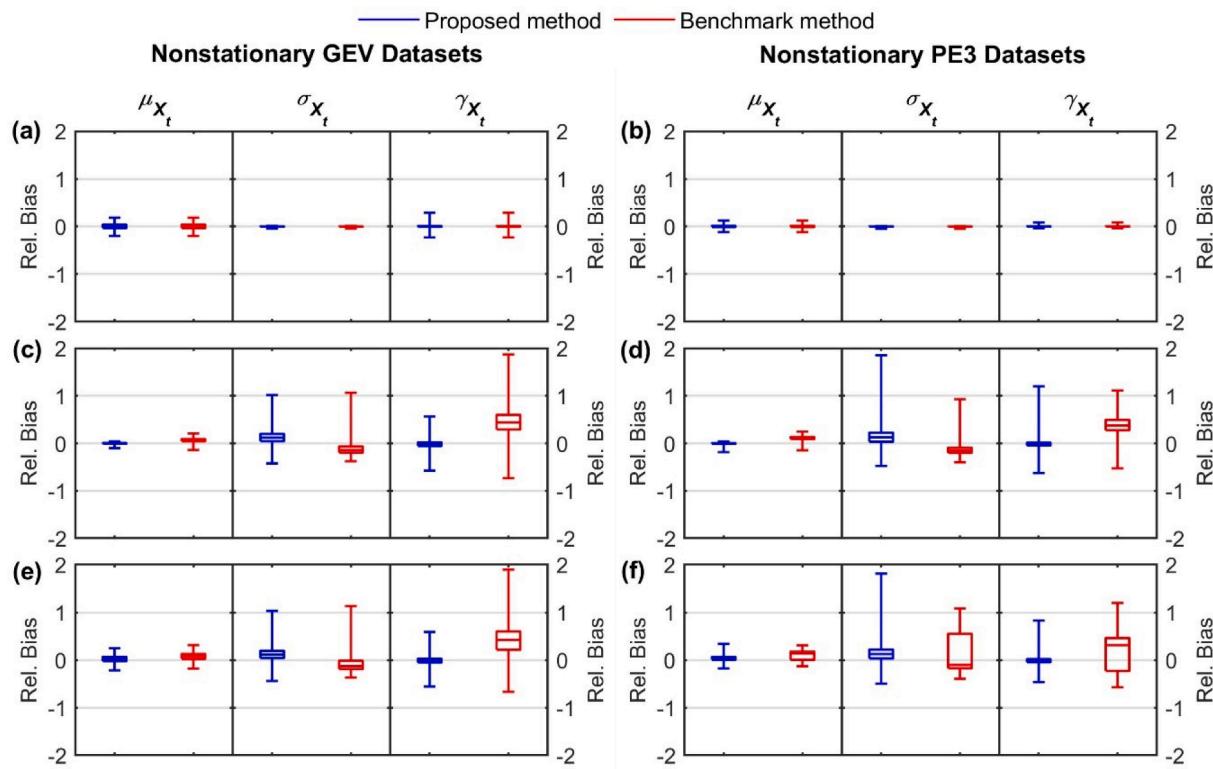


Fig. 2. The relative bias of the estimated moments of the stationary stochastic component derived using the proposed method and the benchmark method for the synthetic nonstationary GEV and PE3 datasets of Class 1 ((a) and (b)), Class 2 ((c) and (d)), and Class 3 ((e) and (f)).

1, C4-2, C4-3, and C4-4) have been also used in previous studies on the NS-FFA (e.g., Ammar et al., 2020; Luke et al., 2017; Obeysekera and Salas, 2014; Salas and Obeysekera, 2014; Villarini et al., 2009). The performance assessment of both approaches was also conducted in terms of *AIC* and *CWI*. In this case, only the empirical assessment was undertaken because the true quantiles are unknown in real practice.

3. Results and discussion

3.1. Simulation study

3.1.1. Advancements of the proposed decomposition procedure

Fig. 2 shows the relative bias of the proposed and benchmark methods. For the synthetic datasets of C1, the two methods yielded the same estimates of μ_{X_t} , σ_{X_t} , and γ_{X_t} , as they use the same procedure to remove the Z_t in μ_{Y_t} . In contrast, the improvements of the proposed method over the benchmark method are apparent for the datasets of C2 and C3. The medians of the estimates of μ_{X_t} , σ_{X_t} , and γ_{X_t} of the proposed method are unbiased or close to unbiased (close to zero). In particular, the proposed method considerably reduced the relative bias in the estimate of γ_{X_t} , which describes the right-tail of the distribution and consequently is particularly important in the distribution selection. In addition, the proposed method offered a narrower or approximately the same interquartile range of the estimates of μ_{X_t} , σ_{X_t} and γ_{X_t} . Similarly, the full ranges of the relative bias in the proposed method are narrower than or approximately equal to those in the benchmark method (except for the estimates of σ_{X_t} for the nonstationary PE3 datasets). These results demonstrate that the proposed method can not only reduce the biases in the estimates, especially of higher-order moments, but also is more robust compared to the benchmark method. Consequently, the proposed method would improve the distribution selection using the derived stationary stochastic component and in turn further advance the decomposition-based NS-FFA.

3.1.2. Performance of the proposed decomposition-based NS-FFA

To further assess the proposed decomposition-based approach, its performance was compared with that of the performance-based approach. As shown in Fig. 3, the decomposition-based approach yielded consistently lower medians and first and third quartiles of *AICs* in all the nonstationarity classes and both parent distributions in the theoretical assessment (i.e., with respect to the true quantiles). In contrast, the performance-based approach produced lower *AICs* (medians, first and third quartiles, and minimum and maximum values) throughout all the datasets in the empirical assessment (i.e., with respect to the empirical quantiles). In terms of the *CWI*, the decomposition-based approach outperformed the performance-based approach in the theoretical assessment, as it yielded lower medians and first and third quartiles, whereas both approaches performed equivalently in the empirical assessment. Moreover, the decomposition-based approach was always superior to the performance-based approach in capturing the true nonstationary structure and distribution, as well as the true NS-FFA model (the combination of the two model components) (Fig. 4). In the simulation study, the decomposition-based approach always selected the true nonstationary structure, as the a priori known temporal trends were directly employed to determine the nonstationary structure. Whereas the true distribution and thus the NS-FFA models were selected on average in 56.7% of the datasets (within the range of 45.6% to 66.7%). The incorrect selections might be ascribed to the sample L-ratios falling approximately in the middle of two theoretical distributions for several datasets. Whereas the performance-based approach selected the true distribution and nonstationary structure on average in 41.4% (within the range of 32.6% to 48.9%) and 17.6% (within the range of 5.7% to 26.4%) of the datasets, respectively, resulting in the selection of true models on average in only 7.8% (within the range of 3.1% to 10.5%) of the datasets.

These results reveal that the two NS-FFA approaches often yielded different models, namely different distributions and/or nonstationary structures, and in turn, different performance. In the performance-based

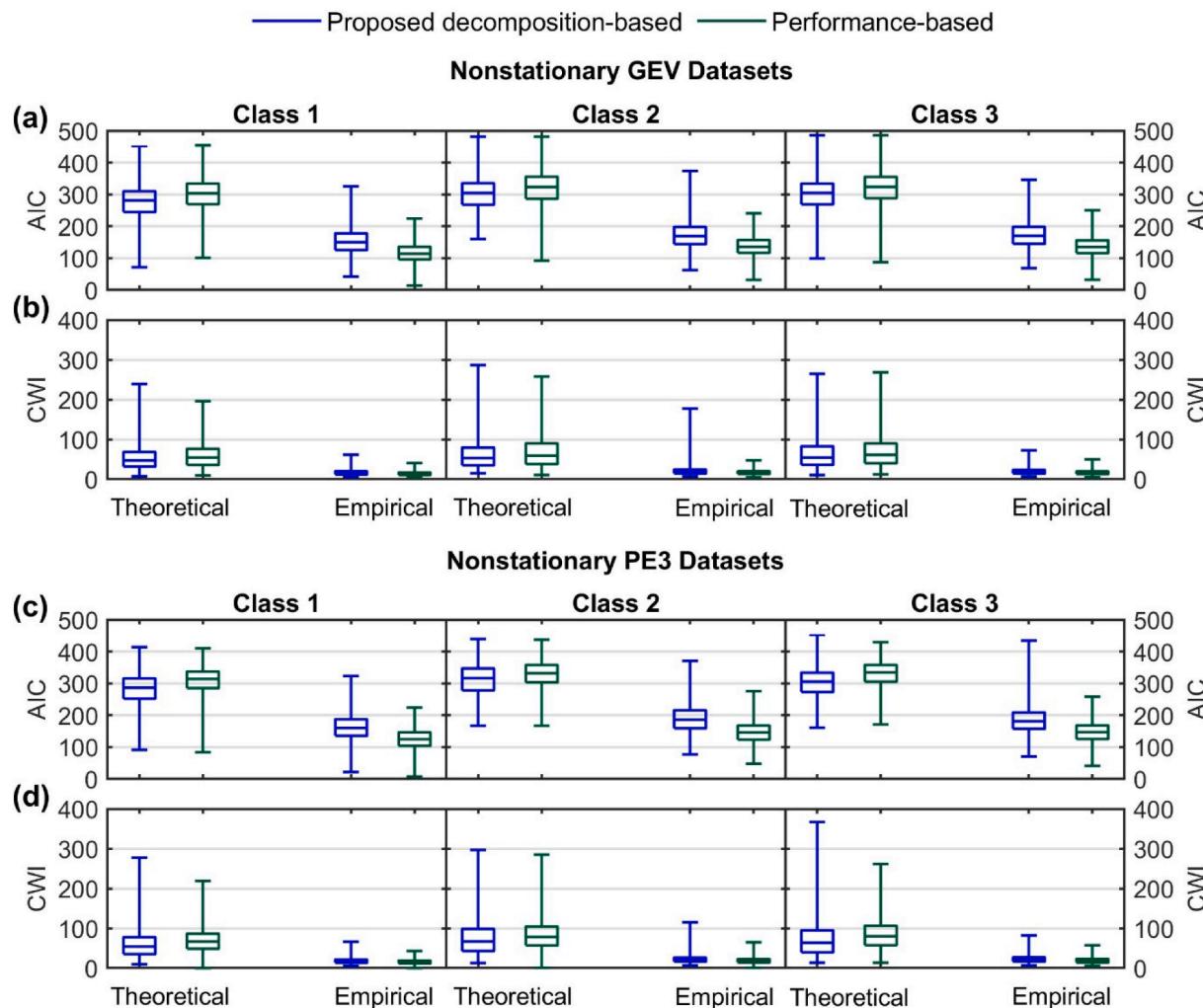


Fig. 3. The performance of the proposed decomposition-based and the performance-based approaches for the synthetic nonstationary GEV and PE3 datasets in terms of fitting efficiency ((a) and (c)) and uncertainty ((b) and (d)) in the theoretical and empirical assessments. The performance-based approach adopted the AIC as the selection criterion herein.

NS-FFA, the ergodicity required for making the statistical inference is completely violated, as both the distribution and the nonstationary structure are selected simultaneously by optimally fitting the nonstationary datasets. Thus, the performance-based approach may result in a NS-FFA model that does not capture the underlying nonstationary stochastic process well despite it might fit the observations adequately according to the pre-selected performance metric. The misspecification of the distribution and/or nonstationary structure can introduce errors in the quantile estimates in the fitting period and could be particularly critical when implementing the model for the out-of-sample predictions. In contrast, the decomposition-based NS-FFA avoids the ergodicity violation, as the distribution is selected from the decomposed stationary stochastic component in which the ergodicity holds, while the nonstationary structure is determined separately according to the available knowledge of the nonstationarity. This fact makes the decomposition-based approach especially advantageous over the performance-based approach from the theoretical point of view.

In addition, the simulation results demonstrate that the performance-based approach outperforms the decomposition-based approach in the empirical assessment, although it is inferior in the theoretical assessment. This argues that the performance-based approach tends to select untrue models due to overfitting the samples irrespective of the actual nonstationary stochastic process. In contrast, the decomposition-based approach is superior in capturing the underlying true nonstationary

process, as reflected by its higher percentage in selecting correct NS-FFA models as well as its higher fitting efficiency and the same or less uncertainty than the performance-based approach in the theoretical assessment. Therefore, both the theory and the simulation results advocate that the decomposition-based NS-FFA is preferable. In addition, the comparison of these two approaches was also conducted when the *BIC* instead of the *AIC* was used as the selection metric in the performance-based approach. The results ([Appendix D](#)) consistently supported the aforementioned findings.

3.2. Real applications

3.2.1. Implementation of the proposed decomposition-based approach

[Fig. 5\(a–c\)](#) show one example of decomposed datasets of Classes 1, 2, and 3, respectively. The proposed decomposition procedure effectively removed the significant monotonic trends in the mean and/or the standard deviation. No significant trends nor change points were detected in the decomposed datasets. These are indicative that the decomposed datasets are stationary. Since the presence of change points was not considered in the NS-FFA, the perceived monotonic trends over the whole observation period were utilized for the decomposition of datasets in Class 4. It is thus not surprising that the change point was still detected in two out of four decomposed datasets of this Class (i.e., C4-3 and C4-4). [Fig. 5\(d and e\)](#) illustrate the decomposed datasets C4-2

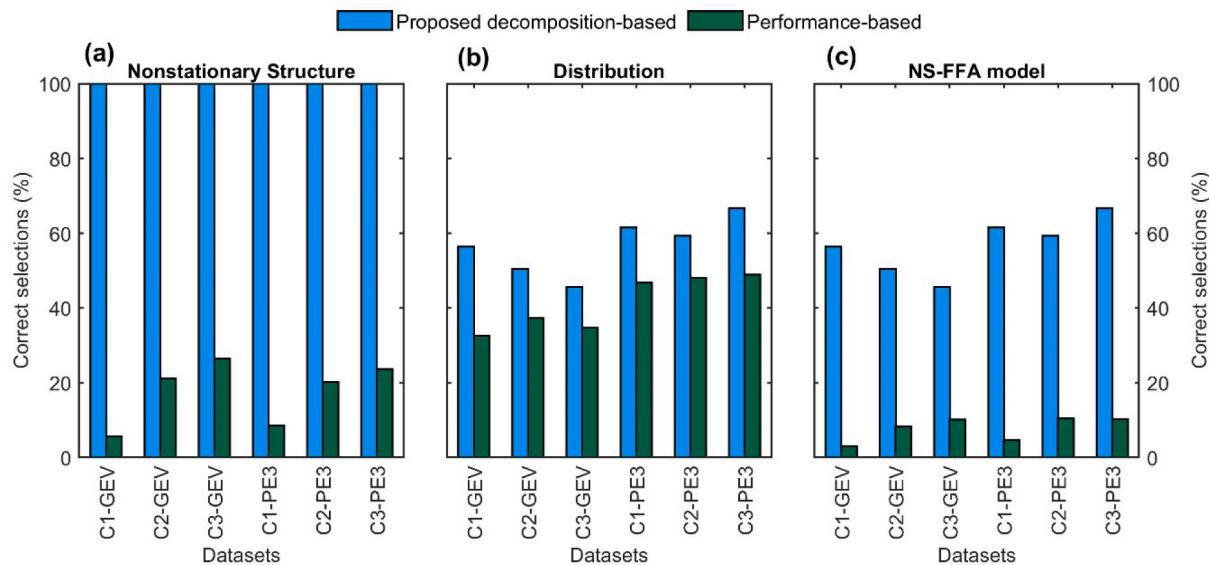


Fig. 4. Comparison of correct selections of (a) nonstationary structures, (b) distributions, and (c) NS-FFA models (i.e., both the nonstationary structures and distributions) between the proposed decomposition-based and the performance-based approaches. The datasets are abbreviated as CX-D, where the CX and D denote the nonstationary class (C1, C2, or C3) and the distribution (GEV or PE3) used for generating the synthetic datasets, respectively. The performance-based approach adopted the AIC as the selection criterion herein.

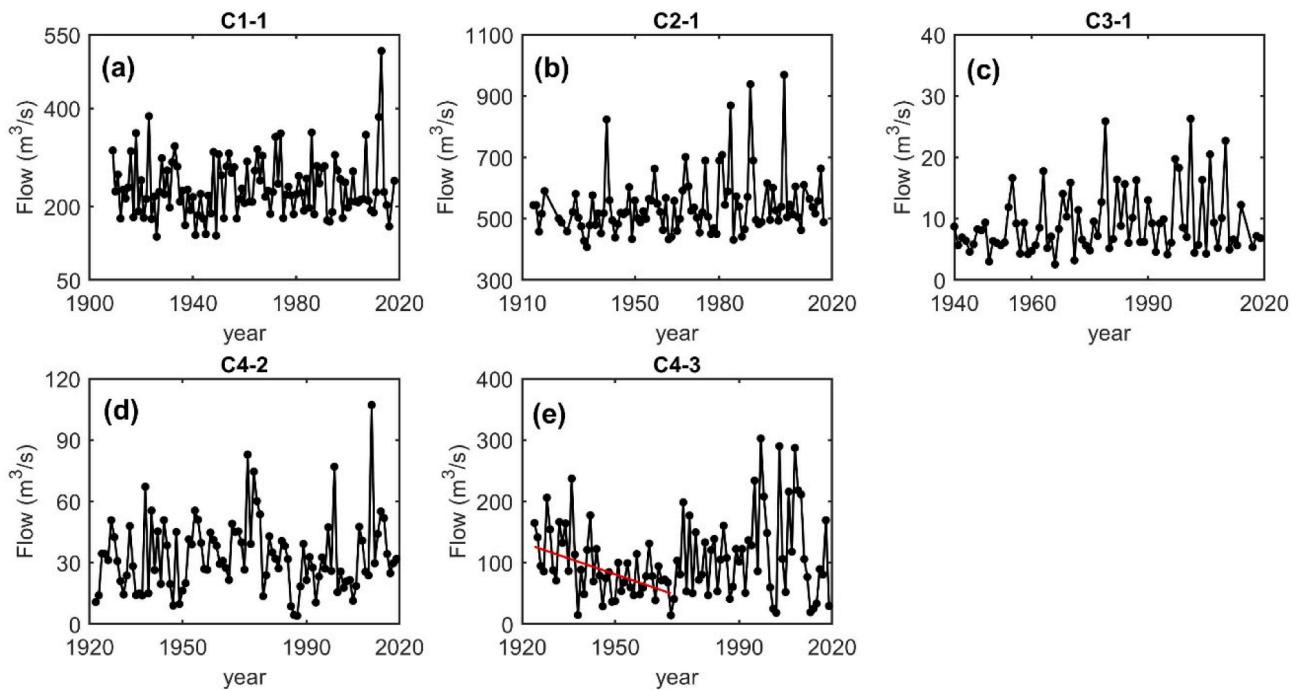


Fig. 5. Sample decomposed datasets derived in the proposed decomposition-based approach according to the perceived monotonic trends over the entire observational period: (a) decomposed C1-1, (b) decomposed C2-1, and (c) decomposed C3-1, which are all free from trends; whereas (d) decomposed C4-2 is free from trends and change points, and (e) decomposed C4-3 still has a significant change point and a significant trend in the mean (the red line) before the change point.

(without a change point) and C4-3 (still with a change point) as examples, respectively. This issue is also present in the performance-based approach, in which no candidate models considering change points were adopted in this paper and most of the literature due to the highly uncertain model parameterization. As a result, the proposed decomposition procedure effectively removed the perceived Z_t in datasets in Classes 1, 2, and 3, and consequently, the approaches commonly employed to select the distribution in the S-FFA (e.g., the L-moment ratio diagram) are applicable.

Table 2 shows the L-kurtosis discrepancies between each

decomposed dataset and the candidate theoretical distributions and highlights the best-fit distributions. Note that in some cases (e.g., C1-1 and C1-2), the difference in the L-kurtosis discrepancy between the best-fit and the second-best distributions is very minor (<0.005). This situation is also common in the S-FFA when the dataset's L-moment ratio locates approximately in the middle of two candidate theoretical distributions in the L-ratio diagram. When more than one distribution fits acceptably well the data, any of them could be a reasonable choice (Hosking and Wallis, 1997), although the distribution with slightly better fitting is usually selected (e.g., Cunderlik and Burn, 2003; Ouarda

Table 2

Selected distributions and nonstationary structures in the decomposition-based approach according to the L-kurtosis discrepancy and the significant perceived trends in each dataset (y_t) or its corresponding log-transformation (y_t^*), respectively. The selected distributions are indicated in boldface.

Dataset	Distribution Selection			Nonstationary structure determination	NS-FFA model		
	L-kurtosis discrepancy						
	GEV	PE3	LPE3				
C1-1	0.004	0.015	0.002	✓	–		
C1-2	0.069	0.152	0.067	✓	–		
C1-3	0.003	0.026	0.020	✓	–		
C2-1	0.043	0.102	0.060	–	✓		
C2-2	0.053	0.019	0.044	–	✓		
C2-3	0.018	0.003	0.023	–	✓		
C3-1	0.061	0.007	0.027	✓	✓		
C3-2	0.074	0.057	0.071	✓	✓		
C3-3	0.009	0.015	0.014	✓	✓		
C4-1	0.001	0.041	0.036	✓	–		
C4-2	0.011	0.025	0.004	✓	–		
C4-3	0.027	0.003	0.010	✓	–		
C4-4	0.086	0.042	0.061	–	–		

a) When GEV or PE3 distribution is selected, the trend(s) in the random variable y_t guide the nonstationary structure determination; when LPE3 distribution is selected, the trend(s) in the random variable y_t^* guide the nonstationary structure determination.

b) Stationary model, \mathcal{M}_S^D , where D denotes the distribution selected.

et al., 2019). Further investigation using additional statistical tests could be conducted if desired. However, this would merely be an additional effort to privilege certain aspects of interest of the fitting (as the *true* distribution is unknown) rather than an issue of statistical inference.

Table 2 also shows the selected nonstationary structures for each dataset according to the detected temporal trends in y_t or its corresponding log-transformation (y_t^*). In several datasets (i.e., C2-3, C3-1, C3-2, C3-3, and C4-2), the trends in the standard deviation of y_t were no longer significant in y_t^* . The possible absence of a trend in the standard deviation due to the log-transformation and consequently the use of a time-invariant scale parameter in the LPE3 distribution have been discussed in Luke et al. (2017) and Stedinger and Griffis (2011). Therefore, when only a trend in the standard deviation is present, the log-transformation may lead to determining a stationary structure for the LPE3 distribution, and consequently might be inadequate for conducting the NS-FFA sometimes. Furthermore, neglecting complex nonstationary patterns (e.g., datasets in Class 4) could yield misleading results in the analysis. For instance, a stationary model \mathcal{M}_S^{PE3} was determined for C4-4 irrespective of the perceived nonstationarity in the dataset (Table 2). Nevertheless, the confident identification of complex nonstationary patterns is difficult due to the sample size of AMSs and the potential noise of natural climate variability occurring on timescales of up to 30 years (François et al., 2019). This along with the high uncertainty in model parameterization constrains the use of complex nonstationary structures in the NS-FFA.

In these applications, the monotonic temporal trends were used as a proxy for Z_t . Despite that theoretically the violation of ergodicity is not completely avoided when using the perceived temporal trends (only) (Serinaldi et al., 2018), this can be resolved by incorporating deductive reasoning and/or the physical causative process(es) or driver(s) into the

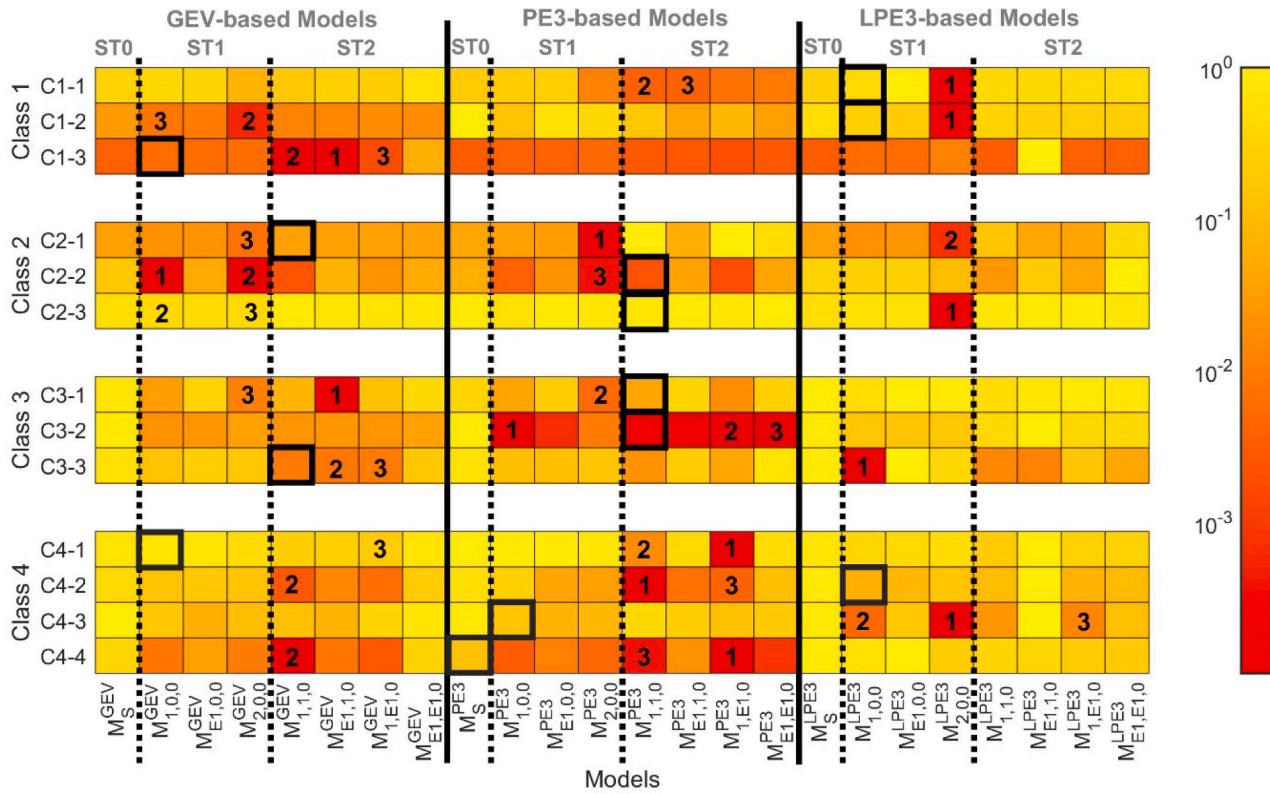


Fig. 6. The normalized AICs (shown by color at a log-scale) of the stationary and all nonstationary candidate models. The numerical indexes (1, 2, and 3) denote the top three models selected according to the AIC in the performance-based NS-FFA, while the solid black boxes highlight the models yielded by the decomposition-based NS-FFA. The vertical solid and dashed lines separate the models adopting the same distribution and time-variant distribution parameter(s). ST0 denotes the stationary structure; ST1 and ST2 are the nonstationary structures with the time-variant location parameter, and the time-variant location and scale parameters, respectively.

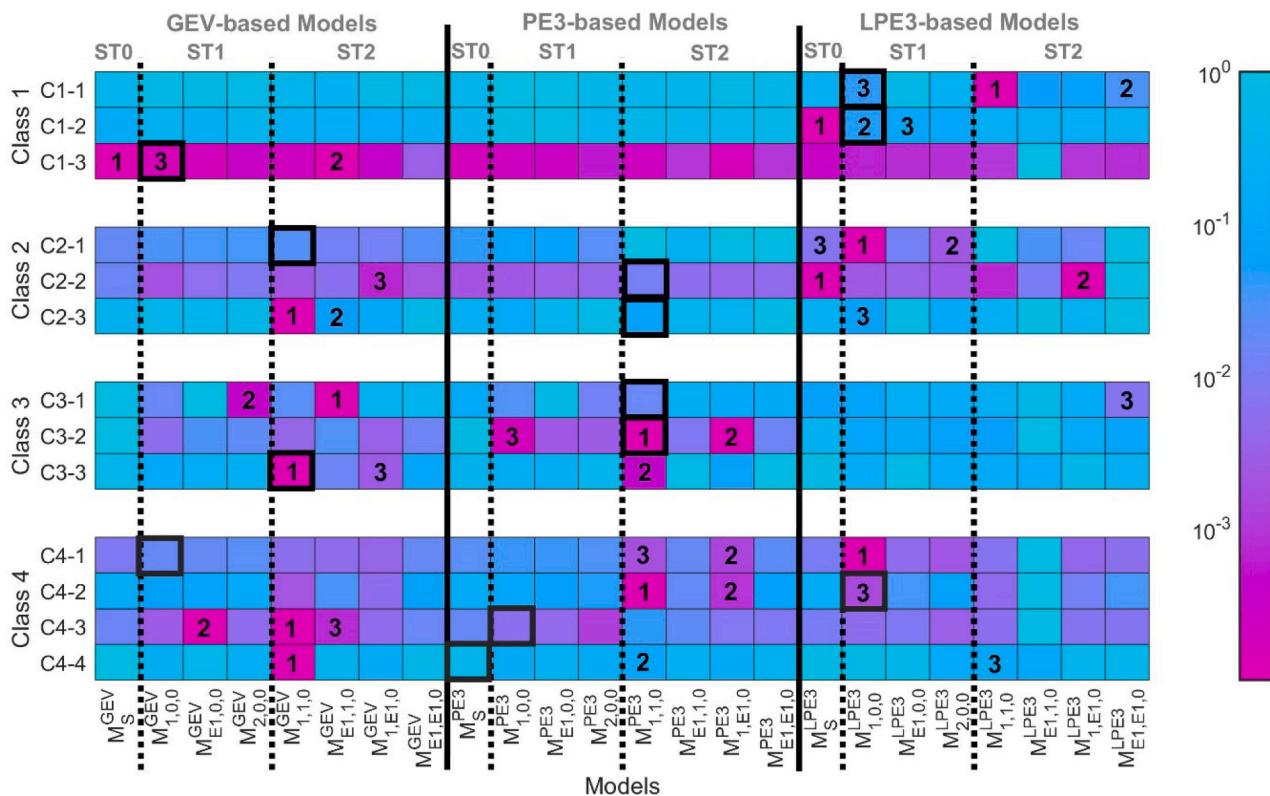


Fig. 7. The normalized CWI (shown by color at a log-scale) of the stationary and all nonstationary candidate models. The numerical indexes (1, 2, and 3) denote the top three models selected according to the CWI in the performance-based NS-FFA, while the solid black boxes highlight the models yielded by the decomposition-based NS-FFA. The vertical solid and dashed lines separate the models adopting the same distribution and time-variant distribution parameter(s). ST0 denotes the stationary structure; ST1 and ST2 are the nonstationary structures with the time-variant location parameter, and the time-variant location and scale parameters, respectively.

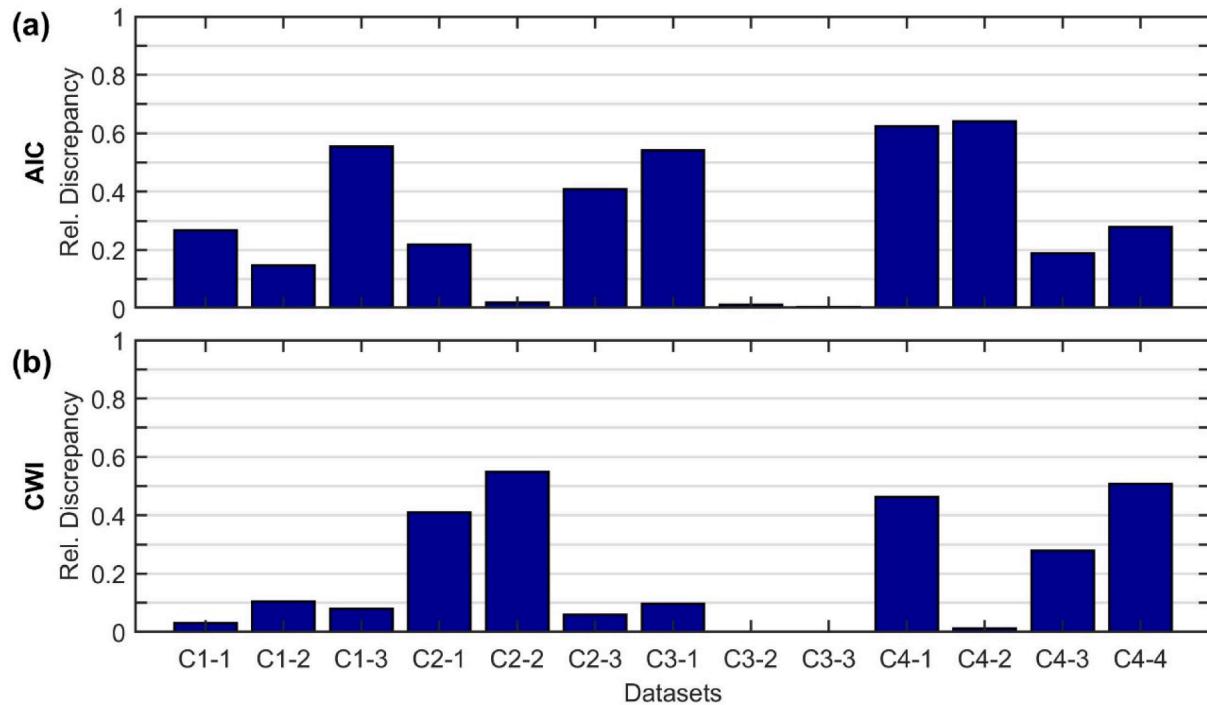


Fig. 8. Relative discrepancies in (a) the fitting efficiency (AIC) and (b) the uncertainty (CWI) of the model selected by the decomposition-based NS-FFA with respect to that by the performance-based NS-FFA. The performance-based approach adopted the AIC and CWI as the selection criteria for (a) and (b), respectively.

analysis whenever they can be properly identified. Case-specific investigations in this regard are beyond the scope of this paper. Furthermore, the determination of the nonstationary structure and the derivation of the stationary component in the proposed decomposition-based approach depends on the methods used to identify Z_t . Although statistical tests such as the MK test can suggest the temporal dependence of the statistical moments and subsequently the corresponding distribution parameter(s), they do not discriminate if the dependence is linear or nonlinear. This paper represented the nonstationarity using linear trends. This is practically feasible and convenient considering the limitations to characterize Z_t in a higher degree of detail. An appropriate and systematic delineation (and attribution) of Z_t is key to reliably implementing the proposed decomposition-based NS-FFA.

3.2.2. Comparison of the proposed decomposition-based and performance-based approaches

Figs. 6 and 7 show the NS-FFA models determined by the proposed decomposition-based approach (with the black box) and selected from the candidate nonstationary models by the performance-based approach (with the numerical index “1”) in terms of *AIC* and *CWI*, respectively, for each dataset. Note that both the *AIC* and *CWI* are normalized into the range of [0 1] for each dataset. The top second and third NS-FFA models selected by the performance-based approach are also indicated with the indexes “2” and “3”, respectively. The corresponding stationary model counterparts (\mathcal{M}_S^{GEV} , \mathcal{M}_S^{PE3} and \mathcal{M}_S^{LPE3}) are also included in these figures. As the nonlinear nonstationary structures were considered in the performance-based approach (Table 1) but not in the decomposition-based approach, the comparison of the selected models determined by the two approaches was conducted in a broad sense; namely, comparing if the models have the same distribution and the same time-variant location parameter or time-variant location and scale parameters in the nonstationary structure. As shown in Figs. 6 and 7, these two approaches generally yielded different NS-FFA models (except in C1-1 and C1-2 in terms of *AIC*, and C3-2 and C3-3 in terms of *CWI*). On one hand, they selected different distributions in five (C2-1, C2-2, C2-3, C3-1, and C3-3) and four (C2-1, C2-2, C2-3, and C3-1) datasets in Classes 1–3 based on the *AIC* and *CWI*, respectively, and all datasets of Class 4 (except C4-4 based on the *AIC*). On the other hand, they also determined different nonstationary structures in five (C1-3, C2-1, C2-2, C2-3, and C3-2) and five (C1-1, C1-2, C1-3, C2-1, and C2-2) datasets in Classes 1–3 in terms of *AIC* and *CWI*, respectively, and half of the datasets of Class 4 (C4-1 and C4-4 in terms of *AIC*, and C4-3 and C4-4 in terms of *CWI*). Recall that different nonstationary structures were identified for y_t and y_t^* in some cases due to the log-transformation. These results evidence that these two NS-FFA approaches are fundamentally different in determining the model, as they are formulated following distinct logics.

Moreover, Fig. 8 shows the relative discrepancies in *AIC* and *CWI* of the model determined by the proposed decomposition-based NS-FFA with respect to that by the performance-based NS-FFA for all datasets. As shown in Fig. 8, the relative discrepancy in *AIC* ranges from zero to moderate (<0.30) in six out of nine datasets of Classes 1–3, whereas it is high (≥ 0.30) in the other three datasets (i.e., C1-3, C2-3, and C3-1). Among these three datasets, the high discrepancy in C1-3 and C3-1 could be ascribed to the small *AICs* (close to zero) of the models selected by the performance-based NS-FFA. In the datasets of Class 4, the discrepancy is high in two cases (C4-1 and C4-2). The discrepancy in *CWI* is in the range from zero to moderate in the datasets of Classes 1–3 overall (except in C2-1 and C2-2); while it is high in two datasets of Class 4 (C4-1 and C4-4). This comparison does not intend to demonstrate which approach outperformed the other. The two approaches often yield different NS-FFA models, and consequently, their different performance is expected. On one hand, the comparison might imply the possible overfitting by the performance-based approach as demonstrated in the simulation study. On the other hand, the zero to moderate discrepancies argue that the decomposition-based approach is in general competitive

to the performance-based approach from the perspective of the performance. Therefore, these results advocate the feasibility and applicability of the decomposition-based approach in real practice.

If using more flexible (e.g., nonlinear) nonstationary structures in the decomposition-based NS-FFA, its performance could be further enhanced in some cases. For instance, the discrepancy in *AIC* can be reduced (e.g., C2-2, C3-1, C3-2, and C3-3) or even eliminated (e.g., C1-1 and C1-2) when a non-linear nonstationary structure is adopted (Fig. 6). However, the use of a flexible and complicated nonstationary structure without justification could lead to the overfitting problem, especially in data-scarce locations. Using complicated nonstationary structures in the decomposition-based NS-FFA is mathematically feasible but requires understanding the physical mechanism(s)/driver(s) behind the nonstationarity to avoid overfitting. When they cannot be properly justified, the use of a parsimonious nonstationary structure (i.e., the linear nonstationary structure) might be preferred.

3.2.3. The challenges in implementing the performance-based NS-FFA

As discussed previously in the simulation study, the performance-based NS-FFA cannot avoid the violation of ergodicity and may lead to spuriously high performance. In this approach, there are two additional issues that deserve attention. As shown in Figs. 6 and 7, when two different distributions are coupled with their corresponding most fitting-efficient nonstationary structures (which are often different), they may yield approximately equal performance. For instance, for C1-1 the optimal model $\mathcal{M}_{2,0,0}^{LPE3}$ is marginally superior to $\mathcal{M}_{1,1,0}^{PE3}$ by 2.3% in *AIC*; while for C3-1, the optimal model $\mathcal{M}_{1,1,0}^{GEV}$ outperforms $\mathcal{M}_{2,0,0}^{PE3}$ by 5.4% (Fig. 6). Hence, a different nonstationary structure may be selected when selecting a different distribution, which in turn yields a completely distinct model in the performance-based NS-FFA. This is analog to the equifinality in hydrological modeling, namely, different model structures may be equally capable of fitting the observations according to a certain or a few pre-selected performance metric(s). These suggest that the selection of the nonstationary structure needs to be separated from the selection of the distribution in order to explicitly use the knowledge of the nonstationarity to guide the model determination, as in the decomposition-based NS-FFA.

The other challenge lies in the selection of the evaluation criterion to determine the optimal model in the performance-based approach. When using different evaluation metrics to assess the models' performance from a specific perspective, such as the fitting efficiency (e.g., using the *AIC* and *BIC*), the results were consistent in general (Appendix D). However, the observed differences reveal that they might lead to determining different NS-FFA models sometimes. When assessing the models' performance from different perspectives, such as the fitting efficiency and uncertainty, the results shown in Figs. 6 and 7 also reaffirm that different models may be selected by this approach when using different evaluation criteria (the *AIC* and *CWI* here). When using the *AIC*, this approach always led to the selection of a nonstationary model. However, it is not uncommon that the stationary models yield the lowest uncertainty when a temporal trend is present in either mean or standard deviation (e.g., C1-2, C1-3, and C2-2). This result supports the argument that the stationary model might outperform the nonstationary model from the uncertainty viewpoint even when the nonstationarity presents (Ouarda et al., 2019; Vidrio-Sahagún et al., 2021). Therefore, the optimal model selected by the performance-based approach is sensitive to the evaluation criterion used. The model has been commonly selected according to the fitting efficiency criterion in the performance-based NS-FFA. However, the uncertainty is also a key when selecting the model in the FFA, as it reflects the reliability of estimates and their applicability for design purposes (Serinaldi and Kilsby, 2015). Hence, the lack of agreement on which selection criterion should be employed challenges the implementation of the performance-based approach, as it would heavily depend on the performance criterion selected according to the preference of the modeler.

3.3. Future research recommendations

In this paper, the proposed decomposition method for the decomposition-based NS-FFA was theoretically justified and practically demonstrated through both the simulation study and practical applications using the datasets in C1, C2, and C3. Yet, more challenges arise when dealing with more complex nonstationary patterns in the NS-FFA, such as those in C4 (containing change point(s), Fig. 1) and/or nonlinear patterns. Mathematically, both the performance-based and decomposition-based approaches are capable of handling any form of nonstationarity (i.e., linear and non-linear, and/or with change point(s)). However, the performance-based method is challenged in practice when incorporating more complicated nonstationary structures, as their use may be difficult to justify and lead to overfitting, as demonstrated in the simulation study. Whereas the decomposition-based approach is challenged to identify a more complex deterministic component behind the nonstationarity to decompose the underlying dataset. To present, it is difficult to reliably characterize complicated nonstationary patterns unless clear mechanisms behind the nonstationarity can be identified. Thus, further research on nonstationarity characterization and attribution is required to ensure the successful practical implementation of the NS-FFA when more complex nonstationary patterns present.

In both the decomposition-based and performance-based approaches, the time was used as the covariate. However, it is worth mentioning that both approaches can employ the physical covariate(s) (e.g., climatic indexes and land use/cover parameters, among others) whenever they and/or their deterministic law of time governing their evolution can be properly identified. This would be achieved by the attribution of the nonstationarity to its physical driver(s)/causative process(es) (e.g., climate change and/or changes in the watershed resulting from human intervention). Hence, the proposed method as well as other NS-FFA approaches (e.g., the performance-based approach) would benefit from the improvement in the nonstationarity attribution, which is beyond the scope of this paper. In the decomposition-based approach, this would be particularly relevant to the identification of the physical driver(s) behind the nonstationarity and thus Z_t . Recall that the existence of a deterministic component Z_t behind the observed changes in the dataset is necessary to justify the implementation of the NS-FFA in general. This requirement does not imply the need for perfect knowledge of Z_t but can be validated by deductive reasoning (Serinaldi et al., 2018). The nonstationary driver(s) can yield a Z_t enveloped by the variability due to chaotic nonlinear internal dynamics and unpredictable natural external forcings (Milly et al., 2015; Montanari and Koutsoyiannis, 2014). Hence, the proposed decomposition-based approach is also applicable for physical drivers showing a deterministic signal of change accompanied by stochastic variability, such as climate variables affected by climate change. The opportunity to explicitly incorporate physical knowledge of the deterministic changes in the system into the model development and the subsequent inference is an attractive feature of the decomposition-based approach.

It is noteworthy that uncertainty originates from distinct sources in different NS-FFA approaches. Differing from the performance-based approach, the decomposition-based method avoids the uncertainty in the model selection due to equifinal models according to the performance metric(s). Whereas uncertainty arises from the specifications of the deterministic component and the distribution selection (as in the S-FFA) in this approach. Thus, improvements in these two aspects would further reduce the uncertainty in the model determination of the decomposition-based approach.

Lastly, the decomposition-based NS-FFA might confront several practical challenges when making the out-of-sample predictions. When implementing the performance-based approach, it can yield such predictions by (a) directly fitting projected floods, (b) extrapolating the NS-FFA models based on the temporal covariate, and (c) extrapolating the models based on the physical covariate(s) and its projections. These

implementations would thus rely on the adequate projections of floods, the validity of the assumption that the temporal trend detected in the historical period remains the same in the future, and the reliable projections of the physical covariate(s) and the validity of its link function with the distribution parameters in the future, respectively. As similar strategies can be implemented when using the decomposition-based NS-FFA to make the out-of-sample predictions, the reliability of this approach is also subject to similar practical challenges as the performance-based approach.

4. Conclusion

Aiming to advance the NS-FFA, this paper proposed a novel decomposition procedure for determining the model in the decomposition-based approach, in which the distribution and the nonstationary structure are selected separately based upon the stationary stochastic component and the available knowledge of the nonstationarity. The proposed method, which strictly follows the theoretical decomposition of nonstationary stochastic processes, was compared with a previous decomposition method in both an analytical deduction and a simulation study. The analytical deduction demonstrated the theoretical advancement of the proposed method, while the simulation results confirmed its effectiveness in reducing the bias in the moment estimates of the derived stationary stochastic component, especially in the skewness that is key in the distribution selection. These demonstrated the improvement in the decomposition-based NS-FFA. Moreover, the proposed decomposition-based approach was further compared with the performance-based approach using both the synthetic and real flow AMSs exhibiting different patterns of nonstationarity. On one hand, the proposed decomposition-based NS-FFA was shown superior to the performance-based approach, as it captured the (known) underlying nonstationary process more efficiently and yielded a higher performance in the simulation. On the other hand, the performance-based approach was found prone to overfitting the synthetic datasets, while the zero to moderate discrepancies between the performance of the two approaches argued that the decomposition-based approach is competitive in general from the performance perspective in real applications. In addition, the performance-based approach was shown to be subject to the equifinality issue, as well as the determination of distinct models when using different performance evaluation metrics. Furthermore, differing from the performance-based approach, the decomposition-based NS-FFA avoided/alleviated the ergodicity violation, and was thus preferable from the theoretical viewpoint. Therefore, this paper demonstrated the superiority and convenience of the proposed decomposition-based NS-FFA.

CRediT authorship contribution statement

Cuauhtémoc Tonatiuh Vidrio-Sahagún: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Visualization. **Jianxun He:** Supervision, Writing – review & editing, Resources, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The first author of this paper is funded by a doctoral Scholarship from the National Council for Science and Technology of Mexico (CONACYT) and the Universidad de Guadalajara. This work is also partially funded by the Discovery Grant of Natural Sciences and Engineering Research Council held by the second author.

Appendix A. – Benchmark detrending procedure (Cunderlik and Burn, 2003)

The detrending methodology applied in Cunderlik and Burn (2003) was taken as the benchmark to assess the decomposition procedure proposed in this paper. The benchmark method deals with the weak/second-order nonstationarity, i.e., the temporal variability of the first two statistical moments of the underlying stochastic process while assuming the temporal-invariance of the higher-order moments (e.g., the skewness) (Lindgren et al., 2013; von Storch and Zwiers, 2002). In this procedure, the y_t is considered in the general form of:

$$y_t = \tau_t + \varepsilon_t \quad (\text{A1})$$

where τ_t is the deterministic time-dependent component in the mean and ε_t is the residual time-dependent stochastic component describing the irregular fluctuations around the τ_t . In practice, the τ_t is determined as a linear function of time (i.e., $\tau_t = \tau_0 + \tau_1 t$, where τ_0 and τ_1 are the regression coefficients) using the non-parametric Sen's estimator. Hence, the ε_t , which is the de-trended series in the mean, is given by:

$$\varepsilon'_t = y_t - \tau_t \quad (\text{A2})$$

The ε_t would contain the trend in the variability of y_t . The trend in the variability would then be approximated using the transformed series ε'_t , which captures the change of ε_t over time and is calculated by:

$$\varepsilon'_t = |\varepsilon_t - \bar{\varepsilon}| \quad (\text{A3})$$

where $\bar{\varepsilon}$ is the average of ε_t and is time-invariant. The deterministic trend in the transformed series ε'_t , which is denoted as φ_t , is then estimated. Similar to τ_t , the φ_t is determined as a linear function of time (i.e., $\varphi_t = \varphi_0 + \varphi_1 t$ where φ_0 and φ_1 are the regression coefficients) using the Sen's estimator. The φ_t is removed from ε_t by:

$$y_t^{st} = \begin{cases} \varepsilon_t - \varphi_t & \forall \varepsilon_t \geq \bar{\varepsilon} \\ \varepsilon_t + \varphi_t & \forall \varepsilon_t < \bar{\varepsilon} \\ \varepsilon_t + \varphi_t & \forall \varepsilon_t \geq \bar{\varepsilon} \\ \varepsilon_t - \varphi_t & \forall \varepsilon_t < \bar{\varepsilon} \end{cases} \quad (\text{A4})$$

where $y_t^{st} = \{y_{t_1}^{st}, y_{t_2}^{st}, \dots, y_{t_n}^{st}\}$ is the resulting second-order stationary series. The y_t^{st} derived by the benchmark procedure is free from trends in the first two statistical moments and has been used to select and/or parameterize the distribution (e.g., Cunderlik and Burn, 2003; Gado and Nguyen, 2016; Zhang et al., 2020).

For datasets in C1 (presenting a temporal trend in the mean only), the benchmark and the proposed methods coincide, as the temporal-invariant ε_t reduces Eq. (A.1) to an equivalent form of Eq. (10) in which $G[\cdot]$ is the identity operator. Thus, Eqs. (12)–(14) are also applicable for the benchmark method and thus the standard deviation and the skewness of the detrend time series are preserved in the benchmark method.

However, the benchmark method does not preserve the stochastic component for datasets in C2 and C3. Let λ_t be the intermediate deterministic variable defined as:

$$\lambda_t = \begin{cases} -\varphi_t & \forall \varepsilon_t \geq \bar{\varepsilon} \\ \varphi_t & \forall \varepsilon_t < \bar{\varepsilon} \\ \varphi_t & \forall \varepsilon_t \geq \bar{\varepsilon} \\ -\varphi_t & \forall \varepsilon_t < \bar{\varepsilon} \end{cases} \quad (\text{A5})$$

This allows to express the derived stationary second-order stationary series as:

$$y_t^{st} = y_t - \tau_t + \lambda_t \quad (\text{A6})$$

For datasets in C2 (presenting a temporal trend in the standard deviation only), as $\tau_t = 0 \ \forall t$, Eq. (A.6) reduces to $y_t^{st} = y_t + \lambda_t$. Therefore, by denoting the deterministic true (yet unknown) counterparts of τ_t and λ_t by their uppercase letters (i.e., T_t and Λ_t), it can be shown that the benchmark method simply shifts the Y_t in its first moment (μ_{Y_t}) without actually detrending its second moment ($\sigma_{Y_t}^2$):

$$\mu_{Y_t^{st}} \equiv E[Y_t^{st}] = E[Y_t + \Lambda_t] = E[Y_t] + E[\Lambda_t] = \mu_{Y_t} + \Lambda_t \quad (\text{A7})$$

$$\begin{aligned} \sigma_{Y_t^{st}}^2 &\equiv E[(Y_t^{st} - E[Y_t^{st}])^2] = E[(\{Y_t + \Lambda_t\} - \{E[Y_t + \Lambda_t]\})^2] \\ &= E[(Y_t + \Lambda_t - \mu_{Y_t} - \Lambda_t)^2] = E[(Y_t - \mu_{Y_t})^2] = \sigma_{Y_t}^2 \end{aligned} \quad (\text{A8})$$

$$\begin{aligned} \gamma_{Y_t^{st}} &\equiv \frac{E[(Y_t^{st} - E[Y_t^{st}])^3]}{\left\{E[(Y_t^{st} - E[Y_t^{st}])^2]\right\}^{\frac{3}{2}}} = \frac{E[(\{Y_t + \Lambda_t\} - \{E[Y_t + \Lambda_t]\})^3]}{\left\{E[(\{Y_t + \Lambda_t\} - \{E[Y_t + \Lambda_t]\})^2]\right\}^{\frac{3}{2}}} \\ &= \frac{E[(Y_t + \Lambda_t - \mu_{Y_t} - \Lambda_t)^3]}{\left\{E[(Y_t + \Lambda_t - \mu_{Y_t} - \Lambda_t)^2]\right\}^{\frac{3}{2}}} = \frac{E[(Y_t - \mu_{Y_t})^3]}{\left\{E[(Y_t - \mu_{Y_t})^2]\right\}^{\frac{3}{2}}} = \gamma_{Y_t} \end{aligned} \quad (\text{A9})$$

Thus, the effects of this procedure on the true nonstationary distribution can be thought of as irregular displacements of the marginal distributions of Y_t controlled by Λ_t without changing their width ($\sigma_{Y_t^{st}}^2 = \sigma_{Y_t}^2$) and shape ($\gamma_{Y_t^{st}} = \gamma_{Y_t}$). As a result, the removal of the trend in the variability of y_t is ascribed to such displacements, and thus, this method fails in removing the actual deterministic component behind the nonstationarity and preserving the underlying stochastic component. Consequently, this procedure would introduce errors in the estimates of the statistical moments of the stochastic

component. For instance, the skewness would not be preserved unless the variability of y_t is symmetrical. This condition is intuitive, as λ_t ($\pm\phi_t$) is derived from the ε'_t , which takes the absolute deviations from the time-invariant mean (Eq. (A.3)), and thus ignores the signs of $\varepsilon_t - \bar{e}$ reflecting the average variability only. The even removal of the average deviations (ignoring whether the series of ε_t is skewed towards the high or low values) would alter the proportionality of the variability ranges above and below the mean, and in turn, distort the asymmetry properties of the underlying dataset. For example, if ε_t is skewed towards the high values, the computed λ_t would remove the average deviations around the mean and over-/under-detrend the variability for below/above the mean, respectively.

For datasets in C3 (presenting temporal trends in both the mean and standard deviation), the Eq (A.6) is directly used. Therefore, it can be shown that the benchmark method also shifts the Y_t in its first moment without actually detrending its second moment and thus does not preserve the stochastic component either.

$$\mu_{Y_t^{st}} \equiv E[Y_t^{st}] = E[Y_t - T_t + \Lambda_t] = E[Y_t] - E[T_t] + E[\Lambda_t] = \mu_{Y_t} - T_t + \Lambda_t \quad (10)$$

$$\begin{aligned} \sigma_{Y_t^{st}}^2 &\equiv E[(Y_t^{st} - E[Y_t^{st}])^2] = E[(\{Y_t - T_t + \Lambda_t\} - \{E[Y_t - T_t + \Lambda_t]\})^2] \\ &= E[(Y_t - T_t + \Lambda_t - \mu_{Y_t} + T_t - \Lambda_t)^2] = E[(Y_t - \mu_{Y_t})^2] = \sigma_{Y_t}^2 \end{aligned} \quad (11)$$

$$\begin{aligned} \gamma_{Y_t^{st}} &\equiv \frac{E[(Y_t^{st} - E[Y_t^{st}])^3]}{\{E[(Y_t^{st} - E[Y_t^{st}])^2]\}^{\frac{3}{2}}} = \frac{E[(\{Y_t - T_t + \Lambda_t\} - \{E[Y_t - T_t + \Lambda_t]\})^3]}{\{E[(\{Y_t - T_t + \Lambda_t\} - \{E[Y_t - T_t + \Lambda_t]\})^2]\}^{\frac{3}{2}}} \\ &= \frac{E[(Y_t - T_t + \Lambda_t - \mu_{Y_t} + T_t - \Lambda_t)^3]}{\{E[(Y_t - T_t + \Lambda_t - \mu_{Y_t} + T_t - \Lambda_t)^2]\}^{\frac{3}{2}}} = \frac{E[(Y_t - E[Y_t])^3]}{\{E[(Y_t - E[Y_t])^2]\}^{\frac{3}{2}}} = \gamma_{Y_t} \end{aligned} \quad (12)$$

Similar to C2, the skewness of datasets in C3 would only be preserved in the special case of symmetrical variability.

Appendix B. – Theoretical preservation of statistical moments of the stochastic component derived in the proposed decomposition procedure

For datasets in C2, Eq. (15) ensures that only the standard deviation is perturbed after detrending the standard deviation, while both the conditional mean and the skewness are preserved:

$$\begin{aligned} \mu_{X_t} &\equiv E[X_t] = E[\mu_{Y_t} + (Y_t - \mu_{Y_t}) \cdot h(t)] = E[E[Y_t] + (Y_t - E[Y_t]) \cdot h(t)] \\ &= E[Y_t] + h(t) \cdot E[(Y_t - E[Y_t])] = E[Y_t] = \mu_{Y_t} \end{aligned} \quad (B1)$$

$$\begin{aligned} \sigma_{X_t}^2 &\equiv E[(X_t - \mu_{X_t})^2] = E[(\{\mu_{Y_t} + (Y_t - \mu_{Y_t}) \cdot h(t)\} - \{\mu_{Y_t}\})^2] \\ &= E[(Y_t - \mu_{Y_t}) \cdot h(t)]^2 = h(t)^2 \cdot E[(Y_t - \mu_{Y_t})^2] = h(t)^2 \cdot \sigma_{Y_t}^2 \end{aligned} \quad (B2)$$

$$\begin{aligned} \gamma_{X_t} &\equiv \frac{E[(X_t - \mu_{X_t})^3]}{\{E[(X_t - \mu_{X_t})^2]\}^{\frac{3}{2}}} = \frac{E[(\{\mu_{Y_t} + (Y_t - \mu_{Y_t}) \cdot h(t)\} - \mu_{Y_t})^3]}{\{E[(\{\mu_{Y_t} + (Y_t - \mu_{Y_t}) \cdot h(t)\} - \mu_{Y_t})^2]\}^{\frac{3}{2}}} \\ &= \frac{E[(Y_t - \mu_{Y_t}) \cdot h(t)]^3}{\{E[(Y_t - \mu_{Y_t}) \cdot h(t)]^2\}^{\frac{3}{2}}} = \frac{h(t)^3 \cdot E[(Y_t - \mu_{Y_t})^3]}{h(t)^3 \cdot \{E[(Y_t - \mu_{Y_t})^2]\}^{\frac{3}{2}}} = \frac{E[(Y_t - \mu_{Y_t})^3]}{\{E[(Y_t - \mu_{Y_t})^2]\}^{\frac{3}{2}}} = \gamma_{Y_t} \end{aligned} \quad (B3)$$

For datasets in C3, Eq. (16) ensures that only the mean and standard deviation are perturbed after detrending both the mean and standard deviation, while the conditional skewness is preserved:

$$\begin{aligned} \mu_{X_t} &\equiv E[X_t] = E[\mu_{Y_t-f(t)} + (\{Y_t - f(t)\} - \mu_{Y_t-f(t)}) \cdot h(t)] \\ &= E[E[Y_t - f(t)] + (\{Y_t - f(t)\} - E[Y_t - f(t)]) \cdot h(t)] \\ &= E[Y_t - f(t)] + h(t) \cdot E[(Y_t - f(t) - E[Y_t] + f(t))] \\ &= E[Y_t] - f(t) + h(t) \cdot E[(Y_t - E[Y_t])] = E[Y_t] - f(t) = \mu_{Y_t} - f(t) \end{aligned} \quad (B4)$$

$$\begin{aligned} \sigma_{X_t}^2 &\equiv E[(X_t - \mu_{X_t})^2] \\ &= E[(\{\mu_{Y_t-f(t)} + (\{Y_t - f(t)\} - \mu_{Y_t-f(t)}) \cdot h(t)\} - \{\mu_{Y_t} - f(t)\})^2] \\ &= E[(\{E[Y_t - f(t)] + (Y_t - f(t) - E[Y_t - f(t)]) \cdot h(t)\} - \{E[Y_t] - f(t)\})^2] \\ &= E[(E[Y_t] - f(t) + (Y_t - f(t) - E[Y_t] + f(t)) \cdot h(t) - E[Y_t] + f(t))^2] \\ &= E[((Y_t - E[Y_t]) \cdot h(t))^2] = h(t)^2 \cdot E[(Y_t - E[Y_t])^2] = h(t)^2 \cdot \sigma_{Y_t}^2 \end{aligned} \quad (B5)$$

$$\begin{aligned}
\gamma_{X_t} &\equiv \frac{E[(X_t - \mu_{X_t})^3]}{\left\{E[(X_t - \mu_{X_t})^2]\right\}^{\frac{3}{2}}} = \frac{E\left[\left(\{\mu_{Y_t-f(t)} + (\{Y_t-f(t)\} - \mu_{Y_t-f(t)}) \cdot h(t)\} - \{\mu_{Y_t} - f(t)\}\right)^3\right]}{\left\{E\left[\left(\{\mu_{Y_t-f(t)} + (\{Y_t-f(t)\} - \mu_{Y_t-f(t)}) \cdot h(t)\} - \{\mu_{Y_t} - f(t)\}\right)^2\right]\right\}^{\frac{3}{2}}} \\
&= \frac{E[(\{E[Y_t-f(t)] + (\{Y_t-f(t)\} - E[Y_t-f(t)]) \cdot h(t)\} - \{E[Y_t] - f(t)\})^3]}{\left\{E[(\{E[Y_t-f(t)] + (\{Y_t-f(t)\} - E[Y_t-f(t)]) \cdot h(t)\} - \{E[Y_t] - f(t)\})^2]\right\}^{\frac{3}{2}}} \\
&= \frac{E[(E[Y_t] - f(t) + (Y_t - f(t) - E[Y_t] + f(t)) \cdot h(t) - E[Y_t] + f(t))^3]}{\left\{E[(E[Y_t] - f(t) + (Y_t - f(t) - E[Y_t] + f(t)) \cdot h(t) - E[Y_t] + f(t))^2]\right\}^{\frac{3}{2}}} \\
&= \frac{E[((Y_t - E[Y_t]) \cdot h(t))^3]}{\left\{E[(Y_t - E[Y_t]) \cdot h(t))^2\right\}^{\frac{3}{2}}} = \frac{h(t)^3 \cdot E[(Y_t - \mu_{Y_t})^3]}{\left\{h(t)^2 \cdot E[(Y_t - \mu_{Y_t})^2]\right\}^{\frac{3}{2}}} = \frac{E[(Y_t - \mu_t)^3]}{\left\{E[(Y_t - \mu_{Y_t})^2]\right\}^{\frac{3}{2}}} = \gamma_{Y_t}
\end{aligned} \tag{B6}$$

Appendix C. – Hyperparameter determination of the moving window method

The influence of the window length and step in the moving window method, which was coupled with the Sen's slope estimator for estimating the trend in the σ_{Y_t} , was investigated. Since the window length and step only influence the estimation of the trend in the σ_{Y_t} , their selections affect the analysis of datasets of C2 and C3. As the synthetic nonstationary datasets of C2 present a trend in the σ_{Y_t} only, they were used for the investigation herein. The assessment was conducted based on the mean absolute relative bias (MARB) of the moment estimates. The MARB was estimated from the absolute bias values of each statistical moment of all the synthetic datasets for each combination of window length and step of a given parent distribution (i.e., GEV or PE3). Fig. C1 shows the MARB of the moment estimates when using window lengths ranging from 10 to 50 years and window steps from 1 year to the window length for the nonstationary GEV and PE3 datasets, respectively. As shown in this figure, the MARBs of μ_{X_t} , σ_{X_t} , and γ_{X_t} are in general higher when using the larger window steps for any window length. It is also worth mentioning that the MARBs slightly increase with the decrease of the window step when using small window steps, especially for σ_{X_t} and γ_{X_t} . These results are not surprising, as for a given window length, the large window step leads to the small sample size of the derived σ_{Y_t} series, and thus deteriorates the estimation of the trend. Whereas the small window step leads to the large window overlapping, and thus the datapoints located in the center of the time series would be counted more times than those located at the bounds and have large influence on the estimates.

Furthermore, an assessment incorporating the biases of all the three estimated statistical moments (μ_{X_t} , σ_{X_t} , and γ_{X_t}) was conducted for each parent distribution. To this end, the MARBs of μ_{X_t} , σ_{X_t} , and γ_{X_t} were averaged, and then used to identify the optimal hyperparameters. Fig. C2 presents the average of the MARBs of μ_{X_t} , σ_{X_t} , and γ_{X_t} , and highlights the minimum average of the MARBs corresponding to the optimal window step for each window length. These results indicate that the lowest average of the MARBs is yielded when using the window length of 10 years and the window step of 5 years. The selected hyperparameters are rational, as they yield a long series of σ_{Y_t} for estimating the trend and also avoid substantial bias of the datapoints located in the middle of the time series. Besides, the window length would also skip potential decadal variability. Therefore, these hyperparameters of the moving window method were adopted throughout this paper.

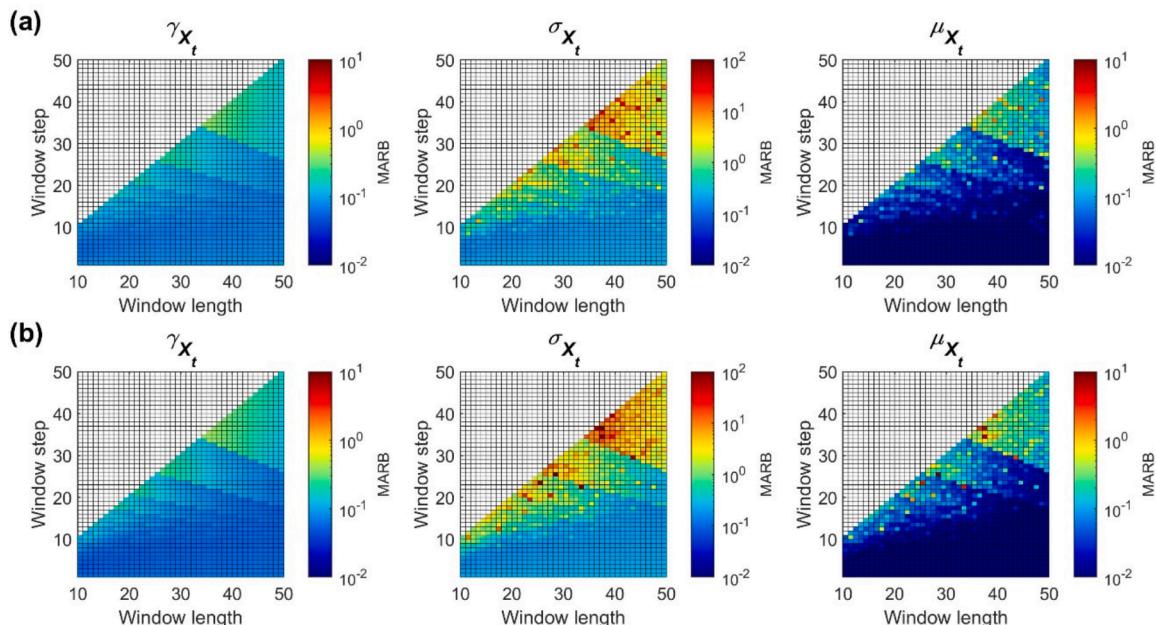


Fig. C1. Mean absolute relative bias (MARB) of the estimated statistical moments (μ_{X_t} , σ_{X_t} , and γ_{X_t}) when using different window lengths and steps for estimating the trend in the σ_{Y_t} for the (a) nonstationary GEV and (b) nonstationary PE3 datasets in C2.

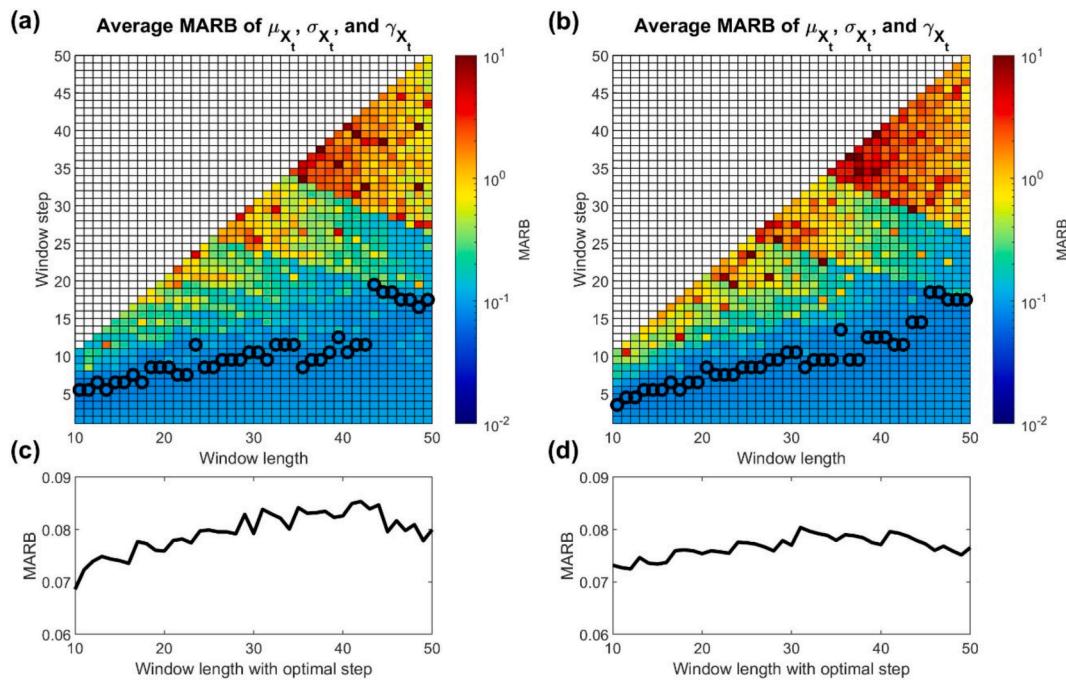


Fig. C2. Average of the mean absolute relative biases (MARB) of the μ_{X_t} , σ_{X_t} , and γ_{X_t} when using different window lengths and steps for estimating the trend in the σ_{Y_t} for the (a) nonstationary GEV and (b) nonstationary PE3 datasets, and the minimum average of the MARBs of the μ_{X_t} , σ_{X_t} , and γ_{X_t} of each window length for the (c) nonstationary GEV and (d) nonstationary PE3 datasets in C2. The black circles in (a) and (b) indicate the identified optimal window step (yielding the minimum average of the MARB) for each window length.

Appendix D. – Results of the performance-based approach when using the BIC as the selection criterion

The potentially different results of the performance-based approach due to using another evaluation metric of fitting efficiency were explored. To this end, the *BIC* was employed in both the simulation study and real applications and was calculated by:

$$BIC = n \log(RMSE) + N_\theta \cdot \log(n) \quad (D1)$$

In the simulation study, similar results were obtained when using the *BIC* and *AIC* as the selection metrics in the performance-based approach. As shown in Figs. D1 and 3, the use of *BIC* or *AIC* yielded practically equivalent results in both empirical and theoretical assessments. When using the *BIC*, the performance-based approach selected the true distribution and nonstationary structure on average in 41.6% (within the range of 32.8% to 48.4%) and 17.8% (within the range of 10.0% to 23.8%) of the datasets, respectively, resulting in the selection of true models on average in only 7.1% (within the range of 3.7% to 9.0%) of the datasets (Fig. D2). Compared to the results when using the *AIC* as the selection metric (Fig. 4), the changes in the results of the performance-based approach were practically negligible. Thus, the results from the performance-based approach when using the *BIC* as the selection metric also support that the performance-based approach tended to select untrue models and overfit the samples compared to the decomposition-based approach.

Fig. D3 shows the NS-FFA models determined using the *BIC* as the selection metric in the performance-based approach in real applications. As shown in this figure, the results based on the *BIC* were in general consistent with those based on the *AIC* (Fig. 6). Comparing the results based on the

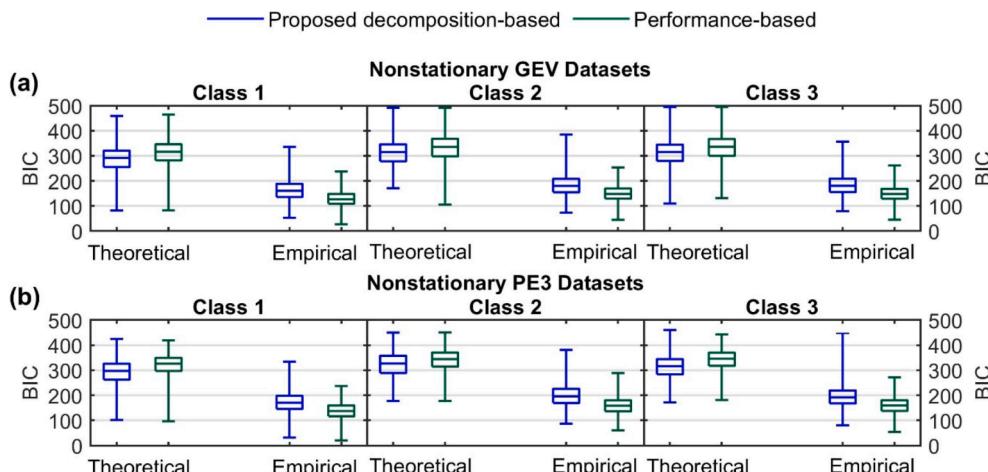


Fig. D1. The performance of the proposed decomposition-based and the performance-based approaches for the synthetic nonstationary (a) GEV and (b) PE3 datasets in terms of fitting efficiency in the theoretical and empirical assessments. The performance-based approach adopted the *BIC* as the selection criterion herein.

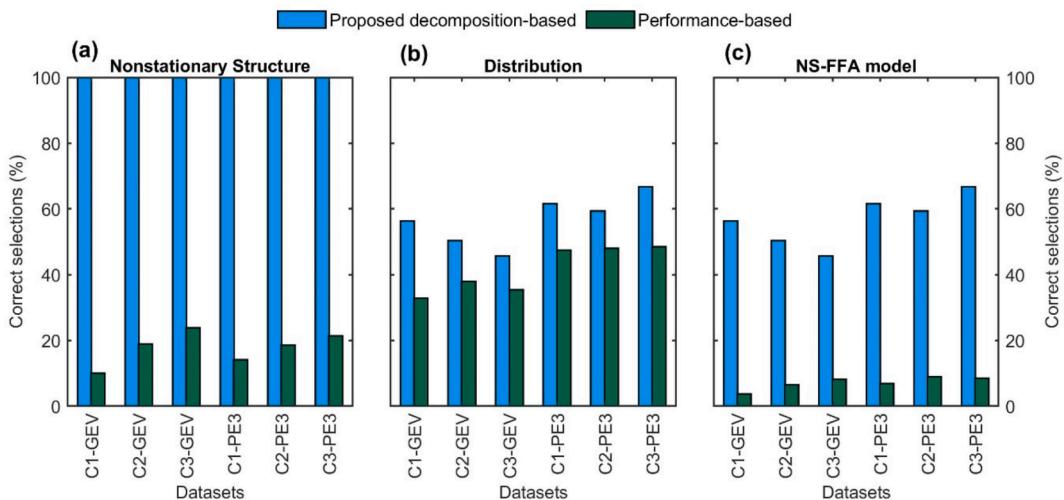


Fig. D2. Comparison of correct selections of (a) nonstationary structures, (b) distributions, and (c) NS-FFA models (i.e., both the nonstationary structures and distributions) between the proposed decomposition-based and the performance-based approaches. The datasets are abbreviated as CX-D, where the CX and D denote the nonstationary class (C1, C2, or C3) and the distribution (GEV or PE3) used for generating the synthetic datasets, respectively. The performance-based approach adopted the *BIC* as the selection criterion herein.

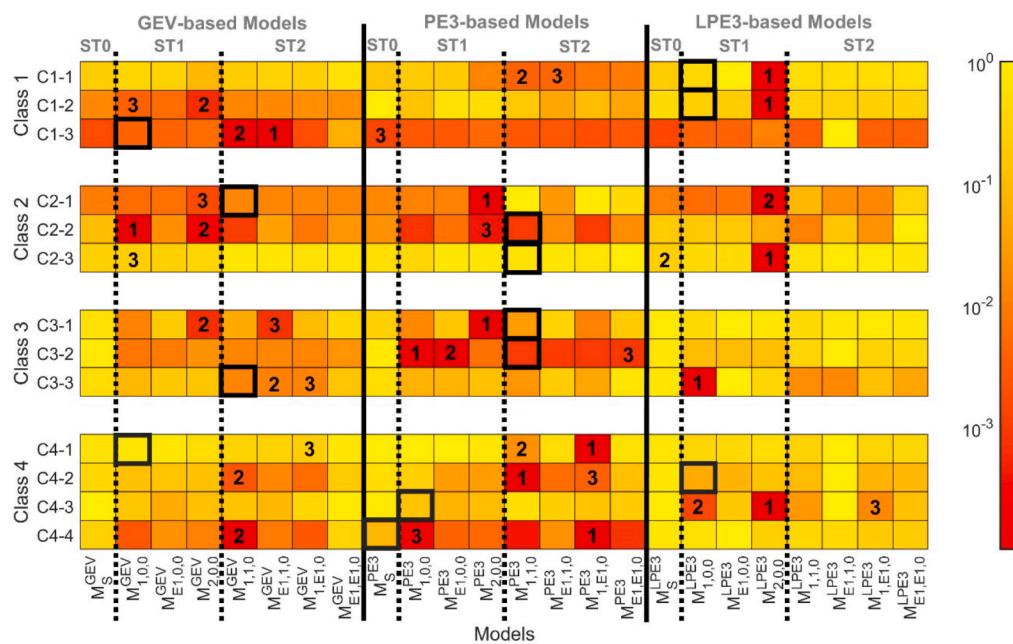


Fig. D3. The normalized *BICs* (shown by color at a log-scale) of the stationary and all nonstationary candidate models. The numerical indexes (1, 2, and 3) denote the top three models selected according to the *BIC* in the performance-based NS-FFA, while the solid black boxes highlight the models yielded by the decomposition-based NS-FFA. The vertical solid and dashed lines separate the models adopting the same distribution and time-variant distribution parameter(s). ST0 denotes the stationary structure; ST1 and ST2 are the nonstationary structures with the time-variant location parameter, and the time-variant location and scale parameters, respectively.

AIC and *BIC*, the use of different evaluation metrics led to selecting a different optimal model in only one dataset (C3-1); while the differences in at least one of the top three models were observed from five out of thirteen datasets (C1-3, C2-3, C3-1, C3-2, and C4-4). In addition, the relative discrepancies in the *BIC* of the model determined by the proposed decomposition-based NS-FFA with respect to that by the performance-based NS-FFA (Fig. D4) were similar to previous results (Fig. 8). Thus, the changes in the performance-based approach were minor in general and the results based on the *BIC* also supported the conclusions drawn based on the results obtained using the *AIC*.

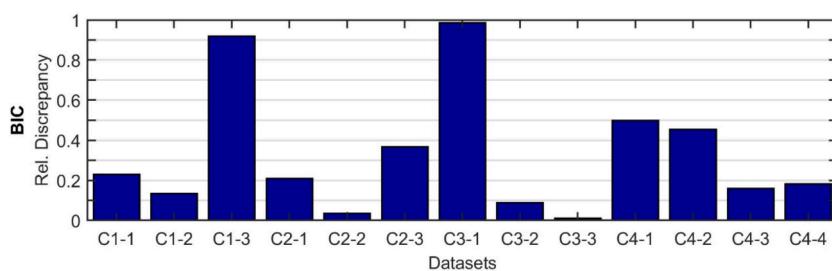


Fig. D4. Relative discrepancies in the fitting efficiency (*BIC*) of the model selected by the decomposition-based NS-FFA with respect to that by the performance-based NS-FFA. The performance-based approach adopted the *BIC* as the selection criterion herein.

Both the *AIC* and *BIC* are used to measure the fitting efficiency but employ a different weighting of the goodness-of-fit and the complexity of the model/length of data. Despite the similar results in general, the observed differences between the results based on the *AIC* and *BIC* demonstrated that the use of different selection metrics might lead to different results in the performance-based approach. These results further argue the issue on the selection of the evaluation criteria in the performance-based approach.

References

- AghaKouchak, A., Chiang, F., Huning, L.S., Love, C.A., Mallakpour, I., Mazdiyasni, O., Moftakhar, H., Papalexiou, S.M., Ragno, E., Sadegh, M., 2020. Climate Extremes and Compound Hazards in a Warming World. *Annual Review of Earth and Planetary Sciences* 48 (1), 519–548.
- Agilan, V., Umamahesh, N.V., 2017. What are the best covariates for developing non-stationary rainfall Intensity-Duration-Frequency relationship? *Advances in Water Resources* 101, 11–22. <https://doi.org/10.1016/j.advwatres.2016.12.016>.
- Ammar, M.E., Gharib, A., Islam, Z., Davies, E.G.R., Seneka, M., Faramarzi, M., 2020. Future floods using hydroclimatic simulations and peaks over threshold: An alternative to nonstationary analysis inferred from trend tests. *Advances in Water Resources* 136, 103463.
- Archfield, S.A., Hirsch, R.M., Viglione, A., Blöschl, G., 2016. Fragmented patterns of flood change across the United States. *Geophysical Research Letters* 43 (19), 10232–10239. <https://doi.org/10.1002/2016GL070590>.
- Beven, K., 2006. A manifesto for the equifinality thesis. *Journal of Hydrology* 320 (1–2), 18–36. <https://doi.org/10.1016/j.jhydrol.2005.07.007>.
- Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* 249 (1–4), 11–29. [https://doi.org/10.1016/S0022-1694\(01\)00421-8](https://doi.org/10.1016/S0022-1694(01)00421-8).
- Blöschl, G., Hall, J., Viglione, A., Perdigão, R.A.P., Parajka, J., Merz, B., Lun, D., Arheimer, B., Aronica, G.T., Bilibashi, A., Boháć, M., Bonacci, O., Borga, M., Čanjevac, I., Castellarin, A., Chirico, G.B., Claps, P., Frolova, N., Ganora, D., Gorbachova, L., Güll, A., Hannaford, J., Harrigan, S., Kireeva, M., Kiss, A., Kjeldsen, T.R., Kohnová, S., Koskela, J.J., Ledvinka, O., Macdonald, N., Mavrova-Guirguanova, M., Mediero, L., Merz, R., Molnar, P., Montanari, A., Murphy, C., Osuch, M., Ovcharuk, V., Radevski, I., Salinas, J.L., Sauquet, E., Šraj, M., Szolgay, J., Volpi, E., Wilson, D., Zaimi, K., Živković, N., 2019. Changing climate both increases and decreases European river floods. *Nature* 573 (7772), 108–111.
- Burn, D.H., Whitfield, P.H., 2017. Changes in cold region flood regimes inferred from long-record reference gauging stations. *Water Resources Research* 53 (4), 2643–2658. <https://doi.org/10.1002/2016WR020108>.
- Cheng, L., AghaKouchak, A., Gilleland, E., Katz, R.W., 2014. Non-stationary extreme value analysis in a changing climate. *Climatic Change* 127 (2), 353–369. <https://doi.org/10.1007/s10584-014-1254-5>.
- Coles, S., 2001. *An introduction to statistical modeling of extreme values*. Springer-Verlag, London, London.
- Cunderlik, J.M., Burn, D.H., 2003. Non-stationary pooled flood frequency analysis. *Journal of Hydrology* 276 (1–4), 210–223. [https://doi.org/10.1016/S0022-1694\(03\)00062-3](https://doi.org/10.1016/S0022-1694(03)00062-3).
- Doucet, A., de Freitas, N., and Gordon, N. (2001). *An Introduction to Sequential Monte Carlo Methods*. New York, USA.: Springer. https://doi.org/10.1007/978-1-4757-3437-9_1.
- England Jr., J. F., Cohn, T. A., Faber, B. A., Stedinger, J. R., Thomas Jr., W. O., Veilleux, A. G., ... Mason Robert, R. J. (2019). *Guidelines for determining flood flow frequency—Bulletin 17C* (Version 1.). *Techniques and Methods*. Reston, VA. <https://doi.org/10.3133/tm4B5>.
- François, B., Schlef, K.E., Wi, S., Brown, C.M., 2019. Design considerations for riverine floods in a changing climate – A review. *Journal of Hydrology* 574, 557–573.
- Gado, T.A., Van Nguyen, V.T., 2016a. An at-site flood estimation method in the context of nonstationarity I. A simulation study. *Journal of Hydrology* 535, 710–721. <https://doi.org/10.1016/j.jhydrol.2015.12.063>.
- Gado, T.A., Van Nguyen, V.T., 2016b. An at-site flood estimation method in the context of nonstationarity II. Statistical analysis of floods in Quebec. *Journal of Hydrology* 535, 722–736. <https://doi.org/10.1016/j.jhydrol.2015.12.064>.
- Gilroy, K.L., McCuen, R.H., 2012. A nonstationary flood frequency analysis method to adjust for future climate change and urbanization. *Journal of Hydrology* 414–415, 40–48. <https://doi.org/10.1016/j.jhydrol.2011.10.009>.
- Giuntoli, I., Prosdocimi, I., Hannah, D.M., 2021. Going Beyond the Ensemble Mean: Assessment of Future Floods From Global Multi-Models. *Water Resources Research* 57 (3). <https://doi.org/10.1029/2020WR027897>.
- Gordon, N.J., Salmon, D.J., Smith, A.F.M., 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEEE Proceedings F Radar and Signal Processing* 140 (2), 107. <https://doi.org/10.1049/ip-f.2.1993.0015>.
- Griffis, V.W., Stedinger, J.R., 2009. Log-Pearson Type 3 Distribution and Its Application in Flood Frequency Analysis. III: Sample Skew and Weighted Skew Estimators. *Journal of Hydrologic Engineering* 14 (2), 121–130. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2009\)14:2\(121\)](https://doi.org/10.1061/(ASCE)1084-0699(2009)14:2(121)).
- Hosking, J.R.M., 1990. L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. Retrieved from Journal of the Royal Statistical Society 52 (1), 105–124. <http://www.jstor.org/stable/2345653>.
- Hosking, J.R.M., Wallis, J.R., 1997. *Regional frequency analysis: an approach based on L-Moments*. Cambridge University Press, New York, USA.
- Kalai, C., Mondal, A., Griffin, A., Stewart, E., 2020. Comparison of Nonstationary Regional Flood Frequency Analysis Techniques Based on the Index-Flood Approach. *Journal of Hydrologic Engineering* 25 (7), 1–7. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001939](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001939).
- Kasiviswanathan, K.S., He, J., Tay, J.-H., Sudheer, K.P., 2019. Enhancement of Model Reliability by Integrating Prediction Interval Optimization into Hydrogeological Modeling. *Water Resources Management* 33 (1), 229–243.
- Katz, R.W., 2013. Statistical methods for nonstationary extremes. In: AghaKouchak, A., Easterling, D., Hsu, K., Schubert, S., Sorooshian, S. (Eds.), *Extremes in a changing climate: detection, analysis and uncertainty*. Springer, Netherlands, pp. 15–37.
- Khatami, S., Peel, M.C., Peterson, T.J., Western, A.W., 2019. Equifinality and Flux Mapping: A New Approach to Model Evaluation and Process Representation Under Uncertainty. *Water Resources Research* 55 (11), 8922–8941. <https://doi.org/10.1029/2018WR023750>.
- Kim, H., Kim, S., Shin, H., Heo, J.H., 2017. Appropriate model selection methods for nonstationary generalized extreme value models. *Journal of Hydrology* 547, 557–574. <https://doi.org/10.1016/j.jhydrol.2017.02.005>.
- Koutsoyiannis, D., Montanari, A., 2015. Negligent killing of scientific concepts: the stationarity case. *Hydrological Sciences Journal* 60 (7–8), 1174–1183. <https://doi.org/10.1080/02626667.2014.950959>.
- Kundzewicz, Z.W., Krysanova, V., Dankers, R., Hirabayashi, Y., Kanae, S., Hattermann, F. F., Huang, S., Milly, P.C.D., Stoffel, M., Driessens, P.P.J., Matczak, P., Quevauviller, P., Schellnhuber, H.-J., 2017. Differences in flood hazard projections in Europe—their causes and consequences for decision making. *Hydrological Sciences Journal*. <https://doi.org/10.1080/02626667.2016.1241398>.
- Li, M., Zhang, T., Feng, P., 2019. A nonstationary runoff frequency analysis for future climate change and its uncertainties. *Hydrological Processes* 33 (21), 2759–2771. <https://doi.org/10.1002/hyp.13526>.
- Lindgren, G., Rootzén, H., Sandsten, M., 2013. *Stationary stochastic processes for scientists and engineers*. Chapman and Hall/CRC.
- López, J., Francés, F., 2013. Non-stationary flood frequency analysis in continental Spanish rivers, using climate and reservoir indices as external covariates. *Hydrology and Earth System Sciences* 17 (8), 3189–3203. <https://doi.org/10.5194/hess-17-3189-2013>.
- Luke, A., Vrugt, J.A., AghaKouchak, A., Matthew, R., Sanders, B.F., 2017. Predicting nonstationary flood frequencies: Evidence supports an updated stationarity thesis in the <scop>U</scop>nited <scop>S</scop>tates. *Water Resources Research* 53 (7), 5469–5494. <https://doi.org/10.1002/2016WR019676>.
- Maraun, D., Shepherd, T.G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J.M., Hagemann, S., Richter, I., Soares, P.M.M., Hall, A., Mearns, L.O., 2017. Towards process-informed bias correction of climate change simulations. *Nature Climate Change* 7 (11), 764–773.
- Maraun, D., Wetterhall, F., Brienen, S., Rust, H.W., Sauter, T., Themeßl, M., Thiele-Eich, I., 2010. Precipitation Downscaling Under Climate Change: Recent Developments To Bridge the Gap Between Dynamical Models and the End User. *Reviews of Geophysics* 48 (2009), 1–34. [https://doi.org/10.1029/2009RG000314.1. INTRODUCTION](https://doi.org/10.1029/2009RG000314.1).
- Milly, A.P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Zbigniew, W., Lettenmaier, D.P., Milly, P.C.D., 2008. Stationarity Is Dead: Whither Water Management? *Science* 319 (5863), 573–574. <https://doi.org/10.1126/science.1151915>.
- Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., Dettinger, M.D., Krysanova, V., 2015. On Critiques of “stationarity is Dead: Whither Water Management?”. *Water Resources Research* 51 (9), 7785–7789.
- Mondal, A., Majumdar, P.P., 2015. Modeling non-stationarity in intensity, duration and frequency of extreme rainfall over India. *Journal of Hydrology* 521, 217–231. <https://doi.org/10.1016/j.jhydrol.2014.11.071>.
- Montanari, A., Koutsoyiannis, D., 2014. Modeling and mitigating natural hazards: Stationarity is immortal! *Water Resources Research* 50 (12), 9748–9756. <https://doi.org/10.1002/2014WR016092>.
- Nguyen, T.H., El Outayek, S., Lim, S.H., Van Nguyen, V.T., 2017. A systematic approach to selecting the best probability models for annual maximum rainfalls – A case study using data in Ontario (Canada). *Journal of Hydrology* 553, 49–58. <https://doi.org/10.1016/j.jhydrol.2017.07.052>.
- Obeysekera, J., Salas, J.D., 2014. Quantifying the Uncertainty of Design Floods under Nonstationary Conditions. *Journal of Hydrologic Engineering* 19 (7), 1438–1446. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000931](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000931).
- Obeysekera, J., Salas, J.D., 2016. Frequency of Recurrent Extremes under Nonstationarity. *Journal of Hydrologic Engineering* 21 (5), 04016005. [https://doi.org/10.1061/\(asce\)he.1943-5584.0001339](https://doi.org/10.1061/(asce)he.1943-5584.0001339).
- Ouarda, T.B.M.J., Charron, C., 2019. Changes in the distribution of hydro-climatic extremes in a non-stationary framework. *Scientific Reports* 9 (1), 1–8. <https://doi.org/10.1038/s41598-019-44603-7>.
- Ouarda, T.B.M.J., Charron, C., St-Hilaire, A., 2019. Uncertainty of stationary and nonstationary models for rainfall frequency analysis. *International Journal of Climatology* September, 1–20. <https://doi.org/10.1002/joc.6339>.
- Ouarda, T.B.M.J., Yousef, L.A., Charron, C., 2018. Non-stationary intensity-duration-frequency curves integrating information concerning teleconnections and climate change. *Int J Climatol* 39 (4), 2306–2323.

- Papalexiou, S.M., Koutsoyiannis, D., 2013. Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research* 49 (1), 187–201. <https://doi.org/10.1029/2012WR012557>.
- Prosdocimi, I., Kjeldsen, T., 2021. Parametrisation of change-permitting extreme value models and its impact on the description of change. *Stochastic Environmental Research and Risk Assessment* 35 (2), 307–324. <https://doi.org/10.1007/s00477-020-01940-8>.
- Ragno, E., AghaKouchak, A., Cheng, L., Sadegh, M., 2019. A generalized framework for process-informed nonstationary extreme value analysis. *Advances in Water Resources* 130, 270–282.
- Ray, L.K., Goel, N.K., 2019. Flood Frequency Analysis of Narmada River Basin in India under Nonstationary Condition. *Journal of Hydrologic Engineering* 24 (8), 1–15. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001808](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001808).
- Salas, J. D., and Obeysekera, J. (2014). Revisiting the Concepts of Return Period and Risk for Nonstationary Hydrologic Extreme Events. *Journal of Hydrologic Engineering*, 19 (3), 554–568. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000820](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000820).
- Salas, J.D., Obeysekera, J., Vogel, R.M., 2018. Techniques for assessing water infrastructure for nonstationary extreme events: a review. *Hydrological Sciences Journal* 63 (3), 325–352. <https://doi.org/10.1080/02626667.2018.1426858>.
- Särkkä, S. (Ed.), 2013. *Bayesian Filtering and Smoothing*. Cambridge University Press.
- Sen, S., He, J., and Kasiviswanathan, K. S. (2020). Uncertainty quantification using the particle filter for non-stationary hydrological frequency analysis. *Journal of Hydrology*, 584. <https://doi.org/10.1016/j.jhydrol.2020.124666>.
- Serago, J.M., Vogel, R.M., 2018. Parsimonious nonstationary flood frequency analysis. *Advances in Water Resources* 112, 1–16.
- Serinaldi, F., Kilsby, C.G., 2015. Stationarity is undead: Uncertainty dominates the distribution of extremes. *Advances in Water Resources* 77, 17–36. <https://doi.org/10.1016/j.advwatres.2014.12.013>.
- Serinaldi, F., Kilsby, C.G., Lombardo, F., 2018. Untenable nonstationarity: An assessment of the fitness for purpose of trend tests in hydrology. *Advances in Water Resources* 111, 132–155.
- Slater, L., Villarini, G., Archfield, S., Faulkner, D., Lamb, R., Khouakhi, A., Yin, J., 2021. Global Changes in 20-Year, 50-Year, and 100-Year River Floods. *Geophysical Research Letters* 48 (6), 1–10. <https://doi.org/10.1029/2020GL091824>.
- Stedinger, J.R., 2017. Chapter 76: Flood Frequency Analysis. In: Singh, V.P. (Ed.), *Handbook of Applied Hydrology*, (2nd ed.). McGraw-Hill Professional.
- Stedinger, J.R., Griffis, V.W., 2011. Getting from here to where? Flood frequency analysis and climate. *Journal of the American Water Resources Association* 47 (3), 506–513. <https://doi.org/10.1111/j.1752-1688.2011.00545.x>.
- Stedinger, J.R., Vogel, R.M., Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events. In: Maidmen, D.R. (Ed.), *Handbook of hydrology*. McGraw-Hill, New York.
- Strupczewski, W.G., Singh, V.P., Mitosek, H.T., 2001. Non-stationary approach to at-site flood frequency modelling. III. Flood analysis of Polish rivers. *Journal of Hydrology* 248 (1–4), 152–167. [https://doi.org/10.1016/S0022-1694\(01\)00399-7](https://doi.org/10.1016/S0022-1694(01)00399-7).
- Sun, P., Wen, Q., Zhang, Q., Singh, V.P., Sun, Y., Li, J., 2018. Nonstationarity-based evaluation of flood frequency and flood risk in the Huai River basin. *China. Journal of Hydrology* 567 (March), 393–404. <https://doi.org/10.1016/j.jhydrol.2018.10.031>.
- Sun, X., Lall, U., Merz, B., Dung, N.V., 2015. Hierarchical Bayesian clustering for nonstationary flood frequency analysis: Application to trends of annual maximum flow in Germany. *Water Resources Research* 51 (8), 6586–6601. <https://doi.org/10.1002/2015WR017117>.
- Sung, J.H., Kim, Y.O., Jeon, J.J., 2018. Application of distribution-free nonstationary regional frequency analysis based on L-moments. *Theoretical and Applied Climatology* 133 (3–4), 1219–1233. <https://doi.org/10.1007/s00704-017-2249-8>.
- Vidrio-Sahagún, C.T., He, J., Kasiviswanathan, K.S., Sen, S., 2021. Stationary hydrological frequency analysis coupled with uncertainty assessment under nonstationary scenarios. *Journal of Hydrology* 598 (July), 125725. <https://doi.org/10.1016/j.jhydrol.2020.125725>.
- Villarini, G., Smith, J.A., 2010. Flood peak distributions for the eastern United States. *Water Resources Research* 46 (6), 1–17. <https://doi.org/10.1029/2009WR008395>.
- Villarini, G., Smith, J.A., Serinaldi, F., Bales, J., Bates, P.D., Krajewski, W.F., 2009. Flood frequency analysis for nonstationary annual peak records in an urban drainage basin. *Advances in Water Resources* 32 (8), 1255–1266. <https://doi.org/10.1016/j.advwatres.2009.05.003>.
- von Storch, H., Zwiers, F.W., 2002. *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge. Retrieved from <http://ezproxy.lib.ucalgary.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=77515&site=ehost-live>.
- Zhang, Z., Stadnyk, T.A., Burn, D.H., 2020. Identification of a preferred statistical distribution for at-site flood frequency analysis in Canada. *Canadian Water Resources Journal/Revue Canadienne Des Ressources Hydrauliques* 45 (1), 43–58. <https://doi.org/10.1080/07011784.2019.1691942>.