Chapter 1
Torwards Visual Analytics for the Exploration of Large Sets of Time Series

Mike Sips, Carl Witt, Tobias Rawald and Norbert Marwan

Abstract In this chapter, we discuss the scientific question whether the clustering of time series based on **RQA** measures leads to an interpretable clustering structure when analyzed by human experts. We are not aware of studies answering this scientific question. Answering it is the crucial first step in the development of a **Visual Analytics** approach that support users to explore large sets of time series.

1.1 Introduction

The advancement of recent technology allows scientists to measure different variables of environmental systems at many positions; e.g., the National Oceanic and Atmospheric Administration (NOAA) [1] measures the surface temperature around the globe. The resulting time series encompass a sequence of data points over long time periods at high sampling rates. To study the temporal behavior of an environmental system, scientists need to detect positions with similar temporal dynamics in large sets of time series.

The number and duration of recurrent states in a system is an important aspect of its temporal behavior. A well-established approach to quantify the number and duration of recurrent states is **Recurrence Quantification Analysis (RQA)** [2]. **RQA**

M. Sips (✉) · T. Rawald
German Research Center for GeoSciences GFZ, Potsdam, Germany
e-mail: sips@gfz-potsdam.de

T. Rawald
e-mail: trawald@gfz-potsdam.de

M. Sips · C. Witt · T. Rawald
Humboldt-Universität zu Berlin, Berlin, Germany
e-mail: wittcarl@googlemail.com

N. Marwan
Potsdam Institute for Climate Impact Research, Potsdam, Germany
e-mail: marwan@pik-potsdam.de

© Springer International Publishing Switzerland 2016

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_1

marwan@pik-potsdam.de

calculates quantitative measures that enable scientists to understand the temporal behavior of systems.

In this chapter, we address the scientific question of whether the clustering of time series based on their **RQA** measures produces a clustering structure that is interpretable by human experts. To study whether the grouping of time series based on their **RQA** measures into clusters produces an interpretable clustering structure, we conduct two independent experiments: one experiment with synthetic signals and one experiment with real-world signals. We utilize the **iVAT** method to analyze the clustering structure in these two experiments. We are not aware of similar experiments. Therefore, a positive answer to this question is a critical first step in the development of a **Visual Analytics** approach to support the exploration of large sets of time series. The main contributions of this chapter follow.

- We present an experiment that demonstrate how the clustering of synthetic time series based on their **RQA** results leads to an interpretable clustering structure (Sect. 1.3).
- We show that the clustering structure of the synthetic time series is robust against noise (Sect. 1.3).
- We compute the clustering structure of nine real-world time series and show that the resulting clustering structure is interpretable by human experts (Sect. 1.4).

1.2 Methodology

In this section, we briefly discuss important methods utilized in our study and define important terms of this chapter.

1.2.1 Recurrence Quantification Analysis

Recurrence plots (RPs) and recurrence quantification analysis (RQA) are powerful methods for analyzing recurrences in measured time series [2]. Their application in many fields have proven their potential for various kinds of analyses [3]. A recurrence plot is a two-dimensional representation of a time series when a m-dimensional phase space trajectory recurs to former (or later) states. Recurrence of a state at time i at a different time j is captured within a squared matrix $\mathbf{r}$ [2]:

$$r_{i,j} = \Theta\left(\varepsilon - \left\|\mathbf{x}_i - \mathbf{x}_j\right\|\right), \quad \mathbf{x}_i \in \mathbb{R}^m, \quad i,j = 1 \ldots N. \tag{1.1}$$

Both of its axes represent the time steps. N is the number of considered states x_i (length of phase space trajectory). ε is a threshold distance, $\|\cdot\|$ a norm, and $\Theta(\cdot)$ the Heaviside function. A pair of states that fulfills the threshold condition is assigned with the value 1 (*recurrence point*), whereas a pair that is considered to be dissimilar

marwan@pik-potsdam.de

is assigned with the value 0. Further details about the reconstruction of phase space vectors from a scalar time series, the recurrence parameters, as well as the typical visual characteristics of RPs can be found in [2].

Small scale structures in the RPs, like diagonal lines, are used to define measures of complexity called **Recurrence Quantification Analysis (RQA)** [2, 4, 5]. As an example, we present the RQA measure *percent determinism (DET)*:

$$DET = \frac{\sum_{l=d_{\min}}^{N} l H_{\mathrm{D}}(l)}{\sum_{i,j=1}^{N} r_{i,j}}. \tag{1.2}$$

It is the fraction of recurrence points that form diagonal lines; $H_{\mathrm{D}}(l)$ is the number of diagonal lines of exactly length l and $d_{\min}$ is a minimal length necessary to be a diagonal line. This measure characterizes the deterministic nature of a dynamical system from a heuristic point of view (further discussions can be found in [2, 6]). Further measures quantify average line lengths or the complexity of the line length frequency distributions $H_{\mathrm{D}}(l)$ (diagonal lines) and $H_{\mathrm{V}}(l)$ (vertical lines).

1.2.2 VAT and iVAT

Visual Assessment of Clustering Tendency [7] is a method to depict the cluster tendency of a set of data objects. The input to the VAT method is usually the pairwise dissimilarity matrix of the data points. The basic idea of the **VAT** method is to rearrange the rows and columns of the dissimilarity matrix in such a manner that it depicts the clustering structure of the data objects. **VAT** visualizes the reordered matrix as intensity image utilizing gray levels. Pure white depicts the largest dissimilarity value and pure black depicts zero dissimilarity in the intensity image. Users utilize the **VAT** method to determine whether clusters are present before applying particular clustering algorithms to their data.

To reorder the rows and columns, VAT computes a permutation of the original row and column indices of the dissimilarity matrix. The algorithm to compute this permutation is similar to Prim's algorithm [8]. The basic idea is to interpret the dissimilarity matrix as a weighted graph of the input data points; i.e., the nodes of the graph are the data points that are connected to each other. The dissimilarity value denotes the distance between the nodes. **VAT** starts with the node v that has the largest distance to another node. The row index of v is the first element of the permutation. It then searches for the node w that has the smallest distance to v. The column index of w is the second element of the permutation. Next, VAT searches for the node x that has the smallest distance to either v or w. The column index of x is the next element of the permutation. This algorithm terminates if all nodes has been visited.

marwan@pik-potsdam.de

This mechanism is responsible for the emergence of black blocks along the diagonal of the reordered matrix. Without this proper reordering of the rows and columns, it is essentially impossible to assess the clustering tendency in the intensity image (we refer to [7] for more detailed discussion).

Improved VAT (iVAT) [9, 10] method is an extension of the **VAT** algorithm, and transforms the dissimilarity matrix using a graph theoretic measure. It utilizes the VAT algorithm on the transformed matrix to improve the rearrangement of rows and columns.

1.2.3 Definitions

The following definitions are valid throughout the chapter.

Clustering Structure Given a set of time series. A clustering structure is a set of distinct groups where (a) each group consists of a subset of time series and (b) each time series is assigned to exactly one group.

Interpretable Clustering Structure We call a clustering structure interpretable by human experts if (a) an expert identifies a cluster as the time series or recurrence plots of a particular system, and (b) the expert considers the group members as similar to each other.

1.3 Experiment with Synthetic Signals

To study whether the grouping of synthetic time series based on their **RQA** measures into clusters produces an interpretable clustering structure, we do not introduce independent variables into the experimental setup. In the next step, we extend the experiment by introducing one independent variable. This independent variable is noise ratio. Note: the clustering structure is always the dependent variable in the experiments.

1.3.1 Experimental Setup

The set of time series is finite and well-known. We utilize nine systems to generate the set of time series and compute the data points for each time series according to the equation and parameters listed in Table 1.2. To generate data points of time series representing the periodic motion system, λ traverses the unit circle three times and each circulation generates at least 16 data points. Figure 1.1 presents representative RPs of the utilized systems.

marwan@pik-potsdam.de

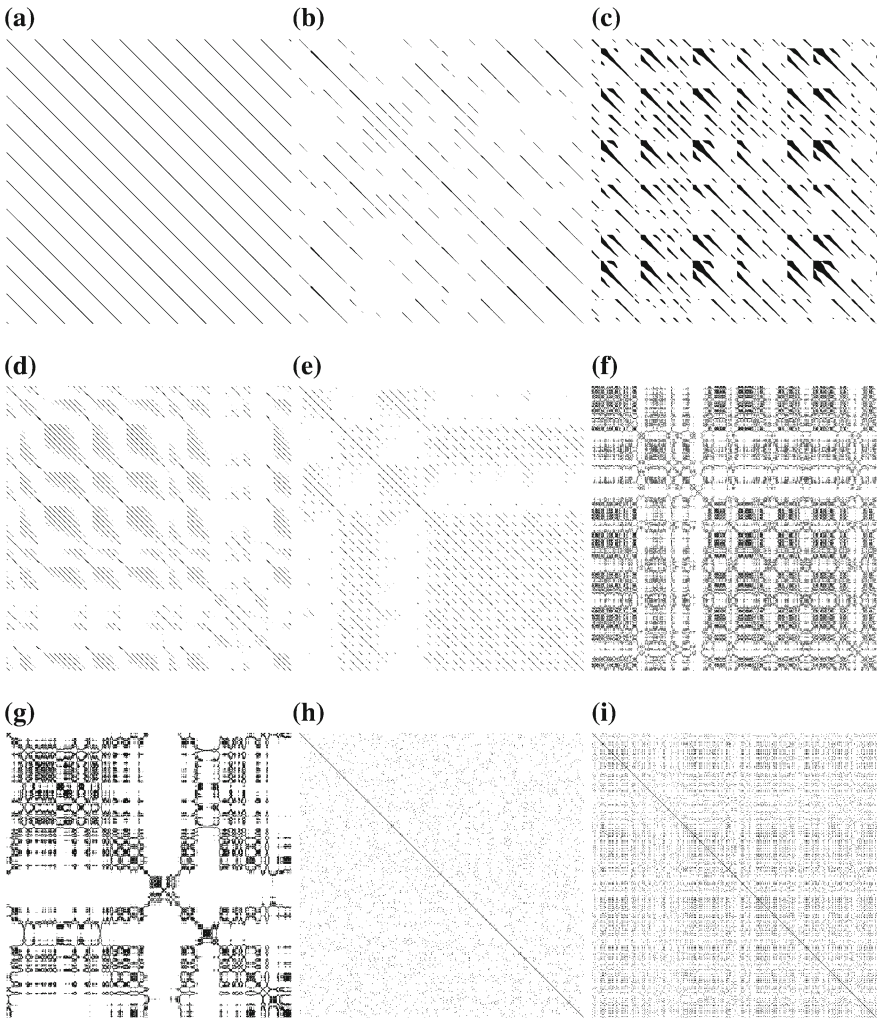


Fig. 1.1 Example RPs of the nine systems utilized in the experiments. **a** OSC. **b** R. **c** RF. **d** L. **e** LO. **f** AR−. **g** AR+. **h** N. **i** NG. Table 1.1 resolves the abbreviations

We generate 100 time series for each system. Hence, the number of time series we use in the experiment is 900. This set represents a broad range of well-known temporal behavior. Each time series has 25,000 data points. To generate the time series for the R, RF, L, LO, we generate $100 \times 25{,}000 + 1000$ data points. We discard the first 1000 data points and group the remaining data points into 100 time series. To compute time series for the other systems, we change the seed of the random number generator for each time series.

marwan@pik-potsdam.de

Table 1.1 Parameter assignments of the RQA method utilized in the experiments

RQA parameter	Assignment
Embedding dimension	1
Time delay	0
Recurrence threshold	5 % of the trajectory diameter
Minimal black diagonal and vertical line lengths	2
Minimal white diagonal and vertical line lengths	2
Theiler corrector	1

To minimize variations in the clustering structure introduced by the **RQA** method, we keep all **RQA** parameters at predefined values. Table 1.1 summarizes the parameter assignments of this experiment. An adaptive choice of the **RQA** parameters may improve the clustering structure. To study the impact of certain **RQA** parameters on the clustering structure, we extend this experiment in the next step and introduce noise ratio as independent variable.

Note that our experiment contains the two noise systems white noise (N) and gamma noise (NG). The reason for including these two noise systems is to see whether the clustering structure also separates the noise systems in two distinct clusters.

1.3.2 Experimental Procedure

First, we compute 16 **RQA** measures (see Table 1.3) for each time series, and store these measures in a 16-dimensional vector. To minimize variations introduced by the different dynamic ranges of the measures, we normalize each value of the 16-dimensional vectors. Let $\{v_i \in \mathbb{R}^{16} \quad i = 1 \ldots 900\}$ be the set of all 16-dimensional vectors. We decided to normalize the i-th value $v_i[i]$ with $\max_{j=1\ldots900}\{v_j[i]\} - \min_{j=1\ldots900}\{v_j[i]\}$.

We utilize the Euclidean distance to determine the dissimilarity between two vectors. This means that small distances denote low dissimilarities between time series, and large distances denote high dissimilarities between time series. Second, we compute all pair-wise dissimilarity values between the **RQA** vectors and store these values in a 900×900 matrix. This matrix is the input to the **iVAT** method. The **iVAT** method determines an ordering of the 900 time series that depicts the clustering structure of the time series.

To see whether the clustering structure is in line with expert opinion, in addition we conducted an experiment with two experts. In this experiment, the expert determines the dissimilarity between time series by comparing their RPs. To this end, we present two RPs to the expert who determined the dissimilarity between them at a scale between zero and 100; zero denotes a low dissimilarity and 100 denotes

marwan@pik-potsdam.de

a high dissimilarity. The expert determined the pair-wise dissimilarity between 45 RPs. We randomly selected 36 pairs of RPs from different systems, and nine pairs from the same system. These RPs were randomly selected from the set of 900 time series described above. We decided to ask the expert to make 45 comparisons of RPs because the attention of the human expert is the limiting factor in this experiment. We stored these pair-wise dissimilarities in a 45×45 matrix. This matrix is the input to the **iVAT** method.

Note: the pair-wise dissimilarity values are different in the experiments, and therefore, the arrangement of rows and columns may differ in the resulting intensity images.

1.3.3 Discussion

Figure 1.2a shows the result of the **iVAT** method computed from **RQA** measures. Dark gray colors denote high similarities between time series and light gray colors denote low similarities between time series.

We see nine well-separated black blocks along the diagonal of the **iVAT** matrix in Fig. 1.2a. Each black block contains 100 time series, and the black color indicates that the members of each block are almost indiscernible based on their **RQA** measures. Furthermore, each block contains time series from the same system only. The labels of the blocks are listed in Table 1.2. The cluster OSC that represents the time series of periodic motion, for example, is composed of the 100 time series of periodic motion only. We conclude from the nine well-separated black blocks in Fig. 1.2a that **RQA** measures consider the time series of one system as similar to each other; each

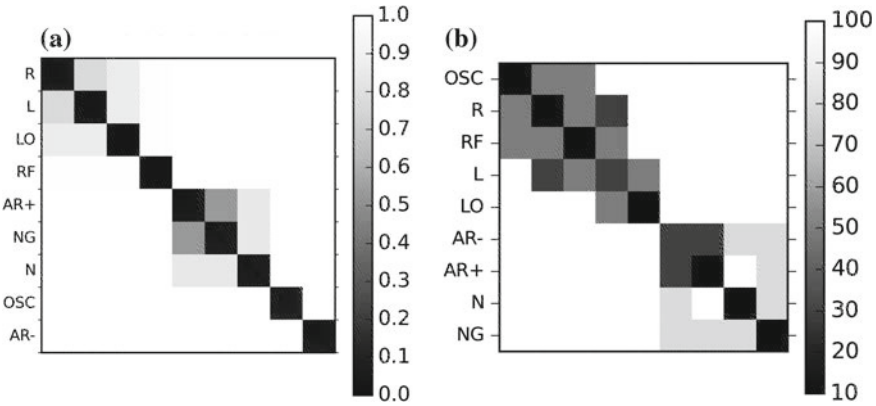


Fig. 1.2 **iVAT** matrix depicting the clustering structure of the experiment. Each *black block* in this matrix contains 100 time series of a particular system. *Note* the different orderings of the systems along the y-axis are due to the **iVAT** method and cannot be changed to maintain the visual cluster tendency. **a** RQA measures. **b** Expert opinion

marwan@pik-potsdam.de

Table 1.2 The nine systems utilized in the experiments and their initial conditions

System	Abbr.	Equations	Parameterization
Periodic motion	OSC	$x(t) = \cos(\lambda t)$, $y(t) = \sin(\lambda t)$	See text
Rössler system, standard parameters	R	Rössler equations	a = 0.2, b = 0.2, c = 5.7, Δt = 0.085
Rössler system, funnel regime	RF	Rössler equations	a = 0.2925, b = 0.1, c = 8.5, Δt = 0.085
Lorenz system, standard parameters	L	Lorenz equations	$\sigma = 10$, $\beta = 2.66$, $\rho = 28$, Δt = 0.025
Lorenz system, noisy oscillation	LO	Lorenz equations	$\sigma = 10$, $\beta = 2.66$, $\rho = 198$, Δt = 0.025,
Autoregressive, negative correlation	AR−	$x(t) = \phi x(t-1) + \mathcal{N}(0, 0.1)$	$\phi = -0.95$
Autoregressive, positive correlation	AR+	$x(t) = \phi x(t-1) + \mathcal{N}(0, 0.1)$	$\phi = 0.95$
Normal noise	N	$x(t) \sim N(\mu, \sigma^2)$	$\mu = 0, \sigma = 1$
Gamma noise	NG	$x(t) \sim \Gamma(k, \theta)$	$k = 0.5, \theta = 1$

block contains the 100 time series of one system only. Furthermore, **RQA** measures consider time series from different systems as dissimilar to each other.

Figure 1.2b shows the result of the expert assessment. We present the result of one expert here; the results among the two experts are similar to each other. Again, dark gray colors denote high similarities between time series and light gray colors denote low similarities between time series. Note, the dissimilarity values range from 0 to 100. We see seven well-separated black blocks along the diagonal of the **iVAT** matrix in Fig. 1.2b. Furthermore, we observe similar properties of these seven clusters in comparison to the nine clusters produced by **RQA** measures. We identify the seven blocks OSC, R, RF, LO, AR+, N, and NG from Fig. 1.2a also in Fig. 1.2b. Each of these blocks contain 100 time series of one system only.

The **iVAT** result suggests that time series of the L system are in the R block, and members of the AR− system are in the AR+ cluster (see Fig. 1.2b). L and R are chaotic systems, and thus, both belong to the same class of systems. AR+ and AR− are correlated noise systems, and thus, both belong to the same class. We argue that the mixing of L with R and AR+ and AR− does not constitute a problem. The interesting point here is that the expert can clearly see the similarity between these systems in the RPs but the quantitative properties of these examples are much more different. This leads to distinct clusters for each of these systems. Nevertheless, there is a high degree of consensus between the two clustering structures of Fig. 1.2a, b. We conclude from the comparison of Fig. 1.2a, b that the clustering structure of time series based on their **RQA** measures is interpretable.

marwan@pik-potsdam.de

1.3.4 Potential of Clustering Algorithms

In this section, we discuss the clustering structure that is likely to be detected by clustering algorithms. Figure 1.2a shows that clustering algorithms experience difficulties to derive exactly the nine clusters as from our visual inspection of the **iVAT** matrix. The reason for this are the light gray blocks around some of the black blocks. The gray box around the R, L, and LO clusters indicate that some time series of the R system are more similar to some time series of L and/or LO systems. This is a reasonable result since these systems belong to the same class, the chaotic systems. A similar observation holds for AR+, NG, and N systems. These systems belong to the same class, the stochastic systems.

To estimate the clustering structure, we assume that clustering algorithms cannot separate clusters in Fig. 1.2a that are close to each other. Hence, the time series of these clusters will be merged into a bigger cluster; together with time series from other systems. Based on this assumption, we conclude from Fig. 1.2a that clustering algorithms are likely to derive the following clusters: c1 = (R, L, LO), c2 = (RF), c3 = (AR+, NG, N), c4 = (OSC) and c5 = (AR−).

To see whether this clustering structure is interpretable, we compare our estimated clustering structure with the clustering structure based on expert opinion (see Fig. 1.2b). We see that experts judge time series of the R, and L systems to be similar to each other. We argue that the merging of these two systems into a common cluster does not constitute a problem since this grouping is in line with expert opinion. The LO system needs a careful discussion. Figure 1.2b shows that LO contains almost 100 time series of the LO system. The ordering along the diagonal of the matrix suggest that the LO cluster is close to the R and L clusters. We conclude from this observation that the cluster c1 is interpretable.

An interesting observation is that clustering algorithms are likely to experience difficulties to separate noise systems N, NG and autoregressive systems (positively correlated). According to expert opinion, the systems AR− and AR+ are similar to each other. In contrast, the system AR+ shares some similarity to the systems NG and N according to expert opinion. We argue that this does not constitute a problem since cluster c3 = (AR+, NG, N) is still in line with expert opinion. We conclude from these observations that the clustering structure computed by clustering algorithms are likely to be interpretable.

1.3.5 Effect of Noise

We extend the experiment to study the effect of noise on the clustering structure of the nine systems. We utilize noise ratio as the independent variable because we expect to see that time series represent certain noise signals. In this experiment, the experimental setup is the same as in the experiment reported above:

marwan@pik-potsdam.de

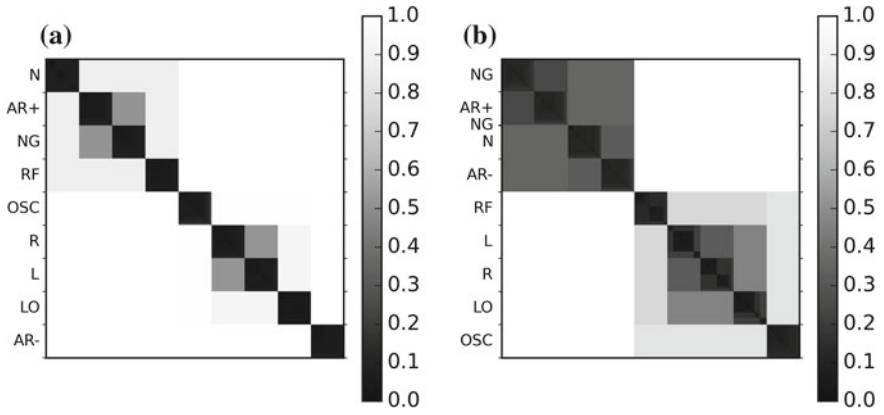


Fig. 1.3 **iVAT** matrix depicting the clustering structure of noise ratio 0.2 and 0.8. Each *black block* in this matrix contains 100 time series of a particular system. *Note* the different orderings of the systems along the y-axis are due to the **iVAT** method and cannot be changed to maintain the visual cluster tendency. The NG label appears two times in the determined ordered sequence in (**b**). It is difficult to see in the **iVAT** matrix that a few time series from the NG system are located between AR+ and N. **a** 0.2. **b** 0.8

- **RQA** parameters at predefined values (see Table 1.1).
- each time series has 25,000 data points. In this experiment, we use the time series of this experiment (see Table 1.2).
- 16 **RQA** measures from Table 1.3.

In contrast to our first experiment reported above, we add Gaussian white noise to each of the 900 time series to generate different noise ratios (our noise model is similar to [11]). We generate six sets of 900 time series representing noise ratios of 0.2, 0.4, 0.6, 0.8, 0.9, 1.0. Hence, we generate 5400 time series. The experimental procedure for each set is the same as in the experiment reported above. Again, the pair-wise dissimilarity values are different in our experiment, and therefore, the arrangement of rows and columns may differ in the resulting intensity images.

Figure 1.3a shows the **iVAT** matrix for noise ratio 0.2. This matrix is similar to the **iVAT** matrix in our initial study (see Fig. 1.2a). We observe similar results for noise ratios 0.4 and 0.6., and we conclude from this result that the clustering structure is robust against low noise ratios, i.e., 0.2, 0.4 and 0.6.

For high noise ratios, the situation becomes more complicated. Figure 1.3b shows the **iVAT** result for noise ratio 0.8. We see a noticeable difference compared to the result of the first experiment. Although we see the nine clusters along the diagonal, the black blocks are not homogeneous as with the first experiment (we see tiny block structures with each system). Furthermore, dark gray blocks capture some of the black blocks. These dark gray blocks make it difficult to decide whether NG, N, AR+, and AR− form distinct clusters. Figure 1.3b indicates that NG, N, AR+, and AR− are likely to be members of the same cluster. This is a reasonable result since

marwan@pik-potsdam.de

Table 1.3 RQA measures of the experiment

Description	Formula	Diagonal	Vertical	White
Filter ratio	$\frac{\sum_{l=l_{\min}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)}$	DET	LAM	–
Average length	$\frac{\sum_{l=l_{\min}}^{N} lP(l)}{\sum_{l=l_{min}}^{N} P(l)}$	L	TT	W_{mean}
Maximum length	$\max(\{l \mid P(l) > 0\})$	$L_{\max}$	$V_{\max}$	$W_{\max}$
1/Maximum length	See above	DIV	–	W_{div}
Entropy	$-\sum_{l=l_{\min}}^{N} p(l) \ln p(l)$	L_{entr}	V_{entr}	W_{entr}

In addition to these measures, we also utilize the recurrence rate (RR) as well as the ratios DET/RR and LAM/DET

these systems belong to the same class, the stochastic systems. We see a similar result for L, LO, and R systems. However, these results are still in line with expert judgment (see Fig. 1.2b). We conclude from these observations that the clustering structure is interpretable for the noise ratio 0.8 (Fig. 1.3b).

1.4 Experiment with Real-World Signals

In this section, we study whether the clustering of real-world time series based on their **RQA** measures produces an interpretable clustering structure, when compared to standard climate classification. In this experiment, we utilize climate time series ("Quality Controlled Local Climatological Data") from the National Oceanic and Atmospheric Administration (NOAA). NOAA provides a variety of time series for stations across the United States.

We select nine stations from three different zones; three stations per zone (see Table 1.4). According to the Köppen climate classification [12], the stations either

Table 1.4 Station names and their climate

Label	Location of station	Climate zone
1	McCarran International Airport, Nevada	BWh
2	Tucson International Airport, Arizona	BWh
3	Phoenix Sky Harbor International Airport, Arizona	BWh
4	Dickinson Municipal Airport, North Dakota	Dfb
5	Minot International Airport, North Dakota	Dfb
6	Bismarck Municipal Airport, North Dakota	Dfb
7	Winkler County Airport, Texas	BSk
8	Lubbock International Airport, Texas	BSk
9	Midland International Airport, Texas	BSk

The corresponding time series are depicted in Fig. 1.4. BWh = hot desert climate, Dfb = humid continental climate and BSk = semi-arid climate

marwan@pik-potsdam.de

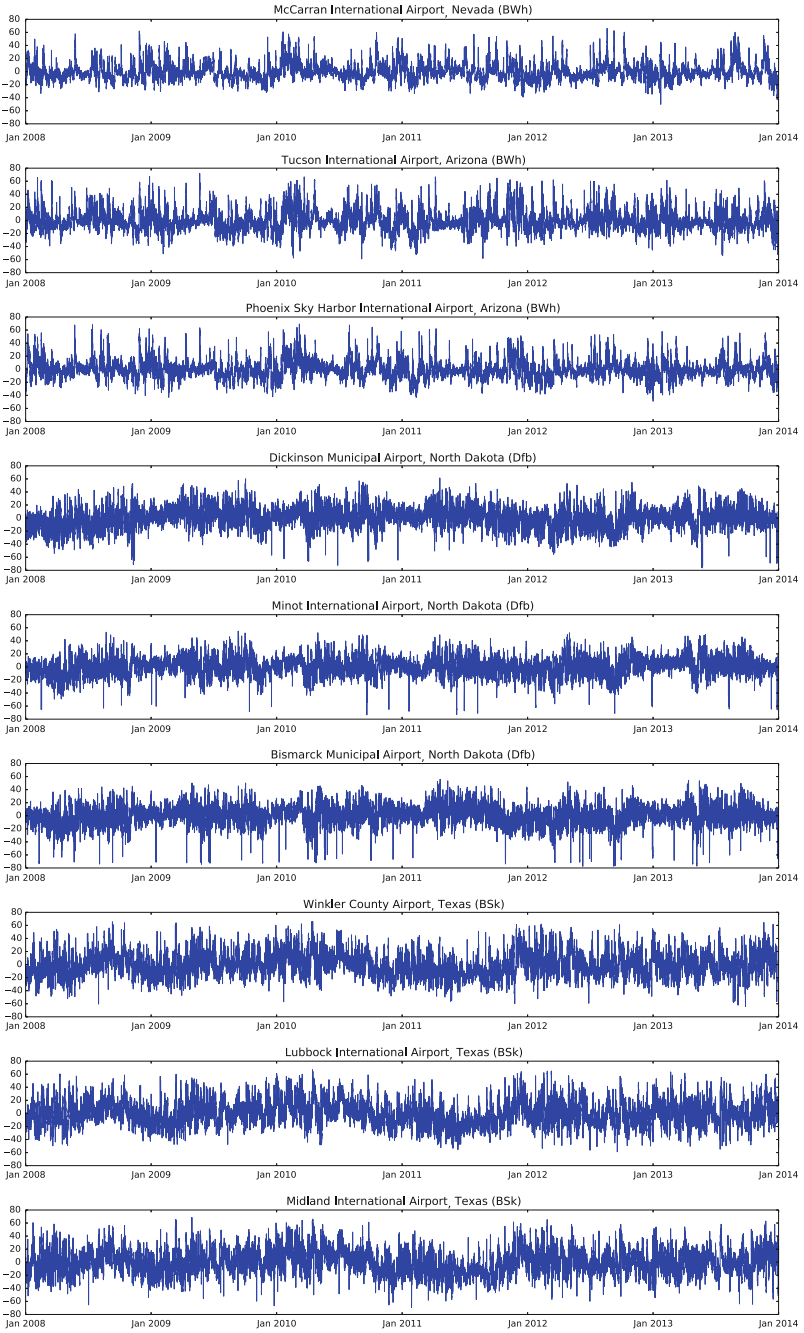


Fig. 1.4 Time series of the NOAA stations utilized in our experiment

marwan@pik-potsdam.de

belong to the humid continental climate (Dfb), semi-arid climate (BSk) or the hot desert climate (BWh). The aim of this experiment is to see whether the **RQA** measures will group stations belonging to the same climate into a common cluster.

In this experiment, we consider time series from January 2008 to December 2013 at an hourly resolution that capture the relative humidity measured at each station. We exclude the annual and the daily humidity cycle from the time series. Figure 1.4 presents the time series utilized in this experiment. To compute the **RQA** measures, we use the parameter assignments (see Table 1.1) and the same normalization procedure of the experiment with synthetic signals.

Figure 1.5 depicts the **iVAT** matrix for the nine stations from Table 1.4. We see three clusters along the diagonal of the **iVAT** matrix (highlighted with red bounding boxes). A close inspection of these clusters revealed that each cluster covers time series from the same climate only. Hence, the clustering structure groups stations belonging to the same climate into the same clusters.

Comparing the similarities of the stations belonging to the BWh climate, it becomes apparent that the temporal dynamics of the time series captured at the McCarran International Airport in Nevada are different to the remaining time series from the same climate. A reason for this behavior may be the differences in geographic location. Generally being a dry area, the airport is located in the Mojave desert. Nevertheless, its dynamics are more similar to the time series belonging the BWh climate than to those belonging to different climate zone, as depicted in the **iVat**.

Based on these findings, we conclude that the clustering structure of the nine stations is interpretable.

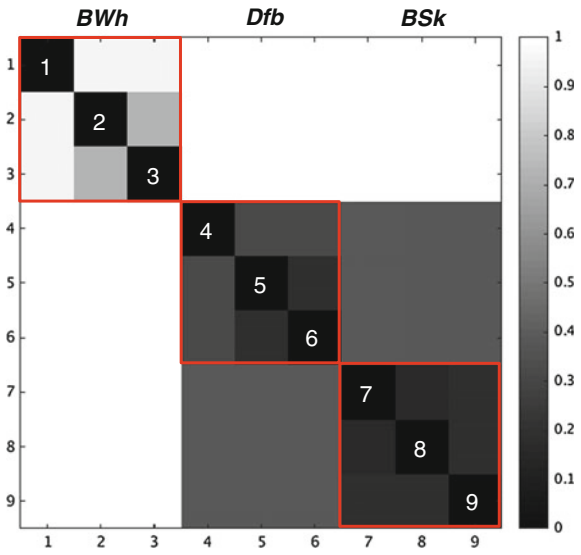


Fig. 1.5 **iVAT** matrix for the nine stations from NOAA. The *three red blocks* along the diagonal represent distinct clusters of the climate zones. Each *red block* contains three time series. Table 1.4 lists the station names for each diagonal item (labels are identical). BWh = hot desert climate, Dfb = humid continental climate and BSk = semi-arid climate

marwan@pik-potsdam.de

1.5 Future Lines of Research

We report these experiments for two reasons. First, answering the scientific question of this chapter is critical for our **Visual Analytics** approach. In close collaboration with experts in the field, we develop a **Visual Analytics** approach that supports users to explore large sets of a time series to identify regions with similar temporal dynamics. **Visual Analytics** has the potential to extent the users toolbox, since users often need to specify in advance what constitutes similar temporal dynamics when trying to study large sets of time series. The positive answer to the scientific question whether the clustering of time series based on their **RQA** measures facilitates the further development of our **Visual Analytics** approach.

We also plan to extend our study to multi-scale algorithms. The basic idea is to group the multi-scale components of time series into clusters of similar temporal dynamics based on their **RQA** measures and to explore this clustering structure. We argue that an exploration of this clustering structure supports users to gain a better understanding of the complex temporal behavior of systems. This study is crucial for the development of a **Visual Analytics** approach for Multi-Scale RQA.

1.6 Conclusion

The scientific question discussed in this chapter is whether the clustering of time series based on their **RQA** measures leads to an interpretable clustering structure when analyzed by human experts. To address this scientific question, we described a first experiment in which we do not introduce independent variables. In this experiment, we utilized nine well-known dynamic systems and 16 **RQA** measures. We generated 100 time series for each of the nine systems, and each time series had 25,000 data points. The dependent variable was the clustering structure.

Furthermore, we calculated the Euclidean distance between the **RQA** vectors created based on those time series and stored these pairwise distances in a matrix. We then utilized the **iVAT** method to uncover the clustering structure. The **iVAT** matrix in the first experiment shows nine distinct and well-separated clusters along the diagonal of the **iVAT** matrix. To see whether this clustering structure is interpretable, we compare our estimated clustering structure with the clustering structure based on expert opinion. We concluded from this comparison that **RQA** measures produce an interpretable clustering structure for a synthetic data set. To estimate the result of clustering algorithms, we assumed that close clusters are likely to be merged into a common cluster. The estimated clusters are in line with expert opinion.

We extended the experiment, and introduced one independent variable. We utilized noise ratio as independent variable. The experimental setup and procedure were similar to the first experiment. The **iVAT** results show that the clustering structure is robust against noise ratios up to 0.8. In addition, we conducted an experiment with the embedding dimension as the independent variable. Again, the experimental

marwan@pik-potsdam.de

setup and procedure were similar to the first experiment. Our result suggested that the embedding dimension has only a minor effect on the clustering structure of the synthetic time series.

In addition, we conducted an experiment with real signals. We select nine stations from three different climate zones; according to the Köppen climate classification. The **iVAT** matrix shows a well-organized clustering structure for these real-world time series. Time series from the same climate are grouped into a common cluster. We concluded from this observation that the clustering structure of the nine stations is interpretable.

Finally, our experiments to determine the similarity between RP's with Hamming distance and Spatiogram distance as alternative approaches to **RQA** vectors did not yield better results.

References

1. National centers for environmental information. Quality controlled local climatological data (2015). http://www.ncdc.noaa.gov/qclcd/QCLCD?prior=N
2. N. Marwan, M.C. Romano, M. Thiel, J. Kurths, Phys. Rep. **438**(5–6), 237 (2007)
3. N. Marwan, Eur. Phys. J. Spec. Top. **164**(1), 3 (2008)
4. J.P. Zbilut, C.L. Webber Jr., Phys. Lett. A **171**(3–4), 199 (1992)
5. C.L. Webber Jr., J.P. Zbilut, J. Appl. Physiol. **76**(2), 965 (1994)
6. N. Marwan, S. Schinkel, J. Kurths, Europhys. Lett. **101**, 1 (2013)
7. J. Bezdek, R. Hathaway, in *Proceedings of the 2002 International Joint Conference on Neural Networks IJCNN '02*, vol. 3, pp. 2225–2230 (2002)
8. R.C. Prim, Bell Syst. Tech. J. **36**(6), 1389 (1957)
9. L. Wang, U. Nguyen, J. Bezdek, C.A. Leckie, K. Ramamohanarao, in Advances, in *Knowledge Discovery and Data Mining*, vol. 6118, Lecture Notes in Computer Science, ed. by M. Zaki, J. Yu, B. Ravindran, V. Pudi (Springer, Berlin Heidelberg, 2010), pp. 16–27
10. T.C. Havens, J.C. Bezdek, IEEE Trans. Knowl. Data Eng. **24**(5), 813 (2012)
11. M. Thiel, M. Romano, J. Kurths, R. Meucci, E. Allaria, F. Arecchi, Physica D **171**(3), 138 (2002)
12. M.C. Peel, B.L. Finlayson, T.A. McMahon, Hydrol. Earth Syst. Sci. **11**(5), 1633 (2007)

marwan@pik-potsdam.de