REGULAR PAPER

Iterative multiscale dynamic time warping (IMs-DTW): a tool for rainfall time series comparison

Mohamed Djallel Dilmi[1] · Laurent Barthès[1] · Cécile Mallet[1] · Aymeric Chazottes[1]

Abstract
In many domains, such as weather forecasting, hydrology or civil protection, it is an important issue to characterize rainfall variability and intermittency in, either or both, a given time period or area. A variety of sensors, for instance, rain gauges, weather radars and satellites, are widely used for this purpose. Techniques to establish the similarity between rainfall time series are commonly based on the comparison of some extracted characteristic parameters (cumulative rainfall height, extreme values, rain occurrence, mean rain rate, etc.). The present study focuses on the development of a tool allowing to compare directly rainfall time series at a fine temporal scale. It allows quantifying the dissimilarity between the time series and determining a nonlinear relationship between their time axes. This study presents an algorithm based on a multiscale dynamic time warping approach, and it is based on the DTW algorithm applied on an iterative multiscale framework called IMs-DTW. This proposed algorithm is well suited for rain time series allowing point-to-point pairing between pairs of rainfall time. It takes the intermittency and the non-stationarity of the precipitation process into account. An application to measurements observed by four pluviometers located in the Paris area makes it possible to interpret the obtained results and to compare the IMs-DTW with more usual statistical features.

Keywords Multiscale dynamic time warping · Rain gauges network · Time series comparison · Precipitations · Warping path · Measure of dissimilarity · Spatiotemporal variability of the rain

1 Introduction

In recent years, a wide range of data mining techniques has been developed and applied in various fields. Many of these techniques are far more flexible than more classical modeling approaches and could be usefully applied to environmental problems [1]. In the field of the water cycle, rainfall observation is essential in many areas such as climate study, weather forecast, urban hydrology [2] or extreme event study. The way these observations are used differs depending on the

✉ Mohamed Djallel Dilmi
  djallel.dilmi@latmos.ipsl.fr

  Laurent Barthès
  laurent.barthes@latmos.ipsl.fr

  Cécile Mallet
  cecile.mallet@latmos.ipsl.fr

  Aymeric Chazottes
  aymeric.chazottes@latmos.ipsl.fr

[1] LATMOS/CNRS/UVSQ/Université Paris-Saclay, 11 Boulevard d'Alembert, 78280 Guyancourt, France

application. Therefore, rain observations are generally used in combination with numerical models representing the physical processes occurring in the atmosphere or in the soil. The basic principle of these numerical models is to represent the state of the atmosphere on a regular 2D or 3D grid composed of pixels of a given size (spatial resolution). Then applying the equations of the model, the temporal evolution is computed at regular time intervals (temporal resolution). In the case of weather forecasts, observational data are meshed to the grid model and averaged at the timescale of the model and then incorporated into the numerical model to improve the prediction (data assimilation). In general, whatever the model considered, the choice of the spatial and temporal scales of the model is a critical point. A too fine spatiotemporal resolution leads to unrealistic computational costs, while a too coarse resolution does not allow to represent the local variability of the geophysical fields. In practice, a compromise has to be found, depending on the study. In the case of global studies, given the coarse temporal and spatial resolutions, no compromise is needed. On the contrary, it is much more complicated for regional studies involving finer scales.

This is especially true for precipitation, as rainfall has high spatiotemporal variability and is an inherently intermittent process. Hence, as soon as the grid size of the model is larger than a few hundred meters, rainy and dry regions can share the same pixel, which leads to a poor representation of the field. From a temporal point of view, the same problem exists as soon as the temporal resolution is greater than a few minutes. In addition, especially for convective rain events, there is such a spatial (temporal) variability in rainfall rate that a grid size (temporal resolution) greater than a few hundred of meters (a few minutes) cannot accurately represent the natural variability of precipitation [3, 4].

Some common techniques for analyzing time series like spectral analysis, autoregressive model, principal component analysis or logistic regression are widely used. They allow modeling some features which summarize the time series. However, these techniques are not necessarily well suited for the modeling of rainfall (or rainfall properties) at a fine scale (see for example Cristiano et al. [2] in the context of urban hydrology). Considering rainfall properties, one can notice that a rainfall rate time series contains dry periods which are composed of a succession of zero values. The percentage of pure zero values in rainfall rate time series is generally quite large (about 95% in the Paris area for an integration period T of a few minutes). This property is not always in agreement with the assumptions (explicit or implicit) made by commonly used models which often assume, for example, the presence of an additive (Gaussian) noise. Moreover, most of these statistical models are only able to model stationary processes, which is not always the case for rainfall processes.

One method to obtain information regarding the characteristics of precipitation at a particular location and for a specific application is the use of the concept called "rain event" [5]. Such a concept is a convenient way to summarize precipitation time series in a small number of macrophysical features so that they make sense for particular applications. However, different points of view exist concerning the concept of "rain event." For weather studies, a "rain event" is associated with a localized atmospheric disturbance in time and space. For hydrological studies, a "rain event" is rather defined from its ground track, materialized by a measured amount of water. A number of definitions of the concept of rain event [6, 7] have been investigated in the literature [8]. There is a wide variety of criteria for dividing precipitation records into rain events. Dunkerley [9, 10] carried out an analysis of the inter-event time (IET) in order to check the influence of this parameter on the definition of rainfall events and its influence on the average rainfall rate. As highlighted in their study, when determining a value for the IET, the compromise between independence of rain events and intra-event variability of rain rates is crucial. The selection of the IET directly impacts the estimated macrophysical features.

In the present study, to overcome the difficulties commonly encountered in the analysis of rainfall data at a given time scale (resolution) or by rain event, we will focus on a tool using a multiscale approach. The main advantage of this approach is that the entire temporal structure is taken into account without any loss of information (unlike averaging), without the arbitrariness of a time scale choice and without the definition of any inter-event time characteristics or macrophysical rain features. The proposed tool, based on the dynamic time warping (DTW) algorithm [11, 12], provides a measure of the dissimilarity between time series. The use of a similarity/dissimilarity measure, looser than a (metric) distance, leads to more appropriate comparisons. Indeed, a metric distance ensures the well-known condition: $d(x, y) = 0 \Leftrightarrow x = y$. This is not appropriate for comparing the time series. Indeed two shifted time series are very similar even though their distance can be very large. Although various similarity/dissimilarity measures have been formulated in the past (see [11, 13, 14] for a complete review), the DTW is particularly interesting for rain studies. Actually, in addition to a measure of dissimilarity between two time series, it provides the temporal lags between them. Such information can be very useful for example to analyze the trajectories of rain cells through a network of rain gauges. In a more general way, time warping techniques could be useful to compare time series but also to deal with pattern recognition, extreme event detection or event clustering. In this context, clustering based on k-medoids or hierarchical clustering seems to be a good approach when combined with a dynamic temporal distortion framework [15, 16]. When applied, for example, on a rain gauge network (urban area for example), this tool could help to deduce some rain spatial features. Finally, it can be noted that, unlike most classical approaches, the DTW does not require a second-order stationarity hypothesis. In the present study, we will apply this tool for the comparison of rain gauge time series located in the same zone of interest. As said previously, a wide variety of similarity measures between time series exists. Among them, the simplest is the Euclidean distance (which is also a measure of dissimilarity). However, this approach is not appropriate when the time series have different time lags or if they have different lengths. In both cases, it leads to a bad estimate of the dissimilarity. Another common technique is the cross-correlation function which is a measure of similarity of the temporal displacement of one time series relative to the other [17]. However, it is only able to catch linear relationships between two stochastic processes and does not allow to take into account some other nonlinear effects like local contraction or dilatation of the time axis (time scaling) of one time series versus the other one. In the framework of precipitation study, these effects can frequently occur. It can be caused by the variation of a rain cell advection velocity (due to horizontal wind) from a rain gauge to another located fur-

ther. Convection or evaporation have the same effect. Subject to the processes of convection (vertical transport), advection (horizontal winds) or evaporation, the precipitation moves at variable speeds while deforming. Rain is a non-stationary phenomenon. Contrarily to classical approaches, the proposed algorithm makes it possible to better take into account precipitation intermittency and non-stationarity.

In this paper, we will focus on a method derived from the original dynamic time warping method (DTW) [12]. This choice as preconized by Sung et al. [18] is conditioned by the good behavior of this method in the presence of nonlinear signal transformations like time scaling and/or time shifting which are present in rain time series. They compared four distances, namely the DTW, the earth mover's distance (EMD) [19], the Fréchet distance [20] and Hausdorff distance [21]. They concluded that the DTW shows the best performances in all of the experiments they conducted.

In this work, we consider rainfall rate time series RR measured by a rain gauge network. Each rain gauge allows measuring the cumulated rain height over an integration time period T. This parameter is usually converted to rainfall rate $RR(mm\,h^{-1})$ by dividing the rain height by the time period T (generally expressed in hour). Rainfall rate is well suited to describe rain variability and thus to classify rain events into different categories like stratiform or convective events. The information provided by rainfall rates is the main input of all hydrological models. A rain gauge time series, of duration D, is a sequence of N samples where $N = \frac{D}{T}$. From this native rainfall rate time series, we can compute a subsequent rainfall time series RR^c which is averaged over consecutive time periods equal to cT. The parameter c called the resolution factor or the compression rate [12] is an integer value ranging between 1 and N. c values greater than 1 correspond to time series with a coarser temporal resolution. Hence, for a given compression rate c, we denote N^c the number of samples of RR^c:

$$N^c = \frac{D}{cT} \tag{1}$$

According to these notations, the time series RR can be expressed as a sequence of rainfall rates (Eq. 2).

$$RR^c = (RR_1^c, RR_2^c, \ldots, RR_{N^c}^c) \tag{2}$$

RR_i^c, the ith element of RR^c is calculated from the native time series RR^1, thanks to the following equation:

$$RR_i^c = \frac{1}{c} \sum_{j=c(i-1)+1}^{ci} RR_j^1 \quad \text{with } i \in [\![1, N^c]\!] \tag{3}$$

The time series RR^c is also known as the temporal aggregation of the native time series. By definition, it represents
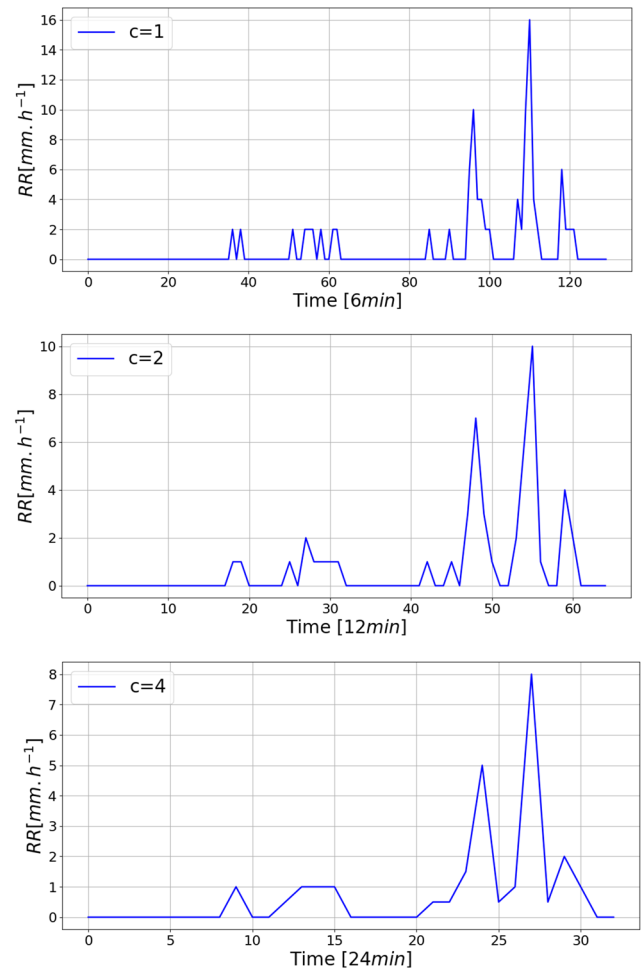

Fig. 1 Example of PAA of rainfall time series for $c = 1, 2$ and 4 obtained from a rain gauge with $T = 6$ min

the piecewise aggregate approximation (PAA) of the finer precipitation time series RR^1. Equation 3 simply states that the precipitation time series is composed of N^c equal-sized "frames" RR_i^c. In the following, we will use equivalently RR^1 and RR as well as N^1 and N. Figure 1 illustrates a precipitation time series at native resolution ($c = 1$)and its PAA for $c = 2$ and 4.

2 The multiscale dynamic time warping algorithms (MsDTW)

Let's first introduce the dynamic time warping (DTW) algorithm and some of its variants. The DTW was introduced in 1978 by Sakaoe and Chiba [22] as the dynamic programming algorithm (DP-Algorithm).

An extensive literature exists on either improving the method (especially to make it faster) or applying it in various fields such as speech recognition. The DTW algorithm computes the time distortion needed to align the two time series.
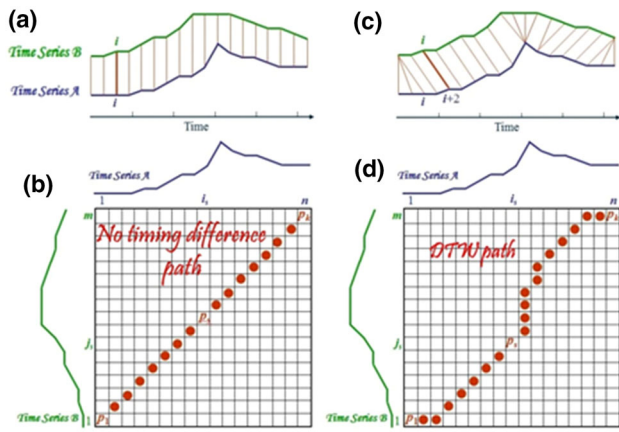
Fig. 2 **a** Comparison between two time series with no timing difference. The gray lines represent the alignments between the two time series. **b** Warping path associated to **a**. **c** Comparison between two time series with some timing differences. **d** The warping path associated to **c** (from [24])

This alignment is calculated in order to minimize a distance between two series. Several variants exist which introduce constraints in order to yield to an algorithm less computationally expensive. For a complete description of the DTW algorithm under constraints see for example Zhanga et al. [23]. Here, we only give a brief description of the algorithm in the framework of precipitation time series alignment.

Let's denote A^1 and B^1 as two rainfall rate time series observed with the same integration time T (Fig. 2a). The two time series do not necessarily share the same number of samples. Hence, we denote, respectively, N_A^1 and N_B^1 as the lengths of the times series A^1 and B^1. As done previously, the two time series can be expressed as sequences:

$$\begin{cases} A^1 = (A_1^1, A_2^1, \ldots A_i^1, \ldots, A_{N_A^1}^1) \\ B^1 = (B_1^1, B_2^1, \ldots B_j^1, \ldots, B_{N_B^1}^1) \end{cases} \quad (4)$$

Sakoe and Chiba [22, 25, 26] proposed to consider an i–j plane (Fig. 2b) where time series A^1 and B^1 are developed, respectively, along the i-axis and the j-axis. This plane is reported in the literature as the distance matrix D [24]. The timing differences between time series A^1 and B^1 can be depicted by a sequence P^1 of K points $p_k^1 = (i_k^1, j_k^1)$ belonging to the distance matrix D:

$$P^1 = (p_1^1, p_2^1, \ldots, p_k^1, \ldots, p_K^1) \quad (5)$$

It is worth noticing that the length of the sequence K is not known at this point.

This sequence represents a mapping from the time axis of time series A^1 onto that of time series B^1. Sakoe and Chiba [25] called this mapping a warping function, and it is currently known as a warping path [24, 27]. When there is no

timing difference between two time series (with the same number of samples), the path coincides with the diagonal line $j_k^1 = i_k^1$ on the distance matrix D as shown in Fig. 2a, b. Otherwise, it deviates further from the diagonal line as the timing difference grows (Fig. 2c, d).

In order to compute an optimal mapping between the two time series, a tool is needed to evaluate how close they are. Let's denote $d(p_k^1)$ the distance corresponding to the kth point of the warping path which corresponds to the distance between two samples $A_{i_k^1}^1$ and $B_{j_k^1}^1$. In our case, we simply chose a quadratic distance:

$$d\left(p_k^1\right) = \left(A_{i_k^1}^1 - B_{j_k^1}^1\right)^2 \quad (6)$$

For the DTW, we denote d_{DTW} the time-normalized dissimilarity between two sequences, here, the two rainfall time series A^1 and B^1. This dissimilarity relies on a corresponding optimal warping path P_{DTW}^1. Both the dissimilarity and the warping path were given by Sakoe and Chiba [22] as the solution of an optimization problem:

$$\begin{cases} d_{\text{DTW}}\left(A^1, B^1\right) = \min_P \left[\sqrt{\frac{\sum_{k=1}^K d(p_k^1) \times w_k}{\sum_{k=1}^K w_k}} \right] \\ P_{\text{DTW}}^1\left(A^1, B^1\right) = \operatorname{Arg} \min_P \left[\frac{\sum_{k=1}^K d(p_k^1) \times w_k}{\sum_{k=1}^K w_k} \right] \end{cases} \quad (7)$$

with w_k the nonnegative weighting coefficients introduced intentionally to allow the dissimilarity flexible characteristics. The denominator $\sum_{k=1}^K w_k$ is employed to make the dissimilarity score independent of K, the length of the warping path P_{DTW}^1. Hence, the dissimilarity d_{DTW} is the minimum weighted average on all the possible warping paths.

All warping paths are not necessarily appropriate, and, since initially the DTW algorithm was developed for speech recognition, some constraints were added by the authors in accordance with speech features:

1. Boundary conditions: $p_1^1 = (1, 1)$ and $p_K^1 = \left(N_A^1, N_B^1\right)$. The path starts at the bottom left and ends at the top right.
2. Monotonicity: $i_k^1 - i_{k-1}^1 \geq 0$ and $j_k^1 - j_{k-1}^1 \geq 0$. The path will not turn back on itself, both the i and j indexes either stay the same or increase, they never decrease.
3. Continuity: $i_k^1 - i_{k-1}^1 \leq 1$ and $j_k^1 - j_{k-1}^1 \leq 1$. The path advances one step at a time. Both i and j can only increase by at most 1 on each step along the path.
   In many cases, searching for the optimal path may result in undesired effects because the global optimal path may not necessarily be the one desired and may even be unrealistic. As an example, in the presence of a succession of constant values (series of null values corresponding to dry periods in our case), a large number of points of one time series is mapped to a single point of the other one. A common way to overcome this problem is to restrict
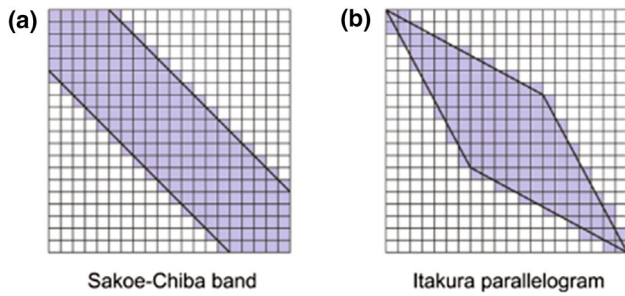
Fig. 3 Examples of global constraints. **a** Sakoe–Chiba band. **b** Itakura parallelogram (from Cassisi et al. [11])

the warping path. This is done in such a way that it has to follow a direction in the neighborhood of the diagonal. To do so, two additional constraints are commonly used:

4. Warping window condition: $\left|i_k^1 - j_k^1\right| \leq \delta$ where δ is a threshold restricting the path. It enforces the recursion to stop at a certain depth. This constraint is known as Sakoe–Chiba band [22] (Fig. 3a). Besides limiting extreme or degenerate mappings, it allows to speed up the DTW distance calculation.
5. Slope constraint condition: $\frac{j_{k+t}^1 - j_k^1}{i_{k+t}^1 - i_k^1} \leq t$ and $\frac{i_{k+s}^1 - i_k^1}{j_{k+s}^1 - j_k^1} \leq s$. For a warping path, the slope should be neither too steep nor too gentle. Here, in a sequence of k consecutive points of the warping path, for one step in the i-direction, we are allowed t steps in the j-direction. Likewise, for one step in the j-direction, we are allowed s steps in the i-direction [22].

The constraints proposed above were used in the first versions of the dynamic time warping algorithm. Since then various modifications have been proposed to speed up the DTW computations as well as to better control the possible routes of the warping paths. Itakura [28] proposed another constraint known as the Itakura parallelogram (Fig. 3b).

With such constraints, the quality of an alignment depends heavily on the choice of one or more parameters, which is quite subjective. Moreover, as pointed by Cassisi et al. [11] the use of such constraints does not guarantee a good alignment between the two time series since it does not allow the optimal warping path to leave the region as defined in steps 4 and 5.

Since the objective function in Eq. 7 is a rational expression, its minimization is an unwieldy problem. However, if the denominator $\sum_{k=1}^{K} w_k$ in Eq. 7 (called normalization coefficient) is independent of the warping path P^1, it can be considered constant and equal to L:

$$\forall P^1, \sum_{k=1}^{K} w_k = L \tag{8}$$

It can be put out of the brackets, resulting in the following equation:

$$d_{\text{DTW}}\left(A^1, B^1\right) = \frac{1}{\sqrt{L}} \min_P \left[\sqrt{\sum_{k=1}^{K} d\left(p_k^1\right) \times w_k}\right] \tag{9}$$

With this simplification [22], the warping path P_{DTW}^1 can be (and is traditionally) computed using dynamic programming [29] with an $O\left(N_A^1 N_B^1\right)$ "quadratic" complexity. To enable this simplification, several weighting coefficient definitions were proposed: symmetric and asymmetric (see a complete discussion in [22, 30]). All these weighting coefficient definitions share the same disadvantage: They favor a type of path to the others.

On the other hand, without any a priori information, we are looking for a weighting coefficient definition that does not suffer from such a disadvantage and evaluates equally the paths. In this case, the obvious definition for the weighting coefficients is:

$$\forall k, w_k = 1 \quad \text{so} \quad \sum_{k=1}^{K} w_k = K \tag{10}$$

This brings the minimization of the objective function back to the problem encountered previously in Eq. 7 (i.e., the dependence of the dissimilarity score on K, the unknown number of points on the warping path).

One way to overcome this problem is to reformulate the objective function keeping the same concept. To simplify the minimization computation, we replace K by a known value of the same magnitude. We chose a normalized quasi-symmetric form of the DTW [24] reformulated below:

$$\begin{cases} d_{\text{DTW}}\left(A^1, B^1\right) = \sqrt{\frac{2}{(N_A^1 + N_B^1)}} \min_P \left[\sqrt{\sum_{k=1}^{K} d\left(p_k^1\right)}\right] \\ P_{\text{DTW}}^1\left(A^1, B^1\right) = \text{Arg} \min_P \left[\sum_{k=1}^{K} d\left(p_k^1\right)\right] \end{cases} \tag{11}$$

This optimization problem can be computed using dynamic programming. It has the following properties:

1. If $N_A^1 = N_B^1 = N$, then $d_{\text{DTW}}\left(A^1, B^1\right) = \sqrt{\frac{1}{N} \sum_{k=1}^{K} d\left(p_k^1\right)}$
2. The d_{DTW} is less sensitive to the difference between the lengths of the two series N_A^1 and N_B^1
3. Consequence of Eq. 11: If $N_A^1 = N_B^1 = 1$, then $d_{\text{DTW}}\left(A^1, B^1\right) = \sqrt{\left(A_1^1 - B_1^1\right)^2} = \left|A_1^1 - B_1^1\right|$

To speedup the DTW, Keogh and Pazzani [12] took advantage of the fact that we can efficiently approximate most time

series by a piecewise aggregate approximation (PAA), so they proposed to apply the DTW on the time series at a coarser resolution $c > 1$ rather than at the native resolution ($c = 1$). They advanced that:

$$d_{\text{DTW}}\left(A^1, B^1\right) \cong d_{\text{DTW}}\left(A^c, B^c\right) \qquad (12)$$

They called their algorithm piecewise dynamic time warping (PDTW). One important limitation of this approach, however, is that the user must carefully choose the compression rate parameter c. Indeed, a too coarse resolution can lead to an inaccurate or even completely useless warping path [27, 31]. Although this approach may be interesting for processes characterized by a low, temporal, or spatial, variability, this is not valid in the case of rain. Indeed rain is well known to exhibit a multifractal behavior [3, 32, 33] characterized by variability that can greatly increase with the resolution.

This limitation motivates other approaches such as the multiscale DTW (hereafter MsDTW) approach which is based on a multilevel mapping achieved through the use of several resolutions. In this regard, Chu et al. [34] presented an algorithm named iterative deepening dynamic time warping (IDDTW). The basic principle is to iteratively apply the piecewise dynamic time warping (PDTW) for different resolutions starting at a very coarse resolution. At each iteration, this algorithm decides whether to apply the PDTW to a higher resolution or to keep the current PDTW approximation. Unfortunately, IDDTW like PDTW is not acceptable for precise alignment of the time series. Later, Salvador et al. [35] proposed the FastDTW algorithm. The authors state that: "A multilevel approach works well if a large problem is difficult to solve all at once, but partial solutions can effectively be refined at different levels of resolution. The dynamic time warping problem can also be solved with a multilevel approach." This approach avoids applying the brute force of the standard DTW algorithm by using the multiscale framework.

The FastDTW algorithm can be divided into four steps:

1. Initialization: The two time series are initially averaged to a low resolution ($c \gg 1$) using a piecewise aggregate approximation (PAA). The DTW algorithm is run (using constraints 1 to 3), and a warping path is found for the current resolution (solid line in Fig. 4a).
2. A projection operator projects this warping path to the next higher resolution giving a list of cells defining a new constraint that only these cells will be evaluated by the DTW algorithm (dark gray cells in Fig. 4b).
3. Aware that the entire optimal warp path may not be contained within the projected path, Salvador and Chan [35] have released a little bit this constraint by allowing an additional number of cells to be evaluated (light gray cells in Fig. 4b). For that, they introduced a radius parameter

r which controls the additional number of cells on each side of the projected path that will also be evaluated.
4. The DTW algorithm is run with this released constraint (and constraints 1 to 3), and an optimal warp path is found. This optimal warp path is then used to find a new released constraint for the next higher resolution. The procedure is repeated (go to step 2) until the full resolution is reached.

Figure 4 (from [35]) provides an illustration of the iterative process for a resolution factor c equal, respectively, to 8, 4, 2 and 1. In this figure, the solid black line represents the optimal warping path for the current resolution factor c. In this example, the optimal path does not move too far from the local diagonal path. However, with a radius r greater than 1, the algorithm allows going far from the local diagonal. This case is illustrated in Fig. 5a, where the FastDTW has been applied to two rain gauges time series recorded in two cities (Trappes and Villacoublay) located near Paris. The FastDTW provides good alignments except for spurious peaks. This is the case for example at the 205 time index where a peak in the Trappes time series is associated with another peak in the Villacoublay time series at the time index 420. The global optimal path P_{DTW}^1 leads to a shift of 215 lags between the two peaks, which correspond to a delay of 215×6 min $= 1290$ min, i.e., more than 21 h. This alignment is unrealistic considering the distance between the two cities (15 km). This kind of situation will occur whenever the rain is observed by a rain gauge but not by the other. In other words, the trajectory of the rain cell meets only one of the two rain gauges.

We want to avoid the unrealistic association of a spurious parasitic peak of one time series to an unrelated rain event from the other. Hence, instead of the global optimal path, we would better choose a local minimum path that respects the independence between distant rain events that should not be linked. Therefore, we only want to consider the intra-rain event deformation mapping a rain event to another. Purposely we set the radius r to 0. This corresponds to removing the third step of the FastDTW algorithm and therefore not to release the constraint defined in step 2. In contrast with Fig. 5a, b shows that the peak at time index 205 in the Trappes time series is now associated with a zero value in the Villacoublay time series. It would be much more realistic if it was associated with the peak located at time index 420 instead. We are aware that the found path is not necessarily the global optimal path but rather an optimal path under multiscale constraints. In the following, this particular configuration will be called IMs-DTW for iterative multiscale DTW. For the sake of simplicity, the optimal warping path will be denoted by P_{DTW} instead of P_{IMSDTW}^1. Finally, it can be noted that even if it was not the main objective, a radius defined to zero considerably reduces the calculation time.
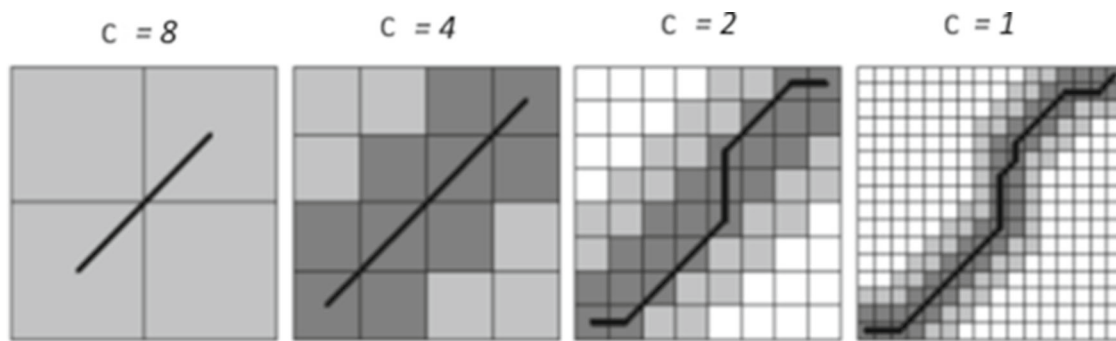
Fig. 4 Illustration of the Salvador et al., algorithm [35]. Dark gray squares are cells that will be evaluated by the constrained DTW. They are derived from the previous step by a "projection" operator. Light gray cells correspond to the released constraint for a radius equal to 1 square. They are also evaluated by the constraint DTW algorithm. White cells are not taken into account by the DTW algorithm. The solid line corresponds to the derived optimal path
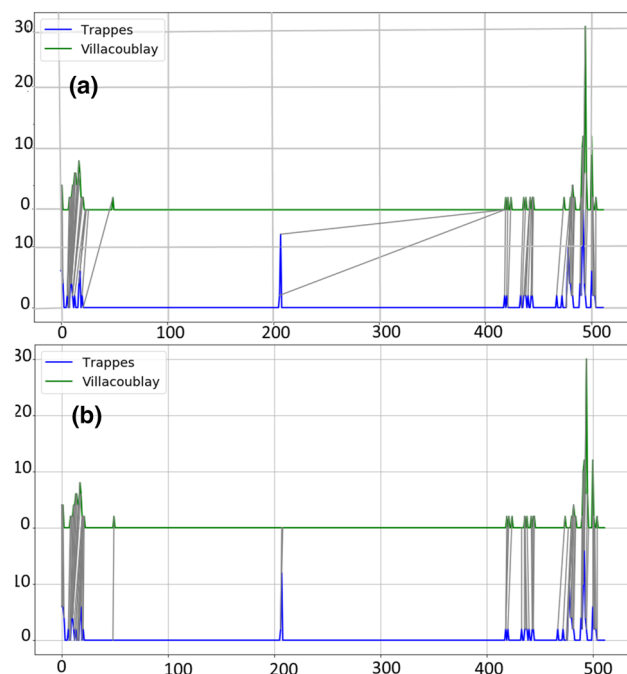

Fig. 5 a FastDTW alignments for two rain gauges time series measured in the cities of Trappes and Villacoublay with a time resolution of $T = 6$ min. The gray lines represent the alignments between the two time series. **b** Same figure but for IMs-DTW alignments

3 The iterative multiscale DTW algorithm

3.1 Experiment outline

In order to evaluate the efficiency of the IMs-DTW algorithm on precipitation time series, several experiments were conducted. Here we present an experiment for a precipitation time series resulting from 10 days of tipping-bucket rain gauge measurements, recorded close to the city of Trappes (27 km southwest of Paris) between June 7, 2009, and June 17, 2009. This time series called reference time series in the following is characterized by an integration time T equals to 6 min and by a tipping-bucket volume corresponding to an equivalent rain height of $h = 0.2$ mm.

This time series of precipitation (Fig. 6a) is characterized by four early rain events between June 7 and June 11. Then a dry period of several days is followed by a last rainy episode. Hereafter, we call this precipitation time series A which corresponds to A^1 of Sect. 2 (for sake of clarity the resolution factor exponent $c = 1$ will be omitted in the following).

This 10-day period is particularly interesting, since it regroups both several short dry periods (few hours or less) and a long dry period (a few days). It will allow us to show the behavior of the IMs-DTW algorithm in different meteorological situations. In the next subsection, this time series is used to illustrate the ability of the algorithm to map rainfall time series through the estimated warping path.

3.2 Robustness of the IMs-DTW on precipitation time series

To demonstrate the robustness of the IMs-DTW algorithm, we simulate a set of transformed time series. They were built from the reference time series A by applying three kinds of transformations, namely adding lags, varying rain rates by multiplying nonzero rainfall rates values by a random value and finally inserting some spurious rain events. In the following, the IMs-DTW algorithm is applied between the time series A and the transformed time series.

To ensure that we cover a wide range of situations, we have simulated 1000 transformed time series with the following three-step procedure:

1. The first transformation is a time shifting. It can be seen has the time lags introduced by a fluctuating wind (advection velocity) transporting a frozen rain event from a place

Fig. 6 **a** The reference precipitations time series A recorded in Trappes between June 7, 2009, and June 17, 2009, with an integration time $T = 6$ min. **b** A realization of the transformed time series B" built from the reference precipitation time series A. Black arrows indicate inserted rain events. The event at time index 2000 is zoomed
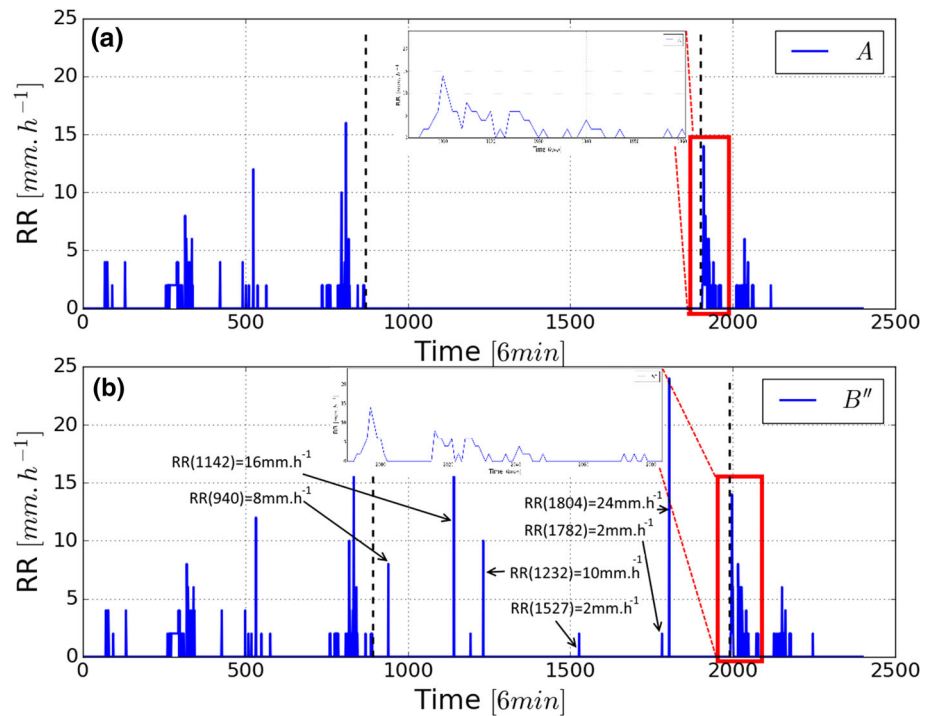
to another. At each time step, the time series B is delayed from A by an offset $h(i)$ corresponding to an accumulation of random time lags α_j. This shifted rainfall time series B is computed by adding zero rain rate values. It can be expressed by Eq. 13:

$$\begin{cases} B_i = A_{i-h(i)} \\ B_{i-h(i)-y(i)} = 0 \text{ when } \alpha_i \neq 0 \end{cases} \tag{13}$$

With the sequence $y(i) = (1, \ldots, \alpha_i)$ and the sum $h(i) = \sum_{j=1}^{i} \alpha_j(\lambda)$ where $\alpha_j(\lambda)$ is drawn according to a Poisson law of parameter λ ($\lambda = 0.1$ in our case).

2. The second transformation is an amplitude modulation. Since a rain event is not frozen when it is transported/advected from a place to another, we have to modulate the values of the rain rates to take the variability of rain into account. Hence, the B' transformed time series is computed by a correction to the nonzero rainfall rates of an already shifted B time series. To exhibit significant rain rate variations from A, *t*he non-null rainfall rate values are multiplied by a random value β_i drawn uniformly between 0.5 and 2. This new type of time series is denoted B':

$$B_i' = \beta_i B_i \quad \text{when } B_i \neq 0 \tag{14}$$

This transformation does not disrupt dry periods since only nonzero rainfall rates are considered.

3. The third transformation adds spurious rain events. Finally, as explained in the previous section, a rain event

can be seen by a rain gauge in a specific area but not by another one located a few kilometers away. For this reason, the transformed time series B'' is derived from B' by inserting spurious rain events. The number of added rain events is drawn according to a uniform law between 2 and 30. Figure 6b shows a realization of a transformed time series B''.

To evaluate the IMs-DTW performance, we first analyzed the warping paths $P_{DTW}(A, B)$ linking the time series A to the different realizations of the shifted time series B. As expected, thanks to the multiscale constraints, the IMs-DTW was able to find the right warping paths whatever the added time delays. The calculated $d_{DTW}(A, B)$ dissimilarity, based on the $P_{DTW}(A, B)$ warping path, manifests two behaviors depending on where the time delays were added:

- *Case 1* A time series B is generated as mentioned above but time shifts only occur inside a dry event (consecutive zero sequences). In other words, the rain event patterns, i.e., non-zero sequences, are preserved. In this situation, as expected the dissimilarity between A and B is always null: $d(A, B) = 0$.
- *Case 2* A time series B is generated but this time one or more zero rain rate values are added inside the rain events. (A zero is added in between consecutive nonzero rain rate values.) It creates intra-event dry periods, therefore, changing event durations, and generally modifying the rain event properties. These intra-event dry periods are expected to be
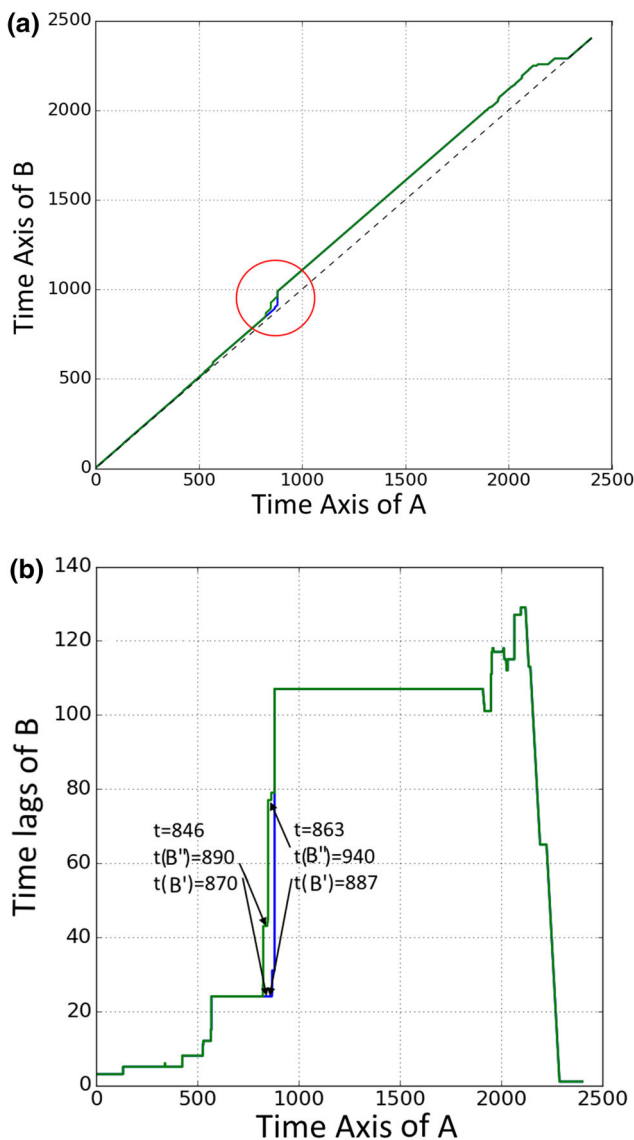
Fig. 7 **a** In blue, the warping path $P_{DTW}(A, B)$ between the time series A time series B. In green $P_{DTW}(A, B'')$ the warping path between time series A and time series B'' shown in Fig. 6b. The dashed line represents the diagonal line. **b** The corresponding time lags (color figure online)

associated to nonzero rain rates. Consequently, the d_{DTW} (A, B) dissimilarity is different from 0. This case is illustrated by the event at the time index 2000 in Fig. 6b. In Fig. 7a, the blue curve shows an example of the estimated $P_{DTW}(A, B)$ warping path. One can see that during inter-event dry periods, the path remains on a local diagonal (see, for example, the time period between 880 and 1900). The associated lags are, consequently, constant (Fig. 7b). Similarly, when the second type of transformed time series B' was tested, the IMs-DTW was also able to find the right paths (not shown).

Now let's consider the third type of transformation (transformed time series B'') in which some spurious rain events were inserted. Again two situations have to be considered:

- *Case 1* In time series B'', the inserted rain event is "far" (few hours) from the rain events in A. In this situation, the inserted rain event is considered by the algorithm to exist only in time series B''. Therefore, it is associated with zero values in time series A. This case is illustrated for example in Fig. 6a, b for index time ranging from 1142 to 1232. (In this case, the rain events are present only in time series B''.) Figure 7b shows that during this period the warping path remains on the local diagonal and consequently the rain event is associated with zero values in A.
- *Case 2* In time series B'', the inserted rain event is "close" to a rain event in A. In this case, the IMS-DTW considers the two rain events as a single event and the time separating them as an intra-event dry period. The inserted rain event is then associated with the closest rain event in A. This case is visible in Fig. 7 in which the event beginning at time index 940 in B'' has been grouped with an event of time series A at time index 863.

Finally, for the 1000 simulations, the IMS-DTW proposed acceptable warping paths P_{DTW} that were consistent with what might be expected from the observation of rainfall. Indeed, two rainy events occurring with a time lag of more than a few hours can be considered independent. The presence of a rainy episode in a time series does not necessarily imply its presence/absence in the other one.

In this subsection, we have shown the good behavior of the IMS-DTW algorithm in various simulated situations based on a real-time series (time series A) and the main transformations encountered in the context of rainfall observation. In the following, the study focuses on how the method can be used to compare two real-time series.

3.3 Case study

In this section, we wish to assess the performance of the IMS-DTW on real rainfall time series observed by rain gauges located in the same urban area. We chose to present an analysis that is performed on four relatively short precipitation time series (13 h) simultaneously recorded by four meteorological stations (Trappes, Le Bourget, Roissy and Nangis) located near Paris. The measurements occurred on June 10, 2009, between 06:00 and 19:00 with an integration time $T = $ 6 min. Figure 8 shows the four time series.

The similarities between the rain patterns on the four series suggest that it could be the same rain event recorded by the four stations. This suggestion is partially confirmed by radar images (not shown) which also indicate that the rain cells at the origin of the four recordings move from Southwest to
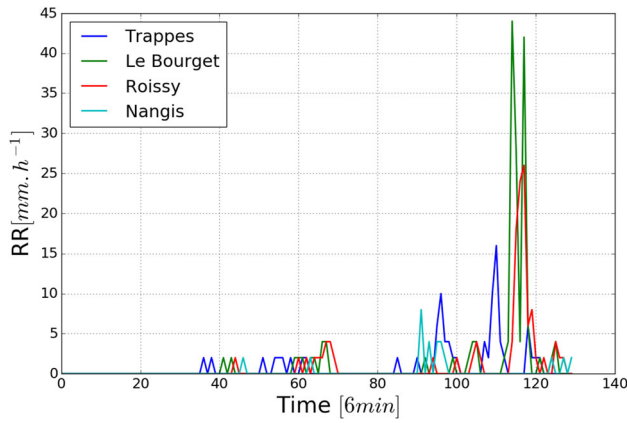
Fig. 8 Precipitations time series recorded in Trappes, Le Bourget, Roissy and Nangis on June 10, 2009, between 06:00 and 19:00 with an integration time $T = 6$ min

Northeast. Radar images also allow the following observations:

1. The same rain cell was at the origin of the time series recorded at Trappes first and then at Le Bourget, and Roissy. The meteorological station of Trappes is located to the southwest of Le Bourget and Roissy. On the radar image, the rain cell is thus observed there a little earlier than for the other two. One notes, moreover, that it seems less intense.
2. A different rain cell was at the origin of the time series recorded at Nangis station which is located in the extreme southeast of the area (60 km from Roissy).

As a result, since Roissy and Le Bourget are close (around 10 km between the two) we expect that their time series should just be slightly delayed. Moreover, the lags obtained between Nangis station and the three others have no physical meaning.

To illustrate the benefit of considering a time warping using the IMS-DTW for real precipitation time series comparison, we first computed the Euclidian distance widely used in the context of precipitation time series comparison [36]. For A and B representing precipitation time series at the native resolution (i.e., for a resolution factor $c = 1$) recorded at the stations S_i and S_j, the Euclidian distance is defined by:

$$d(S_i, S_j) = \sqrt{\sum_{k=1}^{N}(A_k - B_k)^2} \qquad (15)$$

This distance is sensitive to N the length of the time series. To eliminate this undesired effect, we computed a normalized Euclidian distance defined by:

$$d_{\text{NTD}}(S_i, S_j) = d(P_{\text{NTD}}, S_i, S_j) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(A_k - B_k)^2} \qquad (16)$$

In Eq. 16, we add the term P_{NTD}, for the "no timing difference path," which corresponds to a diagonal warping path, i.e., $p_k^1 = (i_k^1, i_k^1)$. In fact, the normalized Euclidean distance is a special case of the dissimilarity defined by Eq. 11, and therefore, it is possible to compare these two distances. Similarly, we will denote the correlation coefficient between the two time series $r_{\text{NTD}}(S_i, S_j) = r(P_{\text{NTD}}, S_i, S_j)$, and the maximum of cross-correlation $r_{\tau,\text{Max}}(S_i, S_j) = r_{\text{Max}}(P_\tau, S_i, S_j)$. P_τ is the warping path associated to the time lag τ_{delay} ($\tau_{\text{delay}} = nT$ with T the integration time), this latter corresponds to the time lag maximizing the cross-correlation function. This means that the warping path P_τ is located on the n-super/sub-diagonal of the distance matrix D, i.e.,

$$p_k^1 = (i_k^1, j_k^1) = (i_k^1, i_k^1 \mp n) \quad \text{with } n = \frac{\tau_{\text{delay}}}{T}$$

The normalized Euclidian distance $d_\tau(S_i, S_j) = d(P_\tau, S_i, S_j)$ considering the time delay τ_{delay} was also performed. Table 1 provides the corresponding values for six (S_i, S_j) station pairs, while Fig. 9 shows the warping paths P_{NTD} (Fig. 9a) and P_τ (Fig. 9b) for the pairs of stations Trappes/Le Bourget and Trappes/Nangis.

The without sliding approach (i.e., based on the P_{NTD} warping path) compares rainy periods to non-rainy periods (see Fig. 9a). The smallest normalized Euclidean distance $d_{\text{NTD}}(S_i, S_j)$ (Table 1 column 3) is obtained for the Trappes/Nangis pair while these two stations are the furthest apart. As mentioned before, they did not record the same rain cell. In fact, the obtained distances seem to be only sensitive to the rain amount of the two considered time series when the stations are sufficiently distant. This example shows that the Euclidian distance is meaningless in the context of rainfall time series comparison. Concerning the $r_{\text{NTD}}(S_i, S_j)$ correlation coefficient (Table 1 column 4), it is close to zero except for the pair Le Bourget/Roissy. There is no linear correlation between the pairs except when the stations are close enough like for the pair previously mentioned. Thus, like the normalized Euclidean distance, the correlation coefficient is ineffective in this context and classical approaches without sliding (P_{NTD}) do not generally allow to identify the similarities/dissimilarities between precipitation time series.

In the second approach, we compare the time series pairs taking into account a shift by a constant time lag τ_{delay} which is equivalent to the use of the warping path P_τ. Very low cross-correlation values (Table 1 column 7) are obtained when Nangis station belongs to a pair. As stated in the without sliding approach, the two recorded rain time series do not

Table 1 Indicators of dissimilarity between the six (S_i, S_j) pairs for P_{NTD}, P_τ and P_{DTW} warping paths

	Distance (km)	P_{NTD} warping path			P_τ warping path		P_{DTW} warping path	
Pair of stations S_i, S_j		$d_{NTD}(S_i, S_j)$ (mm h^{-1})	$r_{NTD}(S_i, S_j)$	τ_{delay} (min)	$d_\tau(S_i, S_j)$ (mm h^{-1})	$r_{\tau,Max}(S_i, S_j)$	$d_{DTW}(S_i, S_j)$ (mm h^{-1})	$r_{DTW}(S_i, S_j)$
Trappes/Le Bourget	34.78	6.35	− 0.05	30	5.05	0.61	3.82	0.81
Trappes/Roissy	44.84	4.24	− 0.03	42	2.58	0.75	1.53	0.91
Le Bourget/Roissy	10.41	4.42	0.65	12	3.84	0.79	3.27	0.83
Trappes/Nangis	77.96	2.24	0.17	− 30	2.14	0.35	1.68	0.64
Le Bourget/Nangis	60.44	6.08	− 0.04	60	6.09	0.21	5.38	0.51
Roissy/Nangis	61.36	3.86	− 0.06	60	3.83	0.15	3.07	0.62
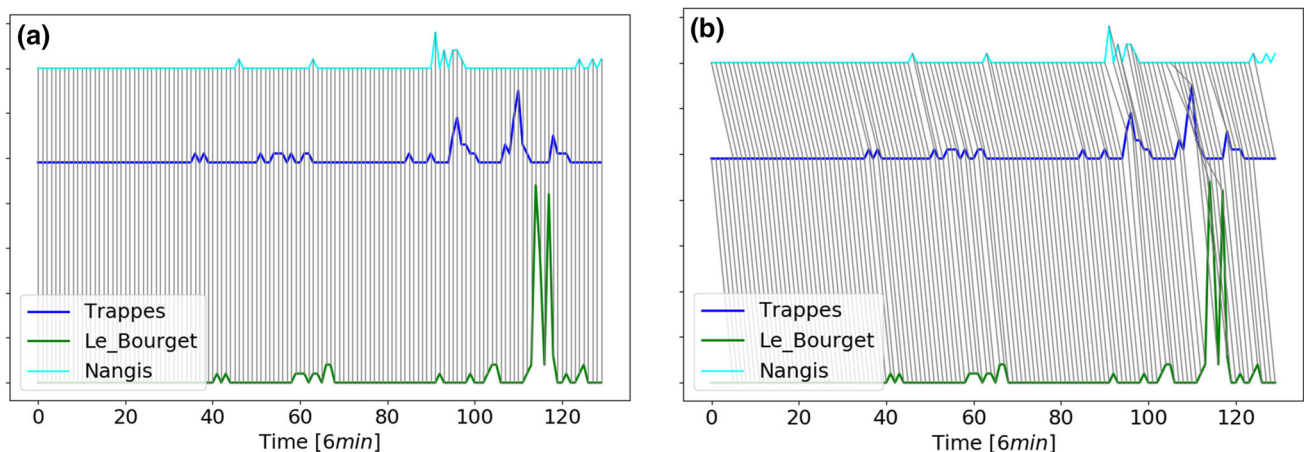
Fig. 9 a Sequences P^1 of points $p_k^1 = (i_k^1, j_k^1)$ defined by Eq. 6 for the P_{NTD} warping path for the pairs of stations Trappes/Le Bourget and Trappes/Nangis. The gray lines represent the alignments between the two time series. **b** Same figure than **a** but with the P_τ warping path

come from the same rain cell and are therefore are not correlated. For the three first pairs of stations Trappes/Le Bourget, Trappes/Roissy and Le Bourget/Roissy, the time delays τ_{delay} are consistent with both the geographic distances (Table 1 column 2) and the advection velocity observed on the radar images (not shown). In these cases, the maximum of cross-correlation $r_{\tau,Max}(S_i, S_j)$ is thus representative. In addition, for these pairs, the normalized Euclidian distance $d_\tau(S_i, S_j)$ decreased compared to $d_{NTD}(S_i, S_j)$ by a factor ranging from 16 to 50%. Trappes and Roissy are not as dissimilar as suggested by the normalized Euclidian distance. On the other hand, for the stations of Le Bourget and Roissy we expected, given their proximity (10 km), a smaller dissimilarity (5.05 mm h^{-1}). When the Nangis station belongs to a pair, the $d_\tau(S_i, S_j)$ distances remain almost unchanged compared to $d_{NTD}(S_i, S_j)$.

Their corresponding $r_{\tau,Max}(S_i, S_j)$ cross-correlations remain meaningless. Finally, we can conclude that the use of a constant time lag (warping path P_τ) does not make it

possible to estimate good indicators representing the dissimilarities between the stations.

After outlining these findings, we performed the IMs-DTW on the six different pairs and we analyzed the warping paths P_{DTW}. Figure 10a, b shows the sequences P^1 of points p_k^1 of P_{DTW} for the pairs Trappes/Le Bourget and Trappes/Nangis.

In the case of the pair Trappes/Le Bourget, the assumption of a dynamic warp improves the associations of the points p_k^1 of the sequence P^1 unlike with P_τ warping path where the matching respects the general behavior even if there are some association mismatches. For example, the second Le Bourget peak recorded at the 118 time index is associated to a zero rain rate value in the time series of Trappes while the P_{DTW} made a better association by associating the peak in Le Bourget time series with a peak in the Trappes time series. In a general way, the IMs-DTW was able to find a better matching of all the couples of points (see Fig. 10c, d). The algorithm realizes good matching, associating together the beginning and the
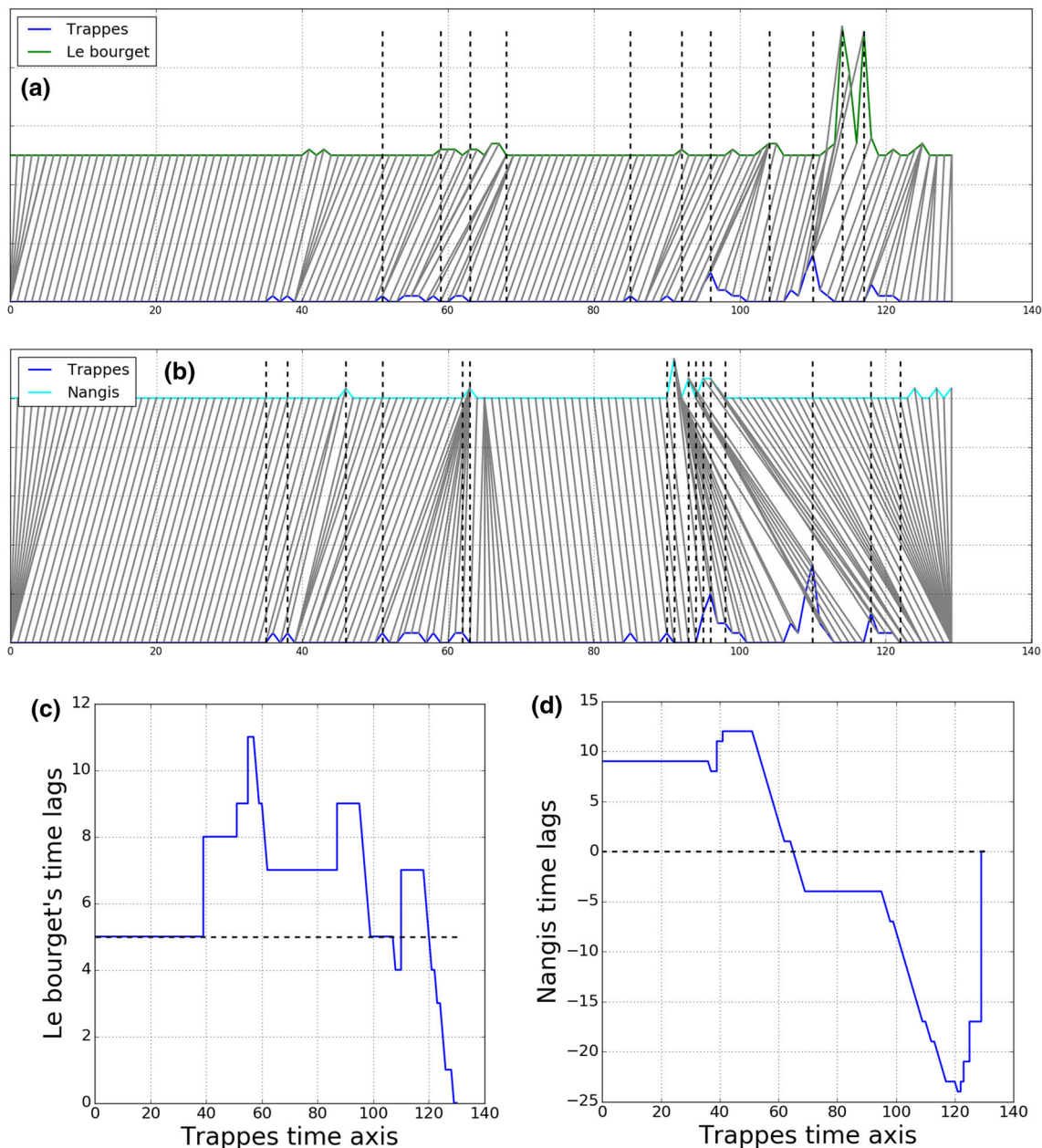
Fig. 10 **a** Warping path for the pair of stations Trappes/Le Bourget. **b** Warping path for the pair of stations Trappes/Nangis. **c** The obtained time lags for the pair of station Trappes and Le Bourget. **d** The obtained time lags for the pair of station Trappes and Nangis

ending times of rainy periods, their peaks and intra/inter-event dry periods independently of the considered situation. In Table 1, column 8 provides the distances $d_{DTW}(S_i, S_j)$ for the six $(S_i, S_j)(S_i, S_j)$ pairs while column 9 provides the correlation coefficients $r_{DTW}(S_i, S_j)$.

When considering a dynamic warping path, the compared time series are not as dissimilar as suggested by the classic approaches. For the six pairs, the dissimilarities $d_{DTW}(S_i, S_j)$ decrease and become smaller than the ones calculated with the two other warping paths (P_{NTD} or P_τ). In the same way, the correlation coefficient increase and

become much more important especially for the three pairs Trappes/Le Bourget, Trappes/Roissy and Le Bourget/Roissy.

For the first three pairs, where the same rain cell is observed, the IMs-DTW was able to spot the similarities between the time series as depicted by the dissimilarities d_{DTW} or the correlation coefficients r_{DTW}. Indeed whatever the considered pair, the dissimilarities are smaller than with P_{NTD} or P_τ warping paths. The same is true for the correlation coefficients, which are higher. For the last three pairs, where the time series represented two different rain cells, the IMs-DTW allows an alignment of the rain patterns. Thus,

Table 2 Time delay τ_{delay} reported from Table 1, the average time difference τ_{ATD}, the corresponding standard deviations τ_{ATD} and the estimated advection velocity for the six pairs (S_i, S_j)

Pair of stations (S_i, S_j)	τ_{delay} (min)	τ_{ATD} (min)	σ_{ATD} (min)	The estimated advection velocity (m s^{-1})
Trappes/Le Bourget	30	37	12.66	15.60
Trappes/Roissy	42	45	12.78	16.29
Le Bourget/Roissy	12	8	4.14	21.90
Trappes/Nangis	− 30	11	66.53	118.98
Le Bourget/Nangis	60	106	51.78	9.49
Roissy/Nangis	60	97	40.08	10.58

Table 3 Main features of the four considered time series

	Trappes	Le Bourget	Roissy	Nangis
Event duration (6 min)	86	85	84	85
Rain amount (mm)	10	17.4	13.2	3.2
Standard deviation of rain rates (mm h^{-1})	0.2536	0.7174	0.44	0.01
Maximum of rain rate (mm h^{-1})	16	44	26	8
Absolute rain rate variation of order 0.5	53.20	66.12	52.46	27.21

the dissimilarities decreased and the correlation coefficients increased but to a lesser extent.

In a general way, the IMs-DTW will find an optimal path minimizing the distance between two time series; however, the question is: Is there a physical sense behind the P_{DTW} warping path? In other words, is it possible to use P_{DTW} to deduce if two rain time series are produced by the same rain cell? In the affirmative, can the deduced delays be linked to the time for the rain cell to move from one station to another? In this case, the dissimilarity $d_{\mathrm{DTW}}(S_i, S_j)$ could be interpreted as a measurement of the spatiotemporal evolution of the same rain cell between the different locations. Indeed, if the assumption of a rain cell displacement is verified, the warping path P_{DTW} should have a mean behavior that is more or less comparable to P_τ. Let us introduce the average time difference delay τ_{ATD} which is the average value of the lags estimated from P_{DTW} during a rain event and σ_{ATD} the associated standard deviation. Table 2 allows comparing τ_{ATD} and τ_{delay} (obtained previously in Table 1) for the six (S_i, S_j) pairs. In addition, the estimated advection velocity is estimated by dividing the geographical distance by τ_{ATD}.

For the three first pairs, as expected the average time difference τ_{ATD} is quite close to the time delays τ_{delay} and the standard deviations σ_{ATD} remains low. In such a way, the coefficients of variations ($\sigma_{\mathrm{ATD}}/\tau_{\mathrm{ATD}}$) are much smaller than unity (respectively, 0.34, 0.28, 0.51). In addition, the estimated advection velocities (Table 2 last column) are quite homogeneous and in agreement with the velocities obtained by radar measurements (not shown). For the last three pairs, τ_{ATD} and τ_{delay} are very different and the standard deviations σ_{ATD} are much higher than previously. This example illustrates that the standard deviation σ_{ATD} and the difference $\tau_{\mathrm{delay}} - \tau_{\mathrm{ATD}}$ are good indicators to decide whether or not the

warping path P_{DTW} is due to the same rain cell and thus can be interpreted as time lags.

In order to characterize and compare the precipitation time series, a large number of features derived from rain rates are commonly used. In Dilmi et al. [37], the authors show that among the many existing parameters a rain event can be fairly well described using only five parameters. Among these parameters, four parameters are commonly used in meteorology or hydrology, namely the event duration, the rain rate peak, the rain event depth and the standard deviation, while the last one called the absolute rain rate variation is less used. The values of these five characteristics are shown in Table 3 for the June 10, 2009, rain event.

Except for the event duration, we observed an important discrepancy of the obtained features. The time series is with the most dissimilar characteristics being that of Nangis. The pair Trappes/Roissy presents the most similar characteristics, and it is interesting to note that the IMs-DTW provides for this pair the lower dissimilarity (1.53 mm h^{-1}) and the higher correlation coefficient r_{DTW}(Trappes,Roissy) (0.91). Trappes and Roissy are the most similar rainfall time series. Concerning the pair Roissy/Le Bourget whose stations are located close to each other (10 km), the maximum rain rate values given by Table 3 suggest that the most intense part of the rain cell has passed over the Bourget station but not above Roissy, giving higher rain rates to the former. Nevertheless, their features remain close as suggested by the IMs-DTW dissimilarity (3.27 mm h^{-1}) and its correlation coefficient r (P_{DTW}, Bourget, Roissy)(0.83).

Contrarily to the previous indicators, commonly used, the IMs-DTW dissimilarity takes the temporality of the rain event into account. Even if, not made explicit, it is well

suited for physical phenomena such as advection (time shift between series) or diffusion (warping of the time series).

Since the rain is a multiscale phenomenon subject to non-stationarity, the time scale is conditioned on the context. The hierarchical fitting of the algorithm enables it to find the similarity where it lies. The algorithm is well balanced since it forbids the released constraint of the FastDTW. Indeed, since rain at a given time scale is conditioned on larger scales, this highlights the coherence through a range of scales. Two time series are similar if they behave coherently both in time scales and in time.

This aspect which is not commonly used is well fitted for rain time series which are known to display multifractal properties.

4 Conclusion

The present study focused on the comparison of rainfall time series through the use of the concept of dissimilarity. Unlike conventional approaches that define a number of features to describe a physical process, the time warping approach provides a measure of dissimilarity between two time series without going through this step. This kind of approach was first developed for signal processing and in particular for speech recognition. The basic concept of dynamic time warping (DTW) is to associate the samples of a time series with those of another by warping the time so that the distance between the two time series is minimal.

Precipitation is characterized by a multifractal behavior leading to strong inhomogeneities and strong variability in rainfall rate. Therefore, at a fine resolution, time series obtained from devices located close enough to each other can differ significantly while maintaining common features such as the overall shape of the event. We showed that the DTW in a multiscale framework (IMs-DTW) by comparing rainfall time series at different time scales allows taking into account (at least partially) these features. Indeed, the multiscale approach of the algorithm that constrains fine-scale associations by the association of sequences on a larger scale is well suited to time series composed of subsequences in which scale relations exist, as in multifractal objects, thus allowing distributions of local statistics not to be identically distributed. As another specificity rain time series encompass a great number of zeros among which a set of non-null rainfall rates belonging to a specific time interval defined as a rainfall event. An important issue for precipitation studies is to ensure that nonzero rainfall rates from a rainfall event are not associated with another event. Again, the multiscale approach helps to distinguish rain events from each other and thus allows to correctly associate the samples in each time series.

When considering intra-event rain samples, the analysis of the warping path provides useful information, especially for rain cells monitoring. Indeed, we have shown that if the same rain cell is at the origin of the two time series the associated warping path has a regular behavior which can be more or less considered as a time shift corresponding to the travel time of the cell rain from one meteorological station to the other. The analysis of the warping path regularity is thus a good indicator to detect a rain cell passing through. When a rain gauge network is available, the calculation of the warping path between each pair of stations allows the identification of the rain cell through the network, and thus, analyzing the warping paths offers a good tool for rain cell tracking. Moreover, the dissimilarity analysis will provide information on the spatiotemporal distortion of the rain cell as it moves.

Finally, since the proposed algorithm is derived from the FastDTW algorithm but with more restrictive constraints, it also significantly reduces the computation time compared to the standard DTW algorithms but also to a certain extent compared to the FastDTW algorithm (radius equal to 1).

This work was primarily dedicated to the ability to use DTW algorithms to provide a pertinent measure of dissimilarity for rainfall time series. In future works, we will focus on applications using the IMs-DTW in the framework of precipitation such as rain cell tracking or rain events clustering [38, 39].

The code source (in Python) and rain gauges data set are available on the GitHub deposit at the following address: https://github.com/djallelDILMI/IMs-DTW.

References

1. Wang, S., Eick, C.F.: A data mining framework for environmental and geo-spatial data analysis. Int. J. Data Sci. Anal. **5**(2–3), 83–98 (2018). https://doi.org/10.1007/s41060-017-0075-9
2. Cristiano, E., Veldhuis, M., Giesen, N.: Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas—a review In Hydrol. Earth Syst. Sci. **21**, 3859–3878 (2017)
3. Verrier, S., de Montera, L., Barthès, L., Mallet, C.: Multifractal analysis of African monsoon rain fields, taking into account the zero rain rates problem. J. Hydrol. **389**(1), 111–120 (2010)
4. Verrier, S., Mallet, C., Barthès, L.: Multiscaling properties of rain in the time domain, taking into account rain support biases. J. Geophys. Res. Atmosp. (2011). https://doi.org/10.1029/2011jd015719
5. Llasat, M.C.: An objective classification of rainfall events on the basis of their convective features. Application to rainfall intensity in the north east of Spain. Int. J. Climatol. **21**, 1385–1400 (2001)
6. Eagleson, P.S.: Dynamic Hydrology. McGraw-Hill, New York (1970)
7. Brown, B.G., Katz, R. W., Murphy, A.H.: Statistical analysis of climatological data to characterize erosion potential: 4. Freezing

events in eastern Oregon/Washington. Oregon Agricultural Experiment Station Spec. Rep. No. 689, Oregon State University (1984)

8. Larsen, M.L., Teves, J. B.: Identifying individual rain events with a dense disdrometer network. Adv. Meteorol. (2015). https://doi.org/10.1155/2015/582782

9. Dunkerley, D.: Rain event properties in nature and in rainfall simulation experiments: a comparative review with recommendations for increasingly systematic study and reporting. Hydrol. Process. **22**(22), 4415–4435 (2008)

10. Dunkerley, D.: Identifying individual rain events from pluviography records: a review with analysis of data from an Australian dryland site. Hydrol. Process. **22**(26), 5024–5036 (2008)

11. Cassisi, C., Montalto, P., Aliotta, M., Cannata, A., Pulvirenti, A.: Similarity measures and dimensionality reduction techniques for time series data mining, advances in data mining knowledge discovery and applications. In: Adem, K. (ed.) InTech. https://doi.org/10.5772/49941 (2012)

12. Keogh, E., Pazzani, M.: Scaling up dynamic time warping for datamining applications. In: Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 285–289. Boston, MA (2000)

13. Goshtasby A.A.: Similarity and dissimilarity measures. Image Registration. Advances in Computer Vision and Pattern Recognition, Book, Chapter 2. Springer-Verlag London Limited (2012). https://doi.org/10.1007/978-1-4471-2458-0_2

14. van Gennip, Y., Hunter, B., Ma, A., et al.: Unsupervised record matching with noisy and incomplete data. Int. J. Data Sci. Anal. **6**(2), 109–129 (2018). https://doi.org/10.1007/s41060-018-0129-7

15. Aghabozorgi, S., Shirkhorshidi, A.S., Wah, T.H.: Time-series clustering—a decade review. Inf. Syst. **53**, 16–38 (2015)

16. Sarda-Espinosa, A.: Comparing Time-Series Clustering Algorithms in R Using the dtwclust Package, R package. https://cran.rproject.Org/web/packages/dtwclust/vignettes/dtwclust.pdf (2017). Accessed 18 June 2018

17. Pearson, K.: Mathematical contributions to the theory of evolution. III. regression, heredity, and panmixia. Philos. Trans. R. Soc. Lond. A Math. Phys. Eng. Sci. 187: 253–318. ISSN 1364-503X. https://doi.org/10.1098/rsta.1896.0007 (1896)

18. Sung, P., Syed, Z., Guttag, J.: Quantifying morphology changes in time series data with skew. In: Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. International Conference on Acoustics, Speech and Signal Processing, pp. 477–480 (2009)

19. Rubner, Y., Tomasi, C., Guibas, L.J.: A metric for distributions with applications to image databases. In: Proceedings of IEEE ICCV, pp. 59–66 (1998)

20. Aronov, B., Har-Peled, S., Knauer, C., Wang, Y., Wenk, C.: "Fréchet distance for curves, revisited. In: ESA'06, London, UK, pp. 52–63, Springer (2006)

21. Huttenlocher, D.P., Klanderman, G.A., Rucklidge, W.J.: Comparing images using the Hausdorff distance. IEEE Trans. PAMI **15**(9), 850–863 (1993)

22. Sakoe, H., Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition. In: IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. ASSP-26 (1978)

23. Zhanga, Z., Tavenardb, R., Baillyb, A., Tangc, X., Tanga, P., Corpetti, T.: Dynamic time warping under limited warping path length. Inf. Sci. **393**, 91–107 (2017)

24. Tsiporkova E.: Dynamic Time Warping Algorithm for. PPT presentation available at: http://www.psb.ugent.be/cbd/papers/gentxwarper/DTWAlgorithm.ppt, date of the last visit (2013)

25. Sakoe, H., Chiba, S.: A similarity evaluation of speech patterns by dynamic programming. Dig. Nat. Meeting, Inst. Electron. Comm. Eng. Japan, p. 136 (1970)

26. Sakoe, H., Chiba, S.: A dynamic programming approach to continuous speech recognition. In: Proceedings of 7th ICA, Paper 20 CI3 (1971)

27. Zinke, A., Mayer, D.: Iterative Multi Scale Dynamic Time Warping. Computer Graphics technical reports, CG-2006/1 (2006)

28. Itakura, F.: Minimum prediction residual principle applied to speech recognition. In: IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. ASSP-23, pp. 52–72 (1975)

29. Bellman, R., Dreyfus, S.: Applied Dynamic Programming. Princeton University Press, New Jersey (1962)

30. Sakoe, H., Chiba, S.: Comparative study of DP-pattern matching techniques for speech recognition. Technical Group Meeting Speech, Acoust.SOC. Japan, Preprints (S73-22) (1973)

31. Muller, M., Mattes, H., Kurth, F.: An efficient multiscale approach to audio synchronization. In: Proceedings of ISMIR, Victoria, Canada, pp. 192–197 (2006)

32. De Montera, L., Barthes, L., Mallet, C., Golé, P.: The Effect of rain-no rain intermittency on the estimation of the universal multifractals model parameters. J. Hydrometeorol. Am. Meteorol. Soc. **10**(2), 493–506 (2009)

33. Akrour, N., Chazottes, A., Verrier, S., Mallet, C., Barthès, L.: Simulation of yearly rainfall time series at microscale resolution with actual properties: Intermittency, scale invariance, and rainfall distribution. Water Resour. Res. Am. Geophys. Union **51**(9), 7417–7435 (2015)

34. Chu, S., Keogh, E., Hart D., Pazzani, M.: iterative deepening dynamic time warping for time series. In: Proceedings of the Second SIAM International Conference on Data Mining, Arlington, Virginia (2002)

35. Salvador, S., Chan, P.: FastDTW: toward accurate dynamic time warping in linear time and space. Intell. Data Anal. **11**(5), 561–580 (2007)

36. Tokay, A., Öztürk, K.: An experimental study of the small-scale variability of rainfall. J. Hydrometeorol. **13**(1), 351–365 (2012)

37. Dilmi, M.D., Mallet, C., Barthes, L., Chazottes, A.: Data-driven clustering of rain events: microphysics information derived from macro-scale observations. Atmos. Meas. Tech. **10**, 1–18 (2017)

38. Truong, C.D., Anh, D.T.: A novel clustering based method for time series motif discovery under time warping measure. Int. J. Data Sci. Anal. **4**(2), 113–126 (2017). https://doi.org/10.1007/s41060-017-0060-3

39. Endo, Y., Toda, H., Nishida, K., et al.: Classifying spatial trajectories using representation learning. Int. J. Data Sci. Anal. **2**(3–4), 107–117 (2016). https://doi.org/10.1007/s41060-016-0014-1