Article

Dynamic Complexity Measures: definition and calculation

José Roberto C. Piqueira

Escola Politécnica da Universidade de São Paulo
Departamento de Engenharia de Telecomunições e Controle
Av. Prof. Luciano Gualberto, travessa 3, n. 158
05508-900, São Paulo, Brasil
*   piqueira@lac.usp.br; Tel.: +55-11-3091-5221

Abstract: This work is a generalization of the López-Ruiz, Mancini and Calbet (LMC); and Shiner, Davison and Landsberg (SDL) complexity measures, considering that the state of a system or process is represented by a dynamical variable during a certain time interval. As the two complexity measures are based on the calculation of informational entropy, an equivalent information source is defined and, as time passes, the individual information associated to the measured parameter is the seed to calculate instantaneous LMC and SDL measures. To show how the methodology works, an example with economic data is presented.

Keywords: Complexity; disequilibrium; equilibrium; individual information; informational entropy.

Key contribution: LMC and SDL complexity measures are generalized for temporal series
Calculations are based on the construction of an informational source
LMC and SDL measures present the same qualitative dynamics
Partition of the range interval depends on the observation time interval

---

1. Introduction

The word complexity, in the common sense meaning, represents systems that are difficult to describe, design or understand. However, since Kolmogorov presented the concept of computational complexity [7], new ideas have been associated to this word, mainly concerning life sciences [1], relating complexity with information [5].

As a consequence, complexity started to be related to open systems and to the emergence of unexpected behaviors, due to nonlinearities [11,13] and, concerning system theory [25], a new meaning was carved, considering that complexity is half way of the equilibrium and disequilibrium [6].

Considering this idea, López-Ruiz, Mancini and Calbet proposed, in a seminal paper [8], the LMC complexity measure by using the informational entropy [22] to evaluate equilibrium, and a quadratic deviation from the uniform distribution to evaluate disequilibrium.

However, there are some criticisms about LMC measure, considering that it is inaccurate for some classes of systems obeying Markovian chains and can not be considered an extensive variable. Feldman and Crutchfield [4] proposed a correction for the disequilibrium term, replacing it by the relative entropy with respect to the uniform distribution.

Following this line, Shiner, Davison and Landsberg proposed another modification of the LMC measure, replacing the disequilibrium term by the complement of the equilibrium term. This measure

is called SDL [23] and presents results qualitatively similar to the obtained by using LMC, for the majority of usual statistical distributions [15].

The main restriction to LMC and SDL complexity measures is due to Crutchfield, Feldman and Shalizi, as they argue that an equilibrium system can be structurally complex [3], but this problem could be solved weighting order and disorder, according to the specific problem to be analyzed [16].

Since the early 2000's, the idea of adapting LMC and SDL to dynamical systems was successfully applied to different types of time evolution problems: sleep-awake cycle [21], bird songs [24], neural plasticity [10,14], interactions between species in ecological systems [1,19], physiognomies of landscapes [9], economical series [12,20], spread depression [18] and quantum information [2,17].

With these ideas in mind, this article presents a systematization of the methodology used in the referred papers, based on LMC and SDL measures, to be applied to temporal series, by defining and calculating dynamical complexity measures.

The procedure, applied to a temporal series representing some organizational or dynamical aspect of a system, provides hints regarding the evolution of its complexity.

As the LMC and SDL dynamical measures are based on informational entropy [12], the first task, described in the next section, is to define an alphabet source, associating a probability distribution to the possible system states.

Following the definition of the probability distribution, a new section defines how dynamical LMC and SDL measures can be calculated at each time, based on the individual information associated to the system state at this time, generating temporal series for LMC and SDL measures.

To illustrate the calculation procedure, an example related to an economical time series is presented and, in the same section, a practical discussion about how to divide the range of the values assumed by the system state is presented.

The work is closed with a conclusion section, emphasizing that the calculations presented can be applied to any kind of temporal real numbers series.

2. Defining source and probability distribution for a temporal series

Considering the Shannon's model [22] for an information source, a time series $x(n)$ is considered to be a function of the integers into a real interval, i.e., $x(n) : Z_+ \to (a, b)$, associating to each time $t_0 + nT$ a real number belonging to (a, b), with $t_0 > 0$ being the initial instant and $T > 0$ an arbitrary period, depending on the data availability.

The first step is to divide the interval (a, b) into N sub-intervals. For the sake of simplicity, N is chosen equal to $2^k, k \in Z_+$.

At this point, it could be asked how to choose N, as there is a compromise between precision (high values of N) and speed of calculation (low values of N). This question will not be addressed theoretically; however in the example section, practical hints about the choice will be presented.

Consequently, the source's alphabet is defined by the intervals $A_i, i = 1, ..., N$, with $\bigcup_{i=1}^{N} = (a, b)$ and $A_i \bigcap A_j = \phi, \forall i \neq j$.

Then, a time interval defined by a given n must be chosen, and for the time sequence $t_0, t_0 + T, ...t_0 + nT$ the values of the variable $x(n)$ must be read and associated to the intervals A_i, containing the respective value.

Therefore, for the whole set $t_0, t_0 + T, ...t_0 + nT$, each interval A_i is associated to $x(n)$ a certain number of times n_i, which defines a relative frequency $p_i = \frac{n_i}{n+1}$, taken as a probability, associated each interval A_i.

Following the definition, for each sub-interval $A_i \in (a, b)$, its individual contribution to the whole information entropy is given by: $S_i = -p_i log_2 p_i$; and the maximum value of the informational entropy for the whole source, $S_{max} = log_2 N = k$, can be calculated [22].

3. Dynamical LMC and SDL

As the source alphabet and individual information were defined, the instantaneous values of $x(n)$ are associated to the respective S_i, allowing the calculation of the instantaneous value of the equilibrium (disorder) term:

$$\Delta(n) = \frac{S_i}{k}. \tag{1}$$

Combining (1) with the different definitions of the disequilibrium (order) terms, dynamical LMC and SDL measures are defined.

3.1. LMC dynamical measure

As indicated by López-Ruiz, Mancini and Calbet [8], the dynamic disequilibrium (order) term can be calculated as the quadratic deviation of the source probability distribution from the uniform distribution and, consequently, the individual contribution of the interval A_i is:

$$D(n) = (p_i - 1/N)^2. \tag{2}$$

Extending the definition of LMC measure, dynamical LMC, calculated in $t_0 + nT$, is given by:

$$C_{LMC}(n) = \Delta(n).D(n). \tag{3}$$

3.2. SDL dynamical measure

As proposed by Shiner, Davison and Landsberg [23], the dynamic disequilibrium (order) term can be calculated as the complement of the dynamic equilibrium term:

$$D(n) = (1 - \Delta(n)) \tag{4}$$

Extending the definition of SDL measure, dynamical SDL, calculated in $t_0 + nT$, is given by:

$$C_{SDL}(n) = \Delta(n).D(n). \tag{5}$$

4. Applying the method: practical hints

In this section, the economic series relative to the conversion of currencies studied in [12,20] is taken as an example, concerning only to the methodological point of view, without any economic conjecture about the results.

The daily conversion rate dollar to Brazilian real (USD/BR) temporal series, from January 1999 to September of 2015, shown in Fig.1 [12,20] represents the value of $x(n)$, which complexity is analyzed.

Consequently, the interval (a, b) related to the excursion of x is $(1.207; 4.178)$. It is divided into $8(k = 3), 16(k = 4), 32(k = 5)$ and $64(k = 6)$ subintervals to build the sources and the respective probability distributions.

Based on these probability distributions, $C_{LMC}(n)$ and $C_{SDL}(n)$ are calculated and plotted giving an idea about how measure choice and interval division affects the results.

4.1. Equivalence between LMC and SDL

Dividing the range of $x(n)$ into 8 parts, the results of the calculation of $C_{LMC}(n)$ and $C_{SDL}(n)$ measures are shown in Fig. 2a and 2b, respectively.

As Fig. 2a and 2b show, in spite of the numerical differences, the time evolution of $C_{LMC}(n)$ and $C_{SDL}(n)$ are qualitatively the same and represented by very similar curves, in the eight-part division case.
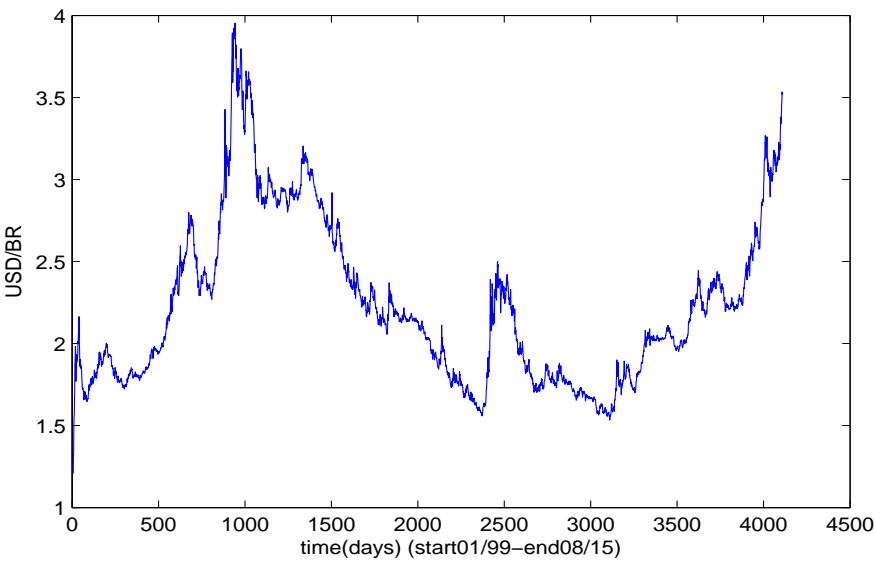
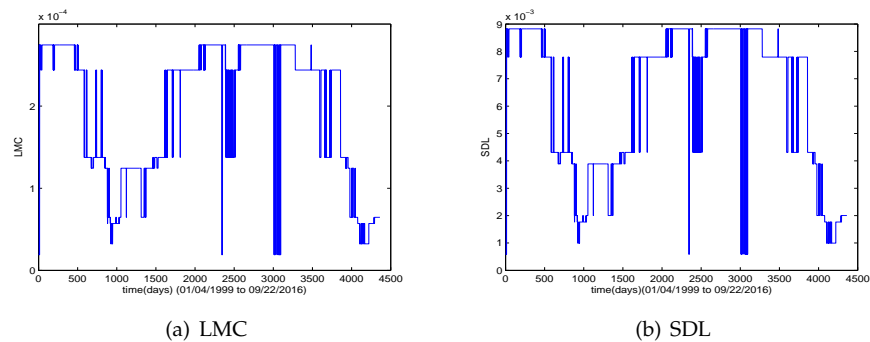**Figure 1.** USA Dollar-Brazilian Real exchanging rate

(a) LMC

(b) SDL

Figure 2. Temporal evolution of complexity (8-part division)

(a) LMC

(b) SDL
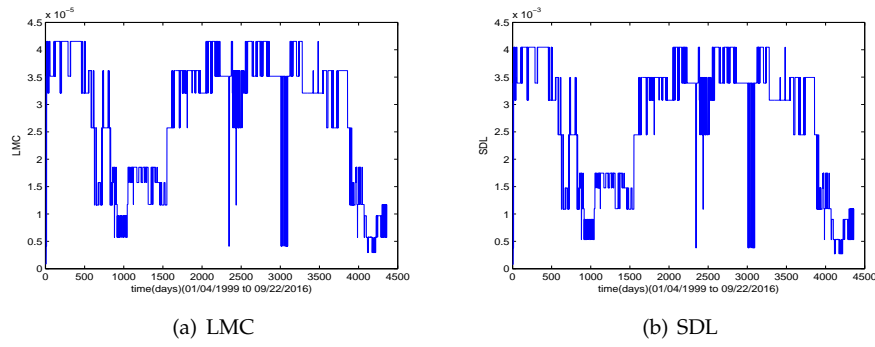
Figure 3. Temporal evolution of complexity (16-part division)
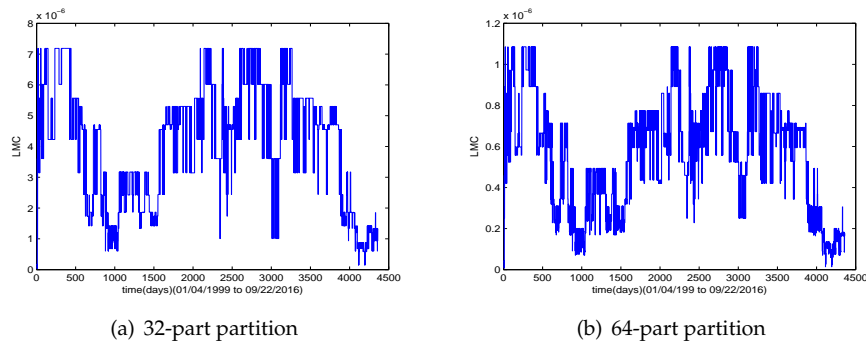

(a) 32-part partition

(b) 64-part partition

Figure 4. Temporal evolution of LMC complexity

If the range of $x(n)$ is divided into 16 parts, Fig. 3a and 3b show the results for $C_{LMC}(n)$ and $C_{SDL}(n)$.

It can be observed that, in this case (sixteen-division case), $C_{LMC}(n)$ and $C_{SDL}(n)$ differ only by a scale factor, with the same qualitative time evolution.

Comparing Fig. 2a and 3a, $C_{LMC}(n)$ for different range partitions, the whole qualitative aspects of the curves are the same and by increasing the number of divisions, the dynamical range of the measures changes, implying some rapid oscillatory variations, similar to noise.

Comparing Fig. 2b and 3b, $C_{SDL}(n)$ for different range partitions, the whole qualitative aspects of the curves are the same and the noisy aspect due to the increasing number of interval divisions is maintained.

Consequently, from now on, only LMC measure will be analyzed, since SDL presents the same qualitative dynamical behavior and partition sensitivity.

4.2. Range interval partition

By increasing the number of intervals of the $x(n)$ and recalculating $C_{LMC}(n)$, the result for a thirty-two partition is shown in Fig. 4a and, for a sixty-four partition, in Fig. 4b.

By analyzing the results from figures 2a, 3a, 4a and 4b it could be observed that by increasing the number of intervals, the maximum value of $C_{LMC}(n)$ decreases improving the precision but, apparently, for this long series the temporal evolution of $C_{LMC}(n)$ maintains its qualitative behavior mixing noise with accuracy.

Attempting to be more precise about how the range interval partition, the $C_{LMC}(n)$ is calculated for the several partitions, but considering a shorter time period for the data. The interval between July and December of 2002 is chosen, because, as explained in [12], it is critical concerning the conversion rates in Brazil.
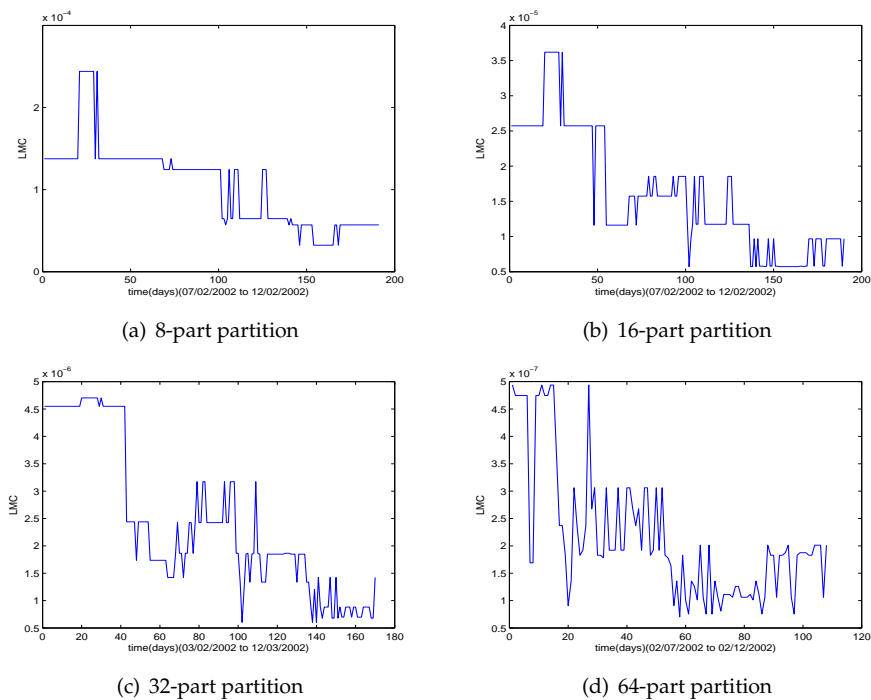
(a) 8-part partition

(b) 16-part partition

(c) 32-part partition

(d) 64-part partition

Figure 5. LMC temporal evolution for shorter time intervals

Figures 5a, 5b, 5c, 5d show the LMC dynamical measure calculated, for the initial set of data, with the range interval divided into 8, 16, 32 and 64 parts, respectively.

It can be observed from these results that, for shorter intervals, the general qualitative characteristics of the time evolution appear, independently on the partition. However, as the number of sub-intervals increases, the instantaneous numerical values change but the precision increases, allowing more accurate analysis.

5. Conclusions

A methodology for calculating LMC and SDL dynamical complexity was developed, starting with the construction of a source and a probability distribution, for any temporal series.

It was observed that LMC and SDL measures, in spite of presenting different numerical results, have very similar temporal evolution, related to the variable $x(n)$.

A point that is always object of discussion is the range interval partition. The choice of the number of sub-intervals is a matter of experience.

Long time intervals are not so sensitive to the increase of the number of divisions; however, for short time intervals, increasing the number of divisions provokes a precision improvement, in spite of the introduction of an apparent noise.

1. M. Anand & L. Orlóci, "Complexity in plant communities: the notion of quantification", *Journal of Theoretical Biology*, **179**, (1996) pp. 179-186.

2. Y. C. Campbell-Borges & J. R. C. Piqueira, "Complexity Measure: A Quantum Information Approach", *International Journal of Quantum Information*, **10**, (2012), pp. 1250047-1-1250047-19.

3. J. P. Crutchfield, D. P. Feldman & C. R. Shalizi, "Comment I on "Simple measure for complexity"", *Physical Review E*, **62**, (2000), 2996.

4. D; P. Feldman & J. P. Crutchfield, "Measures of statistical complexity: Why?", *Physics Letters A*. **238**, (1998) pp. 244-252.

5. H. Haken, *Information and Self-Organization*, (2000), Springer-Verlag: Berlin, Germany.

6. K. Kaneko & I. Tsuda, *Complex Systems: chaos and beyond*, (2001), Springer Verlag: Berlin, Germany.

7. A.N. Kolmogorov, "Three approaches to the definition of the concept "quantity of information"", *Problemy Peredachi Informatsii*, **1** (1965) pp. 3-11.

8. R. López-Ruiz, H. L. Mancini and X. Calbet, "A statistical measure of complexity", *Physics Letters A*, **209** (1995) pp. 321-326.

9. S. H. V. L. de Mattos, L. E. Vicente, A. Perez Filho, J. R. C. Piqueira, "Contributions of the complexity paradigm to the understanding of Cerrado's organization and dynamics", *Anais da Academia Brasileira de Ciências*, **88(4)**, (2016), pp. 2417-2427.

10. M. Mazza, M. Pinho, J. R. C. Piqueira & A. C. Roque, "A Dynamical Model of Fast Cortical Reorganization", *Journal of Computational Neuroscience*, **16(2)**,(2004), pp. 177-201.

11. E. Morin, *On complexity*, (2008), Hampton Press: New York, USA.

12. L. P. D. Mortoza & J. R. C. Piqueira, "Measuring complexity in Brazilian economic crises", *PLOSONE*, **12(3)**, (2017), e0173280.

13. G. Nicolis & I. Prigogine, *Self-Organization in Nonequilibrium Systems*, (1977), John Wiley & Sons: USA.

14. M. Pinho, M. Mazza, J. R. C. Piqueira & A. C. Roque, "Shannon's entropy applied to the analysis of tonotopic reorganization in a computational model of classic conditioning", *Neurocomputing* **44-46** (2002), pp. 923-928.

15. J. R. C. Piqueira, "A comparison of LMC and SDL complexity measures on binomial distributions", *Physica A*, **444**, (2016) pp. 271-275.

16. J. R. C. Piqueira, "Weighting order and disorder on complexity measures", *Journal of Taibah University for Science*, **11(2)**, (2017), pp. 337-343.

17. J. R. C. Piqueira & Y. C. Campbell-Borges, "Extending SDL and LMC complexity measures to quantum states", *Physica. A*, **392**,(2013), pp. 5255-5259.

18. J. R. C. Piqueira, V. M. F, de Lima & C. M. Batistela, "Complexity measures and self-similarity on spreading depression waves", *Physica. A*, **401**, (2014), pp. 271-277.

19. J. R. C. Piqueira, S. H. V. L. Mattos & J. Vasconcelos Neto, "Measuring complexity in three-trophic level systems", *Ecological Modelling*, **220(3)**, (2009), pp. 266-271.

20. J.R. C. Piqueira & L. P. D. Mortoza, "Brazilian Exchange Rate Complexity: financial crisis effects", *Communications in Nonlinear Science & Numerical Simulation*, **17(4)**, (2012), pp. 1690-169.

21. J.R. C. Piqueira & A. A. B. Silva, "Auto-Organizaçç Complexidade: o problema do desenvolvimento do ciclo vigília-sono", *Estudos Avançados*, **12(33)**, (1998), pp. 197-212. (http://dx.doi.org/10.1590/S0103-40141998000200015).

22. C.E. Shannon & W. Weaver, *The Mathematical Theory of Communication*, (1963), Illini Books Edition: Urbana and Chicago, USA.

23. J. Shiner, M. Davison and P. Landsberg,"Simple measure for complexity", *Physical Review E*, **59(2)** (1999) pp. 1459-1464. 8-part partition

24. M. L. Silva, J. R. C. Piqueira & J. M. E. Vielliard, "Using Shannon Entropy on Measuring the Individual Variabilty in the Rufous-bellied Thrush Turdus rufiventris Vocal Communication", *Journal of Theoretical Biology*, **207(1)**, (2000), pp. 57-64.

25. L. Von Bertalanffy - General System Theory: Foundations, Development, Applications- George Braziller Inc.-New York 1968.