A Multiple Changepoint Approach to Hydrological Regions Delineation

Ali Morabbi[a,*], Ahmed Bouziane[a], Ousmane Seidou[b,c], Nabil Habitou[a], Driss Ouazar[a], Taha B.M.J Ouarda[d], Christian Charron[d], Moulay Driss Hasnaoui[a,e], Mounia Benrhanem[f], Ketvara Sittichok[g]

[a] Mohammed V University in Rabat, Mohammadia School of Engineers, Department of Civil Engineering, Hydraulic Systems Analysis Team (EASH), Box765 Agdal, 10000 Rabat, Morocco

[b] Department of Civil Engineering, University of Ottawa, 161 Louis Pasteur Office A113 Ottawa, ON, K1N 6N5, Canada

[c] United Nations University, Institute for Water, Environment and Health (UNU-INWEH), 204-175 Longwood Road South, Hamilton, ON L8P0A1, Canada

[d] Canada Research Chair in Statistical Hydro-Climatology, Institut national de la recherche scientifique, Centre Eau Terre Environnement, Quebec City, QC G1K9A9, Canada

[e] The Ministry of Equipment, Transport, Logistics and Water – Water Department, Rabat, Morocco

[f] Tensift River Basin Agency, 40000, Marrakech, Morocco

[g] Kasetsart University, Department of Irrigation Engineering, Faculty of Engineering Kamphaengsaen campus, Kamphaengsaen, Nakhonpathom, 73140, Tahiland

*Corresponding author: aliucam13@gmail.com

Abstract

Hydrologic regionalization consists of regrouping stations and catchments in pools based on a similarity measure. Regionalization is commonly used to extract a robust signal that can be used to describe the hydrology of the region or extrapolated to a location without measured information. Obviously, the similarity measure used affects the type of hydrological behavior one would expect from stations within a region. Most regionalization methods assume a stable and/or linear relationship between parameters of interests while it is well known that the physical processes driving the behavior of hydrometeorological variables are inherently non-linear and non-stationary. In this paper, we propose a similarity measure that is based on the location of changepoints in hydrological time series. The proposed method has the unique advantage over other hydrological region delineation methods to detect regions where hydrological member stations are non-linearly correlated, and where the strength of the relation varies with time. It therefore has the potential to uncover similarities that would not have been detected by existing regionalization techniques. The proposed method is applied to the Tensift watershed located in Morocco, North Africa. The coherence of the detected regions is checked using wavelet coherence.

Keywords: Hydrologic regionalization; Similarity; Hydrological time series; Changepoints; Wavelet coherence.

1. Introduction

The Tensift Basin (Figure 1), an arid to semi-arid, 20381 km² watershed located in West-Central Morocco. It is limited by the Jbilets desert in the north, the high Atlas Mountains in the south, the Tassaout watershed in the East, and the Atlantic Ocean in the West. The climate of the region is classified as arid to semi-arid (Saidi et al., 2010; Mahé et al., 2012). Precipitation in the watershed is highly variable in space and time. Annual rainfall varies from 650 mm in Mountain areas, to 100 mm in desert plains (Saidi et al., 2003, 2010, 2012). Agriculture is the main socio-economic activity in the watershed. Its expansion has created a pressure on existing water resources, and a depletion of the water table has been observed. The increasing water scarcity is generating fears of recurrent droughts.

The study area is located in a transition zone between Mediterranean and semi-arid climates, a region known to be affected by global changes such as the expansion of the Hadley cell (Peleg et al., 2015). Evidence of the poleward expansion of the Hadley zone has been documented by several authors in past observations as well as CMIP5 and CMIP6 simulations (Hu et al, 2007; Peleg et al., 2015; Xia et al., 2020). The Hadley cell extension will affect the regional climate in the subtropics through the displacement of warm climate coupled with the disruption of deep-water upwelling (Feng and Fu, 2013; Schmidt and Grise, 2017). The poleward extension of the Hadley cell will make its intensity weaker, and ultimately lead to the extension of the subtropical dry zone (Lu et al., 2007). According to the draft IPCC AR6 technical report (Masson-Delmote et al., 2021), there is a high confidence that the total land area subject to increasing drought frequency and severity will expand, and that future aridification will far exceed the magnitude of change seen in the last millennium in several regions of the Mediterranean basin. There is hence a growing interest in understanding regional precipitation variability as it is of major importance for the management of water resources. Decision makers in Morocco are very much interested in understanding whether there is a single or multiple regional climates over the study area in order to implement the appropriate policies. One way of gaining that understanding is through hydrologic regionalization. Hydrologic regionalization consists of regrouping stations and catchments in pools which are expected to have the same hydrological behaviors. Regionalization has been intensively used for flood frequency analysis (Farquharson et al., 1992; Rosbjerg and Madsen, 1994; Durrans and Tomic, 1996; Pandey and Nguyen, 1999; Alila, 1999, 2000; Ouarda et al., 2000; Chokmani and Ouarda, 2004) and hydrological models parameters estimation (Sefton and Howarth, 1998; Mwakalila, 2003; Heuvelmans et al., 2006; Seidou et al., 2006; Hundecha et al., 2008). The delineation of regions is typically based on a similarity measure between watershed physiographic parameters or time-series statistics at the sites of interests. The similarity measure can be geographical distance (GREHYS, 1996b), or the proximity of catchment attributes in projected specific spaces (e.g. Ouarda et al., 1999, 2001; Han et al., 2020). Principal Components Analysis (PCA) is one of the most popular regionalization methods. PCA regionalization approach is defined as a multivariate statistical method that aims to analyze and simplify a data table with M observations and N variables (Basilevsky, 1994). It consists in transforming correlated variables into orthogonal principal components. The variables are projected in a new space where along each axis, the variance is maximized. El Alaoui El Fels et al. (2020) applied PCA to monthly precipitation in the Tensift region to delineate three homogenous regions. Ahattab et al. (2015) applied PCA to monthly precipitation from 23 stations in the Tensift region and found four homogenous regions. One of the limitations of PCA is the assumption that the relationships between time series in a region are linear and constant trough time. At the same time, natural relations tend to be non-linear and non-stationary. Transformations of input time series such as log-transformations allow accounting for a limited amount of nonlinearity. Agarwal et al. (2016) used multiscale wavelet entropy to define the similarity measure between catchments. They studied the monthly streamflow temporal variability of 530 stations during 52 years over the United States and determined homogenous groups where each one has a signature by defining clusters. The idea of using wavelets for regions delineation is interesting as it accounts for relations that are non-stationary and non-linear. One way to account for nonlinearity and non-stationarity in the relations between two time-series is to examine the similarities between the dates of changes in trends in the two time-series. Time series in the same hydrologic region may not be linearly correlated but will have similar locations for the changepoints. While any multiple changepoint approaches can be used to detect the positions of the changepoints, we will be using the multiple changepoint detection procedure developed by Seidou and Ouarda (2007) because of its flexibility. The method is designed to detect multiple changepoints in a multiple linear relation between one or many predictors and a predictand. By changing the predictors, one can detect abrupt shifts, continuous or non-continuous linear segments in analyzed time series. The method also has the advantage of providing a probabilistic description of the changepoint locations.

Another alternative for hydrologic region delineation using non-stationary and non-linear relations, similar to Agarwal et al. (2016), is to define the similarity between two-time series using wavelet cross-coherence as defined by Torrence and Compo (1998). Cross Wavelet Transform (XWT) aims to evaluate the covariance and causality between X and Y. However, there are certain drawbacks associated with its use, because it is not normalized as reported by Maraun and Kurth (2004).

For this reason, the authors believe that the best method to adopt for the examination of causality is the wavelet transform coherence (WTC) in so far as two variables can be dependent without there being a strong link. XWT finds regions in the time-frequency space where the time-series show high common energy. Wavelet coherence is calculated using the continuous wavelet transform (CWT) of two time series. CWT is an alternative approach to the Fourier transform applied in the signal processing.

CWT and XWT has been widely used in many investigational studies (Grinsted et al., 2004; Özger et al., 2009; Keener et al., 2010; Ouachani et al., 2011; Sang., 2013; Naizghi and Ouarda (2016); N Thiombiano et al., 2016; Chang et al., 2017; Santos and al. 2018). The objectives of this paper are to extend the multiple changepoint detection method of Seidou and Ouarda (2007) to hydrological region delineation, apply the methodology to the Tensift watershed and verify the delineated regions using wavelet coherence.

2. Material and methods

2.1. Bayesian multiple changepoint detection approach

The original multiple changepoint detection method proposed by Seidou and Ouarda (2007) is a generalization of the changepoint procedure of Seidou et al. (2007). It is commonly used in the field of water resources (see for instance Ehsanzadeh et al., 2011). The procedure is briefly presented below:

Let $\mathbf{y} = \{y_1, y_2, \ldots, y_n\}$ a time series of observation of length n, m the number of changepoints, $\tau_0 = 0, \tau_1, \ldots, \tau_n = m$ the changepoints and $\mathbf{Y}_{i:j}$ the observations from time i to time j. Let g (.) be probability distribution of the time between changepoints and $g_0(.)$ is the probability distribution of the first changepoint. Assuming the observations are independent conditional on the changepoints and parameters values, Fearnhead (2006) demonstrated that if,

$$\begin{cases} \Pr(\tau_1 | Y_{1:n}) = P(1, \tau 1) Q(\tau 1 + 1) g_0(\tau 1)/Q(1) \\ \Pr(\tau_j | \tau_{j-1}, Y_{1:n}) = P(\tau_{j-1} + 1,\ \tau_1)\ Q(\tau_j + 1) g(\tau_j - \tau_{j-1})/Q(\tau_{j-1} + 1) \end{cases} \tag{1}$$

where

$$P(t,s) = \Pr(Y_{t:s}; t, s \text{ in the same segment}) = \int \prod_{i=t}^{s} f(y_i \,|\boldsymbol{\Phi}) \boldsymbol{\pi}(\boldsymbol{\Phi}) \boldsymbol{d\Phi} \tag{2}$$

and Q(t) is the likelihood of segment $\mathbf{Y}_{t:n}$ given a changepoint at t-1; $Q(t), t = 1, \ldots, n$ and $P(t,s), s \geq t$ are linked by these recursive equations:

$$\begin{cases} Q(1) = \sum_{s=1}^{n-1} P(1,s) Q(s+1) g_0(s) + P(1,n)(1 - G_0\,(n-1)) \\ Q(t) = \sum_{s=1}^{n-1} P(t,s) Q(s+1) g_0(s+1-t) + P(t,n)(1 - G(n-1)) \end{cases} \tag{3}$$

Where $G(t) = \sum_{i=1}^{t} g(i)$ and $G_0(t) = \sum_{i=1}^{t} g_0(i)$.

Seidou and Ouarda (2007) derived an analytical expression for $P(t,s)$ when $\mathbf{Y}$ is linked to a predictor $\mathbf{X}$ by a multiple linear regression equation.

$$y_j = \sum_{k=1}^{d^*} \theta_k x_{ij} + \varepsilon_i \qquad i = 1, \ldots, n \tag{4}$$

or

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \tag{5}$$

$$y_j = \sum_{k=1}^{d^*} \theta_k x_{kj} + \varepsilon_i \qquad = 1, \dots, \quad \boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{\varepsilon}$$

Seidou and Ouarda (2007) used Jeffrey's non-informative prior:

$$\pi_1(\Phi) = \pi_1(\sigma) = p(\sigma|\, a, c) = \frac{\sigma^{-a} exp\left(-\frac{c}{2\sigma^2}\right)}{2^{\frac{a-2}{2}} c^{\frac{a-1}{2}} \Gamma\left(\frac{a-1}{2}\right)} \quad a > 1, c > 0 \qquad (6)$$

$$\pi(\Phi) = \pi(\sigma) = p(\sigma|\, a, c) = \frac{\sigma^{-a} exp\left(-\frac{c}{2\sigma^2}\right)}{2^{\frac{a-3}{2}} c^{-\frac{a-1}{2}} \Gamma\left(\frac{a-1}{2}\right)} \quad a > 1,\ c > 0$$

Where a and c are the parameters of the prior. Using that prior, they demonstrated that:

$$P(t,s) = (2\pi)^{\frac{d^*}{2}} \frac{\left(\pi(\boldsymbol{\varepsilon}_{t:s}^T \boldsymbol{\varepsilon}_{t:s} + c)\right)^{-\frac{(t-s+a)}{2}}}{(c\pi)^{-\frac{a-1}{2}} \left|\mathbf{X}_{t:s}^T \mathbf{X}_{t:s}\right|^{1/2}} \frac{\Gamma\left(\frac{t-s+a}{2}\right)}{\Gamma\left(\frac{a-1}{2}\right)} P(t,s) = \frac{P_1(\{(1-l_1):0\} \cup \{t:s\})}{2P_1(\{(1-l_1):0\})} +$$

$$\frac{P_1(\{t:s\} \cup \{(n+1):(n+l_2)\})}{2P_1(\{n+1):(n+l_2)\})} \qquad (7)$$

$$P_1(\{i_1,\ i_2, \dots, i_{n1}\}) = \int_\Phi \prod_{i \in \{i_1, i_2, \dots, i_{n1}\}} f(y_i|\Phi)\, \pi(\Phi) d\Phi =$$

$$(2\pi)^{\frac{d^*}{2}} . \frac{\left(\pi\left(\varepsilon^T_{\{i_1,i_2,\dots,i_{n1}\}} \varepsilon_{\{i_1,i_2,\dots,i_{n1}\}} + c\right)\right)^{\left(\frac{n_1+a-1}{2}\right)}}{(c\pi)^{-\frac{a-1}{2}} |X^T_{\{i_1,i_2,\dots,i_{n1}\}} X_{\{i_1,i_2,\dots,i_{n1}\}}|^{1/2}} \cdot \frac{\Gamma\left(\frac{n_1+a-d^*}{2}\right)}{\Gamma\left(\frac{a-1}{2}\right)} \qquad (8)$$

Where d^* is the number of explanatory variables (including the intercept if any), $\varepsilon_{t:s}$ $\varepsilon_{\{i_1,i_2,\dots,i_{n1}\}}$ is the vector of residuals in the linear relationship between $\mathbf{X}_{t:s}$ $X_{\{i_1,i_2,\dots,i_{n1}\}}$ and $\mathbf{y}_{t:s}$ $Y_{\{i_1,i_2,\dots,i_{n1}\}}$.

The inference on the position of the changepoints is made by generating a set $E = \{S_k, k = 1:M\}$ of M possible scatter schemes of the changepoints on the segment using the posterior probability mass of the first changepoint and the conditional probability mass of subsequent changepoints. $E = \{S_k, k = 1{:}M\}$. The k^{th} element of E (caller herein changepoint scatter scheme) is a set of m_k changepoints $S_k = \{\tilde{t}_1^k, \tilde{t}_2^k, \dots, \tilde{t}_{\tilde{m}_k}^k\}$. An efficient simulation algorithm for E is given by Fearnhead [2006]:

1. For a sample of size M, initiate M samples with a changepoint at $t = 0$.
2. For $t = 0, \dots, n-2$, repeat the following steps:

a) Compute the number n_t of samples for which the last changepoint was at time *t*;

b) If $n_t > 0$, compute $\Pr(\tau \mid \tau_{j-1} = t, \mathbf{y}_{1:n})$;

c) Sample n_t times from $\Pr(\tau \mid \tau_{j-1} = t, \mathbf{y}_{1:n})$ and use the values to update the n_t samples of changepoints which have a changepoint at time *t*;

In a practical problem, it is unlikely to have two changepoints that are very close. Hence, when sampling from $\Pr(\tau \mid \tau_{j-1} = t, \mathbf{y}_{1:n})$, if the next position is within the length of the time series but is less than a user-defined ($l_{\min}$) from the previous changepoint, it is discarded, and another value is sampled. Inference on the number and positions of the changepoints is readily carried out using the M samples. For instance, the probability of having i changepoints is approximated by:

$$\Pr(m = i) \approx card(\{k \mid card(S_k) = i\}) / M \tag{9}$$

The posterior probability of having the k^th changepoint at position t given m changepoints can be approximated by:

$$\Pr(\tau_i = t \mid m) \approx \frac{card\left(\left\{k \mid \left(card(S_k) = m\right) \& \left(\tilde{t}_i^k = t\right)\right\}\right)}{card(\{k \mid card(S) = m\})} \tag{10}$$

Where $card(S)$ stands for the number of elements of the set S. The estimators of the number and positions of changepoints are the modes of their posterior distributions, i.e.:

$$\hat{m} = \underset{t}{Max}\{card(\{k \mid card_k(S) = t\}) / M\} \tag{11}$$

$$\hat{\tau}_i = \underset{t}{Max}\left\{\frac{card\left(\left\{k \mid \left(card(S_k) = \hat{m}\right) \& \left(\tilde{t}_i^k = t\right)\right\}\right)}{card\left(\{k \mid card(S) = \hat{m}\}\right)}\right\} \tag{12}$$

2.2. Use of changepoints positions as a measure of similarity

We propose to use the positions of the changepoints as a measure of similarity between two time-series. This is a fairly original use of the changepoint procedure. Ouarda et al. (2014) used the same changepoint procedure to identify common dates of change in precipitation series in a number of meteorological stations in the United Arab Emirates with the objective of identifying low frequency climate oscillation indices that influence precipitation in various regions. The procedure is presented below:

Assume an ensemble of n climatic stations where a time series of a particular parameter (e.g., annual precipitation) is available. For the sake of simplicity, let's assume that the time series are available in the same period from y_{start} to y_{end}. Let $\mathbf{y}_i$ the time series at the ith station

$$\mathbf{y}_i = \begin{pmatrix} PCP_{Ystart,i} \\ PCP_{Ystart+1,i} \\ \\ PCP_{Yend,i} \end{pmatrix}$$

Assume that the changepoint method of Seidou and Ouarda (2007) is applied to each station using the following predictand:

$$\mathbf{X} = \begin{pmatrix} 1 & y_{start} \\ 1 & y_{start+1} \\ & \\ 1 & y_{end} \end{pmatrix}$$

The application of the method to station *i* will result in a set $E_i = \{S_{k,i}, k = 1:M\}$ of M possible scatter schemes of the changepoints positions, as illustrated in Figure 2. We define the similarity between n stations as the number changepoints scatter schemes with at least one changepoint that are common to $E_1, E_2, ..., E_n$, divided by the maximum number of scatter schemes with at least one changepoint in any of $E_1, E_2, ..., E_n$. The concept of similarity is illustrated in Figure 3 using M=20 for a set of two and three stations. For the sake of illustration, the changepoints positions scatter schemes $E_i = \{S_{k,i}, k = 1:M\}$ are represented as an array where the columns correspond to the position in the time series, and the rows represent *k*. Black cells represent the changepoints. Lines that are common to all stations are coloured for illustration purposes. In the two-station example (panel a), the first station 19 rows with at least one changepoint, and the second one has 20. The two stations have four similar rows. Hence their similarity is $Sim_{changepoints}(S_1, S_2) = 4 / \min(19, 20) = 4/19 = 0.2105$. In the second example (panel b), the number of lines with at least one changepoint is 19, 17, and 18 for the first, second and third stations, respectively. Given that they have three lines in common, $Sim_{changepoints}(S_1, S_2, S_3) = 3 / \min(19,17,18) = 3/17 = 0.1765$

2.3. Regions delimitation using changepoint-based similarity

The following algorithm is used to delineate regions in a set of n stations using the similarity measure defined in section 3.1:

1. An initial set of n regions containing each one element is created
2. Repeat until each region contains each station:
    a. For each station s
        i. For each region r that does not contain s, calculate the similarity S_r of {region r, station s}
        ii. Assign station s to the region with the highest S_r and record S_r as the entrance score of station s in region r
    b. Remove duplicates in regions
3. At the end of the process, we have m revealed regions (k<=n) containing each of them, includes the n stations in a specific order
4. Iteratively search for the highest threshold t such as if stations with scores higher than t are removed from each region, the union of the remaining elements in the regions contains all n stations.
5. Eliminate all stations with scores above t and obtain m regions that cover the n stations

2.4. Local and regional trend

For each station in the study region, the probability of having one or several changepoints, as well as the most likely position of these changes, can be obtained using equations 9-11. Once the number and positions of the changepoints are determined, a local linear trend can be estimated between the detected changepoints. The modeler can decide whether the trend is continuous or discontinuous at the changepoints. The decision must be guided by the understanding of the physical process under consideration. For instance, hydroclimatic time series tend to vary smoothly because of climate change or climatic oscillations. Discontinuities generally occur only when there is a brutal change in the watershed, such as a forest fire or the construction of a hydraulic infrastructure. Unless the modeler has knowledge of such of a sudden change on the watershed, it is safer to assume continuity at the changepoints. In this paper, the trend is assumed to be continuous as we are dealing with precipitation time series and see no reason for an abrupt change. Once a region is delineated, the same method can be used to have a regional probability of having a changepoint, the position of the changepoints, and a regional trend. The only difference is that the union of the changepoints scatter schemes of all stations in the region is used in equations 9-11. The difference between the regional and local changepoints and trends can be a hint that a local perturbation has occurred, and can be used for hydrologic data homogenization. A number of other elements need to be integrated in the analysis. Future efforts will attempt to add more covariates (physiographic, climatic, land cover, etc.) in the model. This will also explaining a larger portion of the variance and developing a more detailed and complete model, which should translate into improved estimation results. In addition, the spatial and temporal variability, including potential change points, have to be combined and work in parallel in the same modeling logic to cover all aspects. The geomorphological aspect and the study of the prevailing climate remain insufficient in regards to the definition of the homogeneous regions.

2.5. Continuous wavelet transform and wavelet coherence

Wavelet coherence is a mean of evaluating the covariance and causality between X and Y. First, the continuous wavelet transform of X and Y, $W_n^X(.)$ and $W_n^Y(.)$, are calculated. The cross-wavelet transform of X and Y ($W_n^{XY}(s)$) is then calculated to explore high common power in both time-frequency domains between the two time series. Finally, their wavelet coherence $R_n^2(.)$ is calculated using their Wavelet and cross-wavelet transforms. The mathematical expressions of $W_n^X(.)$, $W_n^Y(.)$, and $R_n^2(.)$ Are presented in the next sections:

2.5.1. Continuous Wavelet Transform

The continuous wavelet transform for a discrete time series ($x_n, n = 1, \ldots, N$) is defined by the convolution of x_n with a scaled and translated wavelet, $\psi_0(\eta)$:

$$W_n^X(s) = \left(\frac{\delta t}{s}\right)^{1/2} \sum_{n'}^{N-1} x_{n'} \psi^*[(n-n')\delta t/s] \qquad (13)$$

Where N indicates the length of the time series and the asterisk* is relative to the complex conjugate. The wavelet in its width is ordered by the scale parameter. Whereas, wavelet power spectrum (WPS) is a powerful tool univariate analysis that makes clear the evolution of the variance throughout the time dimension at each frequency as previously reported by Torrence and Compo. (1998). Mathematically, WPS can be described by:

$$|W_n^X(s)|^2 = W_n^X(s)W_n^{X*}(s) \quad (14)$$

Then, the WPS is normalized by the global wavelet spectrum (GWS) which is more suitable measure to describe the variability of the time series in a non-stationary case, by targeting the scale parameter. Hence, interpretation of results becomes easy. It is expressed as:

$$\overline{W_n^2}(s) = \sum_{n=1}^{N} | W_n(s) |^2 \quad (15)$$

While different types of wavelets can be used. The most commonly type of wavelet For a given time series $x_1, x_2, x_3, \ldots, x_n$ of size N, spaced with regular time interval δt, the Morlet wavelet $\psi_0(\eta)$ is a specific wavelet function (with $\omega_0 = 6$), having zero mean and localized simultaneously in time and frequency domain, modulated by a Gaussian envelope. Its choice is crucial, because it is compatible with the criteria required by the CWT, based on maintaining a balance between both dimensions of time and frequency (Torrence and Compo, 1998). It can be expressed as:

$$\psi_0(\eta) = \pi^{-1/4} \exp^{i\omega_0\eta} \exp^{-\eta^2/2} \quad (16)$$

Where η and ω_0 are dimensionless time and the wavenumber respectively.

2.5.2 Cross wavelet transform

In order to explore high common power in both time-frequency domains between two time series X and Y, defined by their corresponding CWTs, $W_n^X(s)$ and $W_n^Y(s)$, $W_n^{XY}(s)$ is obtained by the convolution between CWTs written as:

$$W_n^{XY}(s) = W_n^X(s)W_n^{Y*}(s) \quad (17)$$

However, there are certain drawbacks associated with its use, because it is not normalized as reported by Maraun and Kurth. (2004). For this reason, it was decided that the best method to adopt for the examination of causality is the wavelet transform coherence (WTC) in so far as two variables can be dependent without there being a strong link.

2.5.3 Wavelet transform coherence

According to Torrence and Compo (1998), the magnitude of the covariance between two time-series is given as follows:

$$R_{n,X,Y}^2(s) = \frac{\left|s\left(s^{-1} W_n^{XY}(s)\right)\right|}{s\left(s^{-1} W_n^X(s)\right) . s\left(s^{-1} W_n^Y(s)\right)} \quad (18)$$

Assuming that the time series has a mean power spectrum, Torrence and Compo (1998) proposed to calculate the statistical significance level $SL_n(s)$ of the wavelet coherence as well as its 95% confidence level $CL^{95}_{n,X,Y}(s)$ is estimated using Monte Carlo methods.

2.5.4 Time series similarity based on wavelet coherence

The wavelet coherence provided in the previous section is a two-dimensional surface with high and low values that is hard to interpret. We define the similarity between two time series X and Y as $Sim^{XY}_{wavelet}(s) = \frac{1}{N}\sum_{n=1}^{N}\frac{R^2_{n,X,Y}(s)}{CL^{95}_{n,X,Y}(s)}$

The above definition is a curve of similarity as function of S and is easier to interpret than a wavelet coherence graph (Figure 4). It can easily be extended to n time series by taking the minimum similarity between each pair for each frequency s. In the case of three time series, the formula would be:

$$Sim^{XYZ}_{wavelet}(s) = \min\left(\frac{1}{N}\sum_{n=1}^{N}\frac{R^2_{n,X,Y}(s)}{CL^{95}_{n,X,Y}(s)}, \frac{1}{N}\sum_{n=1}^{N}\frac{R^2_{n,X,Z}(s)}{CL^{95}_{n,X,Z}(s)}, \frac{1}{N}\sum_{n=1}^{N}\frac{R^2_{n,Y,Z}(s)}{CL^{95}_{n,Y,Z}(s)}\right) \quad (19)$$

2.6. Case study

The developed hydrological region delineation method will be applied to the Tensift basin. Unfortunately, the network of climatic and hydrologic stations is usually of low density, so direct information about hydro-climatic conditions at a particular point are usually missing. Daily rainfall time series spanning between 31 and 48 years from 11 hydro-meteorological stations were obtained from the Tensift River Basin Agency (Figure 1). The majority of stations are concentrated in the South-East close to the Atlas Mountains, while vast areas in the North and South-West are not gauged. The statistical characteristics of rainfall time series used in this paper are summarized in Table 1. The highest average total annual rainfall is observed, within the available stations, in Aghbalou in the Atlas Mountains (534.4 mm), while the lowest one is observed in the desert at Abadela (172.2 mm) in the way to the outlet of the basin.

3. Results and discussion

The methodology was applied to the 11 time-series of annual precipitation, using M=1000, and a minimum segment length $l_{\min} = 10$ given that the length of the time series varies from station to station, the data of the common period (1986-2017). The multiple changepoint detection is also applied.

3.1. Delineated regions

A total of 4 regions were detected, containing 3, 6, 1 and 1 stations (Table 2). The local and regional trends are presented in Figure 5. The regional and local probabilities of changes are shown in Figure 6, along with the conditional probability of the position of the first changepoint. It is interesting to note that no regional changepoint was detected in regions 1 and 2, and that a single regional changepoint is detected in regions 3 and 4. No local changepoint was detected at any station in regions 1 and 2, and only one changepoint was detected at each station in region 3 and 4. The position of the local and regional changepoints are very close in region 3 and 4, as the most probable date of change is 1998 for Nkouris and 1996 for Taferiat. For these particular regions, the local and regional trends are the same for all stations, suggesting that the regions are homogeneous. It is however, worth mentioning that the directions of the regional trends are not the same for the two stations in regions 3 and 4 (the two segments of the local and regional trend at Taferiat are all decreasing, while Nkouris has a decreasing trend followed by an increasing trend). Another interesting fact is that two regions were delineated despite both having zero changepoint. Seidou and Ouarda. (2007) assume there is a changepoint only when the probability of not having a changepoint is below 0.5 (or the probability of having one or more changepoints is above 0.5). Regions 1 and 2 have probabilities of having zero changepoint of 0.9145 and 0.8294. However, the regional conditional probability of the position of the changepoints varies between the regions (Figure 6). Figure 7 shows the homogenous groups of regions based on the probabilities of changes and on the conditional probability of the locations of the change. Table 3 presents regions based on the conditional probabilities of the positions of the changepoints. As can be seen, even though stations belong to the same region, there is a conditional probability difference at the date of change between stations. It should be noted that only 3 stations in region 1 show a negative difference between the local and regional conditional probability. These differences led to the delineation of four different regions (Fig. 7) even though no changepoints are expected. It is also worth noticing that while regional and local probabilities of changes and conditional probabilities of changepoint locations are different at each station, the difference between the local and regional conditional probability of the positions of the changepoints is very small, once again suggesting that the delineated regions are homogeneous. As Fig. 8 depicts, a large difference between the local and regional trend at a particular station may be a signal of an inhomogeneity at that station, which means that the absence of abrupt changes in local trend does not imply the verification of this conclusion in a regional one. Consequently, the regional trend could possibly present not only jumps in the average but also in the trend orientation.

3.2. Validation of the delineated regions

Globally, the findings of this study mirror those of Ahattab et al. (2015) and Salama. (2010) who detected 4 regions in the Tensift watershed by applying PCA to monthly precipitation records. The regions detected by each study are presented in Table 4. All methods place the station of Nkouris apart from the others in group 3; 66% (2 out of 3) of the stations in group 1 are the same in the two papers. The merger of groups 2 and 4 contains 7 stations according to each study, and 6 out of the 7 are similar. The main difference is in this study Marrakech was found to be in group 2 and Aghbalou in group 1; Ahattab et al., 2015 classified Aghbalou in group 2 and Marrakech in group 1.

The similarity measure based on wavelet coherence (equation 18) is used to show that the inclusion of any station outside a particular region would significantly reduce the similarity within the region. The similarity (called here intra-region coherence) of each of the 4 delineated regions is shown in Fig. 9. We found that the intra-region coherence is maximal for the 8-10 periods, consistent with the minimum data segment of 10 years used in the Bayesian model. For each of the region, the intra-region coherence is recalculated after the addition of an external station. For instance, the first panel of Figure 10 shows how the intra-region coherence decreases when Marrakech, which belongs to region 2, is added to region 1. The second (resp. third, fourth) shows the decrease of intra-region coherence when Chichaoua (resp. Aghbalou, Abadela) are added to region 2 (resp. region 3, region 4). In all cases, it can be seen that the intra-region coherence drops when an outside station is added to the region, and that the magnitude of the changes varies with the frequency and the region. In this particular example, the decrease in coherence is the most severe in region 3, and the least severe in region 2. The loss of coherence of all regions due to the inclusion of all possible out-of-region stations is presented in Fig. 11. Once again, the loss of coherence is systematically more severe in region 3 and less severe in region 2. This result suggests that region 3 is very different from the rest of stations in the study area, while region 2 has a lot of similarity (at least for this particular measure) with the rest of the study area. The findings from this study make several contributions to the literature.

4. Conclusions and further research

The main goal of the current study was to extend the multiple changepoint detection procedure Seidou and Ouarda (2007) to delineate homogenous hydrological regions. A similarity measure based on the positions of the changepoints and an algorithm for the delineation of homogeneous region were proposed. The method was applied to the Tensift watershed in Morocco, and it was shown that the results are like those of previous studies of the same watershed. Wavelet coherence was used to further confirm the homogeneity of the detected regions. The method presented in this paper relaxes the assumption of a stable and linear relationship between time series at stations within a region and therefore is more general than PCA. These two assumptions are convenient from a mathematical point of view but have no physical basis. The climate system is highly complex and non-linear, and relationships between variables are nonlinear and nonstationary by default. The application of the new method should lead to larger homogeneous regions and uncover hidden relationships between stations. While the method application is limited to one watershed in a semiarid context, there is no limitation to its application to other hydroclimatic regions. The only requirement is the existence of annual time series of the variables of interest. Despite the fact that the method presented in this paper does not provide estimates of hydrological variables, it can improve water management by improving homogeneous regions delineation. A better delineation of homogeneous regions is likely to improve estimation results as the quality of the regional estimation of hydrological variables (e.g., flood quantiles) is directly linked to the homogeneity of the regions.

In this paper, the similarity measure only considers the positions of changepoints and acknowledges the similarity only when the changepoints at the two stations occur at the same date. Possible improvements include a) consider the direction of the trend between changepoints in the similarity measure, b) find a way to account for changepoints which are close but on different dates, and c) combine this similarity measure with a number of other physiographic and climatic similarity measures that are for the instance based on river network configuration.

For further research, it is recommended to apply these approaches to watersheds elsewhere in Morocco and in the world to assess their effectiveness for ensuring sustainable territorial management. Future research can also focus on the use of the identified regions to transfer information within regions and to build a full regional frequency analysis procedure that can be used for quantile estimation at ungauged locations. This can be achieved by combining the regional delineation procedure proposed in the present work with a regional estimation approach. Future research can also look into the potential advantages of using time series of variable lengths in the similarity measure. . The extension of the approach to the multivariate case can also be considered.

Acknowledgments

The authors are grateful to the Tensift Basin River Agency (ABHT, Marrakech, Morocco), for providing us the data throughout the study. The authors also would like to thank the editor András Bárdossy, the associate editor Roger Moussa and the three anonymous reviewers' constructive comments to raise the quality of the manuscript.

References

Agarwal, A., Maheswaran, R., Sehgal, V., Khosa, R., Sivakumar, B., Bernhofer, C., 2016. Hydrologic regionalization using wavelet-based multiscale entropy method. J. Hydrol. 538, 22-32.

Ahattab, J., Serhir, N., Lakhal, E.K., 2015. Vers l'élaboration d'un système d'aide à la décision pour le choix des méthodes d'estimation des débits max des crues : réadaptation aux données hydrologiques récentes. La Houille Blanche (1), 63-70, doi: 10.1051/1hb/2015008.

Alila, Y., 1999. A hierarchical approach for the regionalization of precipitation annual maxima in Canada. J. Geophys. Res. 104(D24), 31,645-31,655.

Alila, Y., 2000. Regional rainfall depth-duration-frequency equations for Canada. Water Resour. Res. 36(7), 1767-1778.

Basilevsky, A., 1994. Statistical Factor Analysis and related methods: Theory and Applications. Wiley Series in Probability and Mathematical Statistics.

Chang, T.P., Liu, F.J., Ko, H.H., Huang, M.C., 2017. Oscillation characteristic study of wind speed, global solar radiation and air temperature using wavelet analysis. App. Energy, 190, 650-657.

Chokmani, K., Ouarda, T.B.M.J., 2004. Physiographical space-based kriging for regional flood frequency estimation at ungauged sites. Water Resour. Res, Vol. 40, W12514, doi: 10.1029/2003WR002983.

Durrans, S.R., Tomic, S., 1996. Regionalisation of low-flow frequency estimates: an Alabama case study. Water Resour. Bull. 32 (1), 23-37.

El Alaoui El Fels, A., Saidi, M.E.M., Bouiji, A., Benrhanem, M., 2020. Rainfall regionalization and variability of extreme precipitation using artificial neural networks: a case study from western central Morocco. Journal of Water and Climate Change. https://doi.org/10.2166/wcc.2020.217.

Ehsanzadeh, E., Ouarda, T.B.M.J., Saley, M.B. 2011. A simultaneous analysis of gradual and abrupt changes in Canadian low streamflows. Hydrol. Process., 25 (5): 727-739. DOI:10.1002/hyp.7861.

Farquharson, F.A.K., Meigh, G.R., Sutcliffe, J.V., 1992. Regional flood frequency analysis in arid and semi-arid areas. J. Hydrol. 138, 487-501.

Fearnhead, P., 2006. Exact and efficient Bayesian inference for multiple changepoint problems. Stat Comput. 16, 203-213.

Feng, S., Fu, Q., 2013. Expansion of drylands under a climate. Atmos. Chem. Phys, 13(19), 10081-10094. doi:10.5194/acp-13-10081-2013.

GREHYS, 1996b. Intercomparaison of flood frequency procedures for Canadian rivers. J. Hydrol. 186 (164), 85-103.

Grinsted, A., Moore, G.C., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. Non-linear Process. Geophys. 11: 561-566.

Grise, K. M., & Davis, S. M. (2020). Hadley cell expansion in CMIP6 models. Atmospheric Chemistry and Physics, 20(9), 5249–5268. https://doi.org/10.5194/acp-20-5249-2020.

Han, X., Ouarda, T.B.M.J., Rahman, A., Haddad, K., Mehrotra, R, Sharma, A., 2020. A Network Approach for Delineating Homogeneous Regions in Regional Flood Frequency Analysis. Water Resour. Res. 56(3): e2019WR025910. doi:10.1029/2019wr025910.

Heuvelmans, G., Muys, B., Feyen, J., 2006. Regionalisation of the parameters of a hydrological model: Comparaison of linear regression models with artificial neural nets. J. Hydrol. 319, 245-265.

Hu, Y., & Fu, Q. (2007). Observed poleward expansion of the Hadley circulation since 1979. *Atmospheric Chemistry and Physics*, *7*(19), 5229–5236. https://doi.org/10.5194/acp-7-5229-2007.

Hundecha, Y., Ouarda, T.B.M.J., Bardossy, A., 2008. Regional estimation of parameters of a rainfall-runoff model at ungauged watersheds using the spatial structures of the parameters within a canonical physiographic-climatic space. Water Resour. Res. Vol. 44, W01427, doi: 10.1029/2006WR005439.

Lu, J., Vecchi, G.A., Reichler, T., 2007. Expansion of the Hadley cell under global warming. Geophys. Res. Lett, 34 (6). doi: 10.1029/2006gl028443.

Masson-Delmote, V., Zhai, P., Pirani, A., Connors, S.L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M.I., Huang, M., Laitzell, K., Lonnoy, E., Matthews, J.B.R., Maycock, T.K., Waterfield, T., Yelekçi, O., Yu, R., Zhou, B (eds)., 2021. IPCC Climate Change: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

Mahé, G., Rouché, N., Dieulin, C., Boyer, J.F., Ibrahim, B., Cres, A., Servat, E., Valton, C., Paturel, J.E., 2012. Carte des pluies annuelles en Afrique. IRD, Bondy, France.

Keener, V.W., Feyereisen, G.W., Lall, U., Jones, J.W., Bosh, D.D., Lowrance, R., 2010. El-Nino/Southern Oscillation (ENSO) influences on monthly NO_3 load and concentration, stream flow and precipitation in the Little River Watershed, Tifton, Georgia (GA). J. Hydrol. 381, 352-363.

Maraun, D., Kurths, J., 2004. Cross wavelet analysis: significance testing and pitfalls. Non-linear Process. Geophys, 11: 505-514.

Mwakalila, S., 2003. Estimation of stream flows of ungauged catchments for river basin management.Phys. Chem. Earth 28, 935-942.

Naizghi, M.S., Ouarda, T.B.M.J., 2016. Teleconnections and analysis of long-term wind speed variability in the UAE. Int. J. Climatol. doi: 10.1002/joc.4700.

Ouachani, R., Bargaoui, Z., Ouarda, T., 2011. Power of teleconnection patterns on precipitation and streamflow variability of upper Medjerda Basin. Int. J. Climatol. 33(1), 58-76, doi: 10. 1002/joc.3407.

Ouarda, T.B.M.J., Charron, C., Niranjan Kumar, K., Marpu, P., Ghedira, H., Molini, A., Khayal, I. 2014. Evolution of rainfall regime in the UAE, Journal of Hydrology, 514: 258–270. doi:10.1016/j.jhydrol.2014.04.032.

Ouarda, T.B.M.J., Girard, C., Cavadia, G.S., Bobée, B., 2001. Regional flood frequency estimation with canonical correlation analysis. J. Hydrol. 254, 157-173.

Ouarda, T.B.M.J., Haché, M., Bruneau, P., Bobée, B., 2000. Regional flood Peak and Volume estimation in northern Canadian basin. J. Cold. Reg. Eng., 14, 176-190.

Ouarda, T.B.M.J., Lang, M., Bobée, B., Bernier, J., Bois, P., 1999. Analysis of regional flood models utilized in France and Québec. Rev. Sci. Eau, 12(1), 155.

Özger, M., Mishra, A.K., Singh, V.P., 2009. Low frequency drought variability associated with climate indices. J. Hydrol. 364(1), 152-162.

Pandey, G.R., Nguyen, V.-T.-V., 1999. A comparative study of regression based methods in regional flood frequency analysis. J. Hydrol. 225, 92-101.

Peleg, N., Bartov, M., & Morin, E. (2015). CMIP5-predicted climate shifts over the East Mediterranean: Implications for the transition region between Mediterranean and semi-arid climates: CMIP5-PREDICTED CLIMATE SHIFTS OVER THE EM. International Journal of Climatology, 35(8), 2144–2153. https://doi.org/10.1002/joc.4114

Rosbjerg, D., Madsen, H., 1994. Uncertainty measures of regional flood frequency estimators. J. Hydrol. 167, 209-224.

Saidi, M.E.M., Boukrim, S., Fniguire, F., Ramromi, A., 2012. LES ECOULEMENTS SUPERFICIELS SUR LE HAUT ATLAS DE MARRAKECH: CAS DES DEBITS EXTREMES. Larhyss. J, ISSN 1112-3680, n° 10, Mars, pp. 75-90.

Saidi, M.E.M., Daoudi, L., Aresmouk, M.E.H., Blali, A., 2003. Rôle du milieu physique dans l'amplification des crues en milieu montagnard : exemple de la crue du 17 août 1995 dans la vallée de l'Ourika (Haut-Atlas, Maroc)*. Séch; 14(2), 1-8.

Saidi, M.E.M., Daoudi, L., Aresmouk, M.E.H., Fniguire, F., Boukrim, Siham., 2010. Les crues de l'Oued Ourika (Haut Atlas, Maroc): Evénements extrêmes en contexte montagnard semi-aride. Comunicaçoes Geologicas, t.97, pp. 113-128.

Sang, Y., 2013. Improved wavelet modeling framework for hydrologic time series forecasting. Water Resour Manag. 27 (8), 2807-2821.

Salama, H., 2010. Régionalisation pluviométrique du bassin versant Tensift. Larhyss. J, ISSN 1112-3680, n° 09, Septembre, pp. 111-119.

Santos, C.A.G., Kisi, O., da Silva, R.M., Zounemat-Kermani, M., 2018. Wavelet-based variability on streamflow at 40-year timescale in the Black Sea region of Turkey. Arab. J. Geosci, 11: 169. https//doi.org/10.1007/s12517-018-3514-6.

Schmidt, D.F., Grise, K.M., 2017. The response of Local Precipitation and Sea Level Pressure to Hadley Cell Expansion. Geophys. Res. Lett, 44 (20), 10, 573-10,582. doi: 10.1002/2017gl075380.

Sefton, C.E.M., Howarth, S.M., 1998. Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales. J. Hydrol. 211, 1-16.

Seidou, O., Asselin, J.J., Ouarda, T.B.M.J., 2007. Bayesian multivariate linear regression with application to changepoint models in hydrometeorological variables. Water Resour. Res, Vol. 43, W08401, doi: 10.1029/2005WR004835.

Seidou, O., Ouarda, T.B.M.J., 2007. Recusion-based multiple changepoint detection in multiple linear regression and application to river streamflows. Water Resour. Res, Vol. 43, W07404, doi: 10.1029/2006WR005021.

Seidou, O., Ouarda, T.B.M.J., Barbet, M., Bruneau, P., Bobée, B. 2006. A parametric Bayesian combination of local and regional information in flood frequency analysis, Water Resource Res. 42, W11408, doi: 10.1029/2005WR004397, 1-21.

Thiombiano, A.N., El Adlouni, S., St-Hillaire, A., Ouarda, T.B.M.J., El-Jabi, N., 2016. Nonstatioanry frequency analysis of extreme daily precipitation amounts in southeastern Canada using a peaks-over –threshold approach. Theor. Appl.Climatol, doi: 10.1007/s00704-016-1789-7.

Torrence, C., Compo, GP., 1998. A pratical guide to wavelet analysis. Bull. Am. Meteorol. Soc. 79(1), 61-78.

Xia, Y., Hu, Y., & Liu, J. (2020). Comparison of trends in the Hadley circulation between CMIP6 and CMIP5. Science Bulletin, 65(19), 1667–1674. https://doi.org/10.1016/j.scib.2020.06.011.

List of tables

Table 1. Summary of rainfall stations and basic statistics of total annual rainfall time series. Minimum, maximum, mean, standard deviation (SD), coefficient of variation (C_V), coefficient of skewness (C_S) and coefficient of kurtosis (C_K)

Table 2. Stations by region with the number of change and the corresponding posterior probability of change

Table 3. Regions based on conditional probability of the positions of the changepoints

Table 4. Comparison of the regions detected in this paper, in Ahattab et al. (2015) and in Salama (2010)

List of figures

Figure 1. Geographical location of the Tensift watershed

Figure 2. Illustration of changepoint scatter schemes and derived changepoint location probability

Figure 3. Similarity calculation based on changepoints locations

Figure 4. Similarity calculation based on wavelet coherence

Figure 5. Detection of trend changes for total annual rainfall for local and regional model. a) Region 1, b) Region 2 and c) Region 3,4

Figure 6. Posterior probability of the number of changepoint, and conditional probability of the date of change for local and regional model. a) Region 1, b) Region 2 and c) Region 3, 4

Figure 7. Homogeneous regions detected in non-linear and non-stationary framework, based on probability of the number of changepoint and on the conditional probability of the locations of the change

Figure 8. Observed changes between local and regional trend at a monthly scale

Figure 9. Minimum coherence intra-region

Figure 10. Outside regions effect on the inside regions coherence

Figure 11. Loss of coherence due to the inclusion of an out-of-region station

Table 1. Summary of rainfall stations and basic statistics of total annual rainfall time series. Minimum, maximum, mean, standard deviation (SD), coefficient of variation (C_V), coefficient of skewness (C_S) and coefficient of kurtosis (C_K).

Stations	Latitude (N)	Longitude (W)	Height (m)	Period record	Years	Min (mm)	Max (mm)	Mean (mm)	SD (mm)	C_V	C_S	C_K
Sidi Rhal	31°37	07°27	690	1970 - 2017	48	176.7	648.2	345.4	107.0	0.3	0.6	0.3
Taferiat	31°32	07°35	760	1987 - 2017	31	185.5	913.4	347.8	24.4	0.1	2.4	9.3
Tahnaout	31°16	07°57	925	1971 - 2017	47	192.2	647.9	372.9	103.3	0.3	0.8	0.3
Aghbalou	31°18	07°45	1070	1970 - 2017	48	276.9	1053.5	534.4	163.8	0.3	0.4	1.2
Nkouris	31°30	08°70	1100	1974 - 2017	44	46.1	443.5	229.3	101.0	0.4	0.7	-0.2
Imin El Hammam	31°13	08°06	770	1970 - 2017	48	161.6	694	380.1	116.4	0.3	0.4	0.2
Chichaoua	31°33	08°45	340	1971 - 2017	47	56.8	298.7	181.9	67.6	0.4	0.1	-1.0
Abadela	31°42	08°33	250	1970 - 2017	48	60.0	329.6	172.2	66.9	0.4	0.4	-0.5
Adamna	31°32	09°40	158	1977 - 2017	41	115.4	757.6	328.3	154.2	0.5	1.1	0.8
Marrakech	31°61	08°01	460	1973 - 2017	45	78.7	351.9	215.6	76.7	0.4	0.1	-1.0
Talmest	31°51	09°16	53	1985 - 2017	33	101.4	604.7	272.0	128.7	0.5	0.8	-0.2

Table 2. Stations by region with the number of change and the corresponding posterior probability of change

Variable	Regions	Stations	Number of detected change	Posterior probability of change
Total rainfall	Region 1	Aghbalou Abadela Chichaoua	0	0.9145
	Region 2	Imin El Hammam Sidi Rhal Tahnaout Marrakech Adamna Talmest	0	0.8294
	Region 3	Nkouris	1	0.6456
	Region 4	Taferiat	1	0.9245

Table 3. Regions based on conditional probability of the locations of the changepoints

	Conditional probability of the positions of the changepoints			
Regions	Stations	Local	Regional	Difference
1	Aghbalou	0.0106	0.0225	-0.0119
	Abadela	0.0322	0.0225	0.0097
	Chichaoua	0.0240	0.0225	0.0015
2	Imin El Hammam	0.0276	0.0609	-0.0333
	Sidi Rhal	0.0626	0.0609	0.0017
	Tahnaout	0.0408	0.0609	-0.0201
	Marrakech	0.0860	0.0609	0.0251
	Adamna	0.0720	0.0609	0.0111
	Talmest	0.0852	0.0609	0.0243
3	Nkouris	0.3808	0.3803	0.0005
4	Taferiat	0.4332	0.4206	0.0126

Table 4. Comparison of the regions detected in this paper, in Ahattab et al. (2015) and in Salama (2010)

	This paper	Ahattab et al. (2015)	Salama (2010)
Group 1	Aghbalou, Abadela and Chichaoua	Abadela, Chichaoua and Marrakech	Abadela, Chichaoua, Marrakech and Sidi Rhal
Group 2	Imin El Hammam, Sidi Rhal, Tahnaout, Marrakech, Adamna and Talmest	Sidi Rhal, Taferiat, Aghbalou, Tahnaout, and Imin El Hammam	Talmest and Adamna
Group 3	Nkouris	Nkouris	Nkouris
Group 4	Taferiat	Talmest and Adamna	Taferiat, Aghbalou, Tahnaout and Imin El Hammam

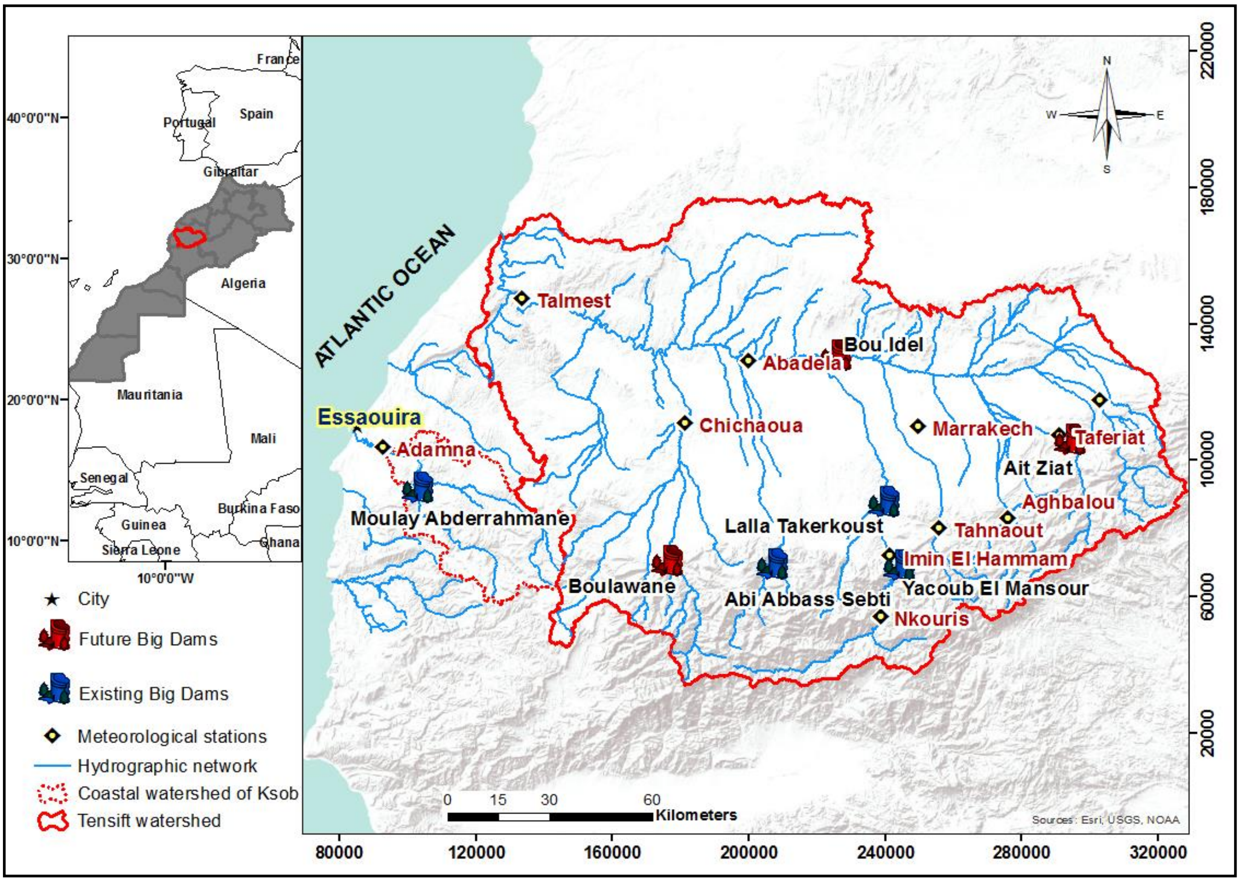


Figure 1.

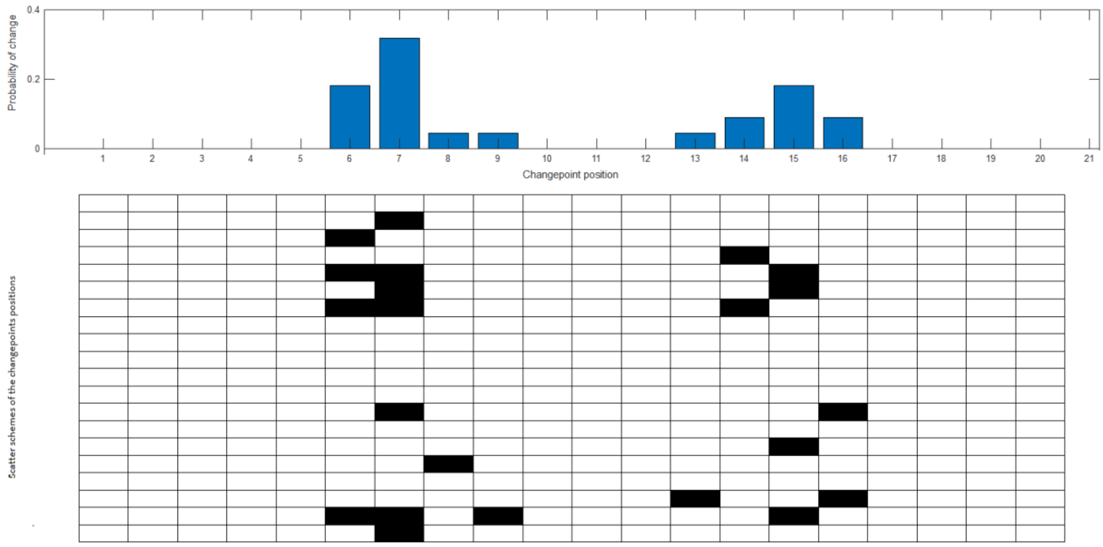


Figure 2.

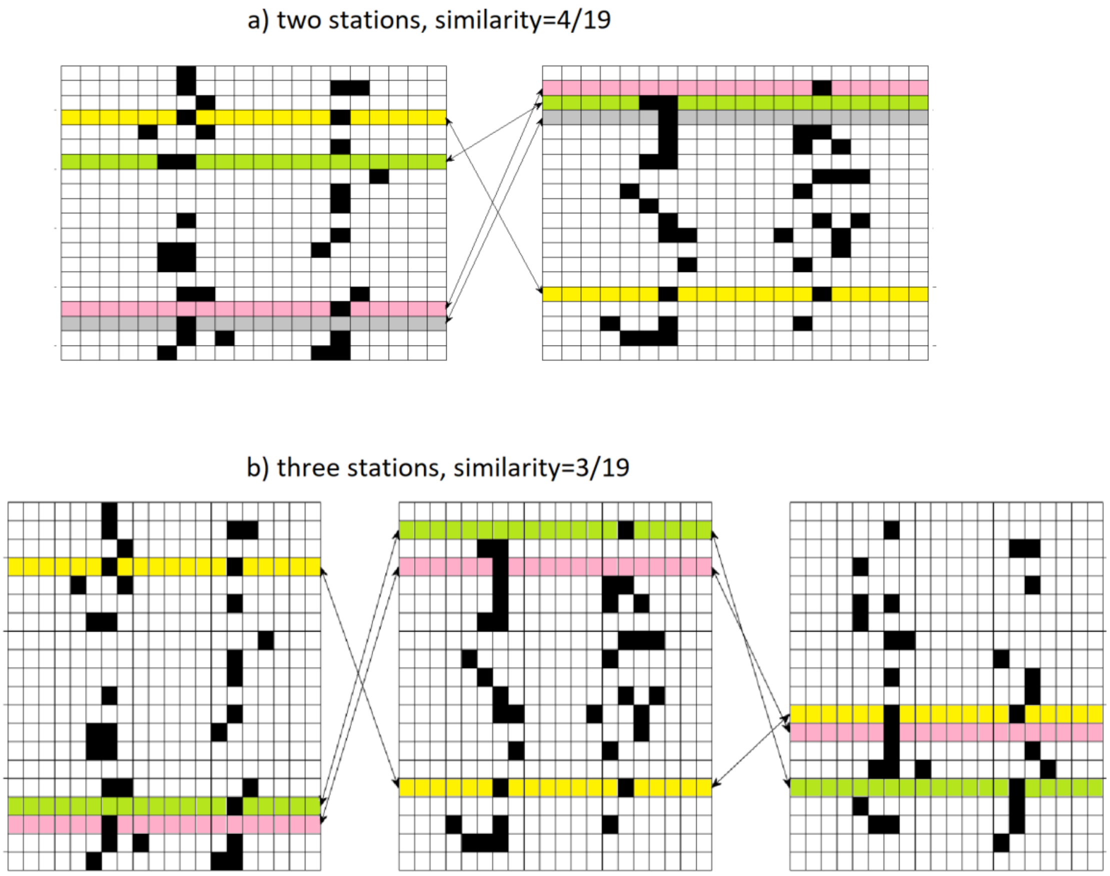


Figure 3.

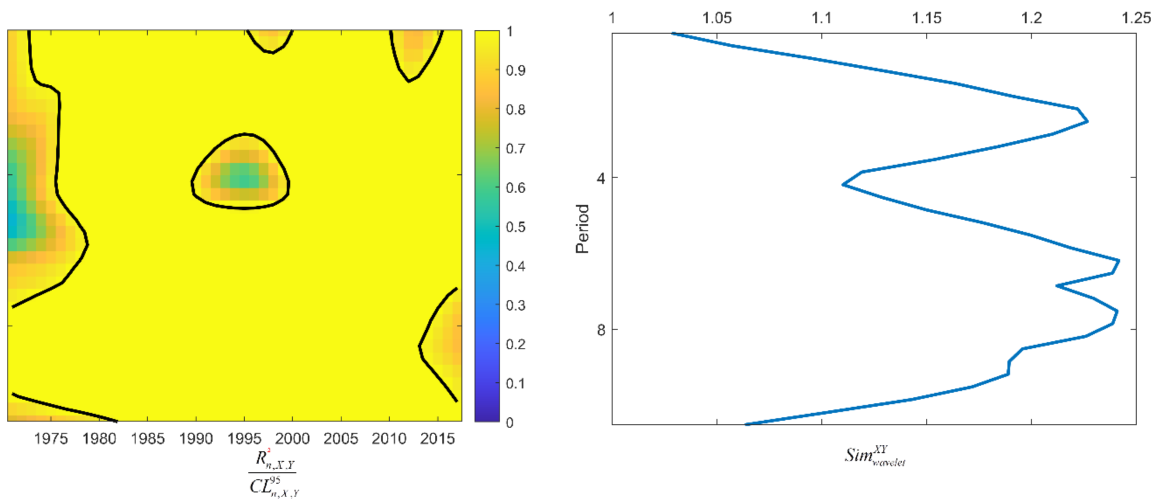


Figure 4.

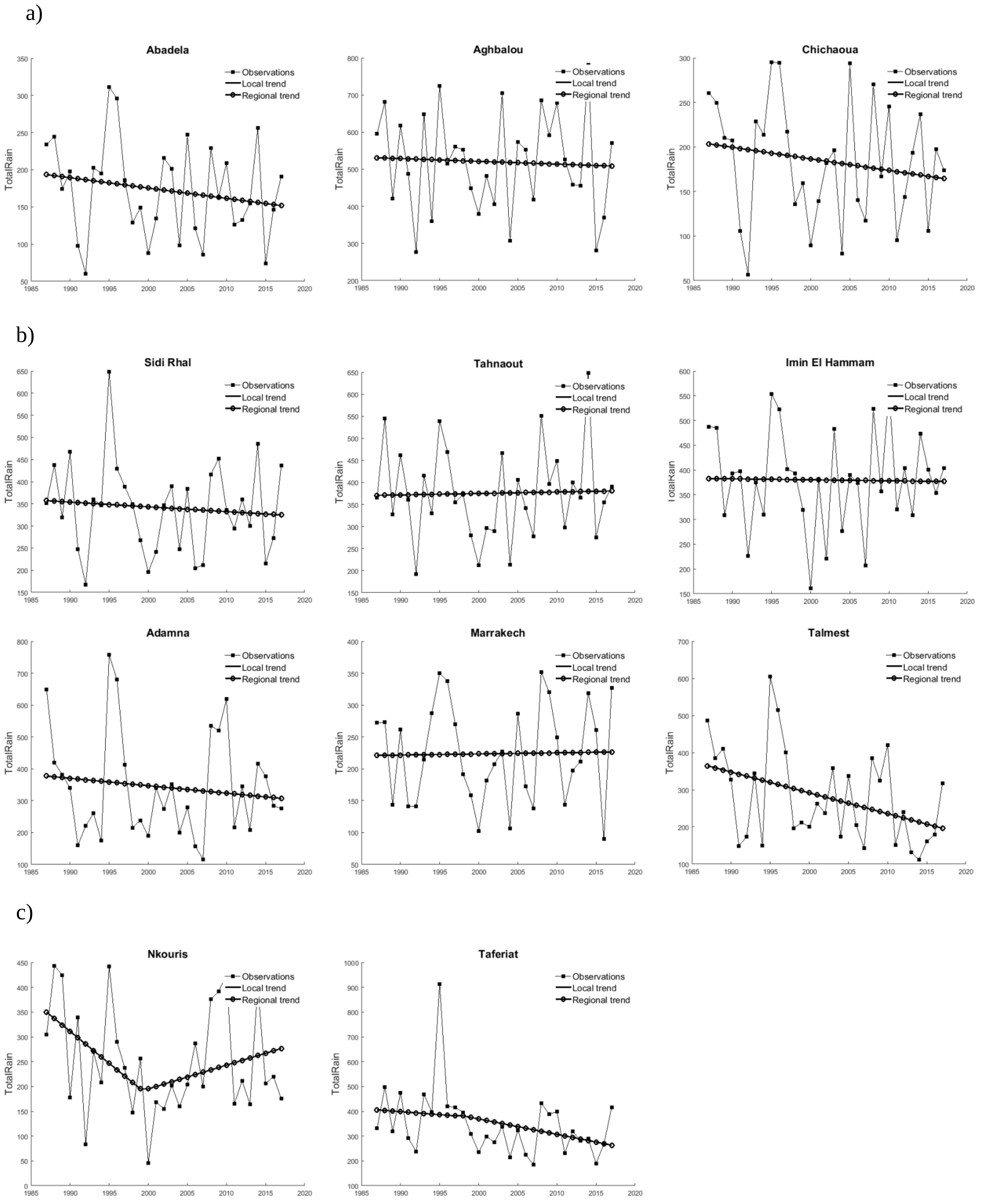
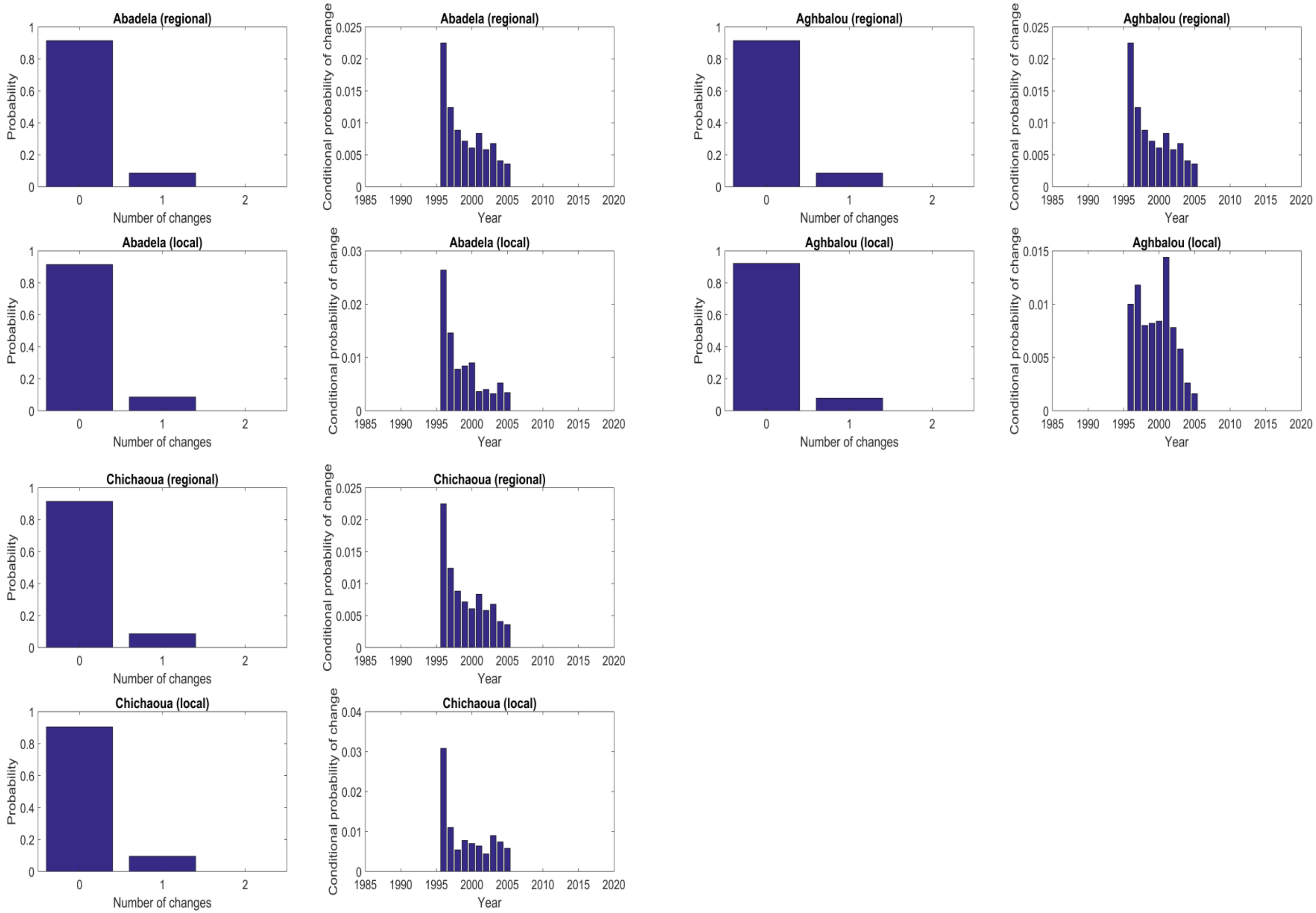

a)
Abadela
Aghbalou
Chichaoua
b)
Sidi Rhal
Tahnaout
Imin El Hammam
Adamna
Marrakech
Talmest
c)
Nkouris
Taferiat
Observations
Local trend
Regional trend
TotalRain

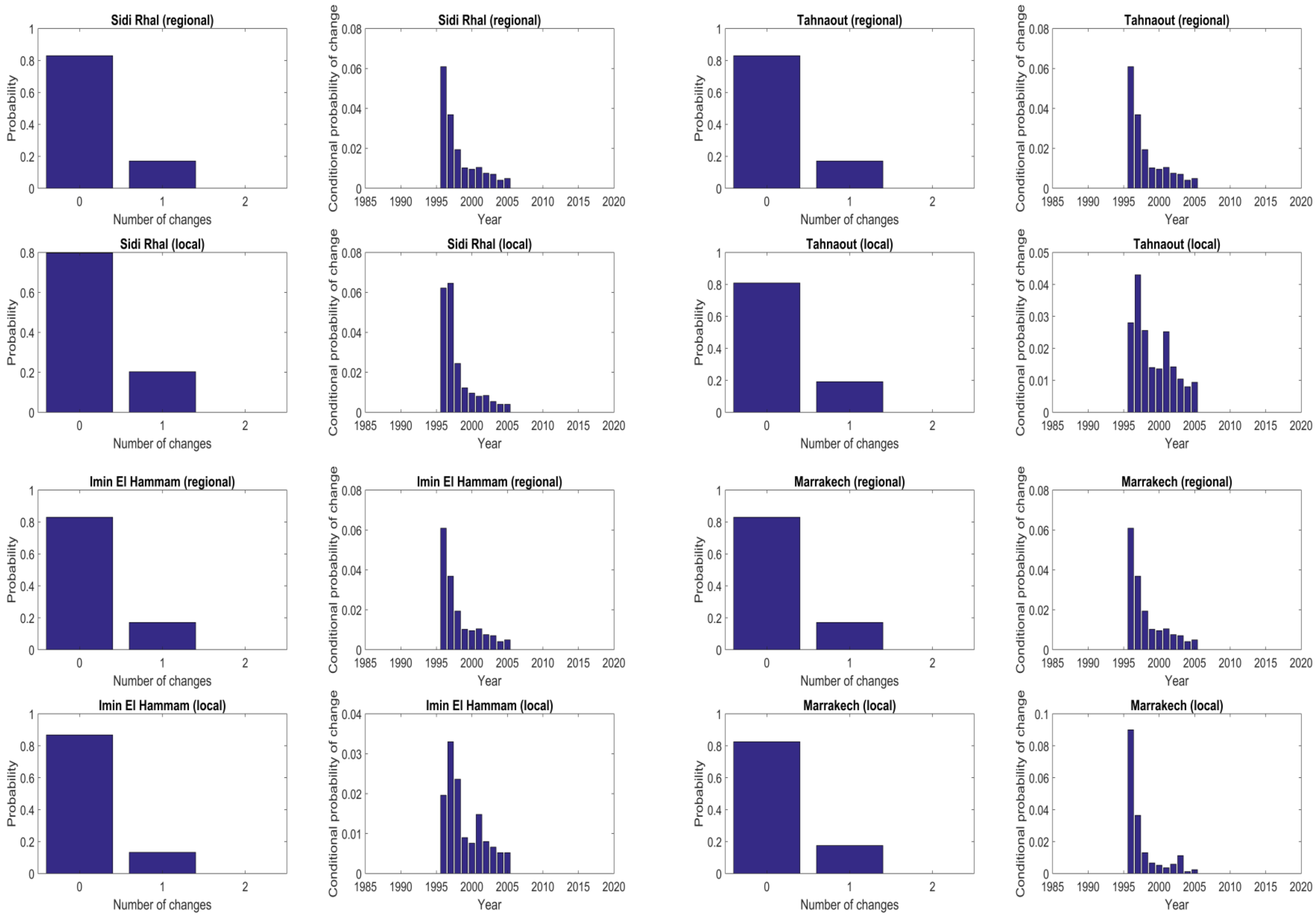
Figure 5.

a)



Abadela (regional)
Abadela (local)
Aghbalou (regional)
Aghbalou (local)
Chichaoua (regional)
Chichaoua (local)
Probability
Number of changes
Conditional probability of change
Year

b)


Sidi Rhal (regional)
Sidi Rhal (local)
Imin El Hammam (regional)
Imin El Hammam (local)
Tahnaout (regional)
Tahnaout (local)
Marrakech (regional)
Marrakech (local)
Probability
Number of changes
Conditional probability of change
Year

c)

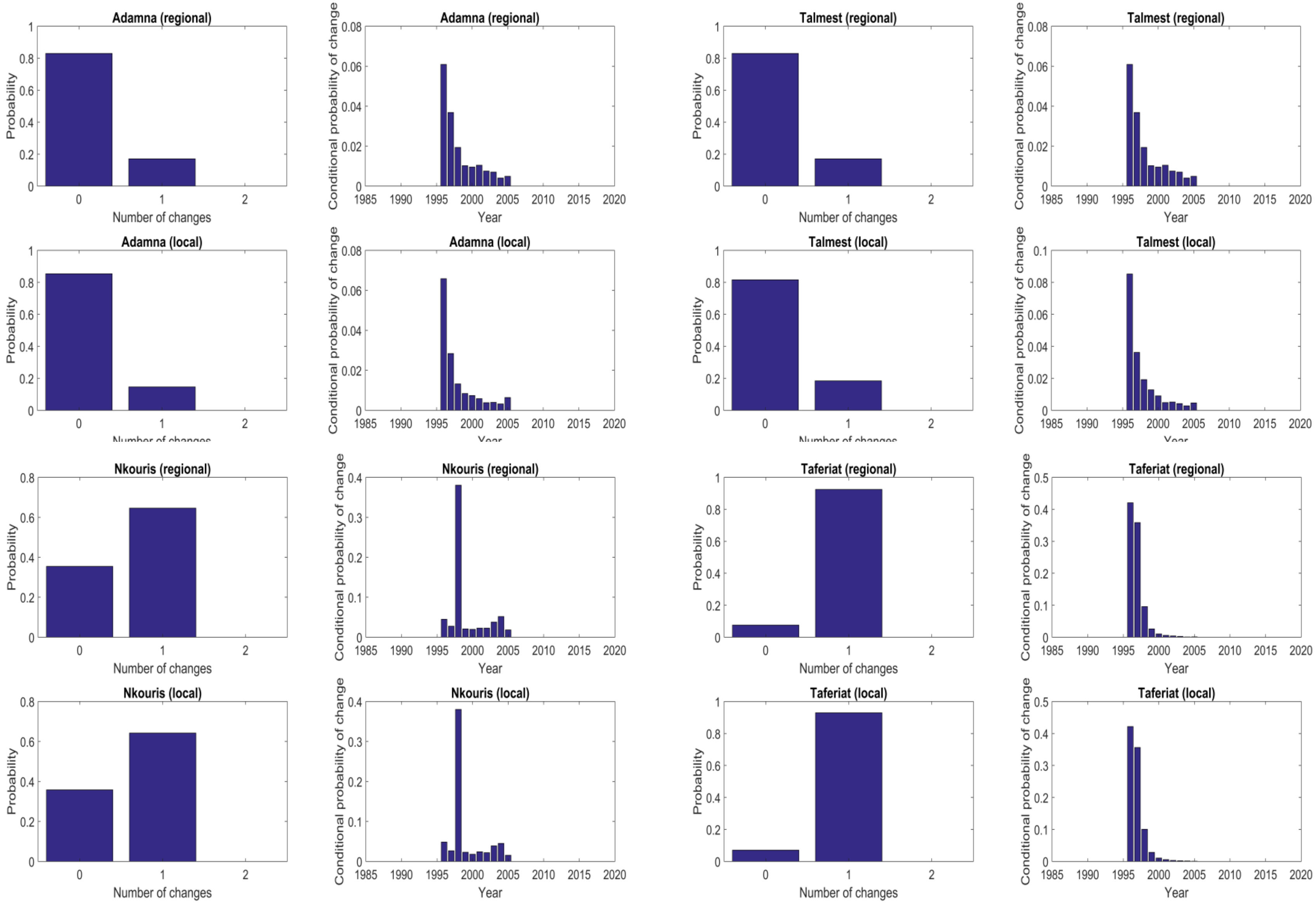

Adamna (regional)
Adamna (local)
Talmest (regional)
Talmest (local)
Nkouris (regional)
Nkouris (local)
Taferiat (regional)
Taferiat (local)
Probability
Number of changes
Conditional probability of change
Year

Figure 6.

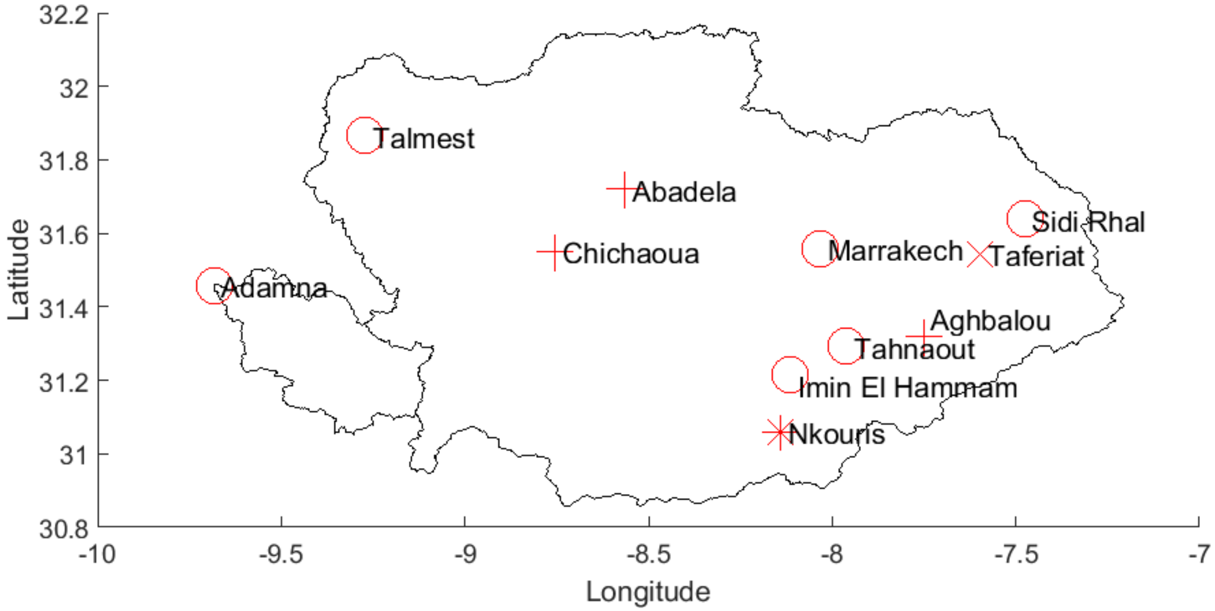

32.2
32
31.8
31.6
31.4
31.2
31
30.8
Latitude
-10
-9.5
-9
-8.5
-8
-7.5
-7
Longitude
Talmest
Abadela
Sidi-Rhal
Chichaoua
Marrakech
Taferiat
Adamna
Aghbalou
Tahnaout
Imin El Hammam
Nkouris

Figure 7.

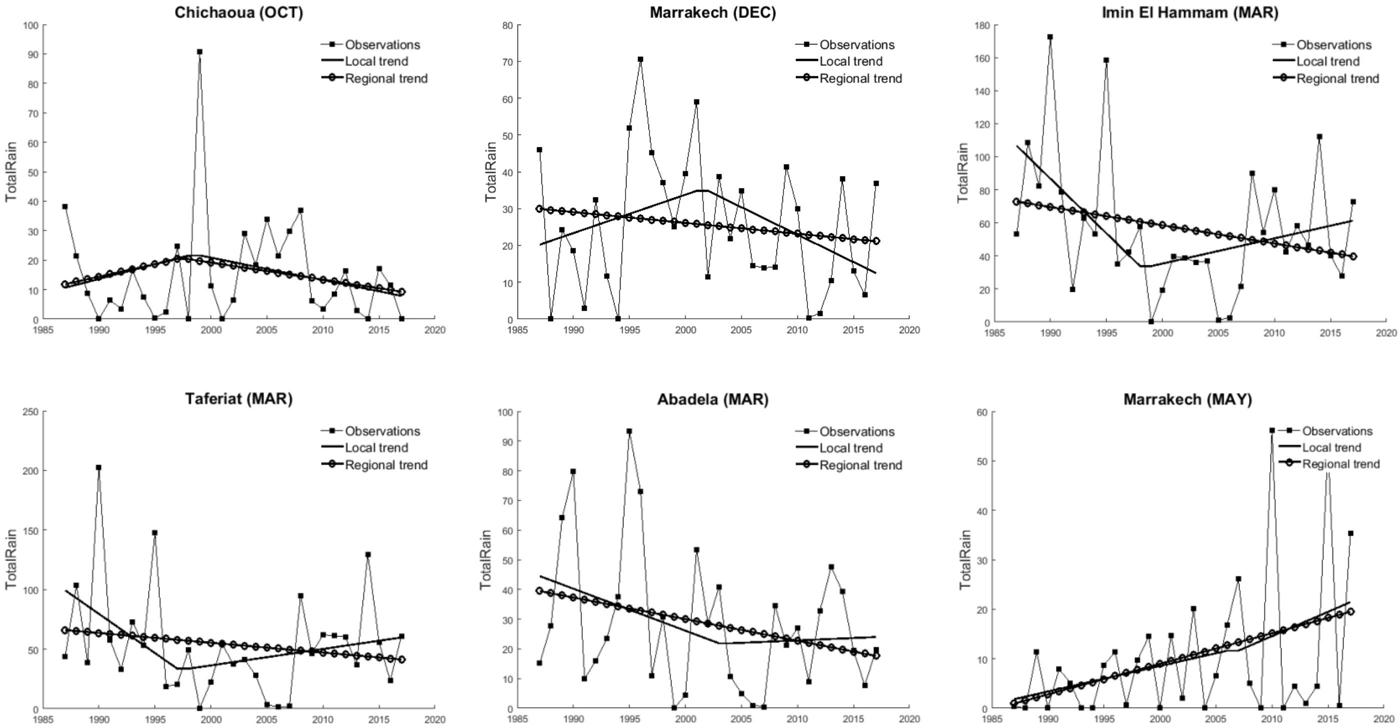

Chichaoua (OCT)
Marrakech (DEC)
Imin El Hammam (MAR)
Taferiat (MAR)
Abadela (MAR)
Marrakech (MAY)
Observations
Local trend
Regional trend
TotalRain

Figure 8.

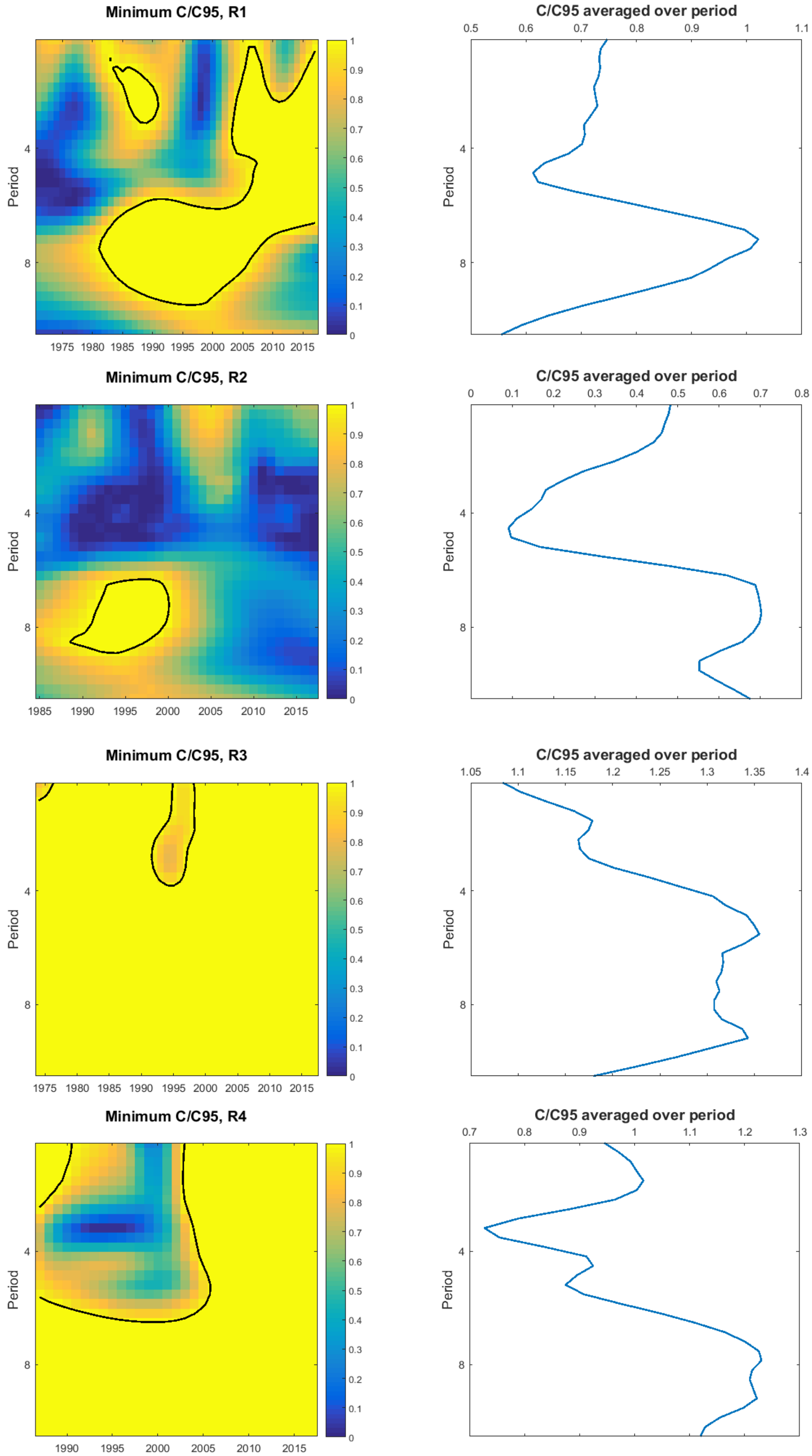
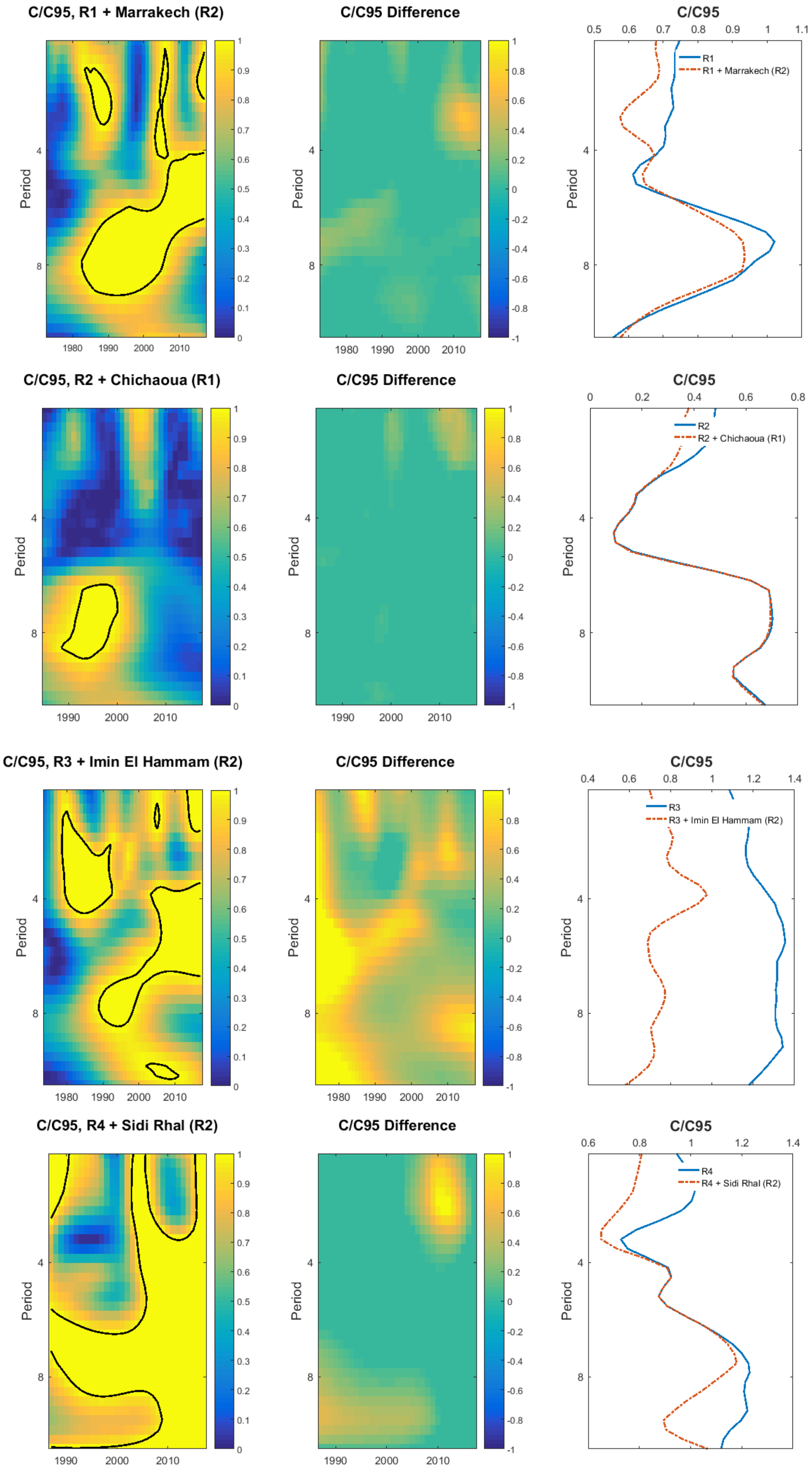

Minimum C/C95, R1
C/C95 averaged over period
Minimum C/C95, R2
C/C95 averaged over period
Minimum C/C95, R3
C/C95 averaged over period
Minimum C/C95, R4
C/C95 averaged over period
Period

Figure 9.



C/C95, R1 + Marrakech (R2)
C/C95 Difference
C/C95
R1
R1 + Marrakech (R2)
C/C95, R2 + Chichaoua (R1)
C/C95 Difference
C/C95
R2
R2 + Chichaoua (R1)
C/C95, R3 + Imin El Hammam (R2)
C/C95 Difference
C/C95
R3
R3 + Imin El Hammam (R2)
C/C95, R4 + Sidi Rhal (R2)
C/C95 Difference
C/C95
R4
R4 + Sidi Rhal (R2)
Period

Figure 10.

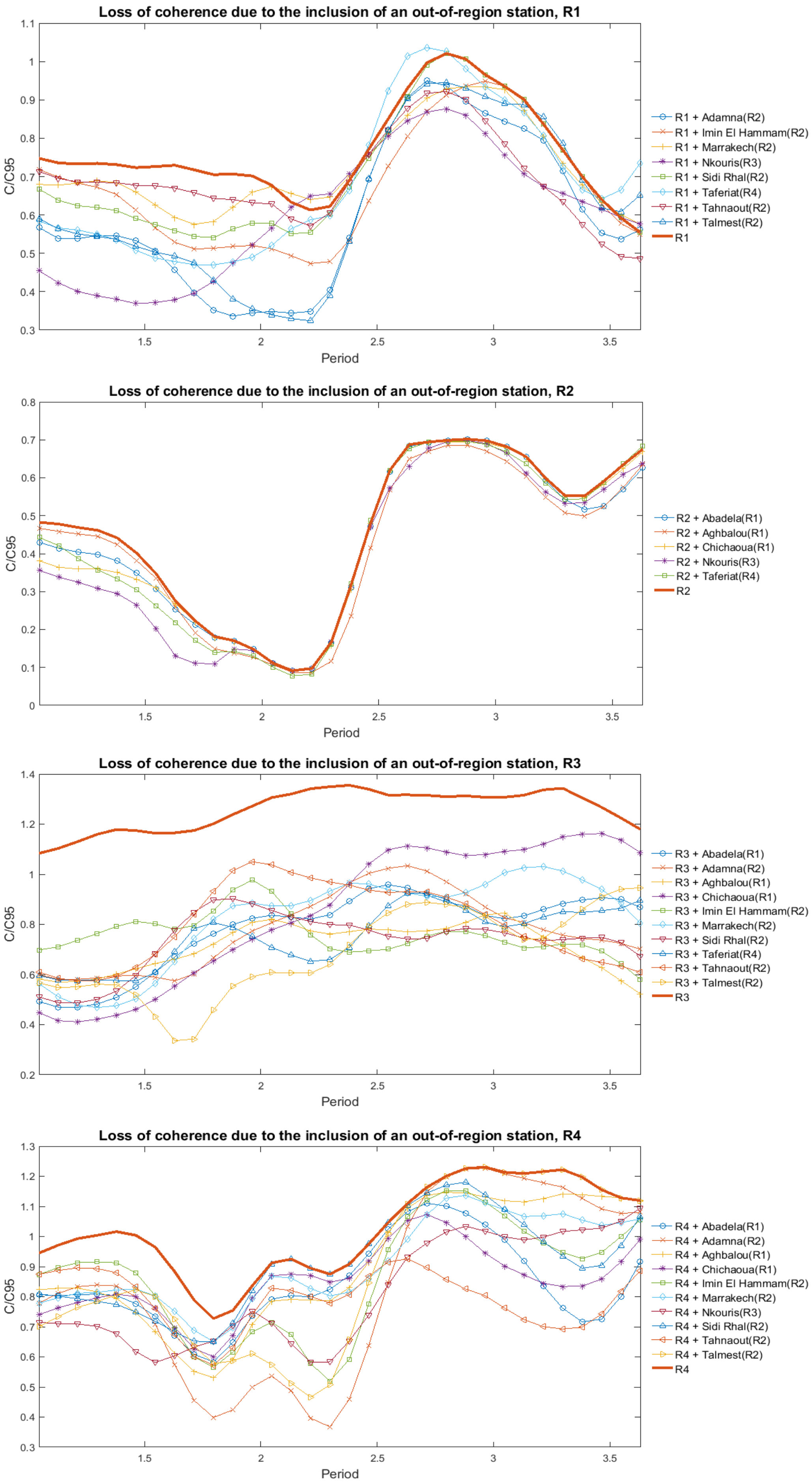

Loss of coherence due to the inclusion of an out-of-region station, R1
C/C95
Period
R1 + Adamna(R2)
R1 + Imin El Hammam(R2)
R1 + Marrakech(R2)
R1 + Nkouris(R3)
R1 + Sidi Rhal(R2)
R1 + Taferiat(R4)
R1 + Tahnaout(R2)
R1 + Talmest(R2)
R1
Loss of coherence due to the inclusion of an out-of-region station, R2
C/C95
Period
R2 + Abadela(R1)
R2 + Aghbalou(R1)
R2 + Chichaoua(R1)
R2 + Nkouris(R3)
R2 + Taferiat(R4)
R2
Loss of coherence due to the inclusion of an out-of-region station, R3
C/C95
Period
R3 + Abadela(R1)
R3 + Adamna(R2)
R3 + Aghbalou(R1)
R3 + Chichaoua(R1)
R3 + Imin El Hammam(R2)
R3 + Marrakech(R2)
R3 + Sidi Rhal(R2)
R3 + Taferiat(R4)
R3 + Tahnaout(R2)
R3 + Talmest(R2)
R3
Loss of coherence due to the inclusion of an out-of-region station, R4
C/C95
Period
R4 + Abadela(R1)
R4 + Adamna(R2)
R4 + Aghbalou(R1)
R4 + Chichaoua(R1)
R4 + Imin El Hammam(R2)
R4 + Marrakech(R2)
R4 + Nkouris(R3)
R4 + Sidi Rhal(R2)
R4 + Tahnaout(R2)
R4 + Talmest(R2)
R4

Figure 11.