Change point detection of flood events using a functional data framework

Mohamed Ali Ben Alaya, Camille Ternyck, Sophie Dabo-Niang, Fateh Chebana, Taha B. M. J. Ouarda

2020

The final publication is available at:

Ben Alaya, M. A., Ternynck, C., Dabo-Niang, S., Chebana, F., & Ouarda, T. B. M. J. (2020). Change point detection of flood events using a functional data framework. *Advances in Water Resources*, *137*, 103522.
https://doi.org/10.1016/j.advwatres.2020.103522

Change point detection of flood events using a functional data framework

Mohamed Ali Ben Alaya[1], Camille Ternynck[2], Sophie Dabo-Niang[3], Fateh Chebana[4], Taha B.M.J. Ouarda[4]

[1]Pacific Climate Impacts Consortium, University of Victoria, PO Box 1700 Stn CSC, Victoria, BC V8W2Y2, Canada

[2]Univ. Lille, CHU Lille, EA 2694 - Santé publique : épidémiologie et qualité des soins, F-59000 Lille, France

[3]Laboratoire LEM CNRS 9221, Université de Lille, INRIA-MODAL, Villeneuve d'Ascq Cedex, France

[4] INRS-ETE, 490 rue de la Couronne, Québec (QC), Canada G1K 9A9

[*]**Corresponding author:** Tel: +1 (250) 472 4682

Email: mohamedalibenalaya@uvic.ca

2020-01-09

Abstract

Change point detection methods have an important role in many hydrological and hydraulic studies of river basins. These methods are very useful to characterize changes in hydrological regimes and can, therefore, lead to better understanding changes in extreme flows behavior. Flood events are generally characterized by a finite number of characteristics that may not include the entire information available in a discharge time series. The aim of the current work is to present a new approach to detect changes in flood events based on a functional data analysis framework. The use of the functional approach allows taking into account the whole information contained in the discharge time series of flood events. The presented methodology is illustrated on a flood analysis case study, from the province of Quebec, Canada. Obtained results using the proposed approach are consistent with those obtained using a traditional change point method, and demonstrate the capability of the functional framework to simultaneously consider several flood features and, therefore, presenting a comprehensive way for a better exploitation of the information contained in a discharge time series.

Keywords: Functional data analysis, Change point detection, Hydrology, Flood.

Introduction

Detection of changes in hydrological data is of interest to better understand hydrological regimes, and separate events. Changes in a series can occur in numerous ways, gradually or abruptly, and can affect the mean, median, variance, autocorrelation, or any other aspect of the data. In the future, regions that are relatively sheltered from wind storms, heat waves, droughts and floods, may no longer be in a warmer climate (Goudie 2006). Detection of changes in long time series of hydrological data is an important and difficult issue, of increasing interest. Change point detection in hydrology are essential to characterize the impacts of the climate disturbances on hydrological regimes (Kingston et al. 2011). It is then very important, particularly where we observe changes in the frequency and/or in the intensity of various forms of extreme weather events. Detection of eventual changes in collected data of hydrologic time series sets is thus obviously an important step before performing any descriptive or predictive analysis.

Literature abounds with studies on change point testing in scalar or vector time series. For example, Kundzewicz and Robson (2004) gave a general guidance on the methodology for change detection in hydrological records. Wong et al. (2006) proposed a relational method for discrete data. Change point analysis is addressed in both classical and Bayesian statistics. Methods in classical statistics usually consist of performing several kinds of tests to either confirm or reject the hypothesis of change. Most of them address slope or intercept change in linear regression models (Solow 1987, Easterling and Peterson 1995, Vincent 1998). Bayesian statistics methods are performed to obtain a statistical distribution for the change point and eventually a distribution for the other model parameters. The inference on parameters was performed using Monte-Carlo

Markov Chain algorithms (MCMC). Seidou and Ouarda (2007) proposed a Bayesian method of multiple change point detection in multiple linear regression. This method is numerically efficient and does not involve the time-consuming Monte-Carlo Markov Chain simulations as opposed to other Bayesian change point methods. The procedure was initially designed to detect a change in the relationship between a set of explanatory variables and the dependent variables. Using the time variable as an explanatory variable, this approach can detect the change point in a given time series.

The flood event is an integration of spatial and temporal variations in water input, storage and transfer processes within a catchment (Hannah et al. 2000). Particularly, discharge (rate of flow) time series is the main source of information for studying flood events. Arguably, the hydrograph of a flood event as a graph showing the discharge versus time has been the cornerstone of statistical hydrology, as it is directly related to the design of hydraulic infrastructures. In spite of considerable progress in the development of new statistical tools for change point analysis, researchers' previous efforts have been mainly focused on a single or few characteristics of the flood hydrograph ignoring the continuous behaviour of the flood event in time. Classical change point detection approaches involve a substantial simplification of the overall extreme hydrological event, through focusing on a single or few characteristics of the flood event such as the peak or the volume, and, therefore, fail to account for the whole information stored in flood hydrographs presented as continuous curves. Despite the extensive literature on change point methods, little recognition appears to have been given to a more general approach considering the entire information contained in the discharge time series. The overall objective of this paper is to present a new approach that attempts to handle this concern by considering the discharge time series of the flood event as a continuous curve using a functional data analysis (FDA) framework.

The first application of FDA to the hydrological context refers to Chebana et al. (2012) introducing an exploratory analysis and outlier detection of hydrographs. Chebana et al. (2012) showed that FDA is more general, flexible and representative of the real hydrological phenomena. For classification of flood events, Ternynck et al. (2016) showed that obtained classes using functional approaches are more representative than those obtained using a traditional multivariate hierarchical classification method. Masselot et al. (2016) adapted a functional regression model for streamflow forecasting. Suhaila and Yusop (2017) employed the functional framework to study the spatial and temporal variability of precipitation in Peninsular Malaysia. More recently, Requena et al. (2018) proposed a functional multiple regression for flow duration curves estimation while Larabi et al. (2018) developed a stepwise multicriteria for rainfall-runoff model calibration defined on the basis of FDA.

A growing research area is being advanced focusing on the development of new statistical tools to analyze functional data. For instance, many existing tools in the univariate and multivariate statistical literature have been adapted to the functional context (Dabo-Niang et al. 2010, Fischer 2010, Chebana et al. 2012). Some authors investigated the change point detection method in the FDA context for testing the assumption of a common functional mean for independent functional data (Aue et al. 2009, Berkes et al. 2009). Thereafter, Zhang et al. (2011) adapted this work to the case of functional dependent data.

The aim of the present paper is to introduce and adapt the FDA framework to change point detection of flood events. The present paper is structured as follows: a brief presentation of the data set and the study area is provided in section 1, the proposed functional change point detection approach is presented in section 2. Results of the application of the proposed method to the case of flood events in two stations from the province of Quebec, Canada, are illustrated in section 3. Discussion and conclusion of the main findings are given in sections 4 and 5, respectively.

1. Data Description

Daily flow data recorded at two hydrological stations in the province of Quebec, the Romaine River and the Moisie River stations, are considered (Figure 1). The available data series for the Romaine river station covers the period from 1961 to 2000 recorded over a drainage area of $13000\ Km^2$. For the Moisie river station, with a drainage area of $19000\ Km^2$, daily flow records between 1968 and 1991 are used. Given the nature of most of the flooding events that characterize the area, mainly caused by snow melting in spring and summer, only flood events occurring between March 1st and August 31st are considered in the current analysis.

The selection of these two stations is mainly based on previous finding about the inhomogeneity of their flood regimes (Ternynck et al. 2016). Furthermore, previous results on flood event behaviour for both Romaine river and Moisie river stations demonstrate an apparent change in annual maxima discharges time series (Seidou and Ouarda 2007). Thus, it is expected that these two case studies may represent comprehensive examples to test and validate the proposed approach.

While the proposed approach is general and can be applied to entire annual discharge series, a prior knowledge about the season on which major flood events occur can be helpful to primarily focus on possible changes in the flood event of interest. This allows avoiding misleading conclusion in change point results that are due to changes affecting streamflow not related to the major flood event. Although climate change might shift the timing of flood events (Blöschl et al. 2017), this should not be a concern since our choice of the spring-summer period is long enough to account for this fact.

2. Functional change point detection method

Consider n years of daily flow series recorded from March 1st to August 31st at a given station corresponding to flood events occurring in the spring-summer period. Let $x_i = \left(x_i(t_1),\ldots,x_i(t_j),\ldots,x_i(t_T)\right),\ i=1,\ldots,n$ be the set of n discrete observations where each $t_j \in \mathbf{T} \subset \Re^+$ and $j=1,\ldots,T$ is the j^{th} record time point corresponding to the day j from time subset corresponding to the $\mathbf{T}$ from March 1st to August 31st which include the set $\{1,\ldots,T\}$. For instance, discrete observations x_i are daily flow within a given i^{th} year for the spring-summer period with $T=181$. For a given year i, each set of measurements $\left(x_i(t_1),\ldots,x_i(t_T)\right)$ will be converted to a functional data denoted $\left\{X_i(t), t\in \mathbf{T} \subset \Re^+\right\}$ using a smoothing technique.

In order to build functions, Ramsay and Silverman (2007) presented two main basis systems namely: the Fourier system and the B-spline system. Those systems are now well-established in the statistical literature of FDA. Actually, most of theoretical developments have been made based on them. As suggested by Ternynck et al. (2016), we use the B-spline basis system for smoothing spring and summer daily discharge data. The Fourier system is commonly used for periodic data, while the B-spline system is rather used for non-periodic data. Fourier basis functions have been used by Chebana et al. (2012) for smoothing daily streamflow that cover the entire year to obtain annual streamflow curves. Since the present application considers only the spring and summer period, the Fourier basis appears, however, to be less suited.

The main idea of the change point detection, here, is to test whether the mean of the functional observations $X_1,\ldots,X_n$ remains constant over time. We assume that $X_i(t)=\mu_i(t)+\varepsilon_i(t), i=1,\ldots,n$ where $\mu_i(t)$ denotes the functional mean and $\varepsilon_i(t)$ is a zero-mean functional sequence. We wish to test the null hypothesis $H_0:\mu_1(t)=\mu_2(t)=\cdots=\mu_n(t)$ against the alternative H_a that there is an unknown change point k^* in the mean, i.e. $H_a:\mu_1(t)=\mu_2(t)=\cdots=\mu_{k^*}(t)\neq\mu_{k^*+1}(t)=\cdots=\mu_n(t)$. The change can occur at any point i and we want to test whether it occurs or not. The existence of change points means that the data can be divided into several consecutive segments, with a constant mean within each segment. Berkes et al. (2009) proposed an approach to test the assumption of a common functional mean for independent data. This approach is based on the following quantity (which measures a deviation between the mean of the functional observations $X_1,\ldots,X_k$ and that of $X_{k+1},\ldots X_n$):

$$P_k(t)=\frac{k(n-k)}{n}\left\{\hat{\mu}_k(t)-\tilde{\mu}_k(t)\right\}, k=1,\ldots,n \qquad (1)$$

where $\hat{\mu}_k(t) = \frac{1}{k}\sum_{1\le i\le k} X_i(t)$ and $\tilde{\mu}_k(t) = \frac{1}{n-k}\sum_{k+1\le i\le n} X_i(t)$. If the mean changes, the difference $P_k(t)$ is large for some values of k and t. To deal with the infinite dimension of the observations (curves), we consider the projections of the functions $P_k(\cdot)$ on the principal components of the data. In fact, principal component analysis represents functional data as $X_i(t) = \mu(t) + \sum_{1\le l\le\infty} \eta_{i,l}\upsilon_l(t)$, where $\mu(t)$ is the functional mean, $\eta_{i,l}$ are the scores and $\upsilon_l(t)$ are the eigen-functions of the covariance operator (Hall and Hosseini-Nasab 2006). These projections can be expressed in terms of functional scores, which can be easily computed using the R package "fda". We consider the estimated scores $\hat{\eta}_{i,l}$ corresponding to the largest L eigenvalues given by:

$$\hat{\eta}_{i,l} = \int \left\{ X_i(t) - \bar{X}_n(t) \right\} \hat{\upsilon}_l(t)dt, \quad i = 1,2,\ldots,n \quad j = 1,2,\ldots,L \tag{2}$$

with $\bar{X}_n(t)$ is the sample mean function and $\hat{\upsilon}_l(t), l = 1,\ldots,L$ are the estimated eigen-functions of the covariance operator. It is supposed that $k = [n\alpha]$ where $\alpha \in (0,1)$ and $[\cdot]$ denotes the integer part. Note that $P_k(t)$ does not change if the $X_i(t)$ are replaced by $X_i(t) - \bar{X}_n(t)$. Hence, $P_k(t)$ can be written as:

$$P_k(t) = \sum_{1\le i\le k} \left(X_i(t) - \bar{X}_n(t) \right) - \frac{k}{n} \sum_{1\le i\le n} \left(X_i(t) - \bar{X}_n(t) \right) \tag{3}$$

Consequently, the projections are defined by $\int P_k(t)\hat{\upsilon}_l(t)dt = \sum_{1\le i\le n\alpha} \hat{\eta}_{i,l} - \frac{[n\alpha]}{n} \sum_{1\le i\le n} \hat{\eta}_{i,l}$ and are used for testing whether the mean function remains constant. For this purpose, the following statistic is considered:

$$S_{n,L} = \frac{1}{n^2}\sum_{l=1}^{L}\lambda_l^{-1}\left(\sum_{1\le i\le nz}\hat{\eta}_{i,l} - \frac{k}{n}\sum_{1\le i\le n}\hat{\eta}_{i,l}\right)^2 \qquad (4)$$

where $\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_L$ denote the L -estimated eigenvalues. The test rejects the hypothesis H_0 if $S_{n,L}$ is greater than the corresponding critical value, tabulated in Berkes et al. (2009).

While this test does not take into account the temporal dependence it will be considered here as a first simple step to introduce the functional change point detection framework in hydrology. Few other researchers recognize this limitation and propose some improvements. For instance, a more complex approach has been proposed by Zhang et al. (2011) in order to take into account the temporal dependence. For sake of simplicity, this latter will not be considered in the current analysis. In the interim, the functional approach being used here may, nevertheless, serves as a stepping-stone towards this more complex approach.

3. Results

Results of application of the proposed method to the above-mentioned data set are compared with those obtained in Seidou and Ouarda (2007). In the latter, the change point detection method has been applied separately to the peak, the duration and the volume of flood events occurring in spring and summer. The first step to apply the functional method consists on performing a functional principal component analysis where the first principal components explaining large part of the data variance are, therefore, to be retained. For the Romaine river station, we retained the first four principal components as they represent 83% of the explained variance. In the hypothesis testing, we set the first type error at 5%. By applying the functional method for the Romaine river station, we obtain a change point at the year 1984. This suggests that we can split the set of curves into the following two segments $\overline{\mathrm{TD}}_1^a$: 1961-1984 and $\overline{\mathrm{TD}}_2^a$: 1985-2000, of size 24 and 16, respectively as shown in Figure 2.a. We can see from Figure 3 that based on the mean, the median and the modal curves, the two obtained segments have two different peaks. The peak of the first segment is significantly higher than that of the second. One can also note that changes affect not only the peak, but also the duration, the volume and the peak date of the flood event as well. Indeed, in both classes, flood events began at the same time, but last longer in $\overline{\mathrm{TD}}_1^a$.

In a second step, we reiterate the procedure on the obtained two segments. We therefore only find a change point on the segment $\overline{\mathrm{TD}}_1^a$ at the year 1968. Consequently, we obtain the three following periods $\overline{\mathrm{TD}}_1^b$: 1961-1968, $\overline{\mathrm{TD}}_2^b$: 1969-1984 and $\overline{\mathrm{TD}}_3^b$: 1985-2000 of respective size 8, 16 and 16. According to the Figure 4, we can see that based on the mean curve, flood events of the segment $\overline{\mathrm{TD}}_2^b$ begin before those of the $\overline{\mathrm{TD}}_1^b$, however floods in both segments end at the same time. Accordingly, the flood durations for the segment $\overline{\mathrm{TD}}_2^b$ are larger than those of the segment $\overline{\mathrm{TD}}_1^b$. While flood events in both $\overline{\mathrm{TD}}_2^b$ and $\overline{\mathrm{TD}}_1^b$ have almost the same peak, the functional approach seems to be able to detect the difference in the duration of the flood events. Flood events of the segment $\overline{\mathrm{TD}}_2^b$ begin at the same time with the flood events of the segment $\overline{\mathrm{TD}}_3^b$, and then they take end at the same time with the segment $\overline{\mathrm{TD}}_1^b$. Moreover, the two segments $\overline{\mathrm{TD}}_1^b$ and $\overline{\mathrm{TD}}_2^b$ have almost the same peak. Consequently, the segment $\overline{\mathrm{TD}}_2^b$ can be considered as an intermediate period that enables the transition from the flood regime of the segment $\overline{\mathrm{TD}}_1^{b}$ to the flood regime of the segment $\overline{\mathrm{TD}}_3^{b}$.

In conclusion, for the Romaine river station, functional change point method, has detected two change points, the first at year 1984 and the second at years 1968 as shown in Figure 2.a. This result has divided flood events for the Romaine river station into three periods: the first with very large floods, which begins later, a second intermediate period, and a third period characterized by less important floods which starts early. For the comparison of the functional change point results with a traditional method approach we applied the Bayesian approach of Seidou and Ouarda (2007) to the peak, the volume and the duration of flood events separately. The method of Seidou and Ouarda (2007) based on the duration detects a change point at the year 1987. The same method, however, based on the volume and the peak detects a change point at the year 1985, which is closer to the first change point detected by the proposed functional approach (at year 1984). The Bayesian approach based on the volume and the peak separately was not able to detect the second change point in the segment $\overline{\mathrm{TD}}_1^{a}$. This is due to the fact that this change does not affect the peak or the volume, but mainly affects the occurrence time of flood events. The functional approach allows detecting this change in the occurring time of flood events because it directly considers a large part of the information contained on the entire discharge series, including information on shape, peak time, duration, etc...

For the Moisie river station, using the functional approach, we choose the first four principal components since they represent 85% of the explained variance. In the hypothesis testing, we set the first type error at 5%. We obtain a change point at the year

1981 which suggests splitting the set of curves into two segments as follows, $\overline{\mathrm{TD}}_1^c$: 1968-1981 and $\overline{\mathrm{TD}}_2^c$: 1982-1991, of size 14 and 10, respectively. We, then, reiterate the procedure on the obtained two segments, but no change point was detected. Therefore, we can conclude that this method allows detecting just one change point at year 1981 as shown in Figure 2.b. Figure 5 shows the mean curve, the median curve and the modal curve of flood hydrograph corresponding to the two obtained segments. This figure shows that flood events in the two segments $\overline{\mathrm{TD}}_1^c$ and $\overline{\mathrm{TD}}_2^c$ occur at the same date, but those of the segment $\overline{\mathrm{TD}}_1^c$, last longer and have a larger peak. For the Moisie river station we test the existence of a change point on the peak, the volume and the duration separately using the method of Seidou and Ouarda (2007). Only, the method based on the peaks detects a change point at the year 1978**.**

4. Discussions

It is worth noting that the purpose of the comparison with the conventional approach is not to show that the functional approach performs better, but rather to check whether this approach gives results consistent with those obtained using a traditional approach. Note that, when the focus is only on one characteristic of the flood event, such as the peak, the volume or the duration, traditional univariate approach preferred. However, the functional approach takes into account all the characteristics of the flood event simultaneously, hence, if no preferences on the flood event characteristic, the functional approach is recommended. Then, the graphical representation of the median curve, the mode curve and the mean curve is helpful to summarize the differences between the different flood periods after the detection of the change point.

It should be borne in mind that numerous caveats apply to our findings. First, the proposed functional framework suffers from an edge effects issue and therefore is unable to identify possible changes near the beginning and the end of the data record. Nevertheless, this is a common issue for the traditional change point approaches. Further theoretical studies using generated (known) functional data sets may help to quantify this issue, as well to answer many other questions such as the determination of the minimum record length in order to detect a change. Secondly, the proposed approach does not allow detecting multiple change points simultaneously, and thus need to be iterated for each segment until no further change point is detected. Finally, as problems in hydrology often involve missing data, the proposed functional approach lacks the ability of handling missing data, and thus unable to take full advantage of the whole data record that may be available. For instance, a complete data records are available for Romaine river station from 1957 to 2012 as well for Moisie river station from 1966 to 2012, while, in contrast, our analysis was mainly limited to data records from 1961 to 2000 for Romaine river station and from 1968 to 1991 which are the longest periods for which there is no missing data.

In change point analysis, if a significant change is detected in hydrological characteristics, then it is important to try to understand the physical reason behind. Change in hydrological characteristics may be caused by climatic factor such as climate variability or climate change, but there may be many other possible explanations, such as anthropogenic change (urbanization, water abstraction etc.), natural catchment changes, and problem linked to data. The best way to improve understanding of change is rather to gather as much information as possible, using, e.g., information about change in the catchment. In addition, related variables, like temperature and precipitation can help to determine whether changes in flow can be explained by climatic factors. Indeed, streamflow depends strongly on the spatial distribution of precipitation in a watershed, and on the interactions between temperature and precipitation which determines whether precipitation falls as rain or snow (Ben Alaya et al. 2014).

In a warming climate it is expected that the atmosphere's water holding capacity will increase with warming according to the Clausius-Clapeyron (C-C) equation (Collins et al. 2013), which may lead to more intense precipitation events that may directly affect streamflow and flood events behaviours. In addition, climate variability through oceanic and atmospheric oscillations on a large scale known as teleconnections, such as the North Atlantic Oscillation (NAO), El Nino-Southern Oscillation (ENSO) and Pacific Decadal Oscillation (PDO), influences the variability and trends in the climate system (Hurrell and Van Loon 1997, Rogers 1997) and thus may in turn affect characteristics of flood events.

Based on the obtained results, the frequency of flood events which occur later has decreased at both Romaine river and Moisie river stations while earlier floods characterized by low peaks and volumes became more frequent. Given the short record length of the data series used, attributing this change to corresponding underlying processes is challenging. Another challenge is that signals such as trends and shifts are superposed on variability arising from the memory within the hydrological system.

While the proposed approach is not able to distinguish between shifts and trends that may be present in the data, the results for the Romaine river station reflect hints about the presence of a trend in functional data. Note, however, that the “trend” terminology in case of the sequence of functional curves is not the same used as in case of random variables where sample elements are points. The definition of a trend in case of the sequence of functional curves requires, first, to define an extended notion of order that tell us in which case a curve can be considered to be higher than another. Such a definition may not be, however, uniquely determined. To the best of our knowledge, a first attempt for functional trend analysis has been proposed by Fraiman et al. (2014). Nevertheless, we think that a more comprehensive way to handle this concern is to account for the notion of autocorrelation in the sequence of functional curves that has been proposed by Zhang et al. (2011).

Note that change point analyses are only descriptive. Hence, they cannot answer questions about how the hydrological system works in a non-stationary climate, and, therefore, cannot be used to predict future conditions. Indeed, seeking answers to those questions requires a hydrological modelling. Nevertheless, the proposed functional change point framework, as a mathematical descriptive tool, can play a very important role for scientific investigations. It can help to get first quantitative clues about what happened in the past. Unlike traditional change point approaches, the conclusions reached from the proposed framework are enhanced by providing reach mathematical pictures summarizing the mean, mode and median curves describing flood regimes. Those pictures are obtained within a mathematical framework that rigorously explains how they were obtained and how the conclusions were reached. Those steps can be easily applied to outputs of hydrological models, whether deterministic or stochastic, to rigorously check and test whether they reproduce similar pictures and same conclusions that have been drawn from original data. As suggested in previous works, the interpretation of hydrological models, particularly in a non-stationary climate is always challenging (Montanari and Koutsoyiannis 2014, Serinaldi and Kilsby 2015, Serinaldi and Kilsby 2018). On the other hand, by following the first clues given by a reach descriptive analysis, we may achieve a deep understanding about the complexity of the real mechanism. Our finding can serve as a starting point toward an effective calibration of hydrological models, or can merely be used for model testing.

As recommended by Koutsoyiannis and Montanari (2015), including additional information from prior physical knowledge about the physical process involved is essential to build a successful hydrological model that can be used to predict the future. In this respect, our approach opens the door to think on how to take full advantage of the functional framework from a modelling perspective. This can be achieved by casting the hydrological modelling in a functional regression framework by including major factors that influence flood events as covariates. The implementation of such approach, however, is not straightforward and is outside the scope of the current paper. Nevertheless, we think that, slowly, over the course of several steps, starting from functional descriptive tools, a pathway can be paved for development of sound functional regression models and perhaps eventual inclusion of additional knowledge from key factors involved in generating flood events.

5. Conclusions

The purpose of the present paper is to propose a new context of the change point detection of flood hydrographs using functional data framework. A functional change point approach is presented and adapted to flood events. An application is performed for two hydrological stations in the province of Quebec, Canada. The presented functional approach is compared to a classical Bayesian univariate approach applied to the peak, the volume and the duration of flood events separately. Based on this comparison, it has been shown that the functional approach gives results that are consistent with the traditional univariate approach. The functional approach has the benefit that it provides a comprehensive way to handle the flood event as a curve within a defined statistical framework and thus an opportunity for a better exploitation of the information contained in a discharge time series.

References

Aue, A., R. Gabrys, L. Horváth and P. Kokoszka (2009). "Estimation of a change-point in the mean function of functional data." Journal of Multivariate Analysis **100**(10): 2254-2269.

Ben Alaya, M., F. Chebana and T. Ouarda (2014). "Probabilistic Gaussian Copula Regression Model for Multisite and Multivariable Downscaling." Journal of Climate **27**(9): 3331-3347.

Berkes, I., R. Gabrys, L. Horváth and P. Kokoszka (2009). "Detecting changes in the mean of functional observations." Journal of the Royal Statistical Society: Series B (Statistical Methodology) **71**(5): 927-946.

Blöschl, G., J. Hall, J. Parajka, R. A. Perdigão, B. Merz, B. Arheimer, G. T. Aronica, A. Bilibashi, O. Bonacci and M. Borga (2017). "Changing climate shifts timing of European floods." Science **357**(6351): 588-590.

Chebana, F., S. Dabo-Niang and T. B. Ouarda (2012). "Exploratory functional flood frequency analysis and outlier detection." Water Resources Research **48**(4).

Collins, M., R. Knutti, J. Arblaster, J.-L. Dufresne, T. Fichefet, P. Friedlingstein, X. Gao, W. Gutowski, T. Johns and G. Krinner (2013). "Long-term climate change: projections, commitments and irreversibility."

Dabo-Niang, S., A.-F. Yao, L. Pischedda, P. Cuny and F. Gilbert (2010). "Spatial mode estimation for functional random fields with application to bioturbation problem." Stochastic Environmental Research and Risk Assessment **24**(4): 487-497.

Easterling, D. R. and T. C. Peterson (1995). "A new method for detecting undocumented discontinuities in climatological time series." International journal of climatology **15**(4): 369-377.

Fischer, A. (2010). "Quantization and clustering with Bregman divergences." Journal of Multivariate Analysis **101**(9): 2207-2221.

Fraiman, R., A. Justel, R. Liu and P. Llop (2014). "Detecting trends in time series of functional data: A study of Antarctic climate change." Canadian Journal of Statistics **42**(4): 597-609.

Goudie, A. S. (2006). "Global warming and fluvial geomorphology." Geomorphology **79**(3): 384-394.

Hall, P. and M. Hosseini-Nasab (2006). "On properties of functional principal components analysis." Journal of the Royal Statistical Society: Series B (Statistical Methodology) **68**(1): 109-126.

Hannah, D. M., B. P. Smith, A. M. Gurnell and G. R. McGregor (2000). "An approach to hydrograph classification." Hydrological processes **14**(2): 317-338.

Hurrell, J. W. and H. Van Loon (1997). Decadal variations in climate associated with the North Atlantic Oscillation. Climatic Change at High Elevation Sites, Springer**:** 69-94.

Kingston, D., J. Thompson and G. Kite (2011). "Uncertainty in climate change projections of discharge for the Mekong River Basin." Hydrology and Earth System Sciences **15**(5): 1459-1471.

Koutsoyiannis, D. and A. Montanari (2015). "Negligent killing of scientific concepts: the stationarity case." Hydrological Sciences Journal **60**(7-8): 1174-1183.

Kundzewicz, Z. W. and A. J. Robson (2004). "Change detection in hydrological records—a review of the methodology/revue méthodologique de la détection de changements dans les chroniques hydrologiques." Hydrological Sciences Journal **49**(1): 7-19.

Larabi, S., A. St-Hilaire, F. Chebana and M. Latraverse (2018). "Multi-Criteria Process-Based Calibration Using Functional Data Analysis to Improve Hydrological Model Realism." Water Resources Management **32**(1): 195-211.

Masselot, P., S. Dabo-Niang, F. Chebana and T. B. Ouarda (2016). "Streamflow forecasting using functional regression." Journal of hydrology **538**: 754-766.

Montanari, A. and D. Koutsoyiannis (2014). "Modeling and mitigating natural hazards: Stationarity is immortal!" Water Resources Research **50**(12): 9748-9756.

Ramsay, J. O. and B. W. Silverman (2007). Applied functional data analysis: methods and case studies, Springer.

Requena, A. I., F. Chebana and T. B. Ouarda (2018). "A functional framework for flow-duration-curve and daily streamflow estimation at ungauged sites." Advances in Water Resources **113**: 328-340.

Rogers, J. C. (1997). "North Atlantic storm track variability and its association to the North Atlantic Oscillation and climate variability of northern Europe." Journal of Climate **10**(7): 1635-1647.

Seidou, O. and T. B. Ouarda (2007). "Recursion-based multiple changepoint detection in multiple linear regression and application to river streamflows." Water Resources Research **43**(7).

Serinaldi, F. and C. G. Kilsby (2015). "Stationarity is undead: Uncertainty dominates the distribution of extremes." Advances in Water Resources **77**: 17-36.

Serinaldi, F. and C. G. Kilsby (2018). "Unsurprising Surprises: The Frequency of Record-breaking and Overthreshold Hydrological Extremes Under Spatial and Temporal Dependence." Water Resources Research **54**(9): 6460-6487.

Solow, A. R. (1987). "Testing for climate change: An application of the two-phase regression model." Journal of Climate and Applied Meteorology **26**(10): 1401-1405.

Suhaila, J. and Z. Yusop (2017). "Spatial and temporal variabilities of rainfall data using functional data analysis." Theoretical and Applied Climatology **129**(1-2): 229-242.

Ternynck, C., M. A. Ben Alaya, F. Chebana, S. Dabo-Niang and T. B. Ouarda (2016). "Streamflow hydrograph classification using functional data analysis." Journal of hydrometeorology **17**(1): 327-344.

Vincent, L. A. (1998). "A technique for the identification of inhomogeneities in Canadian temperature series." Journal of Climate **11**(5): 1094-1104.

Wong, H., B. Hu, W. Ip and J. Xia (2006). "Change-point analysis of hydrological time series using grey relational method." Journal of hydrology **324**(1): 323-338.

Zhang, X., X. Shao, K. Hayhoe and D. J. Wuebbles (2011). "Testing the structural stability of temporally dependent functional observations and application to climate projections." Electronic Journal of Statistics **5**: 1765-1796.

List of Figures

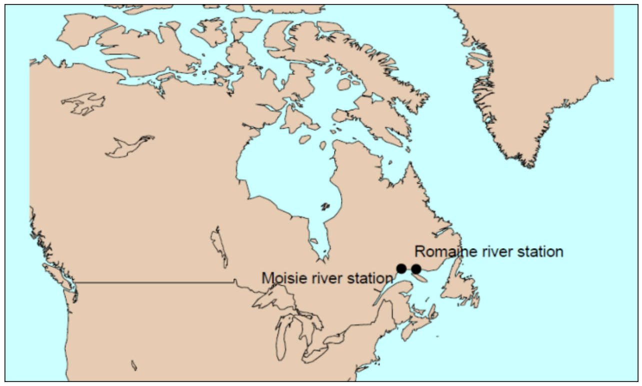

Figure 1. Geographical locations of the Romaine river and the Moisie river stations, in Quebec.

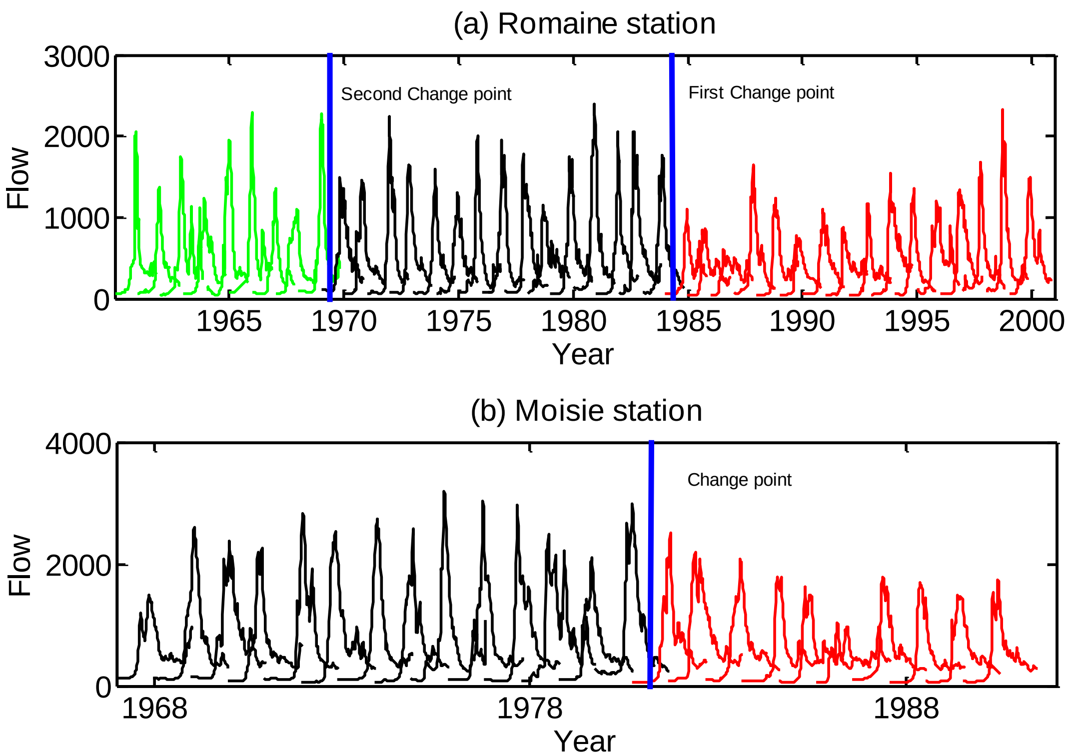

Figure 2. Functional change point results, for romaine river station in (a) and Moisie river station in (b).

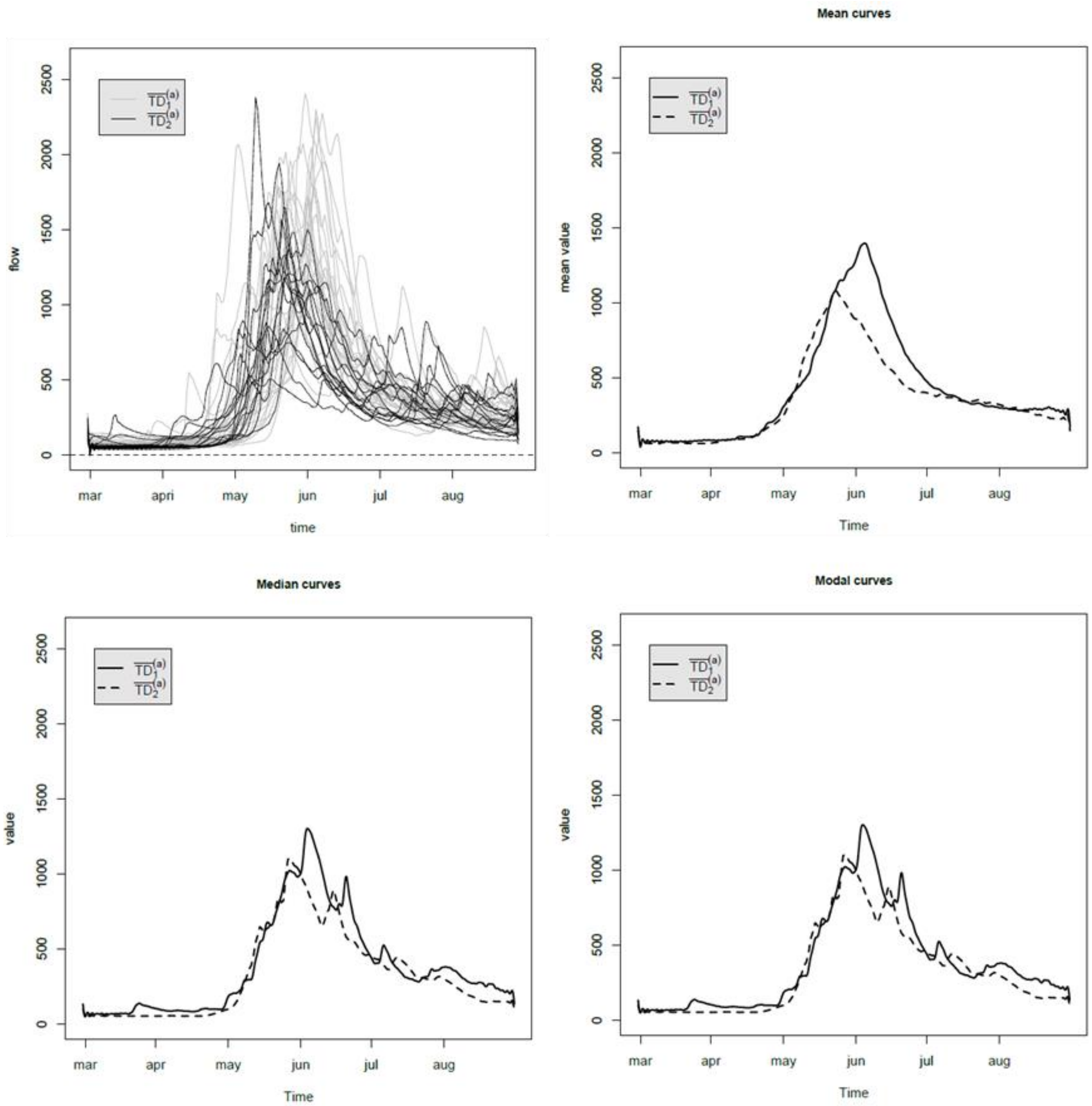

Figure 3. Functional curves of flood events for the two segments obtained using the functional change point detection method for the Romaine river station after the detection of the first change point. All functional curves over the period 1961-2000 are presented in (a). The mean, median and modal curves for each segment are presented respectively in (b), (c) and (d).

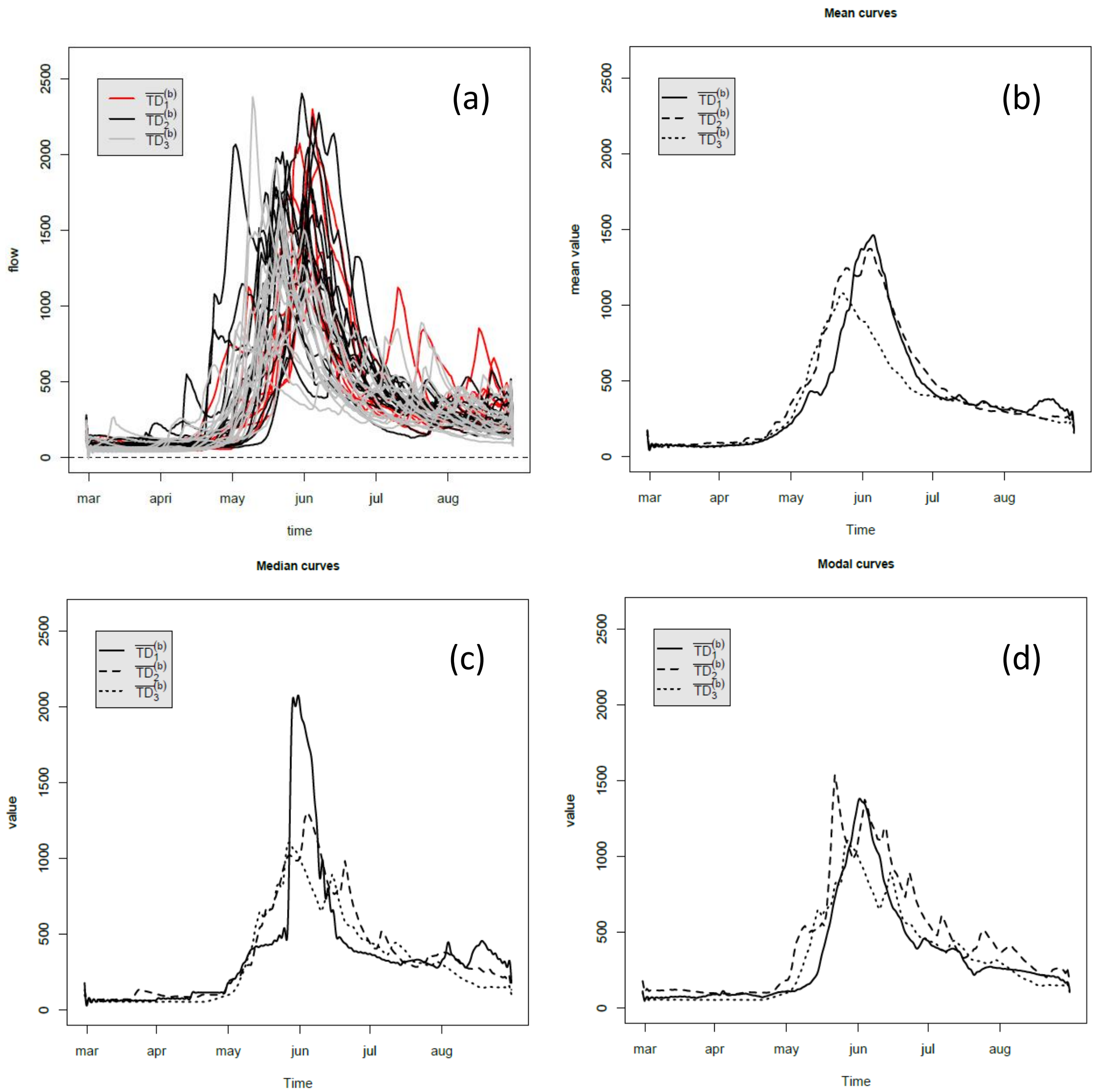

Figure 4. Similar to Figure 3 but for the three segments obtained using the functional change point detection method for the Romaine river station after the detection of the second change point.

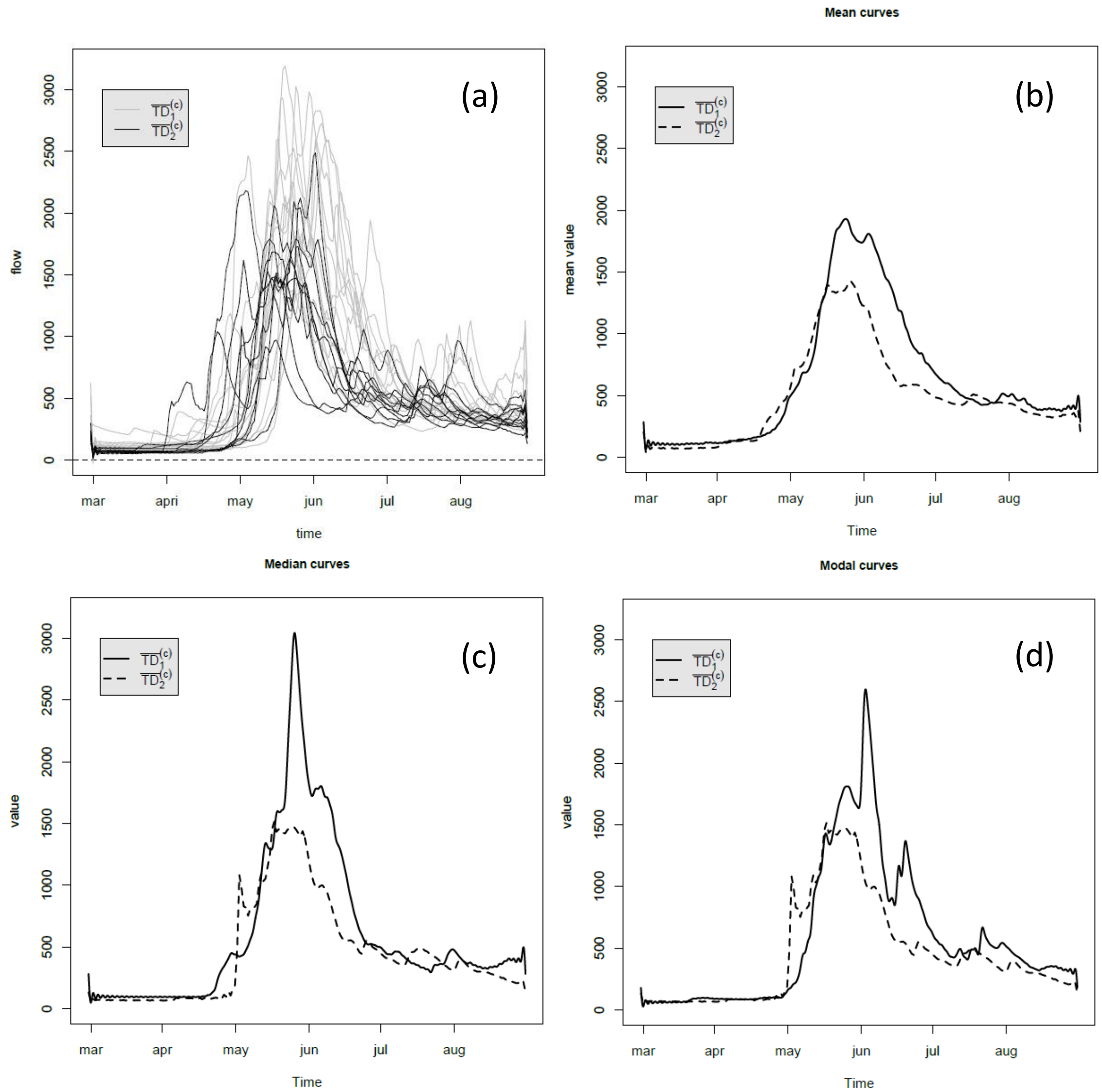

Figure 5. Functional curves of flood events for the two segments obtained using the functional change point detection method for the Moisie river station after the detection of the single change point. All functional curves over the period 1968-1991 are presented in (a). The mean, median and modal curves for each segment are presented respectively in (b), (c) and (d).