Recurrence quantification based Liapunov exponents for monitoring divergence in experimental data

Joseph P. Zbilut a,*, José-Manuel Zaldivar-Comenges b, Fernanda Strozzi c

a Department of Molecular Biophysics and Physiology, Rush University, 1653 W. Congress, Chicago, IL 60612, USA
b European Commission, Joint Research Center, Institute for Environment and Sustainability, 21020 Ispra (VA), Italy
c Engineering Department, Quantitative Methods Group, Carlo Cattaneo University, Castellanza (VA), Italy

Received 1 January 2002; received in revised form 22 March 2002; accepted 22 March 2002
Communicated by C.R. Doering

Abstract

Although the use of positive Liapunov exponents has been emphasized in the context of confirmation of chaotic dynamics, their original conception concerned system stability as a qualitative feature. Using a recurrence-based algorithm, data from experimental reactions are presented for exponent use in this context. Emphasis is placed on the utility for on-line, nonstationary, noisy systems. © 2002 Elsevier Science B.V. All rights reserved.
PACS: 05.45.-a; 05.45.Tp; 07.05.Rm; 82.20.-w; 82.60.-s

Keywords: Liapunov exponent; Recurrence quantification; Nonlinear dynamics; Chaos; Thermal reactions

1. Introduction

In a previous paper [1], the ability of quantification of certain recurrence plot (RP) features to underscore specific aspects of dynamical systems was demonstrated. Specifically, the comment by Eckmann et al. [2], that line segments parallel to the identity line were inversely related to the largest positive Liapunov exponent (PLE) was confirmed. Gao and Cai [3], and Gao [4] subsequently reaffirmed this in the context of chaotic systems, and the utility of using recurrences for transition detection. To be sure, a significant contemporary use of these exponents has been to demonstrate the existence of chaotic dynamics; i.e., sensitive dependence on initial conditions.

As a practical matter, in the case of many experimental systems, confirmation of chaos is doubtful, given the data and stationarity needs for the unambiguous calculation of Liapunov exponents [5]. Certainly, the need to confirm chaos is often questionable, as Ruelle has pointed out [6]. Nonetheless, if considered within the context of their development, PLEs can provide important information apart from notions of chaoticity. It should be recalled that Liapunov exponents were developed to characterize a system’s stability/instability [7,8], which of itself may be an important finding. A drawback has been the computational effort needed. In the present Letter, an example of detection of instability is presented: a relatively simple chemical reaction carried out in a jacketed batch reac-

* Corresponding author.
E-mail address: zbilut@rush.edu (J.P. Zbilut).
tor is monitored by its temperature using an RP-based quantification of line segments as approximations of the largest positive Liapunov exponent.

2. Methods

The model and experiment have been extensively documented [9,10], but briefly: in a series of studies carried out previously, it has been shown, theoretically and experimentally, that chaos theory techniques developed to measure the parametric sensitivity of strange attractors could be used to identify critical regions in which thermal runaway can occur in chemical reactions. The criterion says that a runaway situation occurs when the divergence of the system becomes positive \((\text{div} > 0)\) on a segment of the reaction path; i.e., PLE. The divergence is a scalar quantity defined at each point as the sum of the partial derivatives of the mass and energy balances with relation to the corresponding variables—temperature and conversion, i.e., \(\partial(dT/dt)/\partial T + \partial(dz/dt)/\partial z\). The draw back of this technique is its computational effort based upon the trace of the Jacobian of the system equations. Additionally, the computational delay in real situations may be significant.

The complementary nonlinear technique presented here is based on recurrence quantification analysis (RQA) [11,12]. The technique has several advantages:

1. it is robust against noise in temperature measurement;
2. it is easily implemented;
3. it is easily implemented;
4. it has demonstrated effectiveness in physiological and other chemical systems to detect state transitions [12–14].

2.1. Recurrence quantification analysis (RQA)

The actual methods developed in nonlinear time series analysis assume that the system under analysis is stationary. Furthermore it is assumed that the time series data length is much longer than the characteristic time of the system in question. In the case of temperature data in a batch or semibatch processes (see below, Section 2.3) the assumption of stationarity is no longer valid (there is no steady state in such processes), and for this reason, techniques developed by nonlinear time series analysis are not suitable for direct application [10].

Eckmann et al. [2] introduced the RP as a graphical tool based on the computation of a distance matrix between reconstructed points (delay coordinates) in phase space, i.e., given a scalar time series, \(\{x(i) = 1, 2, 3, \ldots\}\), an embedding procedure will form a vector, \(X_i = (x(i), x(i + L), \ldots, x(i + (m - 1)L))\) with \(m\) the embedding dimension and \(L\) the lag. \(X_i = 1, 2, 3, \ldots, N\) then represents the multi-dimensional process of the time series as a trajectory in \(m\)-dimensional space. RPs are symmetrical \(N \times N\) arrays in which a point is placed at \((i, j)\) whenever a point \(X_i\) on the trajectory is close to another point \(X_j\). The closeness between \(X_i\) and \(X_j\) is expressed by calculating the Euclidian distance between these two normed vectors, i.e., \(\|X_i - X_j\| \leq r\) where \(r\) is a fixed radius. If the distance falls within this radius, the two vectors are considered to be recurrent, and graphically this can be indicated by a dot.

In order to extend the original concept and make it more quantitative, Zbilut and Webber [11] developed RQA. As a result, they defined several variables to quantify RPs. The one of interest here is “maxline”, i.e., the longest diagonal line segment: in the matrix of the RP there exist short line segments parallel to the main diagonal, which correspond to sequences \((i, j), (i + 1, j + 1), (i, j), (i + 1, j + 1), \ldots, (i + k, j + k)\) such that the piece of \(X(j)x(j + 1), \ldots, X(j + k)\), \(X(i + k)\). Eckmann et al. [2] have pointed out that these sequences are inversely related to the largest PLE, while Gao and Cai [3] identify them as a lower bound. Thus, it is clear, this is a rather qualitative estimate of the PLE.

In order to follow changes of maxline in time, a “windowed” version of RQA can be performed, such that for a time series \((s_1, s_2, \ldots, s_n)\), where \((s_j = j t_s)\) and \(t_s\) = sampling time. For a \(N\) point long series

\[
E_1 = (s_1, s_2, \ldots, s_N), \\
E_2 = (s_1+w, s_2+w, \ldots, s_{N+w}), \\
E_3 = (s_1+2w, s_2+2w, \ldots, s_{N+2w}), \\
\vdots \\
E_p = (s_1+(p-1)w, s_2+(p-1)w, \ldots, s_{N+(p-1)w}),
\]  

(1)
with $w = \text{the offset, and the number of epochs (windows), } E_p$, satisfies the relation, $N + (p - 1)w \leq n$.

2.2. Determining parameters for nonstationary series

It has been emphasized, that RQA is useful for understanding nonstationary time series. Yet, since a given system may be changing state, i.e., as the relevant degrees of freedom may change, the choice of $m$, $L$ and $r$ can become confounding. Unfortunately, most algorithms for such choices are based upon computer simulations of well-known, stationary examples. Typically, as has been pointed out, real systems are rarely stationary, and often exhibit rather sudden changes of state. Nonetheless, some guidelines can be established, based upon available research, and a careful consideration of the import of nonstationarity.

2.2.1. Choice of embedding

In the context of nonstationarity, the notion of a “correct” embedding or delay is inappropriate as has been demonstrated by Grassberger et al. [15]. Instead it becomes important to remember that a sufficiently large embedding be chosen which will “contain” the relevant dynamics (as it may change from one dimensionality to another) as well as account for the effects of noise, which tend to inflate dimension. There are no clear guidelines relative to this question, except from what can be inferred from studies of noise. In this respect Ding et al. [16] have indicated that noise will tend to require higher dimensions, even in the case of stationary dynamics. Gao and Cai [3] have studied this question in the context of a noisy Lorenz attractor, and concluded that an embedding of 6 is required to provide reasonable clarification of the dynamics. Hegger et al. [17] justifies the approach we have used for some time, namely, to “overembed”. Care, however, must be made to ensure that the system is not excessively noisy, since embedding will amplify such noise to the detriment of the real dynamics. As a practical matter, it is often difficult a priori to determine relevant degrees of freedom and strength of noise. A heuristic response to the problem would embed the time series in a relatively high dimension (say 20–25), and then progressively decrease the embedding while observing the behavior of the $R_P$. If there is minimal change, the decrease is continued until there is an obvious change in the character of the plot. Presumably

minimal changes are due to noise and the effects of the “curse of dimensionality”, i.e., the shortening of distances in high dimensions with resulting “false nearest neighbors” [16]. However, major topological changes should not result unless the embedding is insufficient. Thus it would seem to be prudent to stop decreasing embedding when such an obvious change occurs. Adding a few dimensions to this minimum would insure against such vagaries as local changes (especially when windowing is used) [18].

2.2.2. Choice of lag

Choice of lag is governed by similar considerations. As a system changes from one dimension to another the effects of the lag are perforce changed. Thus, a so-called “optimal” lag in one embedding, becomes less so as the relevant dimension changes [18].

Although there have been numerous proposals for choice of lag, chief among them the first local minimum of the autocorrelation or mutual information, they all are presented with the assumption of stationarity [19–21]. What would appear to be more important is an understanding of the data acquisition apparatus (typically A/D), as well as the system studied. Specifically, what is the sampling, the bandwidth, the precision, the nonlinearity?

Since there is no fully developed theory of nonlinear systems as there is for linear systems, the problem would seem to be complex. This is to say that a nonlinear system, is often sampled by guidelines developed for linear systems, i.e., the sampling theorem. As a matter of fact, the delay reconstruction theorem indicates that practically any delay is appropriate [11]. In practice, actual sampling determines whether a nonlinear system is sufficiently sampled to capture the important dynamics. Thus, it would seem to be reasonable to “over” sample a system to determine if there are important small, but putatively significant features in the time series. Clearly, however, care must be made to balance this need for adequate sampling against the possibility of introducing greater amounts of noise associated with a larger bandwidth.

Once this has been determined, use of autocorrelation or mutual information as a guide may be used to provide a reconstruction which may maximally “unfold” the dynamics, while keeping in mind the limitations. (An alternative would be to ramp up the embedding and delay and check if the number of recurrences
plateau, but this most likely would not happen in the case of nonstationarity [1].) If the embedding is high enough to contain the nonstationary dynamics, and the lag constant, the change in recurrences simply fulfill the objective of identifying state changes. The critical point is that the method of choice of lag be consistent for a given method of data acquisition.

2.2.3. Choice of radius

The object of RQA is to view the recurrences in a locally defined (linear) region of phase space. Practically speaking, however, because of intrinsic and extrinsic noise, too small a value of \( r \) results in quantification of the noise only; whereas too large a value captures values which can no longer be considered recurrent. To get to the dynamics proper, a strategy is to calculate the total percentage of recurring points (a variable calculated under the RQA algorithm) for several increasing values of \( r \) and to plot the results on a log-log plot to determine a “scaling” region, i.e., where the dynamics are found. If the data are extremely nonstationary, a scaling region may not be found. The guide then is the percent recurrence. A critical factor is that there be sufficient numbers of recurrences so as to make sense for computation of the other variables such as maxline. A value of approximately 1% recurrence tends to fulfill this criterion. Further verification can be obtained from an inspection of the recurrence plot: too sparse a matrix suggests a modest increase. Windowed RQA is especially informative. If a given window fails to achieve 1% recurrence, the radius should be increased, or the window enlarged.

2.3. Experiments

So-called “runaway” and “non-runaway” esterification experiments were performed to test the idea that RP-based PLEs could effectively monitor the reactions as a “proof of concept”.\(^1\) The experiments were based on typical reactions performed on large-scale bases in industrial settings, and are a constant threat to safe operation, redounding also to concerns for environmental pollution. The term “runaway” refers to loss of control of the process to the point that the temperature increases suddenly with possible explosive results (traditionally, the divergence in such a scenario has been monitored by the second derivative). Considerable effort has been expended to devise methods to provide early warning for such events. The experiments were performed in “batch” (reactants added simultaneously), and “semibatch” (reactant added during the experiment at a constant rate) conditions. One runaway experiment was performed under “isothermal” conditions, i.e., an attempt was made to control the temperature inside the reactor by adjustment of the reactor jacket temperature. Additionally, intermediate profiles were carried out, as well as controlled experiments. In all cases the variable of interest was temperature as monitored by sensors within the reactors.\(^2\) Embedding was set at 10, delay was set at 5, and radius of 3 normalized units (i.e., on the unit interval). A 50 or 100 point window was chosen to obtain optimal sensitivity. Thus, the maximal line segment measurable depends upon this window length criterion.

3. Results and discussion

What was noteworthy was that most of the data had minimal noise, such that the 2nd derivative could easily register a transition. (The maximal noise remained constant across the different experiments as registered by the 2nd derivative.) Important also is the interpretation of the maxline variable relative to the dynamics: recall that a positive Liapunov exponent suggests instability. In this case, a large maxline indicates less instability than a smaller one (albeit it is still instability), relative to the local (windowed) dynamics. Thus, the maxline cannot inform an observer if the dynamics are “dangerous”—only that they are changing. And it is the direction of the measured time series that defines such a condition. Thus Fig. 1 depicts the results of a batch runaway experiment under isoperibolic (constant jacket temperature) conditions (2 L reactor).

---

\(^1\) Specific details of the experiment will be set forth in a separate article.

\(^2\) The choice for embedding was based on previous studies demonstrating that typically only three variables were needed to model the dynamics [9,10], but that given the effects of noise, an embedding of 10 was required. See also Section 2.2.1 above and Ref. [17].
Clearly the temperature increases, however, the second derivative becomes significant (registering a large divergence) only at the last moment. The maxline describes an interesting scenario: there appears to be some noise-based variation until approximately 1100 s, at which point, the value becomes very low indicating (relatively) high instability proceeding until approximately 1600 s (first solid bar). At this point, the maxline increases and stays approximately the same for the next few hundred seconds (next solid bar). The interpretation may follow along these lines: shortly after the beginning of the reaction, the temperature begins to rise slowly (in an unstable way), i.e., no clear indication of runaway. However, it suddenly “stabilizes” in this temperature increase because the deterministic recurrent points “pile-on” quickly because of its acceleration. After 2000 s the temperature decelerates in a “stable” way (punctuated by some noise). In
Fig. 3. No reaction. Operator manipulation in heating/cooling experiment. (Dashed line denotes jacket temperature.)

Fig. 4. Semibatch runaway experiment in 100 L reactor.

contradistinction to this, the 2nd derivative registers a change only at the last few seconds as the temperature maximum is achieved. In Fig. 2, a batch runaway (2 L reactor) cannot be controlled by changing jacket temperature (dashed line).

The second derivative momentarily registers a small change at the beginning, and again just before the maximum. The maxline, however, is sensitive initially to the increasing temperature (low maxline), as well as, apparently, to the subtle changes involved in changing jacket temperature. This is contrasted with Fig. 3 which depicts a simple experiment without a reaction. Instead, the jacket temperature controls the reactor temperature. Here there is no subtle increase in temperature, and in the acceleration phase, the maxline is very uneven and promptly registers the fact that the reactor temperature has stabilized. Note that there is no maxline even indicated...
Fig. 5. Intermediate profile of semibatch experiment. Initial runaway is controlled.

Fig. 6. As in Fig. 5, but the runaway is greater.

Fig. 4 is similar to the process in Fig. 1 except that this is a semibatch experiment, and the reactor was much larger (100 L). Thus the method is robust with rescaling.

Figs. 5 and 6 present the results for intermediate profiles of semibatch experiments in 100 L reactors. What is clearly of interest is that in neither of these experiments, does the second derivative hint of any accelerating process. No doubt this is due to the control of the process at the last few seconds. However, if the maxline criterion is invoked, the interpretation could easily be that runaway was incipient. Clearly, this experiment was contrived to create a scenario of runaway controlled at the last minute. Note also that as the processes change (decrease temperature), small maxline increases are registered.
Finally, Fig. 7 demonstrates results in a controlled reaction (semibatch). The maxline briefly increases, is not sustained, and quickly dies out, indicating absence of runaway.

In all these examples, it should be underscored that as with all experimental algorithms involved in the calculation of PLEs, there is no comparison of the phase-space volume. This is especially obviously the case with nonstationary, transient data. Thus, if there is a change in the exponent, no evaluation can be made to determine multi-dimensional volume direction. To be useful, then, as has been previously indicated, it is important to track simultaneously the direction of the temperature (or other variable of interest). Related to this last point is the question of significance of the maxline changes. Clearly some noise can trigger maxline changes, and in the case of industrial processes, false alarms could be costly. Since these experiments were designed as ”proof of concept”, no clear criteria were developed to determine level changes of maxline to sound warnings. This was to recognize that real conditions would be scaled up and require further evaluation. Nonetheless, previous experience with other systems suggest that an appropriate algorithm is feasible. Specifically, baseline 95% confidence limits for maxline can be obtained and used as a boundary for significant increases/decreases and compared to temperature direction change [12,22].

4. Conclusions

Although, numerous studies have attempted to document that a system is chaotic, it should be recalled that PLEs have a separate utility; namely, to determine the stability of a system. The present Letter documents this utility by using a simple RP-based algorithm which can be implemented on-line. Clearly, other systems which are typically noisy and nonstationary, such as biological, financial and physical may also profit from such analyzes.

Acknowledgements

This research has been supported by the EU funded project AWARD (contract G1RD-CT-2001-00499) of the GROWTH programme of the European Commission.

References