Chapter 56
Novel Quantification Method for Hydrograph Similarity

Dadiyorto Wendi, Bruno Merz, and Norbert Marwan

Abstract We propose an additional elaborate hydrological signature index to quantify similarity (and dissimilarity) between recurring flood dynamics and between observation and model simulation as implied by their phase space trajectories. These phase space trajectories are reconstructed from their corresponding hydrographs (i.e., event time series) using Taken's time delay embedding method. This reconstructed phase space allows multi-dimensional relationship between observation points (i.e., at different time of the event) to be analyzed. Such approach considers the relationships of set of magnitude points in their unique time sequence that are relevant to the complex temporal cascading processes in flood. In a simpler terms, the new index considers the characteristics shape dynamics of a hydrograph and optionally the antecedent discharge conditions that may implicitly cascade to the subsequent rainfall-runoff event and cause an extreme or unusual hydrograph shape. This new similarity index can be used to comprehensively assess the recurrence of extreme event characteristics, change of flood dynamics, shift of seasonality, and as additional metric or objective function to evaluate and calibrate hydrological and hydraulics models.

Keywords Comprehensive hydrograph similarity index · Time sequence characteristics · Shape dynamics · Model performance objective function

D. Wendi (✉)
University of Potsdam, Potsdam, Germany
e-mail: dwendi@gfz-potsdam.de

D. Wendi · B. Merz
GFZ German Research Centre for Geosciences, Section Hydrology, Potsdam, Germany

D. Wendi · N. Marwan
Potsdam Institute for Climate Impact Research, Potsdam, Germany

P. Gourbesville and G. Caignaert (eds.), *Advances in Hydroinformatics*, Springer Water,
https://doi.org/10.1007/978-981-15-5436-0_56

56.1 Introduction

The shape of flood event hydrographs can vary substantially in space (regions) and time (seasons), depending on catchment, river and event characteristics associated with their drivers, e.g. climate factors. For instance, short-rain floods [2]. The dynamic of discharge magnitudes during a rainfall-runoff process would also change depending on its antecedent conditions. For instance if the presence of subsurface water content is high due to the previous event and therefore has already saturated the soil moisture and hydraulics capacity of the catchment, the incoming rainfall would no longer infiltrate the ground but instead run as surface flow (see Fig. 56.1). In such condition, the rising slope of the hydrograph would be steeper due to the flashier nature of surface flow. The hydrograph is therefore a fingerprint of the catchment processes involved in the rainfall-runoff event [2].

The scientific questions whether extreme events (e.g. floods) are increasing and if their regime are changing, are often challenged by the erratic nature of the flood, as one big event does not indicate the trend very clearly. Common data based approach to detect flood changes is the trend analysis of flood discharge peak frequency, should this analysis focuses on the individual station time series, ensemble of stations or regional composite. However, flood peak is merely a single point information found within the event hydrograph and therefore do not describe all necessary information or signatures contained in the process footprint. There are also other hydrological signatures that can be derived from hydrographs such as discharge volume (V), event duration (td), time to peak (tp), recession time (tf), base flow index (BFI), or the rising and falling limb slope of the hydrograph (ΔQrise and ΔQrise). However we think that these signatures are not elaborate either, as each only represents a part or segment

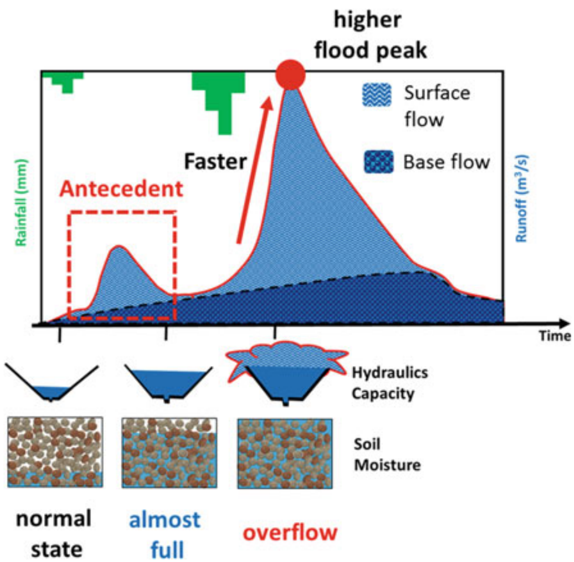

Fig. 56.1 Illustration diagram of the impact of antecedent condition towards flood. Antecedent event (e.g. from the rainfall event prior to the flood) could already saturate the soil moisture and hydraulics capacity and therefore exacerbate the subsequent rainfall input to result as overflow and flashier surface runoff

of the hydrograph, and importantly they are derived through statistical aggregation where uncertainty can be high and may lead to misleading information. An example of this is when deriving a slope index of a multi-peaks hydrograph.

In this study we investigate a more elaborate characteristics of hydrograph based on its time sequence property as an additional hydrological signature in order to compare and distinguish event hydrographs further. Our signature is based on the time delay embedded phase space representation of discharge hydrographs. Instead of using a single element index or a joint considerations of statistical aggregates of the hydrograph signatures, it considers the entire, continuous hydrograph shape, i.e. the time sequence and additionally its dependence on the antecedent conditions of the flood event. It is also to note that the proposed approach is also able to compare hydrographs of different durations. The compared phase space trajectories, i.e. between 2 hydrographs, would then be summarized with a similarity index using cross recurrence plot (CRP) and one of its quantification measure called Determinism (DET) which will be elaborated in the methodology section.

This proposed similarity index can be used to comprehensively assess the recurrence of extreme event characteristics, change of flood dynamics, shift of seasonality, and as additional metric or objective function to evaluate and calibrate hydrological and hydraulics models [5].

56.2 Method

In our approach a phase space trajectory is reconstructed from the corresponding hydrograph using Taken's time delay embedding method (see Eq. 56.1). Please note that the official term 'reconstruct' is used instead of 'construct' because the theory claims that this embedding allows recreating the system behavior, represented by the phase space geometry, by just using the time series of one of the system variables [4].

$$\overrightarrow{X_t} = Q_t, Q_{t+\tau}, \ldots, Q_{t+(m-1)\tau} \tag{56.1}$$

where:

$\overrightarrow{X_t}$ is the reconstructed phase space trajectory of the hydrograph
Q_t is the discharge value at time t
τ and m correspond to the required embedding parameters as time delay and number of embedding dimension

However, we apply the embedding method in a more practical manner, such that the reconstructed phase space trajectory allows the analysis of the multi-dimensional relationship between discharge values in different points in time. This means that we can implicitly consider the cascading impact of antecedent conditions of the flood, i.e. discharge values prior to the flood peak. Figure 56.2 shows the example of such phase space reconstruction, where 3 dimensional phase space is created to represent

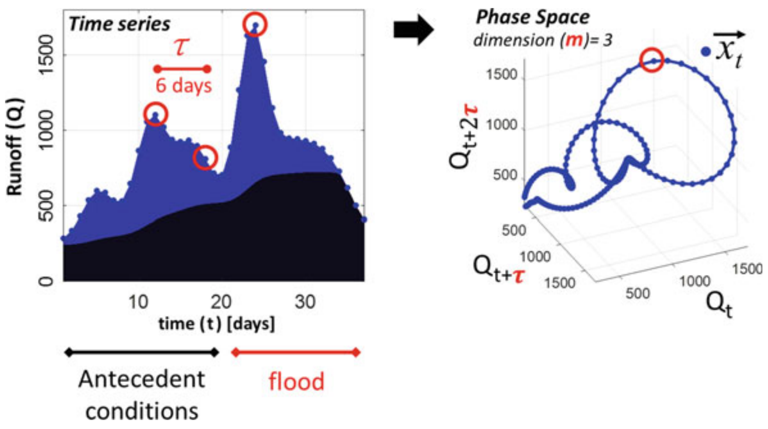

Fig. 56.2 An example of the phase space trajectory of an event hydrograph with daily sampling frequency using 3 dimensional phase space ($m = 3$) and time delay (τ) of 6 days. A single point (red circle) in the phase space in this case represents the relationship of 3 discharge magnitudes of flood peak, 6 and 12 days prior to the peak in the hydrograph

relationship of magnitude sets at 3 different time [5]. The result of this phase space reconstruction is that the geometrical structure representing relationship of discharge magnitudes at different time is highly non-linear and non-monotonic and therefore should not be aggregated in any linear function like autocorrelation. Moreover, the consideration of such relationship is important for flood analysis, as for instance, a moderate rainfall prior to a flood may partially saturate the catchment and lead to a high flood event peak—much higher than would be expected from the event precipitation alone. Further, when no embedding is being applied, i.e. m = 1, the phase space in this case represents purely the 1 dimensional shape of the hydrograph.

In our hydrograph pairwise comparison approach, two phase space trajectories can therefore be reconstructed from two event hydrographs within the same graphs corresponding to different flood periods (see Fig. 56.3 left). These 2 phase space trajectories (i.e. $\overrightarrow{X_i}$ and $\overrightarrow{Y_j}$ corresponding to time i and j) can therefore be directly compared for their similarity or dissimilarity. In order to compare the two trajectories, CRP can be used to identify and visualize the time where the trajectories are considered to be similar [3, 5]. CRP is basically 2-dimensional plot that summarizes the similar trajectories in time of the two flood phase space (see Fig. 56.3 right). This tool is especially useful when working on high dimensional phase space that can no longer be directly visualized. Similar trajectories are defined based on a user-defined phase space distance threshold (ε) and translated into cross recurrence points (CRi,j) corresponding to the time i and j in CRP following the Eq. 56.2.

$$CR_{i,j}(\varepsilon) = \begin{cases} 1, \text{ if } \left\|\vec{x}_i - \vec{y}_j\right\|_2 < \varepsilon \\ 0, \text{ otherwise} \end{cases} \tag{56.2}$$

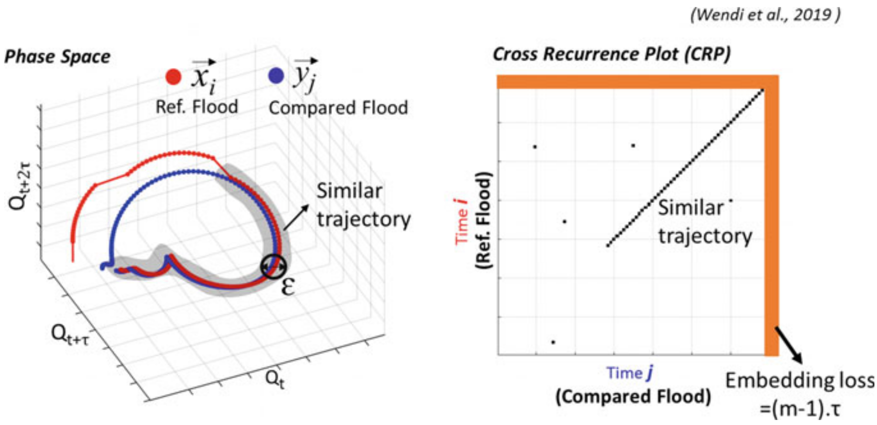

Fig. 56.3 Comparing 2 phase space space trajectories corresponding to different flood events (left), and the CRP summarized similarity (diagonal lines) corresponds to time *i* and *j* where the distances of the two phase space trajectories are within the threshold ε (right)

where $\| \|_2$ is defined as Euclidian Norm.

The degree of similarity found (diagonal lines) represented in the CRP can also be summarized using a DET index. This DET index is calculated as the ratio of relative frequency of diagonal lines in the CRP over all the points in the plot (see Eq. 56.3). This index ranges from 0 to 1 where 0 value indicates no similarity found at all, while the value 1 indicates the full similarity (identical). This index is therefore adapted as our hydrograph similarity index.

$$DET = \frac{\sum_{l=l_{\min}}^{N} l P(l)}{\sum_{i,j}^{N} CR_{i,j}} \tag{56.3}$$

where

l is diagonal line length, $P(l)$ is the relative freq. of l

56.3 Showcase Example

To exemplify the concept of CRP and to demonstrate its differences to the widely used tools scatter plot and cross correlation analysis, Fig. 56.4 compares two hydrographs using scatter plots and CRP. One of the differences between the CRP and the scatter plot is that the scatter plot axes represent magnitudes, while in the CRP they represent the time of occurrence of the two series. When two identical time series are compared, CRP shows a single diagonal line (45° angle) that divides the CRP matrix symmetrically (Fig. 56.4.1c, green line). Further, it is worth to note that the high-dimensional embedding results in embedding loss with a size of (m − 1)τ.

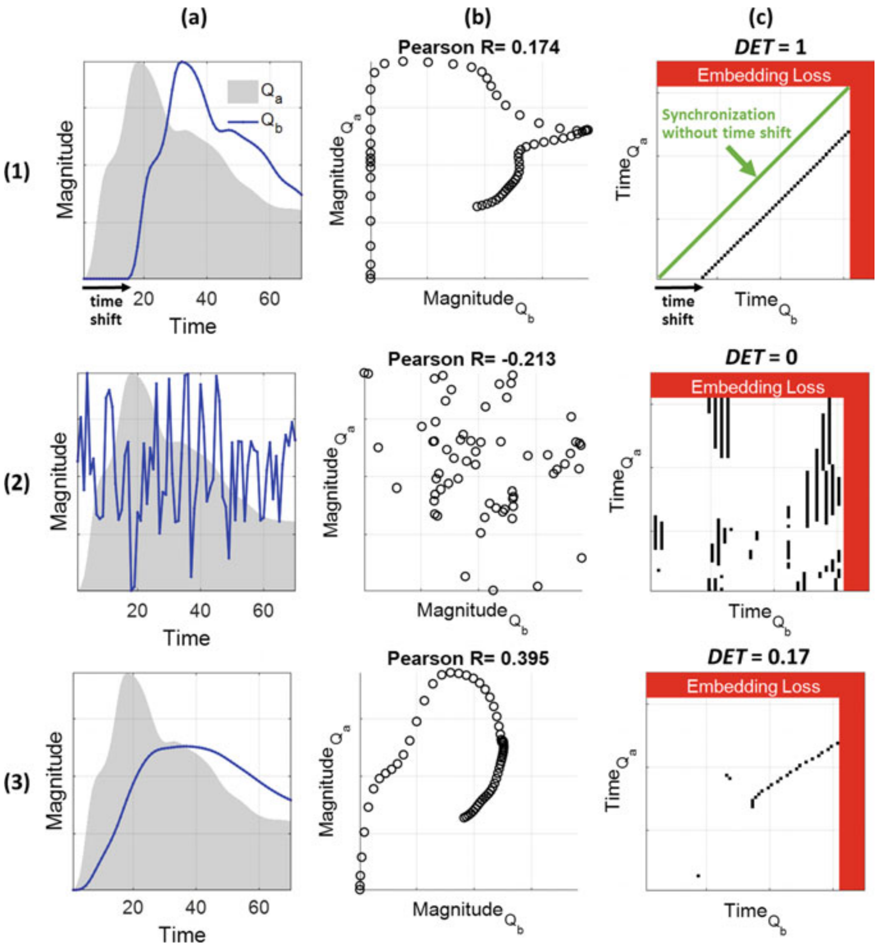

Fig. 56.4 Comparison of two discharge time series (a), with scatter plot and correlation coefficient R (b), and CRP and DET (c). Example (1) compares two identical hydrographs with a shift in their timing, example (2) compares a hydrograph with a randomly shuffled version of the same hydrograph, and example (3) compares two different runoff dynamics where Qb results from a storage-based Muskingum transformation of Qa representing an increased storage capacity in the catchment. The embedding losses are shaded in red. This figure is extracted from Wendi et al. [5]

The first example shows the CRP comparing two identical hydrographs Qa and Qb which occur at different times, i.e. Qb is shifted by 17 time units (Fig. 56.4.1a). Despite the time shift, the CRP still indicates the similarity of the two time series as indicated by the diagonal line and DET value of 1. In addition, it shows where the time shift has occurred. This similarity cannot be derived from the scatter plot without knowing the time shift in advance. Hence, the scatter plot and correlation analysis (Pearson correlation coefficient R = 0.17) could lead to the wrong conclusion that there is no similarity between Qa and Qb. The CRP approach is useful for detecting recurring runoff dynamics, possibly related to the same causative mechanism, which do not necessarily happen at the same season. It should also be noted that the CRP is useful for comparing hydrographs regardless of their, possibly dissimilar, duration.

The second example compares Qa with a random system dynamics where Qb is a randomly shuffled time sequence of Qa (Fig. 56.4.2a–c). In this case, the CRP does not show any diagonal lines and hence a DET value of 0. Similarly, the scatter plot does not indicate a relationship, but it should be cautioned that the correlation analysis suggests a substantial anti-correlation (Pearson correlation coefficient $R = -0.28$).

The third example (Fig. 56.4.3a–c) compares two hydrographs with different runoff dynamics where Qb represents an increased storage capacity in the catchment that dampens the flow. This could result from a perturbation to the catchment such as dam construction. Qb is obtained by a storage-based Muskingum transformation that is commonly used for flow routing in hydrological modelling [1]. The parameters of the Muskingum transformation are set as: storage constant $K = 15$ time units; weighting factor $x = 0.01$. Due to the different runoff dynamics, the resulting CRP shows a rather high dissimilarity in contrast to the substantial correlation coefficient R ~0.4. Although this CRP contains an inclined line, this line is rather broken and not tilted with 45° and therefore summarized with low DET value of 0.17 that does not indicate similarity.

56.4 Conclusions

In this study we propose a novel and elaborate hydrograph similarity measure considering their runoff dynamics. Each event runoff dynamics is characterized by its time delay embedded phase space trajectories. Since the phase space vector is reconstructed from discharge series using time delay embedding, each point of the phase space trajectory contains the relation of several points in time within the event, including for example, the initial conditions caused by antecedent rainfall. The phase space vectors of two events are then analyzed for their similarity using Cross Recurrence Plots (CRP) and evaluated using one of the Recurrence Quantification Analysis measures called Determinism (DET).

The closest concept to this similarity assessment using CRP and DET is a scatter plot between two time series and its correlation coefficient. The comparison between these two concepts demonstrates the benefit of the proposed method. In contrast to scatter plots and correlation analysis, the CRP-based method allows comparing time series of different duration, and it detects similar or identical signals that are shifted in time. The most important benefit stems, however, from the fundamentally different approach of the CRP-based method to quantify similarity based on the multi-dimensional relation of different points in time within an event.

To our knowledge, this is the first application of (Cross) Recurrence Plots and Recurrence Quantification Analysis in hydrology. We believe that these methods have a large application potential in hydrology. These methods could be used to comprehensively compare event hydrographs over time and space, and also calibrate and validate hydrological and hydrodynamic simulation models, by applying them as measure to quantify the agreement between simulation results and observations.

Acknowledgements This research was carried out within the Research Training Group "Natural Hazards and Risks in a Changing World" (NatRiskChange; GRK 2043/1) funded by the Deutsche Forschungsgemeinschaft (DFG) and additionally Helmholtz "Digital Earth" project initiative.

References

1. Hattermann FF, Huang S, Burghoff O, Willems W, Österle H, Büchner M, Kundzewicz Z (2014) Modelling flood damages under climate change conditions – a case study for Germany. Nat Hazards Earth Syst Sci 14:3151–3169. https://doi.org/10.5194/nhess-14-3151-2014
2. Merz R, Blöschl G (2003) A process typology of regional floods. Water Resour Res 39(12). https://doi.org/10.1029/2002WR001952
3. Marwan N, Thiel M, Nowaczyk NR (2002) Cross recurrence plot based synchronization of time series. Nonlinear Process Geophys 9(3/4):325–331. https://doi.org/10.5194/npg-9-325-2002
4. Takens F (1981) Detecting strange attractors in turbulence. In: Dynamical systems and turbulence, Warwick 1980. Lecture notes in mathematics, vol 898, pp 366–381. Springer-Verlag. ISBN 978-3-540-11171-9. https://doi.org/10.1007/BFb0091924
5. Wendi D, Merz B, Marwan N (2019) Assessing hydrograph similarity and rare runoff dynamics by cross recurrence plots. Water Resour Res 55:4704–4726