Recovering Dynamics of Latent Variables Using Recurrence Quantification Analysis

Maas van Steenbergen [1,*]

[1] *Faculty of Behavioural and Social Sciences, Methodology & Statistics, Utrecht University, the Netherlands*

Correspondence*:
Corresponding Author
m.vansteenbergen@uu.nl

> "Eh bien!"–exlaimed Walras characteristically–"this difficulty is not insurmountable. Let us suppose that this measure exists, and we shall be able to give an exact and mathematical account of it". [. . .] In view of the fact that theoretical science is a living organism, it would not be exaggerating to say that this attitude is tantamount to planning a fish hatchery in a moist flower bed.
> – Nicholas Georgescu-Roegen

1 INTRODUCTION

Intensive longitudinal methods are useful for capturing and analyzing dynamic, real-time variations in individuals' behaviors and experiences over a period of time (3). They come with a unique set of methodological challenges. Psychological variables are generally latent: they are not directly observable and our knowledge of their mechanisms is incomplete at best (4). Psychological researchers rely on participants that estimate values for psychological variables at a set number of time points (5). Whereas between-person methods rely on averaging out the effects of time and within-person variation to deal with the complications this causes, (quantitative) dynamical within-person methods rely on that variation to make inferences about the underlying trajectory: how the process changes over time (2, 23, 15).

To introduce our topic, we make a number of assumptions about the nature of psychological constructs that are studied in longitudinal methods. It is important to make these assumptions explicit to increase the ability to reject their tenets if they turn out not to hold (21). We will use these assumptions to introduce the topic and embed the study in the literature. To our understanding, however, these assumptions are not made explicit all that often. We invite the reader to evaluate them critically, and even think of arguments or experiments to disprove them. To help this process along, we present some questions that can function as a starting point to think about the validity of these assumptions.

The first assumption is our working definition of psychological processes or attributes. It is a truism that it is impossible to measure a psychological attribute as one would measure height or temperature. They are exclusively accessible to the participant themselves, and we therefore have to rely on participants of studies to estimate values based on an ordinal scale at a set number of time points (10, 17).

Questions for assumption one:

- Can you quantify happiness?
- What about community spirit?
- Can you be happy without knowing it?
- Can you make errors when judging your own feelings?
- If so, how do we define feelings if we cannot judge them ourselves?

Secondly, we posit that psychological attributes are quantities and fulfill the same criteria as measurement does in the physical science unless otherwise specified. While often held for true, deciding whether an attribute is a quantity is not a trivial matter. The classical point of view within psychology is that you can quantify any variable, as long as it has a natural order, but this has long been shown to be false (14) (For an extensive discussion of the history of this refer to (22)). Deciding whether an attribute is a quantity is a point of active discussion. Some say that quantifiability of a construct can only be empirically settled by using techniques such as Conjoint Measurement Theory (16). Others argue that psychological processes are not quantifiable at all (27), or that that the question itself rests on conceptual confusion (9, 26). Some, shockingly, even go as far as saying everything is fine because Item Response Theory solved these issues ages ago (6). We will avoid this contention by blatantly assuming it away. In fact, we will it is a settled manner that the attribute in question has a quantitative structure, that this has to be proven empirically, *is* proven empirically, and it has been resolved for our theoretical construct.

We will separate the 'real' value from its likert-scale representation, and

Let X, Y & Z by any three values of the mapping. Then the projection of the quantity to an ordinal scale, which we will refer to as P, holds to the following conditions:

- if $X \geq Y$ & $Y \geq Z$, then $X \geq Z$ (transitivity);
- if $X \geq Y$ & $Y \geq X$, then $X = Z$ (antisymmetry);
- either $X \geq Y \parallel Y \geq X$ (strong connexity).

This means that only order is maintained, and nothing else: a number-score on an item cannot be seen as a quantity in the classical sense. A score in P can only indicate that a score is higher or lower than another score in P. Based on the measure only, beyond that ordering, we know nothing about the real value of P (assuming a scaling S).

As for the quantitative attribute Q, we will assume that the following additional characteristics (above the ones for the projection of the score) will hold:

- $X + (X + Y) = (X + Y) + Z$ (associativity);
- $X + Y = Y + X$ (commutativity);
- $X \geq Y$ iff $X + Z \geq Y + Z$ (monotonicity);
- if $X > Y$ then there exists a Z such that $X = Y + Z$ (solvability);
- $X + Y > X$ (positivity).
- there exists a number n such that $nX \geq Y$ (where $1X = 1$ and $(n + 1)X = nX + X$) (Archimeadean condition).

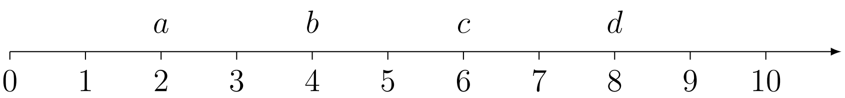
This means essentially that a value of Q in terms of another value of Q always stand in ratios between the ,

By pulling the scores away from the ordinal measurement instrument, we can formalize their relationship. This allows us to provide an unambiguous classification mechanism of scores in Q to scores in P. This lack of ambiguity is only a result of this formalization. We may find that the relationship of the 'real' variable to its ordinal estimate is not quite so straightforward.

Assume that instrument P has n elements. Then we can define a series of $n-1$ 'pegs' R as $/a, b, c \ldots z/$. These pegs are the where a score q in Q classification in P jumps to another classification in P on the basis of a score q in Q. We then define a score p in P for each q in Q as follows, assuming we have an n-level measurement scale for P.

$$\begin{cases} 1 & q \leq a \\ 2 & a \leq q \leq b \\ 3 & b \leq q \leq c \\ \ldots & \ldots \\ n & z \leq q \end{cases}$$

This relationship can also be visualized on a number line. Imagine a five-point instrument is used to measure a quantitative psychological variable. Then, at a particular point in time t, we assume that the measurement device is divided into five segments of equal lengths (excluding the last element, which goes on infinitely). It should be noted that the scaling is arbitrary, but it has been set to 10 for our convenience.



The relationship between P and Q can then be described as

Sadly, life is never easy. It would be perfectly agreeable if nature chose to let each ordinal measuring instrument correspond to

> Questions for assumption two:
>
> - How depressed are you when you are asleep?
> - How agreeable are you when you are very focused on something?

Third, we assume that those psychological processes which are part of Q are time-dependent. These measures are shaped by different forces and the previous states of the construct (24). Their values are continuous and are related to each other in a structured manner (2). We also assume that their values are differentiable over time, changing smoothly. Their values can drop or increase very quickly, but not instantaneously. They are 'complex' measures, meaning that they come to be through the interdependencies of the numerous non-trivially interacting forces that influence the system (25). This allows us

The last assumption is that ordinal likert-type scales are approximations of this underlying continuous measure (12). We cannot measure the underlying immediately, because we have to rely on . This means that if a person

Questions for assumption four:

- Did you ever feel there were not enough answering options to answer a likert-type questionnaire fully? (Did you ever feel a 3.5 out of five about your shopping experience?)
- Would you have trouble answering a question with too many (ordinal) answer options?

Likert-type scales were not initially envisioned as mappings, and do not just include the 5-point scale we are all familiar with. Treating the

The overarching goal is to model errors in . While full recovery of trajectory is impossible, it may be possible to recover some relevant aspects of the system under study. The research question is 'Given that a psychological construct has a real-valued continuous trajectory, can we recover elements of it using RQA from limited sampling occurrences on an ordinal scale?'. The major elements to be examined include the stability of several recurrence indicators under degradation, the implications for measurement and analysis of time series of latent variable constructs, and the weaknesses and oversights that we found when we tried to simulate a theoretical trajectory and degrade it.

A strength of this project is that it explicates normally tacit assumptions, and uses these assumptions to model the entire process: we model the underlying trajectory, the latent variable that estimates this trajectory, and we make an explicit prediction for its relationship if these assumptions are met. We use computational methods to generate the data because it is impossible to answer this question in the same way empirically as the real underlying trajectory is unknown. There is, however, an important trade-off being made: because we simulate our data, many of the complicating aspects that would come up during empirical studies are overlooked. We do not use real data, and that means that the inferences are only correct when our assumptions are correct. Serious objections to these assumptions can be found in the discussion section. Another weakness is that the performance of recurrence methods can be sensitive to the parameter settings of the computational model. In the current project, we treat only four different trajectories.

2 MATERIALS AND METHODS

2.1 Stage 1: Data generation

In the first stage, we used a toy model to simulate the data based on the *3 + 1 Dimensions Model* introduced by (11). This model captures clinical observations found in psychiatric symptomatology by modeling internal factors (y), environmental noise (z), temporal specificities (f), and symptomatology (x) using coupled differential equations. x is the basis of the time series. The original aim of this toy model is to simulate the trajectory of symptomatology of psychiatric symptoms over time. It is suitable for our project because the model creates realistic looking trajectories for psychological phenomena. The model uses four coupled differential equations to model the effect of time on symptom intensity. Symptom intensity is a subset of latent constructs, and we see its behaviour over time as similar to other latent constructs. It is of note that many different systems could have let to similar results to the ones outputted here. The data generating process is of secondary importance: it should result in somewhat plausible trajectories. We chose this model over more traditional choices, such as the Lorenz attractor, because we prioritized its flexibility in capturing realistic trajectories based on latent variable constructs in the social sciences.

2.1.1 Symptom intensity

$$\tau_x \frac{dx}{dt} = \frac{S_{\max}}{1 + \exp(\frac{Rs - y}{\lambda_s})} - x \tag{1}$$

2.1.2 Modelling of internal elements

$$\tau_y \frac{dy}{dt} = \frac{P}{1 + \exp(\frac{R_b - y}{\lambda_b})} + L - xy - z \tag{2}$$

2.1.3 Modelling of perceived environment

$$\tau_z \frac{dz}{dt} = S(ax + \beta y)\,\zeta(t) - z \tag{3}$$

2.1.4 Temporal specificities

$$\tau_f \frac{df}{dt} = y - \lambda_f f \tag{4}$$

2.1.4.1 Parameter definitions

Parameter definitions and parameter settings are shortly mentioned here. For a more in-depth treatment, see Gauld and Depannemaecker (11). α & β are the weight of the effect of variables x and y on environmental perception. $\tau_{x,y,z,f}$ are the different time scales the equations operate on. $S_{\max}$ is the maximum level of the symptoms. $R_{s,b}$ is the sensitivity to triggering the system. $\lambda_{s,b}$ are the slopes of the internal and symptom curves. P is the maximal rate of internal elements of the systems. S is the overall sensitivity to the environment. L is the level of predisposing factors. λ_f is the scaling factor of the slow evolution of fluctuations affecting L. $\zeta(t)$ is a point in the normal distribution where $\sigma = 0.5$. It is calculated at each $0.01t$, and is clamped between -1 and 1.

2.1.4.2 Parameter settings

There are four initial parameter settings that we have taken from the same source (11). They represent four different disorders. Their initial conditions are given at page 12. Each time series is representative of a different kind of chaotic behaviour. The time series of the 'healthy'-trajectory moves randomly around 0.1. The time series of the 'schizophrenia' time series moves close to 8, before dropping for intervals. Both 'bereavement' and 'bipolar' oscillate quickly in symptom strength, covering the full total range. For visualizations, see page 10.

2.1.5 Solvers

We used the Tsitouras 5/4 Runge-Kutta method as the solver for the differential equations, as implemented in the DifferentialEquations.jl package (28). We used standard settings for all of the parameters, aside from a higher number of maximum iterations ($1e^7$). The baseline is calculated for $0.01t$, where t represents one day in the model.

2.2 Stage 2: Binning data and removing time points

Afterwards, we systematically reduced the quality of the data. We binned the range of the width of the data into n intervals of equal length, where n stands for the number of bins. Moreover, we removed time points from the data by keeping the first and every k^{th} observation of the simulated data. We systematically decreased the number of bins and the number of time points, and re-analyze the data. k at 1 is set at the baseline. This implies no reduction. The other k-values include 2, 4, and 8. For binning, n = 100 is set at the baseline, and is equivalent to a visual analog scale. Other n-values include 20, 7, 6, 5, 4, 3, and 2. These were chosen to reflect different types of measuring instruments, such as several types of likert and forced-choice scales.

2.3 Stage 3: Data analysis

We judged the sensitivity of the data by deriving the recurrence indicators introduced before for each time series in each state of degradation. We calculated the deviation of each of these values from the baseline, which are the recurrence values derived for the intact dataset. We mapped the changes as the deviation for these indicators between the baseline and a set of degraded data, adjusting indicators based on line length by multiplying them by the reduction factor.

2.3.1 Recurrence Quantification Analysis

RQA is a method that is based on the identification of recurrent points in a time series. A point recurs if it is within the recurrence threshold of another point in time (29). Indicators can then be derived from this matrix. The development of these indicators has seen considerable development (19). We chose to focus on the core set of indicators, as described by Marwan and Webber (20). The recurrence threshold was set at the size of the bins of the degraded data set. E.g., if the range of the trajectory was 0 to 2, and the number of bins is 7 (data is degraded so that it is similar to likert-scale data), then the recurrence threshold would have been set at $\frac{2-0}{7} = \frac{2}{7}$. A visualization of these recurrences for the four trajectories can be found on page 11.

2.3.2 Recurrence Indicators

The *recurrence rate* is the proportion of points in the phase space that reoccur at later times (29). Higher recurrence rates indicate that an underlying function is more periodic. *Determinism* is the share of recurrent points that are part of diagonal lines, which indicate that the structure might be deterministic. It should be noted that it is a necessary condition, not sufficient by itself, to indicate determinism (18). *Average and maximum length of diagonal structures* are also given. A longer average length means more predictable dynamics. A longer maximum indicates the longest segment. *Entropy of diagonal structures* concerns the Shannon entropy of diagonal line lengths (13). It is an indicator of the amount of randomness, or information, in the data. *Trapping time* is the average length of vertical lines in the plot. It is a measure of how long a system stays in a particular state. *Most probable recurrence time*, similarly, is the mode of the length of the vertical lines in the plot.

2.4 Software

We used the Julia language, and in particular the 'DynamicalSystems.jl', 'RecurrenceAnalysis.jl', and 'Statistics.jl' packages to implement the toy model and run the recurrence analyses (1, 7, 8). Analyses were run on a personal computer. Full information about dependencies and version numbers can be found in a machine-readable format in the Manifest.toml file in the Github-repository. Instructions for running the analysis through a sandboxed project environment identical to our system can be found on the main page of this repository.

3 DISCUSSION

The Romanian economist and mathematician Georgescu-Roegen has played a similar role of aligning

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

FUNDING

No external funding was used for this project.

ACKNOWLEDGMENTS

I acknowledge the work of my thesis supervisors, who introduced me to the method and left me free to pursue the project as I imagined it. The great help of the Julia community was also appreciated, as they have been pushing me forward where I got stuck and took the time to respond to my stupid questions. Finally, I would like to acknowledge the feedback and conversations between me and my thesis group, who have been working through my text and made sure that it is easy to follow and well-written. Special thanks to Giuliana Orizzonte, Daniel Anadria, dr. Derksen, and dr. Bringmann for fruitful discussions and feedback about my topic. Lastly, I would like to thank my girlfriend, family, and friends for the mental support throughout.

DATA AVAILABILITY STATEMENT

The code, additional material, and generated data for this study can be found on GitHub.

REFERENCES

1 .Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM review* 59, 65–98

2 .Boker, S. M. (2002). Consequences of Continuity: The Hunt for Intrinsic Properties within Parameters of Dynamics in Psychological Processes. *Multivariate Behavioral Research* 37, 405–422. doi:10.1207/S15327906MBR3703_5

3 .Bolger, N. and Laurenceau, J.-P. (2013). *Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research*. Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research (New York, NY, US: Guilford Press)

4 .Bollen, K. A. (2002). Latent Variables in Psychology and the Social Sciences. *Annual Review of Psychology* 53, 605–634. doi:10.1146/annurev.psych.53.100901.135239

5 .Borsboom, D. (2008). Latent Variable Theory. *Measurement: Interdisciplinary Research and Perspectives* 6, 25–53. doi:10.1080/15366360802035497

6 .Borsboom, D. and Mellenbergh, G. J. (2004). Why Psychometrics is Not Pathological: A Comment on Michell. *Theory & Psychology* 14, 105–120. doi:10.1177/0959354304040200

7 .Datseris, G. (2018). DynamicalSystems.jl: A Julia software library for chaos and nonlinear dynamics. *Journal of Open Source Software* 3, 598. doi:10.21105/joss.00598

8 .Datseris, G. and Parlitz, U. (2022). *Nonlinear Dynamics: A Concise Introduction Interlaced with Code* (Cham, Switzerland: Springer Nature). doi:10.1007/978-3-030-91032-7

9 .Franz, D. J. (2022). "Are psychological attributes quantitative?" is not an empirical question: Conceptual confusions in the measurement debate. *Theory & Psychology* 32, 131–150. doi:10.1177/09593543211045340

10 .Fried, E. I. (2017). What are psychological constructs? On the nature and statistical modelling of emotions, intelligence, personality traits and mental disorders. *Health Psychology Review* 11, 130–134. doi:10.1080/17437199.2017.1306718

11 .Gauld, C. and Depannemaecker, D. (2023). Dynamical systems in computational psychiatry: A toy-model to apprehend the dynamics of psychiatric symptoms. *Frontiers in Psychology* 14

12 .Haslbeck, J. M. B. and Ryan, O. (2022). Recovering within-person dynamics from psychological time series. *Multivariate Behavioral Research* 57, 735–766. doi:10.1080/00273171.2021.1896353

13 .Kraemer, K. H., Donner, R. V., Heitzig, J., and Marwan, N. (2018). Recurrence threshold selection for obtaining robust recurrence characteristics in different embedding dimensions. *Chaos (Woodbury, N.Y.)* 28, 085720. doi:10.1063/1.5024914

14 .Krantz, D. H. (1972). Measurement Structures and Psychological Laws. *Science* 175, 1427–1435. doi:10.1126/science.175.4029.1427

15 .Lamiell, J. T. (2019). Statistical Thinking in Psychology: Some Needed Critical Perspective on What 'Everyone Knows'. In *Psychology's Misuse of Statistics and Persistent Dismissal of Its Critics*, ed. J. T. Lamiell (Cham: Springer International Publishing), Palgrave Studies in the Theory and History of Psychology. 99–121. doi:10.1007/978-3-030-12131-0_5

16 .Luce, R. D. and Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology* 1, 1–27. doi:10.1016/0022-2496(64)90015-X

17 .Maraun, M. D., Slaney, K. L., and Gabriel, S. M. (2009). The Augustinian methodological family of psychology. *New Ideas in Psychology* 27, 148–162. doi:10.1016/j.newideapsych.2008.04.011

18 .Marwan, N. (2011). How to avoid potential pitfalls in recurrence plot based data analysis. *International Journal of Bifurcation and Chaos* 21, 1003–1017. doi:10.1142/S0218127411029008

19 .Marwan, N. and Kraemer, K. H. (2023). Trends in recurrence analysis of dynamical systems. *The European Physical Journal Special Topics* 232, 5–27. doi:10.1140/epjs/s11734-022-00739-8

20 .Marwan, N. and Webber, C. L. (2015). Mathematical and Computational Foundations of Recurrence Quantifications. In *Recurrence Quantification Analysis: Theory and Best Practices*, eds. Jr. Webber, Charles L. and N. Marwan (Cham: Springer International Publishing), Understanding Complex Systems. 3–43. doi:10.1007/978-3-319-07155-8_1

21 .Meehl, P. E. (2004). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Applied and Preventive Psychology* 11, 1. doi:10.1016/j.appsy.2004.02.001

22 .Michell, J. (1999). *Measurement in Psychology: A Critical History of a Methodological Concept*. Ideas in Context (Cambridge: Cambridge University Press). doi:10.1017/CBO9780511490040

23 .Molenaar, P. C. M. (2004). A Manifesto on Psychology as Idiographic Science: Bringing the Person Back Into Scientific Psychology, This Time Forever. *Measurement: Interdisciplinary Research and Perspectives* 2, 201–218. doi:10.1207/s15366359mea0204_1

24 .Olthof, M., Hasselman, F., and Lichtwarck-Aschoff, A. (2020). Complexity in psychological self-ratings: Implications for research and practice. *BMC Medicine* 18, 317. doi:10.1186/s12916-020-01727-2

25 .Olthof, M., Hasselman, F., Oude Maatman, F., Bosman, A. M. T., and Lichtwarck-Aschoff, A. (2023). Complexity theory of psychopathology. *Journal of Psychopathology and Clinical Science* 132, 314–323. doi:10.1037/abn0000740

26 .Tafreshi, D. (2022). Sense and nonsense in psychological measurement: A case of problem and method passing one another by. *Theory & Psychology* 32, 158–163. doi:10.1177/09593543211049371

27 .Trendler, G. (2019). Conjoint measurement undone. *Theory & Psychology* 29, 100–128. doi:10.1177/0959354318788729

28 .Tsitouras, Ch. (2011). Runge–Kutta pairs of order 5(4) satisfying only the first column simplifying assumption. *Computers & Mathematics with Applications* 62, 770–775. doi:10.1016/j.camwa.2011.06.002

29 .Webber Jr, C. L. and Zbilut, J. P. (2005). Recurrence quantification analysis of nonlinear dynamical systems. *Tutorials in contemporary nonlinear methods for the behavioral sciences* 94, 26–94

FIGURES

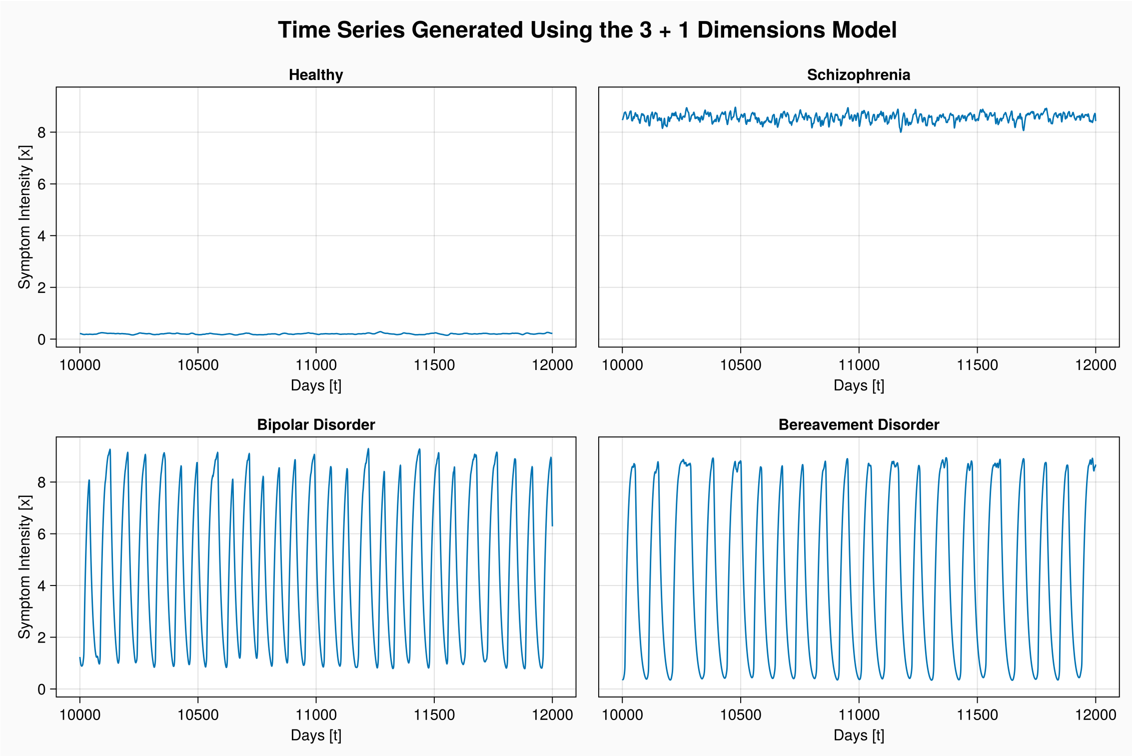


Figure 1. A section of the time series created using the coupled differential equations and parameter settings specified in section 2.1.1. This is the intact data, before degradation takes place.

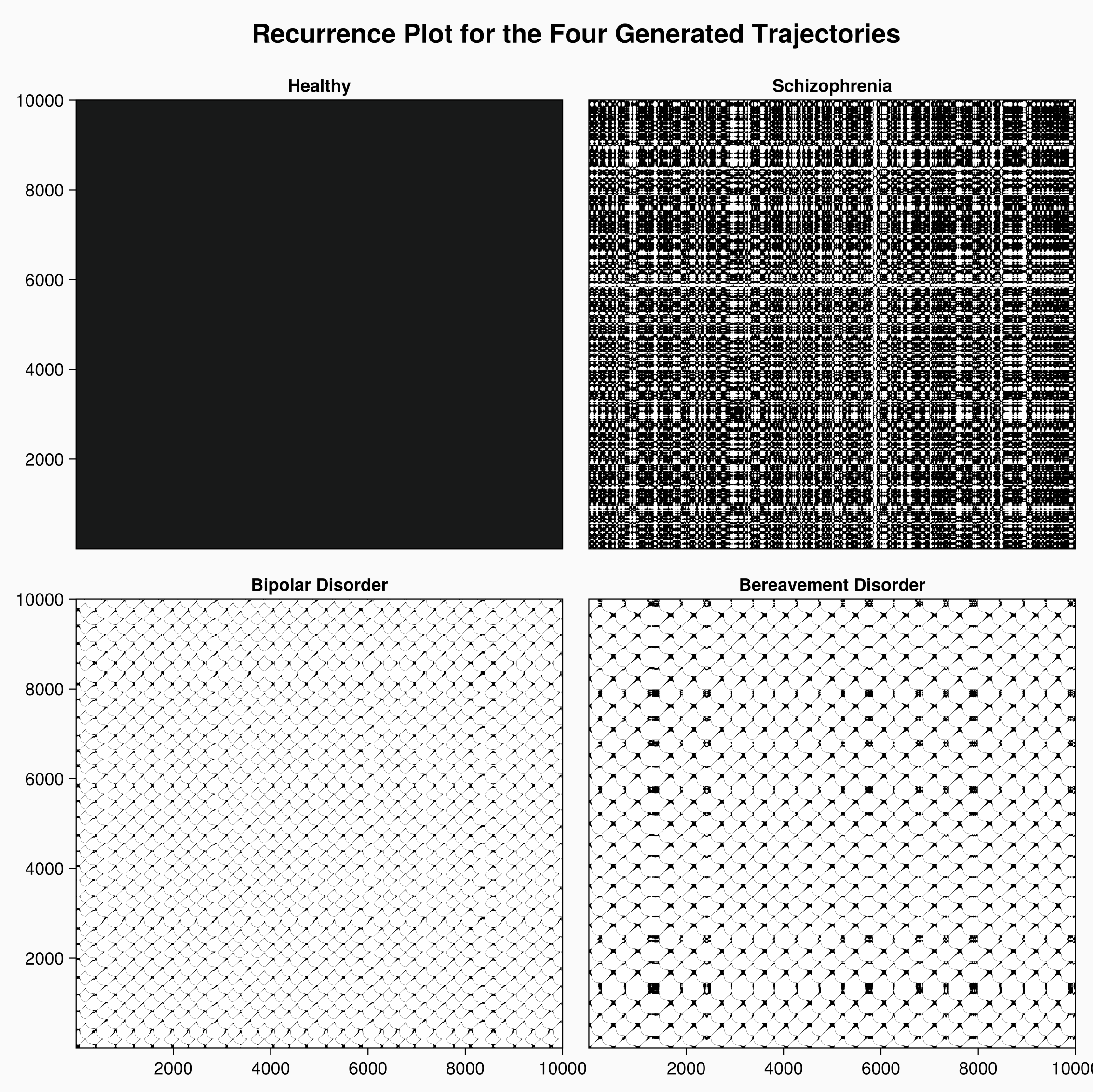


Figure 2. Recurrence plot for the four time series generated using the coupled differential equations and parameter settings specified in section 2.1.1. A point recurs when it is within the recurrence threshold of another point. Recurrent points are black, non-recurrent points are white. The axes represent time points, each location on the matrix represents a combination of time points. The recurrence threshold is set at 0.2 for illustration purpose. Note that the plot for the 'healthy' trajectory is completely black: this is because every point in the plot falls within the recurrence threshold. Also note the black 'boxes' where the bottom two trajectories are stagnant.

TABLES

Parameter	S_{max}	R_s	λ_s	τ_x	P	R_b	λ_b	L	τ_y	S	α	β	τ_z	λ_d	τ_f
Healthy	10	1	0.1	14	10	1.04	0.05	0.2	14	4	0.5	0.5	1	1	720
Schizophrenia	10	1	0.1	14	10	0.904	0.05	0.2	14	4	0.5	0.5	1	1	720
Bipolar	10	1	0.1	14	10	1.04	0.05	1.01	14	10	0.5	0.5	1	1	720
Bereavement	10	1	0.1	14	10	1	0.05	0.6	14	4.5	0.5	0.5	1	1	720

Table 1. The parameter settings used as initial parameter settings for the coupled differential equations specified in paragraph 2.1.1