# Reply for interactive comment on "Finding recurrence networks threshold adaptively for a specific time series" by D. Eroglu et al.

We followed both Referees suggestions very carefully to improve the quality of our manuscript. We hope that the paper is now ready for publication.

**Reply for Referee #1**

We are grateful that Referee #1 states that "The paper discusses the important problem when we use a recurrence plot, and thus I will support its publication after the following minor points are revised".

In what follows, we address the remarks of the Referee #1 item by item.

1. **Referee #1 comment:** It is very interesting that the authors finally have reached the proposed condition, which (accidently?) coincides with the mathematical condition that one can reproduce the rough shape of the original time series from a recurrence plot (see Thiel et al., Phys. Lett. A 330, 343-349 (2004); Hirata et al., Eur. Phys. J. Spec. Top. 164, 13-22 (2008)). Therefore, it would be nice if the authors can discuss the relationship between their intention for the proposed condition and this mathematical condition for the reproducibility of the original time series from a recurrence plot.

   **Answer of authors:** We thank the Referee for his/her evaluation of our work and for the critical reading of the manuscript. As the Referee said that the adaptive threshold selection technique coincides with the condition for the reconstruction of time series from a RP (Thiel et al., Phys. Lett. A 330, 343-349 (2004) and Hirata et al., Eur. Phys. J. Spec. Top. 164, 13-22 (2008)). Note, however, the motivation for this condition is different for the reconstruction and for our aim. For the reconstruction it is necessary that all time points are connected by their overlapping neighborhoods. For example, in the approach as suggested by Hirata et al., the RP is transformed to a matrix of weights corresponding to the shortest path length of connected neighborhoods and then multidimensional scaling is applied. This requires that there are no isolated components in the network. In contrast, our first motivation in this study is to apply network measures on the RP for its quantification and to ensure that the measures will be calculated for connected networks without isolated components. The proposed condition is a trade-off between the smallest threshold (what is desired) and that the network does not have isolated components. Moreover, as one aim of the time series reconstruction from RPs is to show that the RP contains the dynamics of the underlying process (because of its reconstructability) appropriately. Our second motivation is that by choosing such an optimal threshold, the applied recurrence measures will quantify a large amount of dynamics encoded by the RP. For a too small threshold the RP can lose information about the dynamics and, therefore, the recurrence measures might be biased.

   We have now included the previous stated motivation to the paper (p. 3 of the present draft) as,

   "An even more important motivation for avoiding isolated components in the RN is that the RN provides a large amount of information about the dynamics of the underlying system although it contains only binary information. This has been demonstrated by reconstructing time series from RPs (Thiel, M. et al: How much information is contained in a recurrence plot?, Physics Letters A, 330, 343–349, (2004); Hirata, Y. et al: Reproduction of distance matrices from recurrence plots and its applications, European Physical Journal – Special Topics, 164, 13–22, (2008)). The condition for reconstructing a time series from a RP is that all points are connected by their neighborhoods, i.e., there are no isolated components. By applying recurrence measures we would like to quantify the dynamics encoded by the RN. This can be ensured by the above mentioned condition.

   To find a *sufficiently* small threshold $\epsilon$ that fulfills the desired condition of connected neighborhoods, we will use the connectivity properties of the network. In particular, we choose the value for $\epsilon$ that

is the smallest one for the RN to be connected.".

2. **Referee #1 comment:** How long did it take for a computer to obtain $\epsilon_c$ in Eq. (4) for the examples of Figs. 2-4?

   **Answer of authors:** The calculation time for determining the critical threshold $\epsilon_c$ depends strongly on the computer configuration, implementation of the algorithm, etc. Using one compute node of our IBM iDataPlex cluster (Intel Harpertown), it takes 1.5–2 mins to find a critical threshold $\epsilon_c$ for each control (bifurcation) parameter of the logistic map.

3. **Referee #1 comment:** Because Figs. 2 and 3 look to me that the results of constant thresholds are more consistent with the behavior of Lyapunov exponent, it might be more appropriate to say that the proposed threshold should be used especially for identifying some bifurcations where the maximal Lyapunov exponent keeps non-positive.

   **Answer of authors:** We followed the suggestion by the Referee, and now emphasized which technique (constant threshold or adaptive threshold) is better to detect what kind of dynamical transitions. This change can be found in p.4 of the new draft as,

   Both threshold selection methods could detect transitions between dynamical regimes (periodic-chaos or chaos-periodic). Transitivity gives large values for the chaotic regime and small values for periodic. In the betweenness centrality case, it is contrary to transitivity, large values for periodic and small values for the chaotic regimes. Although the constant threshold selection detects the periodic windows (chaos-period transitions) more sharply than the adaptive threshold case, the transitivity and betweenness centrality for the constant threshold selection case (in the constant threshold case, as general, threshold arbitrarily chosen by $\epsilon = 5\%$), $T_{\text{constant}}$ and $BC_{\text{constant}}$, cannot distinguish between different periodic dynamics, i.e., cannot detect certain bifurcation points such as for period doublings, e.g., at $a \approx 3.544, 3.564, 3.84$. Contrary, in the adaptively chosen threshold case, $T_{\text{adaptive}}$ and $BC_{\text{adaptive}}$ are sensitive to these bifurcations (Figs. 2, 3). Thus, using the adaptive threshold allows also the detection of period-period transitions (i.e., the study of bifurcation points where the maximal Lyapunov exponent keeps non-positive). .

4. **Referee #1 comment:** It was great that the authors demonstrated the difference between the results obtained from a standard criterion for the threshold and the ones obtained from the proposed condition. It would be greater if the authors can show further difference between their results and the ones possibly obtained from conventional recurrence quantification measures.

   **Answer of authors:** We thank the Referee for his/her important comment of including recurrence quantification analysis (RQA). We agree that the proposed threshold selection method should also be tested for RQA, which is currently work in progress. The initial motivation of the threshold selection method was to improve the recurrence network analyses by ensuring that the RN has no unconnected components. For the RQA, a further systematic study is required which is future work and would also be out of the initial scope of the paper. We mentioned this future study on p. 6 of the revised version as,

   "Moreover, the proposed threshold selection can also be useful for the recurrence quantification analysis. A systematic investigation of the different threshold selections remains future work.".

5. **Referee #1 comment:** How computationally expensive to calculate the network transitivity and the average betweenness centrality? Because the former contains the triple sums and the latter needs to count the number of shortest paths, I suspect that the computational demands could be huge. In such a case, I imagine that the authors can state the advantages for computing the

network transitivity and/or the average betweenness centrality more strongly if they compare it with computing the Lyapunov exponents.

**Answer of authors:** We agree with the Referee, who is completely right about the computation cost of the model (logistic map) application. Calculating the second smallest eigenvalue for large matrices is computationally expensive. But in real world application, we are using a sliding window analysis technique. That means, in general, we have very less number of points in the sliding windows and the calculation of the Lyapunov exponent of such a short time series segment is very difficult or even impossible. Therefore, we need to analyse such short time series with other techniques providing good estimators and quantifiers of the dynamical properties. The advantage of the recurrence network measures over the Lyapunov exponent is that they allow good results even for short time series.

**Reply for Referee #2**
In what follows, we address the remarks of the Referee #2 item by item.

1. **Referee #2 comment:** p806 l13-14 I don't understand 'extreme points in the time series could break the connected components in the network', and what 'biasing the recurrence analysis' means. Does 'standardisation method' refer here to what was called 'normalisation' above?

   **Answer of authors:** We thank the Referee for his/her evaluation of our work and for the detailed and critical reading of the manuscript. We added some extra information to be more understandable as,

   "When considering the time series by a RN representation, extreme points (very high jumps or falls in the fluctuation of time series) in the time series could break the connected components in the network since the distance between an extreme point and other points would be larger than the threshold value. The normalization method would then result in non-optimal recurrence thresholds biasing the recurrence analysis."

2. **Referee #2 comment:** l17 What is a 'spectral property' of a network? Please provide reference at least.

   **Answer of authors:** We thank to Referee for his/her suggestion to add a reference for this issue, spectral property of a graph is very important part of our work. The reference is added as the Referee suggested.

3. **Referee #2 comment:** p808 l18 Do you do your calculations for a range of epsilon's, or you find $\epsilon_c$ more efficiently by, say, iterative bisection?

   **Answer of authors:** We have very much appreciated this question of the Referee since we miss out to add which method we use when we are finding the threshold value adaptively in the manuscript. Yes, we use the iterative bisection method to find the critical threshold $\epsilon_c$. In the manuscript, we mentioned the method as,

   "In order to find such an adaptive threshold, we start from very small values of the threshold and vary the $\epsilon$ parameter until we get a connected network. In order to apply this approach efficiently, we use iterative bisection method in the simulations. ".

4. **Referee #2 comment:** l20-21 Please provide a reference for the theorem 'If the network is connected. . .'

   **Answer of authors:** We thank to the Referee again for the same reason with 2nd comment of the Referee. The reference is added as the Referee proposed.

5. **Referee #2 comment:** eq. (4) Instead of ∈ (by latex notation for 'element of a set') I believe that = is the correct symbol for defining the set T of thresholds.

   **Answer of authors:** We thank the Referee for the detailed check of the equations. We agree with the Referee and changed the Eq.(4) as she/he suggested.

   $$\epsilon_c = \min(\mathbf{T}) \text{ with } \mathbf{T} = \{T_i \,|\, \forall i : \lambda_2(T_i) > 0\}. \tag{1}$$

6. **Referee #2 comment:** p809 l15-19 The result of what is called a 'regular threshold selection method' is called e.g. in Fig. 2 and throughout the paper a 'constant threshold'. This is contrasted to the 'adaptive threshold' (calculated the way newly proposed in this article). However, the first one can also be called an adaptive threshold: adaptively chosen based on the requirement of $RR = 5\%$. Please reconsider your terminology that you want to introduce in this paper.

   **Answer of authors:** We followed the Referee suggestion and added extra information for the constant threshold selection case. In the case of constant threshold selection, we fix our threshold distance arbitrarily with a number or a percentage for the normalized threshold case. In the adaptive case, we find the critical threshold number to get a connected recurrence network. We have changed the following paragraphs as,

   "As a first application we compare some RN measures for using first the adaptive and then the constant threshold (arbitrarily chosen threshold value) approach by analysing the logistic map.."

   "Both threshold selection methods could detect transitions between dynamical regimes (periodic-chaos or chaos-periodic). Transitivity gives large values for the chaotic regime and small values for periodic. In the betweenness centrality case, it is contrary to transitivity, large values for periodic and small values for the chaotic regimes. Although the constant threshold selection detects the periodic windows (chaos-period transitions) more sharply than the adaptive threshold case, the transitivity and betweenness centrality for the constant threshold selection case (in the constant threshold case, as we have mentioned before, threshold arbitrarily chosen by $RR = 5\%$), $T_{\text{constant}}$ and $BC_{\text{constant}}$, cannot distinguish between different periodic dynamics, i.e., cannot detect certain bifurcation points such as for period doublings, e.g., at $a \approx 3.544, 3.564, 3.84$. Contrary, in the adaptively chosen threshold case, $T_{\text{adaptive}}$ and $BC_{\text{adaptive}}$ are sensitive to these bifurcations (Figs. 2, 3). Thus, using the adaptive threshold allows also the detection of period-period transitions (i.e., the study of bifurcation points where the maximal Lyapunov exponent keeps non-positive).".

7. **Referee #2 comment:** p812 l 7-15 Please reconsider the use of terms 'confidence interval' and 'statistical test', both of which have a standard meaning different from what they are used for here.

   **Answer of authors:** We thank the Referee for pointing out this terminology issue. We have now relaxed the use of the term 'statistical test' by modifying the manuscript as,

   "We apply a ~~statistical~~ rather simple test in order to see whether the characteristics of the dynamics at a certain time statistically differs from the general characteristics of the dynamics. In order to apply this ~~statistical~~ test, we use the following approach.'

   However, we believe that we are using the term 'confidence interval' in the standard meaning: we find an empirical test distribution from thousands of realization and consider 90% of the distribution as our confidence interval.

8. **Referee #2 comment:** p813 l6-13 Perhaps the following is the main point regarding the merit of the paper. I believe that the regularity and irregularity of segments of the time series can be 'seen', or can be quantified by FFT, or by the method of Prasad et al. (2004). Why not apply those methods first to identify the events missed (?) by Prasad et al. – instead of using recurrence networks?

**Answer of authors:** We thank the referee for this suggestion. First of all we would like to emphasize that we do not consider our method as a replacement of other time series methods such as power spectrum estimation but as a method that provides complementary aspects of the data. For example, the power spectrum analysis would allow to estimate major frequencies in the signal but the transitivity coefficients provide a quantitative measure of regularity, in a more strict sense, of the dimensionality of the dynamics (cf. Donner et al, The Geometry of Chaotic Dynamics – A Complex Network Perspective, European Physical Journal B 84, 2011). Moreover, our analysis approach is based on sliding windows. The available data within these windows is too less to get reliable estimates from a power spectrum analysis, but the recurrence approach is still working. We are, therefore, convinced that the recurrence based methods are useful and complimentary tools that can enrich the toolbox of palaeoclimate research. We would also like to note that the used recurrence method is not only applicable to geoscientific data but also has great potential in other disciplines, like medicine or engineering. The power of this method was already demonstrated by numerous applications in the diverse scientific disciplines. Finally we would like to note that all this is clearly in agreement with our co-author S. Prasad who is also the first author of the paper Prasad et al. (2004) you mentioned.