Recapitulation of Recurrence Theory and Practice

Charles L. Webber, Jr., Ph.D
Professor Emeritus
Department of Cell and Molecular Physiology
Loyola University Chicago, Health Sciences Division
2160 South First Avenue
Maywood, IL 60153
U.S.A.

cwebber@LUC.edu
http://homepages.LUC/~cwebber/

Keywords: Recurrence plots; recurrence quantification analysis, principle component analysis

Abstract: In broad strokes this chapter examines the roots of recurrence quantification analysis (RQA) and looks forward into its future. Much has already been written about the theory and application of recurrence methodology to real-world and complex systems. The overarching purpose of this present contribution is to introduce recurrence quantification as a unifying tool that has the timely potential of bringing together numerous and diverse fields of science under a common rubric of dynamical systems. The simple idea is that all scientists are studying “linguistic” systems that possess unique “vocabularies” and varied “grammars.” But RQA can cut through these differences and encourage researchers in different research fields and scattered geographical locals to start talking seriously about the similarities of otherwise apparently diverse systems.

1. Context

The concept of recurrence has a long mathematical history that can be traced back at least to Poincare (1880) of the late 19th century and more recently to Feller (1950) of the mid 20th century. Poincare may be considered the “father” of nonlinear dynamics, whereas Feller was a mathematician focusing on probability theory. As we will see, recurrence analysis is actually a statistical tool that can “diagnose” any moving system, but we are getting ahead of ourselves. Systems can be considered as recurrent if dynamic trajectories describing those systems return exactly to (periodic) or nearby (chaotic) their starting states. According this definition, recurrent systems possess periodicities or quasi-periodicities which keep such systems intact and functioning. Any active system with strong recurrences can be considered to be in some type of steady state or quasi-steady state. However, any system exhibiting weak recurrences is either transitioning to another steady state or is moving toward dissolution (cessation). And theoretically, systems possessing no or at best unstructured recurrences are no systems at all in the classical sense, and are better understood as random processes (which nonetheless may have high-dimensional structuring that are undetermined as yet).

In the second half of the 20th century three key mathematical breakthroughs (greatly aided by the rise of digital computers) have preceded the inception of modern recurrence quantification analysis (RQA). First, in 1963 Lorenz (1963) (13,559 citations) while modeling weather systems found that small changes in initial conditions in his three-variable system led to unpredictable dynamic outcomes. This modern finding was reminiscent of Newton’s three-body problem which Poincare proved had no analytical solution. Thus was born the famous Lorenz strange attractor which digital and analog computers could easily display. Besides displaying sensitivity to initial conditions, strange attractors have return trajectories that are very close but non-exact. Second, in 1981 Takens (1981) (8,206 citations) published results from his work on strange attractors in fluid turbulence. His huge contribution was the mathematical proof that dynamical systems in N dimensions could be reproduced typologically
by embedding just one of the system’s variables by the method of time delays. Another way of stating this is that a single variable in a multiple-variable system can serve as the surrogate of all other coupled variables present. Third, in 1987 Eckmann, Kamphorst, and Ruelle (1987) (1,506 citations) famously merged recurrence theory with embedding theory and showed “hidden” patterns in the Lorenz system of equations. They called their new graphical tool the recurrence plot (RP) and showed how this visualization tool could be used to assess time constancy (stationarity) of dynamical systems. To them the RP was a “new diagnostic tool” as indeed it was and has proven to be multiple times over. The critical idea was that the distance between all N-point vectors derived from the input stream could be mapped into two dimensions (at i, j coordinates) using a threshold that separated small distances (close, recurrent trajectories) from large distances (far, non-recurrent trajectories). For further details of the Ruelle-Takens scenario see Eckmann (1981). So this was the state of affairs prior to the last decade of the 20th century.

Evidences of recurrence are prevalent in nature if not the entire natural world at all scales. Surely the ancients including those living in pre-history times, clearly noticed the recurrent motion of the sun which rose in the east and set in the west. Peterson (2002) paraphrases the words of King Solomon of old (Ecclesiastes 1:4):

\[
\text{The sun comes up and the sun goes down,} \\
\text{then it does it again, and again – the same old round.}
\]

If we were to put words into Solomon’s mouth we could accurately say, “The sun recurs every day, but not exactly so.” Indeed, annual observations of the sun shows that the points of its rising and setting shift horizontally (north and south) along the horizon. This is due, of course, to earth’s axis of rotation which is tilted 23.5 degrees from the perpendicular to the ecliptic plane of revolution around the sun. As archeoastronomy shows, the ancients used these (subtle) motions of the sun to determine the proper timing for crop plantings. To us moderns, the dynamic trajectories of the sun during daylight hours are not exact (non-overlapping), but approximate each other (approximately recurrent). Indeed, to capture the full uninterrupted trajectory for an entire day, one needs to be located above the arctic in the summer as in the northern latitudes of Norway, the land of the midnight sun. Here, at midnight the summer sun “kisses” the horizon, but never sinks below it (Figure 1).

![Figure 1. The midnight sun at Nordkapp, Norway.](http://en.wikipedia.org/wiki/Midnight_sun#mediaviewer/File:Midnight_sun.jpg)

Dynamical systems of all flavors share one thing is common: movement in time or variations in space. For example, the following definition is as catchy as it is instructive: “If it wiggles, it’s physiology; if it stops wiggling, it’s anatomy!” (Webber, 2005). Physiology, of course, is a life science which is defined by a multitude of rhythms. But cessation of motion bespeaks death, leaving only the lifeless corpse. In this case, anatomy is a death science with its “dynamics” relegated solely to the spatial domain. And as an aside, the cadaver state may be the best example of a steady-state in the temporal domain due to its lack of all body motions.

The recurrence plot of a flat-line system (no temporal motion or no spatial structure) is simply one large black square. It may look like a black box, but the truth of the matter is that no hidden information is being held captive within. Rather, any flat-line system stigmatized by black box recurrences is fully known by its total lack of information and no activity. On the other hand, the presence of non-flat-line structures in any system no matter how...
small will be accompanied by patchy, non-uniform recurrence plots possessing much more information. In this sense, recurrence plots can be viewed as linguistic systems converting dynamical motions and structures into visual patterns with texture (or language) that informs about the system at hand. What this means scientifically is that motions/structures in nature at the mesoscopic scale as well as those at the macroscopic (large) scale to microscopic (small) and quantum (ultra-small) scales are broadcasting information than can be viewed as languages without sound as it were. Even the heterogeneous display of the cosmic background radiation in whole sky maps “screams” something instead of nothing (Fig. 2). What such diverse systems might be communicating is the subject of recurrence quantifications.

Figure 2. Whole sky map of fluctuations in the Cosmic Microwave Background (CMB) Radiation, the oldest light in universe, as recorded by the COBE spacecraft. Public source: http://en.wikipedia.org/wiki/Cosmic_Background_Explorer#mediaviewer/File:COBE cmb fluctuations.gif

2. Codings

Matrices are ideal structures in which meaningful information can be stored as well as retrieved. One excellent example is the Quick Response Code which is nothing more than a two-dimensional barcode. QRs, as they are called, appear as light and dark squares distributed within a large square box (or window) with three of the four corners designated with identical target patterns (for alignment). As shown in Figure 3, QR patterns look exactly like cross-recurrence plots (non-symmetrical) which contain meaningful information (text, WebSite links, etc.). The patterns can be conveniently captured via any cell phone camera and decoded with any QR apt. The big idea being conveyed here is that dynamical patterns in nature possess recurrence structures that can be projected or revealed as recurrence plots (like QR codes) that await reading (like QR apt). As we will see, recurrence plots provide the patterns, but recurrence quantifications provide the interpretations.

Figure 3. Quick Response coding used to embed detailed information and Web Page links. Public source: http://en.wikipedia.org/wiki/QR_code#mediaviewer/File:QR_Code_Structure_Example_3.svg
3. Calculations

Digital computers and matrix mathematics have made it possible to generate recurrence plots of dynamical systems in time or space with ease. Signals in the temporal or spatial domain can first be embedded by the method of time delays (Takens, 1981) and then mapped to recurrence space as recurrence plots (Eckmann et al., 1987) to yield beautiful two-dimensional plots (curious QRs at best). But just what are the meanings of these displays? Zbilut and Webber (1992) (514 citations) and Webber and Zbilut (1994) (769 citations) tackled this problem by introducing recurrence quantification analysis (RQA) into both the physics literature and physiological literature, respectively. Simply put, Webber and Zbilut defined five recurrence variables that were unambiguously extracted (excluding personal bias) from the recurrence plot including: recurrence (REC, density of recurrent points); determinism (DET, diagonal line structures); diagonal maximum (DMAX, longest diagonal line); entropy (ENT, distribution of diagonal line structures); trend (TND, homogeneity of recurrent points). To these five variables, in 2002 Marwan et al. (2002) (409 citations) added three more RQA variables: laminarity (LAM, vertical line structures); vertical maximum (VMAX, longest vertical line); trapping time (TT, average vertical line length).

A major 14-chapter volume edited by Webber and Marwan (2014) has just been published that describes all the mathematical details and practical implementations of recurrence plots and recurrence quantifications. And a second volume follows closely behind (Marwan, Riley, Giuliani and Webber, 2014). So there is no need to reiterate the facts in the restricted space here. Suffice it to say that each extracted recurrence variable has a dynamical interpretation that imparts meaning not only to the recurrence plots, but to the input signals themselves. Metaphorically, RQA can be viewed like a Swiss army knife with eight different purposes (Fig. 4). Thus, recurrence indicates the density of recurrence structures present, determinism reveals the level of rule-obeying structures in operation, maximal diagonal line reports on the chaoticity of the signal, entropy captures the complexity of the signal, trend is sensitive to the stationarity of the system, maximum vertical line is sensitive to the stochasticity of the signal, laminarity reveals the level of rule-defying dynamics, and trapping time is an average stochasticity measure. The experimenter coming upon a difficult and complex dynamic now has a sophisticated 8-pronged tool with which to approach the problem. Maybe one or two prongs are sufficient to solve the problem; maybe all eight prongs are required to unlock the mystery (if we can call it that).

Figure 4. Swiss army knife metaphor for the 8 recurrence variables, 5 from Webber and Zbilut (red labels) and 3 from Marwan (blue labels).

The original RQA software was programmed by Webber which is now in its 14th version (Webber, 2012). The programming language used has always been the C language and complied programs still run under the emulated and antiquated DOS system (e.g. DosBox). Officially DOS stands for Disk Operating System, but Webber (2009) dubbed it as “Dinosaur or Something” years ago. Since URL hits were first counted starting on August 22, 1996 almost 30,000 hits have been registered to Webber’s site. But there are some interesting, hidden facts about the development of this software. For example, initial program forays into the world of recurrence were named
Recurrence Plot Analysis (RPA). This was a steep learning curve of trial and error, find-tuning and error correction. Early programs were distributed by snail-mail to interested users on 3.5 inch diskettes (freeware).

Despite all cautions, in the mid 1990’s I discovered a heart-breaking error in the software which I now call a confession. Close examination of recurrence plots displayed antiparallel lines even what the embedding dimension was greater than 1. Such palindromic structures were simply impossible for embedded time series in linear time. The problem was quickly isolated to an incorrect calculation of vector distances. That is, what was believed to be the distance between vector pairs (VD) was actually the difference in vector magnitudes (VM)! Here, VM and VD are equivalent only for embedding dimensions (ED) of 1. The misapplied and incorrect formula for VM (Eq. 1) was immediately replaced by the appropriate formula for VD (Eq. 2), and warning notes were sent out to all known RPA users. Subsequently, the entire software package was renamed as RQA. In the meantime numerous recomputations on previously published time series were done to assess the damage done. Fortuitously, the results confirmed that the conclusions of those studies were basically correct. Thus in one key example, muscle fatigue was detected earlier by both RQA vector distances (Fig. 2.9 in Webber and Zbilut, 2005) and RPA vector magnitude differences (Fig. 5 in Webber et al., 1995) both compared against the linear Fast Fourier Transforms (FFT).

\[
VM = \text{ABS} [\text{SQRT} (P_{i1}^2 + P_{i2}^2 + \ldots + P_{i\text{ED}}^2) - \text{SQRT} (P_{j1}^2 + P_{j2}^2 + \ldots + P_{j\text{ED}}^2)] \quad (1)
\]

\[
VD = \text{SQRT} [(P_{i1} - P_{j1})^2 + (P_{i2} - P_{j2})^2 + (P_{i3} - P_{j3})^2 + \ldots + (P_{i\text{ED}} - P_{j\text{ED}})^2] \quad (2)
\]

The next recurrence software to be freely distributed was Visual Recurrence Analysis (VRA) as programmed in C++ for Windows computers by Kononov in 1996. His latest version of VRA was 4.9 published a decade later (Kononov, 2006). This software is very easy to use and focuses on unthresholded recurrence plots. Still, VRA can compute recurrence variables with exact correspondence to RQA ensured by the fact that Webber shared RQA code with Kononov. Kononov is one impressive programmer who works in financial markets.

The third software package offered for free use was Cross Recurrence Plots (CRP) as programmed in M-files for MatLab platforms by Marwan (2013). CRP is now up to version 5.17 in 2013. Marwan has enjoyed up to almost 24,000 hits to his CRPtoolbox since 2005. MatLab, of course, is a very popular platform among different scientific disciplines. Marwan also shared his M files to users can understand how computations are made. He is another clever programmer.

As a last thought on calculations, one of my longstanding pet peeves as a professional reviewer for scientific journals has been the misuse of the term parameter. Parameters, of course, are adjustable constants. Consider the famous logistic difference equation of May (1976) which possesses three specific components: the constant “1”; the changing variable “x”; the parameter or adjustable constant “a.”

\[
x_{n+1} = a \cdot x_n \cdot (1 - x_n) \quad (3)
\]

Here “a” is a tuning parameter which when set to low values (a < 3.0) results in monotonic dynamics in iterated variable “x_n”. With higher fixed values of “a” (3.1 < a < 3.5) the dynamics of “x_n” become periodic with successive period doublings. And with the highest values of “a” (36. < a <= 4.0) the dynamics of “x_n” are in the chaotic mode (with some narrow periodic windows of calm). Recurrence properties of the logistic equation were first described by Trulla et al. (1996), unfortunately with the incorrect calculation of distances between vector pairs.

Looking back on my early publications I note the misnaming of physiological variables as parameters in in the titles of two abstracts (Pleschka et al., 1975; Speck and Webber, 1978) and the text of at least two full publications (Webber and Peiss, 1975, 1979). But as “misery loves company,” Eckmann et al. (1987) made the same blunder when they wrote in their famous recurrence plot paper: “In recent years a number of methods have been devised to compute dynamical parameters from time series.” Again, what these mathematicians refer to as “parameters” are really “variables.” Just to be clear on this point, to date there are 8 recurrence variables in use and they should never be construed as dynamical parameters. Fixed parameters are adjustable but are held constant during any calculations. But the question is how does one set these parameters properly? This challenging issue is now discussed.
4. Challenges

There are two levels of difficulty surrounding the proper implementation of recurrence plots and recurrence quantifications which have been addressed thoroughly elsewhere (Webber and Marwan, 2012; Marwan, Riley, Giuliani and Webber, 2014, Webber and Zbilut, 2005; Marwan et al. (2007). First is the challenge of setting the RQA parameters correctly; and second this is the challenge of interpreting the recurrence plots and their quantifications wisely. On the input side of things, there are a full 8 recurrence parameters each of which must be set properly and this is no small task. These 8 adjustable parameters include: embedding dimension (EMB), delay (DLY), window (WIN), norm (NORM), rescaling (RESCALE), radius (RAD), line (LINE) and shift (SHIFT). In the RQA software of Webber (2012) these parameters are entered sequentially, one by one, and illegal combinations are prohibited. The two most important parameters are the embedding dimension and radius (threshold). The embedding is important because it represents the dimensionality into which the dynamic system under study must be contained. In addition, real-world systems are noisy and noise inflated the apparent dimension (Parker and Chua, 1989). To estimate the best embedding dimension (it is better to over-estimate than under-estimate the dimension) one can apply the false nearest neighbor approach (Kennel et al., 1992).

The second parameter that can be tricky to set is the radius which determines the boundary between vector distances (D_{i,j}) that are defined as being recurrent (D_{i,j} <= radius) or non-recurrent (D_{i,j} > radius). An unthresholded recurrence plot employs no radius (or sets radius > max D_{i,j}), but color codes each recurrent point depending upon the distance at that point as was classically implemented in VRA software (Kononov, 2006). Mathematically speaking, insertion of the radius discriminator actually converts the distance matrix (table of distances) into the recurrence matrix (table of recurrent 1’s and non-recurrent 0’s). Black and white recurrence plots are simply visualizations of the recurrence matrix at correspondent locations (black points for the 1’s; white points for the 0’s). If the radius is set too low, the recurrence density can be too sparse for statistical utility. If the radius is set too high, paired vectors that are “far” from one another are nonetheless declared “close” but incorrectly so. This is not the place to give all the ins and outs of how to practically solve these issues. But one piece of advice may prove useful. Pretend that you are a photographer and that the embedding dimension is the shutter speed and the radius is the F-stop. To “take a photo” of your dynamic why not “bracket your exposures”? That is, one can study the same dynamic on multiple passes using different combinations of embedding dimensions (e.g. 5 10, 15, 20) and radii (1%, 2%, 5%, 10%, etc.). Then the results can be studied to see which output combinations give results that are most robust. This is moving from theory to practicality, a necessary and sufficient practice of scientists no matter what their field of inquiry.

Now on the output side of things, there is a set of 8 recurrence variables that are computed for each recurrence window. As inferred above, these output values critically depend upon the parameter settings. This fact alone drives the wisdom of trying different combinations of parameters (within reason) on the same input data set to observe what works best. Back in 1994 Webber and Zbilut (1994) demonstrated how the recurrence variables could be rendered as dynamic variables by introducing the system of moving windows. Here is where the shift parameter comes into play. After computations were completed within one window, the frame of reference could be shifted into the future to capture a second set of computations. Large shifts or steps could be computed quickly, but at the loss of resolution. Small shifts (e.g. with window steps as low as 1) the highest resolution could be acquired, but at expense of long computation times.

Staying with the linguist theme, let us now consider the famous children’s poem by Dr. Seuss, “Green Eggs and Ham” (Geisel, 1960). This example, by the way, has proven to be remarkable for introducing newcomers to the world of recurrence plots and quantifications as illustrated by Webber and Zbilut (1996). This rhythmical poem consists of 812 words, but the vocabulary is restricted to only 50 different words. This means that the same few words must be reused to build the entire poem. In recurrence terminology, reused words are simply recurrent words. Three steps are taken to compute recurrence quantifiers for this poem: first, encode each new word in the poem with a unique integer (1-50) to construct a vector of integers (N = 812); second, compute recurrence quantifications (program RQE) selecting an embedding dimension of 3 (short 3-word vectors), a radius of 0 (only identical 3-word sequences recur), a window of 80 (one tenth length of poem) and a shift of 1 (maximum resolution). Third, plot the 5 traditional recurrence variables over the 731 sliding-window epics as is shown in Figure 5 (the 3 variables of LAM, VMAX and TT show no dynamics in this example). Close examination of variables REC, DET, DMAX, ENT and TND reveals five unique rhythmical structures. The structuring of the poem can be completely destroyed by random shuffling of the words order. All words are represented, but their unique sequencing (along with the
story they tell) is lost. In this case, the recurrence variables are greatly changed (e.g. decreases in REC and DET, etc.) suggesting that shuffled points in any input series may serve as a control for the native ordering of those same points.

It is important to remember that these five variables are all being computed from the same input series, but each from its own perspective. From this illustration it is good to think of the 5 (may 8) recurrence variables as separate “observers” of the same “event” from their own particular point of view. This is exactly what happens when 8 witness attempt to describe the details of the same car accident. Each has a truth to tell, but only within individualized frameworks, contexts and biases.

Analyzing text by matching word for word (embedding dimension = 1) or chain of words for chain of word (embedding dimension > 1) is based on similarity of terms. Recently, Argus et al. (2012) have introduced what they refer to as conceptual recurrence plots. Here recurrence similarity is not based on exact words (or letters) but on conceptual similarities within speech patterns. This approach is particularly useful when analyzing a complex conversation between two subjects, like doctor and patient or interviewer and interviewee.

Figure 5. Recurrence quantification variables computed within 731 moving windows, each 80 words long, sliding over the entire poem, “Green Eggs and Ham.” This poem has 8212 words and recurrences are scored only on exactly matching 3-word sequences (e.g. “I am Sam” recurs with “I am Sam” but not with “Sam I am”).

5. Connectome

The linguistic motif easily extends to the biological world of genomics and proteomics. Deoxyribonucleic acid (DNA) is the language of life and consists of only four base pairs (A for adenine, C for cytosine, G for guanine, T for thymine). Unique triplets (24 permutations) of bases (embedding of 3) code for each of the 20 naturally occurring amino acids in the construction of linear proteins on the ribosomes. These protein strings are subsequently folded into 3-dimensional structures (knots?) on the endoplasmic reticulum. For example, the recurrence structure of the entire genome of bacteriophage lambda (48,502 nucleotides) has been studied with a 1000-point sliding window (Webber and Zbilut, 1998). Likewise, the recurrence structure of entire proteins can similarly be examined using recurrence quantifies. However, beyond coding individual amino acids in their native sequence by arbitrary integer coding, actual hydrophobicity values for each protein can be substituted, allowing non-zero RQA radii to be used. Moving into the real physical domain has been exceptional fruitful in describing different nonlinear aspects of proteins (Zbilut et al., 2004).
The entire natural world is replete with rhythms, each of which is conveying a language with meaning. The tides rise and fall, weather patterns swirl around the globe, earthquakes occur aperiodically in different locals. In physics there is the entire electromagnetic spectrum which encompasses concurrent oscillations in numerous waveforms at different frequencies. Indeed, string theory proposes that the fundamental basis of physical reality rests on vibrating, one dimensional strings each with a length, but no width (dimensionless). But more than theory is the periodic table of the elements which cluster according to physical properties attributed to the unique mixing and matching subatomic particles of which they consist (protons, neutrons, electrons, quarks, etc.). So no matter which system one is examining, each and every system is amenable to recurrence analysis. Thus, to study “wiggling” dynamics implies that recurrence analysis is a linguistic interpreter.

Another linguistic system of great importance is the central nervous system of the human brain. If one thinks the estimated number of neurons at 86 billion is staggering (Azevedo et al., 2009), consider the 100 trillion synaptic connections with great diversity among all those neurons (O’Rourke et al., 2012)! Initial attempts to map the somatosensory homunculus have yielded static pictures of the human brain. But much more recently capturing the topological and spatial organization of the brain in terms of synaptic connections or connectomes gives a much more dynamic view of the brain. Indeed, one gets a better feel for the flow of information one brain regions to another (Kaiser, 2011). One researcher boasts that we are more than our genome; we are our connectome (Seung, 2012).

Figure 6 is a functional connectivity matrix of the human brain from The Human Connectome Project sponsored by the National Institutes of Health (NIH). The matrix presents as a correlation matrix similar to an unthresholded recurrence plot. It is now posited that adding recurrence quantifications to brain activities would be a powerful contribution to understanding the functional rhythms of the brain in time and space. The gross electroencephalogram has been studied in this manner (Thomasson et al., 2002).

Figure 6. Functional connectivity matrix, fictive of an unthresholded recurrence plot.
http://www.neuroscienceblueprint.nih.gov/connectome/

6. Combinations

The case has been built herein that dynamical systems in time and space possess rhythms or patterns, respectively, that are amenable (ripe) for recurrence analysis. RQA is a nonlinear tool which makes no assumptions regarding the type, length or even quality of the data. Input data need not be Gaussian distributed, can be relatively short (as low as 30 scalars in the vector), can be either nonstationary or stationary, and corrupted by noise (Zbilut et
These things stated, RQA is not an end-all or be-all of tools. This is why the concept of mixing and matching methodologies provide the best approach to problem solving and system understanding.

To illustrate this key point, I would like to introduce what is now called the “linguistiscope” as shown in Figure 7. This is a very unique instrument that combines the power of the RP and RQA with Principle Component Analysis (PCA). It is very different from the “retrospectroscope” of Comroe (1977) which is of historical importance. The design this linguistiscope is both like a telescope (magnifying distant objects) and a microscope (magnifying small objects) using three lenses. The user points this device at his/her system of interest which most likely is multidimensional, nonlinear, and noisy. The first lens (RP) collapses the N-dimensional signal down into a two-dimensional and thresholded recurrence plot. The second lens (RQA) expands the compressed signal back up into an eight-dimensional vector of recurrence quantifiers (for each moving window). The third lens (PCA) compresses the matrix of RQA variables into the first three principle components (PC1, PC2, PC3) which typically capture 95% of the variance. This third lens frees the investigator from having to pick and choose which one RQA variable is most sensitive for the dynamic at hand. When all is said and done, the observer “eyes” the data graphically in 3-D plots (PC1 vs. PC2 vs. PC3) or 2-D plots (PC1 vs. PC2 or PC1 vs. PC3 or PC2 vs. PC3). By this means, different sets of nonlinear experimental data can be sorted and classified linearly. This idea of “marrying” RQA and PCA and proof of concept on it power in analyzing complex systems can be read about elsewhere (Zbilut et al., 1998; Giuliani et al., 2001).

![Figure 7. The Linguistiscope combines the advantages of recurrence plots (RP) and recurrence quantification analysis (RQA) with principle component analysis (PCA) to focus in on dynamical systems of any scale. Data flowing forth from these structures are treated as linguistic strings which are communicating dynamical details of the behavior of any system on any scale in time and space.](image)

7. **Conversations**

It is intentional that this chapter ends not with conclusions, for the conversations about recurrence plots and recurrence quantification are continuing and open-ended. With some scrutiny, it can be observed that experimenters and theoreticians representing a vast variety of scientific disciplines and fields are basically asking rather similar questions, but do so by employing concepts and vocabularies that are very different and foreign to one another. What is suggested is that RQA is a unifying tool that can bring together scientists of very different persuasions under the unity umbrella of a common recurrence vocabulary. To the extent that dynamical inputs from any field are understood to be linguistic systems (multilingual albeit language generating), recurrence analyses can uncover similar patterns or dynamics in systems that are otherwise very different. For example, personal interactions in psychology (Shockley and Riley, 2014) are not all that different from coupling properties of strange attractors in mathematics (Chelidze and Matcharashvili, 2014). Demonstration that RQA brings researchers together is seen in...
the continuing success of bi-annual International Symposia on Recurrence Plots: Potsdam (2005); Siena (2005); Montreal (2009); Hong Kong (2011); Chicago (2013); Grenoble (2015 projected).

Recurrence plots are very beautiful to look at, so much so that initiates can stall at this gestalt-feeling level and miss the deeper meanings as carried in the recurrence quantifiers. It is critically important to move from the pretty qualitative images to the necessary quantitative descriptions. This said, as one studies recurrence plots, different patterns can be observed which may stimulate the definition of new recurrence variables. As an example, take a look at the pixelated painting in Figure 8. Here the artist depicts large-scale curved structures which can only be appreciated by observing the full collection of 2,304 pixels, but not by any of the 8 RQA descriptors. Curved structures have indeed been observed in recurrence plots (Facchini and Kantz, 2007) which can be attributed to nonstationary signals (Facchini et al., 2005) or synchronization phenomena of two time series (Marwan et al., 2002). But how can such structures be defined mathematically? This challenge is left for future researches to solve.

Figure 8. The painting “Resurrection” by Dale Olsen consisting of 2,304 pixels arranged in a square matrix [48, 48]. Used with written permission from the artist: [http://daleolsen.fineartstudioonline.com/works](http://daleolsen.fineartstudioonline.com/works)

One final thought is offered to unify researchers and keep the conversations going across the greater scientific community. One experimental design that seems to ubiquitous across disciplines is the desire/need to have early warnings for catastrophic events. In medicine it would be of great advantage to be able to forecast heart attacks or seizures. In industry it would be fiscally responsible to predict system failures compromising product qualities. In meteorology it would be lifesaving to warm the masses of impending severe storms such as tornados, monsoons, floods, etc. In the old days it is said that farmers needed not look to the sky to sense an approaching thunder storm, but rather to watch the changing behavior of their barnyard chickens (which purportedly sense the electrostatic charges in the atmosphere). In more recent times animal behaviors have been found to be different and unusual prior to earthquakes (Buskirk et al., 2010). Might the “linguisticscope” be seen as a better “chicken” sensitive to that which is beyond human sensitivities?

To illustrate this important point, let us pursue the earthquake theme. In the earth sciences earthquake prediction is as critical as it is illusive. How important it is to understand the movement of the earth’s crust again to warm and evacuate large populations of people and get them out of harm’s way. Earthquakes are indeed being studied using recurrence strategies on seismographic signals (Chelidze and Matcharashvili, 2014). Figure 8 outlines
one way to apply RQA to earthquake dynamics, but this example extends to any early-forecast system. The idea is this. Seismographs measure the tremors in the earth’s crust. It is possible that there are subtle changes in the recorded signals preceding “the big event” that can be magnified by RQA and give ample warning. The experimental design would simply be to continuously feed the seismographic data string(s) into the Linguistiscope and analyze the dynamics occurring within each moving window. Hopefully, early detection would be realized.

8. References


