Research Article

Proposal of Two Measures of Complexity Based on Lempel-Ziv for Dynamic Systems: An Application for Manufacturing Systems

Miguel Alfaro,1 Manuel Vargas,2 Guillermo Fuertes,1,3 and Juan Pedro Sepúlveda-Rojas1

1 Industrial Engineering Department, University of Santiago de Chile, Avenida Ecuador 3769, Santiago de Chile, Chile
2 Facultad de Ingeniería y Tecnología, Universidad San Sebastian, Bellavista 7, Santiago de Chile, Chile
3 Facultad de Ingeniería, Ciencia y Tecnología, Universidad Bernardo O’Higgins, Avenida Viel 1497, Ruta 5 Sur, Santiago de Chile, Chile

Correspondence should be addressed to Guillermo Fuertes; guillermo.fuertes@usach.cl

Received 18 August 2017; Revised 23 November 2017; Accepted 5 December 2017; Published 29 January 2018

Academic Editor: Jonathan N. Blakely

Copyright © 2018 Miguel Alfaro et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The measure of complexity of Lempel-Ziv (LZC) is used for the analysis of time series generated by dynamic systems with the objective of determining its complexity. This work measures LZC of different series coming from periodic functions, probabilistic functions, and chaotic systems. Later, these metrics of complexity are applied to the average of the number of parts in the waiting line in a manufacturing workshop. The results demonstrate the efficiency of the LZC to establish the level of the dynamic behavior of the manufacturing systems, through the time series.

1. Introduction

For Efthymiou et al. [1] a complex system is composed of a structure defined for a great amount of elements, with different types of interactions and characterized by a behavior with uncertainty. Preliminary studies published by Alfaro and Sepulveda [2] consider the manufacturing systems as complex systems. A manufacture system consists of many entities: flexible machines, hundreds of products, thousands of operations, location of the operations, variable context in the political, economic, and environmental factors, and uncertainty of the demands; these features define the structure of a complex system. In Elmaraghy et al. [3] other sources of complexities in the manufacturing systems are defined: customer requirements, products complexity, control systems, product structure, and manufacturing technology.

Complex systems are also found in natural, social, and economic systems. Nonlinear relations of the variables composing the system structured its complex behavior. Abarbanel and Lall [4] studied the time series of the aquifer volume of the Great Salt Lake of Utah, using the theory of nonlinear systems, and obtained as a result the chaotic behavior of the series. Other natural complex systems are presented in the patterns of natural formation of lands [5], in systems consisting of multiple natural and human subsystems [6] and Holling [7] develops the term panarchy to expose the evolution of adaptive systems formed by interconnected natural, human, and socioecological systems.

The measures of complexity in the manufacturing system can be separated into two types of metrics: (1) metric oriented to the measure of the static or structural complexity and (2) metric directed to determine the dynamic complexity of the system. According to Modrak et al. [8] the static complexity is defined by the configuration of the system and the number of components. The dynamic complexity studies the uncertainty of the behavior of the system. The research is centered in the metric of the complexity of a manufacture system and of a dynamic system, independent of the variables and the structure of the system.

To measure the structural complexity, [9–12] have based their studies on the entropy of Shannon, deriving mathematical expressions to measure static complexity, through
the study of the amount of information needed to describe the system. Zhu et al. [10] study the complexity through entropy and Wang and Hu [11] develop the relation between the static complexity and the different configurations of the assembling system and its productivity, using the entropy. Zhu et al. [12] propose a dynamic discrete optimization model using the entropy to minimize the complexity in an assembling system using the entropy. ElMaraghy et al. [13] present a measure of static complexity based on entropy to determine the amount and variety of information related to the machines, equipment, inventories, and material handling. In Modrak and Marton [14] the entropy is used to measure the static complexity and they utilized it in the design of the structure of a supply chain. Phukan et al. [15] use a derivative of the measures of complexity of the entropy to determine the static complexity of manufacturing systems in the design of production systems. Samy et al. [16] present a measure of structural complexity based on the granularity model allowing achieving a balance between the complexity of the machines and the layout of the productive system. Elmaraghy et al. [17] presented a model based on a set of indicators that allow measuring the performance associated with the structural complexity of the layout.

With regard to the measures of the dynamic complexity applied to the manufacturing systems, in [18] the fractal dimension of Lyapunov are described, obtained as of the Kaplan-Yorke conjecture, and the relation between the dynamic complexity and the flexibility in a production system is studied. Wu et al. [19] analyzed the relation between the operational complexity and the costs of the inventories. The authors define the operational complexity as the amount of information needed to describe the state of deviation with respect to the ideal system, measuring the entropy of the waiting line of the system. Papakostas et al. [20] propose a measure of dynamic complexity, inspired by the definition of the Lyapunov exponents, to determine the sensitivity of the variables with the complexity of the manufacture system. Zhang [21] proposes measures based on the entropy to calculate the static and dynamic complexity. Efthymiou et al. [22] suggest three inconveniences of the metrics of entropy for the dynamic complexity: (1) it is difficult to determine the state of an element or subsystem; (2) the condition of Independence of events is not accomplished; (3) it is difficult to identify the parties of the system to reduce the complexity.

The theory of nonlinear dynamic systems gives tools to study the systems through the analysis of a time series [23]. A system with chaotic behavior can be rebuilt and studied through the attractor of the system. This methodology is particularly useful when the equations that describe the system are unknown. However, this reconstruction is particularly difficult and requires special conditions of nonstationary time series and determinism of the system. According to [2] the flexible and automatized manufacture systems under certain operation conditions show conditions of chaotic behavior. In [24, 25] the relation between chaos and the manufacturing systems is studied. For Fuertes et al. [26] the chaotic behavior is a form of complex behavior; nevertheless the use of chaos behavior to characterize the dynamic complexity is restricted to systems with chaotic behavior. Donner et al. [24] propose the analysis of Recurrence Quantification Analysis (RQA) to characterized the dynamic of a manufacturing network. March et al. [25] have studied the relation between the dimension of inlay and the measures that gives the RQA for some systems of low dimension which are not modified regarding the inlay dimension. However, Hasson et al. [27] present the relation of the sensitivity of the RQA variables to the inlay dimension. The method RQA analyzes the predictability of the behavior of a system of manufacture. Efthymiou et al. [28] propose to measure the dynamic complexity of a manufacturing system through the LZC measure in the automotive industry. Kedadouche et al. [29] compare the methods of processing signals based on the LZC complexity, approximate entropy (ApEn), and the simple entropy (SampEn), for the early detection of failures or defects in the industrial gearboxes. The symbols used in this research are listed in Notations.

The study searches a complexity metric to study the dynamic complexity of a system. The complexity metric uses the times series called work in process (WIP) from the system of manufacture. The measurement of the dynamic complexity of the studied productive systems delivers the sensitivity of the systems to variables related to the productive structure of the system. Productive systems with a high level of dynamic complexity with values closer to one (1) are related to stochastic structures of unpredictable behavior [30].

The complexity measures of dynamic complexity are mainly based on the maximal Lyapunov exponents. For the validation of the metric based on the maximal Lyapunov exponents, it is required to characterize the time series to confirm its chaotic behavior. The two LZC metrics presented in the research can be used in any time series and each one presents different levels of sensitivity allowing a more accurate analysis of the dynamic system.

The proposed application develops two metrics based on the Lempel-Ziv complexity; these two metrics allow studying the behavior of stochastic series, chaotic, of natural phenomenon, such as the level of volume of the Great Salt Lake of Utah, and periodic series to establish a reference frame of study for time series of comparable productive dynamic systems. The studied productive system corresponds to a flexible system of manufacturing and the variable that will be analyzed is the average of the amount of the pieces in the waiting line of flexible manufacturing. The manufacturing systems can process three types of products and the machine shop is formed by three different machines; each operation can be executed on any machine. The experiment modifies the setup times and the variability of the demand of the products to evaluate the impact of these variables in the system of manufacture and its dynamic complexity.

The research set two metrics based on Lempel-Ziv analyzed and classified complex systems represented by time series. The paper is organized as follows: section two develops the Lempel-Ziv indicator, section three details the methodology, section four presents the results of the indicators and their validation as measures of complexity of time series of the work in process of a manufacturing system and the aquifer volume of The Great Salt Lake in Utah, and section five sets the final conclusions of the study.
2. Lempel-Ziv Complexity

The calculation of the Lempel-Ziv complexity [31] is done through the transformation of the time series in a sequence of symbols $P$ [32, 33]. This $P$ sequence is examined from left to right. A counter $c(n)$ is defined that increased in a unit each time a new subsequence of different consecutive characters is found, reflecting the apparition of new patterns along $P$.

The algorithm uses only two simple operations: comparison and accumulation, which allows calculating the complexity in a simple way and with a reduced computing cost [32]. It is possible to demonstrate the superior limit of $n$ given by

$$\lim_{n \to \infty} c(n) = b(n) = \frac{n}{\log_2(n)},$$ (1)

where $\alpha$ is the number of symbols. Therefore, the counter of complexity $n$ normalized is defined as

$$C(n) = \frac{c(n)}{b(n)},$$ (2)

where $n$ represents the apparition rate of new patterns along the sequence. Therefore, it is verified that this apparition rate tends to zero when the number of elements $n$ tends to infinity. The patterns are repeated when the series is extended. The counter $n$ will be smaller in relation to the amount of characters of the analyzed series. As a consequence, to compare the complexity of different systems, it is required to set an equal size for all the series. The interpretation of the extension of the series is the loss of uncertainty as a product of the increase of information. This main characteristic corresponds to a measure of dynamic complexity.

3. Methodology

For all time series two different LZC measures are calculated, through the conversion of the original series, using a binary series and a series based on three symbols (base 3).

For converting the $X$ original series to a binary series $S_{01}$, the mathematical formula described in (3) is used: the first element of the series $S_0$ is always 0; if the differential of $x_{i+1} - x_i$ is greater than or equal to zero, then the element $s_{01,i+1}$ is 1; on the contrary it is 0. Once the series $S_0$ is obtained we proceed to the calculation of $LZC_{01}$, associated with the variation of the elements of the binary series.

$$s_{01,1} = 0,$$

$$s_{01,i+1} = \begin{cases} 0 & x_{i+1} - x_i < 0, \\ 1 & x_{i+1} - x_i \geq 0, \end{cases}$$ (3)

$$i = 1, 2, \ldots, N - 1.$$

The second method to develop the series $S_{012}$ is started given an $X$ time series and the function of density of probability $f(x)$ is estimated through the estimation based on the normal kernel function [33], assuming a defined series defined by a stochastic system and a probability $P$ is calculated given by

$$P = \frac{1}{3 + 10 \max(f(x))}.$$ (4)

Once this probability is calculated, the limits of the ranks $x_l, x_m$ are calculated, through the functions of accumulation of inverse probability $F^{-1}(p)$, through

$$x_l = F^{-1}(p),$$

$$x_m = F^{-1}(1 - p).$$ (5)

The original $X$ series becomes the series $S_{012}$ of three symbols $(0, 1, 2)$, according to the mathematical formula described in (6). Obtaining the series of three symbols we proceed to the calculation of $LZC_{012}$, this indicator measures the complexity of the series based on the behavior of the elements according to their concentration.

$$s_{012,i} = \begin{cases} 0 & x_i < x_l, \\ 1 & x_l \leq x_i \leq x_m, \\ 2 & x_i > x_m. \end{cases}$$ (6)

3.1. Time Series. The analyzed deterministic series are the cosine functions and the time series of the chaotic series of Mackey-Glass and Rössler and Lorenz. The length of the series is set in six different magnitudes (from 2001 to 12001 elements) and a step of evaluation and integration $1/t$ with $t$ being variable from 1 to 100. The initial condition for the cosine series is $t = 0$, for the equation of Mackey-Glass is $x(0) = 0$, for the Rössler system is $(x(0), (0), z(0)) = (-10, -10, 0)$, and for the Lorenz system is $(x(0), y(0), z(0)) = (5, 5, 10)$.

The stochastic series studied are the normal, exponential, and rectangular distributions, with length of series identical to the determinism series. The same number of experiments are done varying the parameters of each distribution from 1 to 100. The deviation standard for the normal distribution, the average for the exponential distribution, and the superior limit for the rectangular distribution are the parameters modified in the range of whole values $[1 : 100]$. The average for the standard deviation and the lower limit for the rectangular distribution keep constant with zero value.

The study is developed in three phases: (1) each experiment was simulated 100 times in MatLab® to obtain the average of each indicator of complexity; (2) the comparison of the three stochastic distributions with series of 4001 length and equal dispersion of data was done. The variance is set according to Table 1 and takes the whole values from 1 to 100; (3)

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>$\sigma^2$</td>
</tr>
<tr>
<td>Exponential</td>
<td>$b^2$</td>
</tr>
<tr>
<td>Rectangular</td>
<td>$(B - A)^2$</td>
</tr>
<tr>
<td></td>
<td>12</td>
</tr>
</tbody>
</table>

The table is defined as in Table 1 and takes the whole values from 1 to 100; (3)
the analysis and classification of the time series of a flexible manufacturing line and the series of the total volume of The Great Salt Lake in Utah were performed.

3.2. Flexible Manufacturing System. Authors as Alfaro et al. [34] developed a flexible manufacturing line, with three machines or productive cells where three products are manufactured. The flexibility of the manufacturing line allows executing the manufacturing operations in any cell of production. The operation of the system is built based on the first-in first-out (FIFO) productive model; the assignment to the production cell for the new product is the machine with low waiting level and the system is balanced (rates of arrivals of the products never tend to zero or infinity). The rate of arrival of each product corresponds to the vector $B_j (7.28, 7.94, 15.81)$ with unit (amount of products/for time unit); the parameters of manufacturing time are deterministic (Table 2).

The variable in study of the manufacturing system is work in process. The characterization of the time series WIP confirms a chaotic behavior associated with a dynamic system. The maximum exponent of Lyapunov has a positive value confirming the chaotic character of the time series (Figure 1).

To compare and analyze the relation between the complexity metric of the time series WIP and the parameters of the flexible manufacturing system, three experiments are performed: case 1 corresponds to the system with the parameters established in the literature of [34]; in case 2 the times of manufacturing of each product are increased; in case 3 the times of arrivals of the products are reduced. The rates of increase and reduction are similar in both cases (30%). Finally establishing the simulations of the experiments we proceed to calculate the complexity metric of each series of time of the WIP.

4. Results

The results of the metrics of $LZC_{012}$ and $LZC_{01}$ of the eight following functions: (1) cosine function, (2) Lorenz system, (3) Rössler system, (4) Mackey-Glass equation, (5) distribution of the normal probability, (6) distribution of rectangular probability, (7) exponential function, and (8) stochastic distributions of equal variance, are studied through the LZC graphics (Figures 2–17).

Figures 2 and 3 present the $LZC_{012}$ and $LZC_{01}$ average value for cosine function. The tendency presented by the series is a decrease in the $LZC_{012}$ and $LZC_{01}$ complexity values.
with the increase of the number of elements of the series. This variation is independent of the evaluation step.

Figures 4 and 5 show the LZC₀₁₂ and LZC₀₁ average values for the Lorenz system. The study for the system of Lorenz does not converge for lower integration steps 1/13 and the value LZC₀₁₂ decreases together with the reduction of the step or the increase of the size of the series. The LZC₀₁ value has a similar behavior to LZC₀₁₂, having only one difference in the highest value; the series present more complexity due to the data variation.

Figures 6 and 7 show the LZC₀₁₂ and LZC₀₁ average value for the Rössler system. The Rössler system has a similar behavior to the Lorenz system, with a superior step at 1/7; the
system converges at one point. The LZC₀₀ value for the same step is lower than the LZC₀₀₂ value; this indicates a behavior of the more complex series for the concentration of data.

Figures 8 and 9 show the LZC₀₀₂ and LZC₀₀ average value for the series corresponding to the variable x(t) of the Mackey-Glass equation. For the two measures of complexity a reverse exponential development is observed as the step of integration decreases. The series of the Mackey-Glass equation present more complexity in their behavior, based on the variation of the size of the data.

Figures 10 and 11 showed the LZC₀₀₂ and LZC₀₀ average values for the stochastic series of normal distribution. Each
LZC_{002} and LZC_{01} complexity value is the average of 100 simulations of series with medium value of the distribution equal to zero and standard deviation of range [0–100]. This calculation is repeated for each size of the series. The complexity value of LZC_{002} increases exponentially when the standard deviation increases. The behavior of LZC_{01} has constant tendency regarding the change of the standard deviation and decreases its value with the increase of the elements of the series. The behavior of the two complexity values allows setting an increase of the complexity associated with the increase of the standard deviation and the decrease of the size of the series.
Table 3: Summary of minimum and maximum values of $\text{LZC}_{012}$ and $\text{LZC}_{01}$ for series of 4001 elements coming from different time series.

<table>
<thead>
<tr>
<th>Type</th>
<th>Series</th>
<th>$\text{LZC}_{012}$ Minimum</th>
<th>$\text{LZC}_{012}$ Maximum</th>
<th>$\text{LZC}_{01}$ Minimum</th>
<th>$\text{LZC}_{01}$ Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stochastic series</td>
<td>Rectangular</td>
<td>0.6829</td>
<td>0.9801</td>
<td>0.9587</td>
<td>0.9662</td>
</tr>
<tr>
<td></td>
<td>Normal</td>
<td>0.6043</td>
<td>0.9615</td>
<td>0.9589</td>
<td>0.9658</td>
</tr>
<tr>
<td></td>
<td>Exponential</td>
<td>0.4301</td>
<td>0.9228</td>
<td>0.9585</td>
<td>0.9654</td>
</tr>
<tr>
<td>Deterministic series</td>
<td>Lorenz</td>
<td>0.1472</td>
<td>0.3604</td>
<td>0.1974</td>
<td>0.4516</td>
</tr>
<tr>
<td></td>
<td>Rössler</td>
<td>0.0717</td>
<td>0.1837</td>
<td>0.0658</td>
<td>0.1196</td>
</tr>
<tr>
<td></td>
<td>Mackey-Glass</td>
<td>0.0249</td>
<td>0.1132</td>
<td>0.0271</td>
<td>0.3260</td>
</tr>
<tr>
<td></td>
<td>Periodic</td>
<td>0.0642</td>
<td>0.0925</td>
<td>0.0479</td>
<td>0.0658</td>
</tr>
</tbody>
</table>

Table 4: $\text{LZC}_{012}$ and $\text{LZC}_{01}$ values for the test series.

<table>
<thead>
<tr>
<th>Series</th>
<th>$\text{LZC}_{012}$</th>
<th>$\text{LZC}_{01}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average number of pieces in waiting line, case 1</td>
<td>0.4566</td>
<td>0.5952</td>
</tr>
<tr>
<td>Average number of pieces in waiting line, case 2</td>
<td>0.5234</td>
<td>0.6437</td>
</tr>
<tr>
<td>Average number of pieces in waiting line, case 3</td>
<td>0.5182</td>
<td>0.6322</td>
</tr>
<tr>
<td>Total volume of Great Salt Lake</td>
<td>0.0793</td>
<td>0.4067</td>
</tr>
</tbody>
</table>

The results of the $\text{LZC}_{012}$ complexity value for series consisting of 4001 elements coming from the normal, exponential, and rectangular distribution are shown in Figure 16, with the same variance of data and value 0 of the media of the normal distribution and the lower limit of the rectangular distribution. In Figure 17 the results obtained for the indicator $\text{LZC}_{01}$ for the same series are observed. With the increase of the variance of the series the average value of $\text{LZC}_{012}$ increases in all the distributions; for any variance the lowest value corresponds to the exponential distribution and the highest value to the rectangular distribution.

The simulations described in Table 3 allow setting the maximum values and the minimum of the measures of complexity for each series of stochastic origin or determinist. To establish a model of comparison, the measures of complexity of the time series with length of 4001 elements and equal dispersion of data for the stochastic function were calculated.

Table 4 presents the $\text{LZC}_{012}$ and $\text{LZC}_{01}$ values for the two series in study with a length of 4001 elements in the two cases of study: average number of parts in the waiting line and total volume of the Great Salt Lake. For the case of the average number of parts in the waiting line of a flexible manufacturing system the indicator $\text{LZC}_{012}$ is similar to the maximum of the series of Lorenz and the minimum of the stochastic series. The indicator $\text{LZC}_{01}$ is found between the minimum and maximum of the values for series of exponential distributions. The indicator $\text{LZC}_{01}$ increases a 30% regarding the indicator $\text{LZC}_{012}$ and the complexity of the series is associated with the variation of the size of the series. For the case of the volume of the Great Salt Lake the $\text{LZC}_{012}$ metric is found between the minimum and maximum values of the Cosine series, Mackey-Glass, and Rössler. The value of the $\text{LZC}_{01}$ is located between the minimum and maximum values for the series $x(t)$ of the Lorenz system.
The LZC$_{01}$ indicator increases 413% regarding the LZC$_{012}$ metric and the complexity of the system is associated with the variation of the size of the elements constitutive of the series.

The results of the simulations of the three experiments of the system of flexible manufacturing are presented in Table 4. The complexity metrics LZC$_{01}$, LZC$_{012}$ are sensitive to the change of the parameters of the system. The metrics have correlation respect to the variability of the demand or times of manufacturing.

### 5. Conclusions

This paper is based on two measures of LZC and presents a methodology to measure the dynamic complexity of a system through the analysis of a time series associated with the system. The two measures of LZC allow (1) comparing a system with other known behavior systems (stochastic functions, chaotic maps, or deterministic systems) to establish its complexity level and (2) comparing two or more systems through the classification per level of uncertainty or dynamic complexity.

The manufacturing system presented and the study of the complexity metric of the productive variable WIP allow establishing a relation between the level of complexity of the variable WIP with the setup times and the variability of the demand of the products. The importance of the capacity to reduce the uncertainty of the variable WIP in a productive system is the reduction of costs of production. A lower level of complexity of the variable WIP reduces the uncertainty of the productive process, reduces the unit cost of production, and increases the capacity of satisfying the demand. Making decisions in a working productive plant, the design of a new plant, or the comparison between plants can be improved using the complexity metric of the proposed variable WIP.

A perspective of work is to use the metrics developed in this research to address the question about the predictability of time series. For series with metric values similar to stochastic function their level of predictability is lower regarding series with metrics equivalent to chaotic or deterministic functions. Future works should establish the relation between the predictability level and the value of the LZC metric of a time series.

Finally, it can be stated that the measures of dynamic complexity proposed require low computing efforts in comparison to other metrics used for the same objectives.

### Notations

- RQA: Recurrence Quantification Analysis
- ApEn: Approximate entropy
- SampEn: Simple entropy
- WIP: Work in process
- FIFO: First-in first-out
- LZC$_{01}$: Complexity of Lempel-Ziv base 2
- LZC$_{012}$: Complexity of Lempel-Ziv base 3
- $S_{01}$: Binary series
- $S_{012}$: Three-symbol series.
Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research has been supported by DICYT (Scientific and Technological Research Bureau) of the University of Santiago of Chile (USACH) and Department of Industrial Engineering.

References


