Making Self-Report Ready for Dynamics: the Impact of Low Sampling Frequency and Bandwidth on Recurrence Quantification Analysis in Within-Person Ecological Momentary Assessment

Maas van Steenbergen [1,*]

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

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

ABSTRACT

For full guidelines regarding your manuscript please refer to Author Guidelines.

As a primary goal, the abstract should render the general significance and conceptual advance of the work clearly accessible to a broad readership. References should not be cited in the abstract. Leave the Abstract empty if your article does not require one, please see Summary Table for details according to article type.

Keywords: Complex Dynamics, Data Quality, Idiographic Methods, Recurrence Quantification Analysis

1 INTRODUCTION

Self-report scales have a long historical precedent in psychology. Ecological momentary assessment (EMA) is a technique meant to construct time series based on self-report instruments, allowing for ‘idiographic’ inference on the basis of self-report data (henceforth referred to as within-person) (2). While traditional statistical methods are frequently and fruitfully employed to analyze data generated using EMA, these methods are not suitable for capturing the dynamics some complex temporal within-person patterns (16).

Methods to study within-person trajectories of psychological constructs are still in infancy within a psychological context (14). One of the reasons for this is that group-comparison research is often incorrectly equated with finding general laws in psychology, and thus with scientific rigour, which leads to the marginalization of other approaches (6, 9). Another reason is that the statistical tools we predominantly rely on are a natural fit for between-person research and are both well-researched and well-understood for relatively course measurement devices. Statistics has a built-in relience on aggregation to offset the problems that are caused by such devices. Besides, many devices exist to deal with course measurement devices or relatively low accuracy in statistics when these methods fail: think of missing data imputation, or reliability measures. While dynamical systems research has a long history, it has been developed in places that have more accurate measuring devices (such as physics) or fields that do not really initially concern themselves with ‘real’ measurement at all (such as some subfields of mathematics).

Non-statistical within-person methods are often imported from dynamical systems theory, which is an area of mathematics that concerns itself with the study of the time-dependent dynamics of complex systems. A popular analysis technique is called Recurrence Quantification Analysis (RQA). It results in the identification of recurrent patterns, or repetitions, in time series analysis (18). One can then derive several indicators of the stability, predictability, and dynamical behavior of data from these recurrences. This method was developed in the physical sciences under the assumption that measurements can be retrieved at great frequency and at high resolution, to an extent that is impossible when relying on self-report scales. Hence, it is necessary to systematically assess the consequences of utilizing EMA data on the quality of RQA output (7).

We assume throughout this paper that there is an actually existing underlying continuous trajectory of the constructs that EMA devices aim to measure. This is not universally agreed upon, and contrasts with representational measurement theory (13). In other words, we make the explicit assumption that ordinal likert type-scales are approximations of an underlying continuous measure that is real, not just the assignment of numbers to objects. We also assume that the trajectory of the time series is chaotic, which means that the data generating mechanism is very sensitive to changes in the initial parameters, it is only predictable in the short-term, and observations are dependent on the state of the system and its externalities at earlier time points (15). Hence, chaotic behaviour can be easily mistaken for random behaviour, but can be fully deterministic. Finally, we suppose that self-report measures are accurate measures at a certain timepoint of this continuous measure, but courser. Working from these assumptions, we first simulate a potential continuous trajectory of a dynamical system, coursegrain and ordinalize it, and then test the deviation of recurrence measures from a segment of the full dataset to its degraded alternatives.

We chose to generate the data through a set of coupled differential equations that lead to chaotic behaviour. Note that while randomness does have a part to play in the generation of the data, most of the variance is created deterministically through four broadly modeled influences. These influences consist of symptom intensity, the modelling of the person's internal factors, influences of the perceived environment, and the influence of time. Many alternative models could be formulated that could result in similar behavior or would be as or more realistic. For our purposes a suitably broad model would work best, however, as many different trajectories can be reconstructed using parameter tuning.

Our choice for this way of modelling is based on a number of considerations. The most important of these is that self-report data cannot be measured consistently and continuously with high validity in all use cases where self-report is used, meaning that measured baseline data is currently impossible. We note that if it were possible to measure these variables directly, then researchers would not have to rely on the instrument in the first place. Therefore, theory-based simulation is a necessary starting point for further research. We hope that the models, in turn, can be refined using empirical data that takes into account time dependence. Furthermore, we chose the '3 + 1 Dimension Model' because it is one of the only ones that has the stated goal of simulating the trajectory of psychological constructs over time (in the form of symptomatology of psychiatric illness), thus being a good candidate for real empirical feedback studies.

This brings us to a natural place to introduce the goal of this project. Right now, there is often no empirical feedback system for theoretical models that aim to recover the time-dependent characteristics of constructs that are measured using EMA. We hope that studying the stability of complexity characteristics under data degredation can help us recover certain aspects empirically of the trajectory of the system under study. E.g., incompatibility of the theoretical time development of disease symptomatology contrasted with the recovered complexity characteristics of a patient can be used as a counterfactual against the choice of one model over another. A more concrete example: say that a patient is postulated to have a disorder where a particular symptom oscillates predictably and repeatedly for a period of 12 hours. A more fine-tuned understanding of that trajectory can then be used to optimize treatment times. If their symptoms were to oscillate more irratically, then this would be reflected in the recurrence indicators of the patient. If it would show dampened oscillations instead where the bandwidth becomes lower with the passage of time, then that would be reflected in the recurrence inidicators in a different manner. This is, of course, only true if those characteristics can be picked up reliably from the degraded data.

2 MATERIALS AND METHODS

We have chosen to add a detailed overview of parameter settings and design choices in the materials section. This leads into a clear and concise explanation of the basic structure of our experiment.

2.1 Materials

2.1.1 Software

We used the Julia languages, and in particular the 'DynamicalSystems.jl', 'RecurrenceAnalysis.jl', and 'Statistics.jl' packages to implement the toy model and run the recurrence analyses (1, 3, 4). All 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 the github repository.

2.1.2 Toy Model

For this study, we will use the '3 + 1 Dimensions Model' introduced by (5). The original aim of this toy model is to simulate the trajectory of symptomatology of psychiatric symptoms over time, but it can be used for our project because it has a plausible segmentation of external effects, internal effects, time, and symptoms and is flexible enough to capture a wide range of plausible trajectories. It uses four coupled differential equations to model the effect of time on symptom intensity. We give a basic explanation of each equation, and note some interesting behaviour that might be more or less suitable for use in this project. It is of note that many different systems could have let to similar results to the ones outputted here. For our purposes, we focused on the flexibility of the model in capturing different interesting plausible suitably realistic trajectories in a relatively straightforward manner as our main motivation for choosing this model over more traditional choices such as the Lorenz attractor.

2.1.3 Symptom intensity

The first equation is supposed to represent symptom intensity. For our study,

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

2.1.4 Modelling of internal elements

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

2.1.5 Modelling of perceived environment

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

2.1.6 Temporal specificities

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

2.1.7 Solvers

We used the Tsitouras 5/4 Runge-Kutta method as the solver for the differential equations, as implemented in the DifferentialEquations.jl package (17). We used standard settings for all of the parameters, aside from a higher number of maximum iterations ($1e^7$).

2.2 Recurrence Quantification Analysis

There are many different recurrence indicators, and developing new ones has been an area of considerable development (11). We chose to focus on the core set of indicators, as described by Marwan & Webber (12). 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 $\frac{1--1}{7} = \frac{2}{7}$. Because the data is discrete, the embedding dimension is set to the amount of time that is covered by one data point.

2.2.1 Recurrence Indicators

The *recurrence rate* is the proportion of points in the phase space that reoccur at later times (18). 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 (10).

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 (8). It quantifies 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.

2.3 Analysis

2.4 Methods

2.4.1 Stage 1: Data generation

In the first stage, we use a toy model to simulate the data based on the *3 + 1 Dimensions Model* introduced by (5). 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. Fluctuations will be the outcome variable of this study. We modeled four. We save each one of these models as a separate time series.

2.4.2 Stage 2: Binning data and removing time points

Afterwards, we aim to systematically reduce the quality of the data. We bin a range of the width of the data into n intervals of equal length, where n stands for the number of bins. We also vary the minimum (min) and maximum (max) value of this range to simulate ceiling and floor-effects. Moreover, we remove time points from the data by keeping the first and every k^{th} observation of the simulated data. We systematically decrease the number of bins, the range, number of time points, and re-analyze the data.

2.4.3 Stage 3: Data analysis

We will judge the sensitivity of the data by deriving the recurrence indicators introduced before for each time series in each state of degradation. We judge sensitivity to degradation by calculating the deviation of each of these values from the baseline, which are the recurrence values derived for the intact dataset. We will map the changes as the deviation for these indicators between the baseline and a set of degraded data.

3 RESULTS

4 DISCUSSION

The results of this study suggest that applying recurrence methods to

The limitations of this study are both in scope and in . First of all,

5 ARTICLE TYPES

For requirements for a specific article type please refer to the Article Types on any Frontiers journal page. Please also refer to Author Guidelines for further information on how to organize your manuscript in the required sections or their equivalents for your field

6 MANUSCRIPT FORMATTING

6.1 Heading Levels

6.2 Level 2

6.2.1 Level 3

6.2.1.1 Level 4

6.2.1.1.1 Level 5

6.3 Equations

Equations should be inserted in editable format from the equation editor.

$$\sum x + y = Z \tag{5}$$

6.4 Figures

Frontiers requires figures to be submitted individually, in the same order as they are referred to in the manuscript. Figures will then be automatically embedded at the bottom of the submitted manuscript. Kindly ensure that each table and figure is mentioned in the text and in numerical order. Figures must be of sufficient resolution for publication see here for examples and minimum requirements. Figures which are not according to the guidelines will cause substantial delay during the production process. Please see here for full figure guidelines. Cite figures with subfigures as figure 2a and 2b.

6.4.1 Permission to Reuse and Copyright

Figures, tables, and images will be published under a Creative Commons CC-BY licence and permission must be obtained for use of copyrighted material from other sources (including re-published/adapted/modified/partial figures and images from the internet). It is the responsibility of the authors to acquire the licenses, to follow any citation instructions requested by third-party rights holders, and cover any supplementary charges.

6.5 Tables

Tables should be inserted at the end of the manuscript. Please build your table directly in LaTeX. Tables provided as jpeg/tiff files will not be accepted. Please note that very large tables (covering several pages) cannot be included in the final PDF for reasons of space. These tables will be published as Supplementary Material on the online article page at the time of acceptance. The author will be notified during the typesetting of the final article if this is the case.

7 ADDITIONAL REQUIREMENTS

For additional requirements for specific article types and further information please refer to Author Guidelines.

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.

AUTHOR CONTRIBUTIONS

The Author Contributions section is mandatory for all articles, including articles by sole authors. If an appropriate statement is not provided on submission, a standard one will be inserted during the production process. The Author Contributions statement must describe the contributions of individual authors referred to by their initials and, in doing so, all authors agree to be accountable for the content of the work. Please see here for full authorship criteria.

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 persue the project as I imagined it. I also acknowledge the great help of the Julia community, which has been helping me with programming where I got stuck and which took the time to respond to my stupid questions. Finally, I'd 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 followable.

SUPPLEMENTAL DATA

Supplementary Material should be uploaded separately on submission, if there are Supplementary Figures, please include the caption in the same file as the figure. LaTeX Supplementary Material templates can be found in the Frontiers LaTeX folder.

DATA AVAILABILITY STATEMENT

The code, all additional material, and generated data for this study can be found in the [NAME OF REPOSITORY].

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 .Conner, T. S., Tennen, H., Fleeson, W., and Barrett, L. F. (2009). Experience Sampling Methods: A Modern Idiographic Approach to Personality Research. *Social and Personality Psychology Compass* 3, 292–313. doi:10.1111/j.1751-9004.2009.00170.x

3 .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

4 .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

5 .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

6 .Hamaker, E. L. (2012). Why researchers should think ”within-person”: A paradigmatic rationale. In *Handbook of Research Methods for Studying Daily Life* (New York, NY, US: The Guilford Press). 43–61

7 .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

8 .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

9 .Lamiell, J. (2021). The 'Problem of Individuality' in Scientific Psychology. 111–131. doi:10.1007/978-3-030-67734-3_6

10 .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

11 .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

12 .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

13 .Michell, J. (2021). Representational measurement theory: Is its number up? *Theory & Psychology* 31, 3–23. doi:10.1177/0959354320930817

14 .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

15 .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

16 .Olthof, M., Hasselman, F., Wijnants, M., and Lichtwarck-Aschoff, A. (2020). Psychological dynamics are complex: A comparison of scaling, variance, and dynamic complexity in simulated and observed data. In *Selbstorganisation – Ein Paradigma Für Die Humanwissenschaften*, eds. K. Viol, H. Schöller, and W. Aichhorn (Wiesbaden: Springer Fachmedien Wiesbaden). 303–316. doi:10.1007/978-3-658-29906-4_17

17 .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

18 .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

FIGURE CAPTIONS

Figure 1. Enter the caption for your figure here. Repeat as necessary for each of your figures

Figure 2a. This is Subfigure 1.

Figure 2b. This is Subfigure 2.

Figure 2. Enter the caption for your subfigure here. **(A)** This is the caption for Subfigure 1. **(B)** This is the caption for Subfigure 2.