Pesquisa Operacional (2019) 39(1): 205-223
© 2019 Brazilian Operations Research Society
Printed version ISSN 0101-7438 / Online version ISSN 1678-5142
www.scielo.br/pope
doi: 10.1590/0101-7438.2019.039.01.0205

HIGH-ORDER MULTIVARIATE MARKOV CHAIN APPLIED IN DOW JONES AND IBOVESPA INDEXES

Rafaela Boeira Cechin and Leandro Luís Corso*

Received July 13, 2018 / Accepted March 18, 2019

ABSTRACT. In this paper we analyzed the probabilities of transitions of state between Ibovespa and Dow Jones indexes using High-order Multivariate Markov Chain. While the stock market may be profitable, the existence of risks can lead to large losses. A mathematical model capable of considering different sources can aid in decision making. This model can work with stochastic data, causing different databases to be transformed into transitional matrices between states. For this, a set of a daily variation data were used between January 2008 and March 2018. Through this application, it was possible to show an interaction between the indexes and that the highest frequency of events was of the variation of –0.49 to 0.5% in Dow Jones to –0.49 to 0.5% in Ibovespa, with 428 cases, and the probability of this situation occurring again, of Dow Jones at time t to Ibovespa at time $t + 2$, is 27.21%. Empirical results suggest that this application can help investors make decisions based on transition probabilities.

Keywords: Markov Chain, High-order Multivariate Markov Chain, stock market index.

1 INTRODUCTION

To invest in the stock market it is important to calculate the risks. Related to this idea, it is important to understand that smaller markets are affected by larger markets (Achcar et al. (2012)). The authors comment that the Brazilian stock exchange is affected by the American stock exchange, since the values of transitions of the second are considerably higher.

Shephard (2008) presents a collection of relevant models and studies. In this study is presented the ARCH models (autoregressive conditional heteroscedasticity) applied to Univariate stochastic volatility (SV) and multivariate ARCH models (MARCH). This type of problem is often studied using Econometric Theory (Yu & Meier (2006)). Although these models exist, the relationship between the two stock exchanges was not presented simultaneously. In order to analyze this relation, we are proposing an application of the High-order Multivariate Markov Chain. The

*Corresponding author – https://orcid.org/0000-0001-9962-9578
Universidade de Caxias do Sul, Rua Francisco Getúlio Vargas, 11300, Petrópolis, 95070-560 Caxias do Sul, RS, Brasil.
E-mails: rbcechin@ucs.br; llcorso@ucs.br

concept of Markov Chain, which belongs to the area of Operational Research, is a statistical analysis of a process, in which future state depends solely on its current state.

The authors Doubleday & Esunge (2011) used Markov Chain to determine the relationship between a diverse portfolio of stocks and the market as a whole, where, the authors comment on the feasibility of applying the tool to support decision making. In the same line of application, Fitriyanto & Lestari (2018) presented an article with the application of Markov Chain in PT HM Sampoerna stock price. These mentioned articles used Markov Chains of 1st order.

Although the applications with 1st order is more widespread in the literature, some articles of higher order began to gain space in different areas. For example, the articles of Ching et al. (2004), Ching et al. (2008) and Yang et al. (2011). However, the investment area is also attractive for these models, as the publication of Ky & Tuyen (2018). The authors emphasize the importance of this technique for forecasting time series. They also comment that while there are ARIMA methods, exponential smoothing and even artificial intelligence, these methods demand a hardly fit nonlinear data. The authors present a novel High-order Markov model for time series forecasting where the state space of the Markov chain was constructed from different levels of changes of the time series.

The objective of this paper is to calculate the probability of variation of the main index of the Brazilian and American stock exchanges, being respectively Ibovespa and Dow Jones, using the High-order Multivariate Markov Chain. It also was studied the behavior of these two indexes separately, to obtain the expected recurrence time, in order to be useful information for investors.

2 THEORETICAL REFERENCE

2.1 Financial market volatility

Interest in forecasting stock market volatility has been around for some time (Bollerslev et al. (1992)). According to Mello (2009), volatility, in the financial area, shows the intensity and the frequency of fluctuations in the prices of a financial asset, which may be stock, bond, investment fund or stock market indices, in a given period of time. The more the price of a stock fluctuates over a short period of time, the greater the risk of gaining or losing money by trading this stock, and therefore volatility is one of the parameters most frequently used as a way of measuring the risk of an asset, according to Ishizawa (2008).

The more volatile an asset is, the more significant its variation in relation to market fluctuations. In other words, it is a riskier investment. Savings are at the extreme of low volatility and profitability, while the derivatives market as options and futures are at the other extreme, with much volatility and possibility of return. Stochastic data analyzes can help reduce investment risk (Yu & Meyer (2006)), although it can involve a computer complexity (Souza et al. (2011)).

Ribeiro (2009) comments that the measurement of a risk is according to the history of the asset. According to Mello (2009) and Almeida (2013), one of the ways to analyze the volatility of an asset is by measuring its oscillation or the standard deviation of its value or profitability. It can

be used in many ways, for example in the stock market, it is possible to analyze the levels of overbought/oversold, and the definition of targets for gains and losses.

As the financial world is influenced by several external factors, such as politics itself, it is also subject to lesser or greater volatility. Chiu et al. (2018) comment that long-term volatility is related to macroeconomic fundamentals associated with future cash flows and rates, while short-term volatility is related to transitory determinants, such as investor sentiment. This is related to the hypothesis that volatility reflects both market expectations of future cash flows and rates, and the short-term behavioral effects that are not directly linked to economic activity.

For Fornari & Mele (2013), one characteristic of the capital markets is the countercyclical behavior of asset price volatilities. For example, in the last 50 years, the annualized return volatility of the S&P 500 was 14.18% on average. However, during recessions, this figure increased to 17.39%, 23% more than the overall average. During the expansions, however, this same volatility reached an average of 13.5%, being 4% below the general average.

The work of Schwert (1989) studies the changes in stock market volatility over time. In particular, it relates stock market volatility to time-varying volatility of various economic variables. For the period 1857-1987, volatility was remarkably high from 1929 to 1939 for many economic series, including inflation, monetary growth, industrial production, and other measures of economic activity. This is because stock market volatility increases during recessions.

Schwert (1989) also comments that the value of corporate equity depends on the health of the economy. If discount rates are constant over time, the price movement of securities is proportional to the variation of expected future cash flows. It is possible that a change in the level of uncertainty of future macroeconomic conditions would cause a proportional change in the volatility of stock returns.

Christiansen et al. (2012) and Zhu & Singh (2016) comment that financial volatility is a fundamental study for risk management, asset pricing and portfolio management. It can have important repercussions on the economy as a whole, evidenced by the financial crises. Therefore, according to the authors, it is extremely important to learn more about the economic drivers of financial market volatility.

According to Ribeiro (2009), Fávero & Confortini (2010) and Chiu et al. (2018), a good understanding of how the asset, its profitability and its volatility interact with the economy, will enable policymakers and finance and investment professionals to obtain more accurate forecasts, influenced by macroeconomic factors.

2.2 Markov Chain

According to Hillier & Lieberman (2005), Ho & Quinino (2012) and Roshan & Nastos (2018), a stochastic process describes the behavior of variables $\{X_{t_n}\} = \{x_1, x_2, \ldots, x_n\}$ within a certain period of time T and that such a process is a Markov chain if the probability of occurrence of a future state depends only on the present state, that is, if it is independent of past events.

Taha (2007), Andersen et al. (2017) and Hunter (2018) present the conditional probability P in Equations 1 and 2.

$$p_{ij} = P\{X_{t+1} = j \,|\, X_t = i\} \tag{1}$$

$$p_{ij}^{n} = P\{X_{t+n} = j \,|\, X_t = i\} \tag{2}$$

Which, X_t represents a variable at time t, and p_{ij}^{n} is the probability that a process passes from state i to state j in n steps at time t. As it is conditional probability, Hillier & Lierberman (2005) affirm that these values cannot be negative. Staudt et al. (2011), Ho & Quinino (2012), Andersen et al. (2017) and Roshan & Nastos (2018) comment that one way of presenting these transitions is with the use of the Transition Matrix, shown in Equation 3.

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1j} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2j} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{i1} & p_{i2} & p_{i3} & \cdots & p_{ij} \end{pmatrix} \tag{3}$$

The authors demonstrate that the transition of states occurs from the row-to-column index, that is, the probability p_{ij} corresponds to the transition of state i to state j. Taha (2007) comments that the sum of each row of the matrix must be equal to 1, besides that an array is classified as ergodic when it is possible to go from any state to another any in n time steps.

Taha (2007), Hunter (2018) and Roshan & Nastos (2018) then present the stable state probability π_j for an ergodic Markov chain in Equation 4 and its property in Equation 5.

$$\pi_j = \sum_{i=0}^{M} \pi_i \, p_{ij}, \quad \text{para} \quad j = 0, 1, \ldots, M \tag{4}$$

$$\sum_{j=0}^{M} \pi_j = 1 \tag{5}$$

Taha (2007) explains that, after a large number of transitions, the probability of finding the process in a given state, for example j, tends to the value π_j, independent of the probability distribution of the initial state. Hillier & Lierberman (2005) comment that because there is $M + 2$ equations, $M + 1$ unknowns and a single solution, it is necessary to exclude one of the equations, but it cannot be Equation 5, because $\pi_j = 0$ for any j will satisfy $M + 1$ equations.

Furthermore, it is possible to analyze the expected recurrence time, represented by μ_{ii} for when $j = i$, which, according to Taha (2007), is the expected number of transitions until the process returns to the initial state i, shown in Equation 6.

$$\mu_{ii} = \frac{1}{\pi_i}, \quad \text{para} \quad i = 0, 1, \ldots, M \tag{6}$$

2.2.1 High-order Multivariate Markov Chains

As mentioned before, the probabilities associated with various state changes are called transition probabilities. The process is characterized by a transition matrix describing the probabilities of transitions from one initial state to another. The High-order Multivariate Markov Chain is used when there is more than one time series of data and the transition occurs for more than one-time step, that is, it can be considered of order n.

In order to facilitate the understanding of this concept, this section has been divided into three subsections.

2.2.1.1 High-order Markov Chains

According to Ching et al. (2008), a High-order Markov Chain (nth-order) is used to model time series x_t of m categories, which can be represented by a sequence of vectors $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots\}$ in canonical form. If this system is in state $j \in M$ at time $(t+i)$, then the probability distribution vector of the state is given by Equation 7.

$$\mathbf{x}_{t+i} = \left(0, \ldots, 0, \underbrace{1}_{j\text{th entry}}, 0, \ldots, 0\right)^T \tag{7}$$

Kárný (2016) assumes that the state probability distribution at time $t = r+1$ depends on the sequence state probability distribution at times $t = r, r-1, \ldots, r-n+1$. Then, Yang et al. (2011) presents Equation 8.

$$\mathbf{x}_{r+1} = \sum_{h=1}^{n} \lambda_h P_h \mathbf{x}_{r-h+1}, \quad \text{for} \quad r = n-1, n, \ldots \tag{8}$$

Which λ_h represents the weight and is a non-negative real number, which summation is equal to 1; $\mathbf{x}_r$ is the probability distribution of the state at time r; P_h is the h step transition matrix.

2.2.1.2 First-order Multivariate Markov Chains

Ching et al. (2008) explain that a First-order Multivariate Markov Chain models the behavior of multiple data sequences for a one step in time. Ching et al. (2004) understand that there may be s sequences and each with m possible states. Then, it is assumed that the probability of the state of the jth sequence at time $t = r+1$ depends on the state probabilities of all sequences (including itself) at time $t = r$. Yang et al. (2011) presents Equation 9 to represent the First-order Multivariate Markov Chain model.

$$\mathbf{x}_{r+1}^{(j)} = \sum_{k=1}^{s} \lambda_{jk} P^{(jk)} \mathbf{x}_r^{(k)}, \quad \text{for} \quad r = 0, 1, \ldots \tag{9}$$

Which $\mathbf{x}_0^{(j)}$ is the initial probability distribution of the sequence jth; λ_{jk} is a non-negative real number, its sum being equal to 1; and $1 \leq j, k \leq s$.

The probability of state of the sequence jth, $\mathbf{x}_{r+1}^{(j)}$ at time $r+1$, depends on the weight average of $P^{(jk)}\mathbf{x}_r^{(k)}$. Then, $P^{(jk)}$ is the transition matrix for a one step in time, from the state at time t in the sequence kth to the state at time $t+1$ in the sequence jth, and $\mathbf{x}_r^{(k)}$ is a distribution of the probability of states at time r in the sequence kth. Ching et al. (2004) presents in Equation 10 the matrix form.

$$\mathbf{X}_{r+1} = \begin{pmatrix} \mathbf{x}_{r+1}^{(1)} \\ \mathbf{x}_{r+1}^{(2)} \\ \vdots \\ \mathbf{x}_{r+1}^{(s)} \end{pmatrix} = \begin{pmatrix} \lambda_{11}P^{(11)} & \lambda_{12}P^{(12)} & \dots & \lambda_{1s}P^{(1s)} \\ \lambda_{21}P^{(21)} & \lambda_{22}P^{(22)} & \dots & \lambda_{2s}P^{(2s)} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{s1}P^{(s1)} & \lambda_{s2}P^{(s2)} & \dots & \lambda_{ss}P^{(ss)} \end{pmatrix} \begin{pmatrix} \mathbf{x}_r^{(1)} \\ \mathbf{x}_r^{(2)} \\ \vdots \\ \mathbf{x}_r^{(s)} \end{pmatrix} \tag{10}$$

2.2.1.3 High-order Multivariate Markov Chains

The High-order Multivariate Markov Chain is an addition of the concepts presented earlier, that is, models the behavior of multiple data sequences for more than one step in time, in the nth order. Kárný (2016) understands that the state probability distribution of the sequence jth at time $r+1$ depends on the state probability distribution of all sequences, including itself, at times $t = r, r-1, \dots, r-n+1$. Then, the High-order Multivariate Markov Chain (nth order) model is presented in Equation 11, according to Ching et al. (2008).

$$\mathbf{x}_{r+1}^{(j)} = \sum_{k=1}^{s}\sum_{h=1}^{n} \lambda_{jk}^{(h)} P_h^{(jk)} \mathbf{x}_{r-h+1}^{(k)}, \quad \text{for} \quad r = n-1, n, \dots \tag{11}$$

which λ_{jk} is a non-negative real number, its sum being equal to 1; and $1 \leq j, k \leq s$.

The probability of state of the sequence jth, $\mathbf{x}_{r+1}^{(j)}$ at time $r+1$, depends on the average of the weights of $P_h^{(jk)}\mathbf{x}_{r-h+1}^{(k)}$. Then, $P_h^{(jk)}$ is the transition matrix of the time step hth, which describes the transition from h steps, from the state at time $r-h+1$ in the sequence kth to the state at time $r+1$ in the sequence jth, and $\lambda_{jk}^{(h)}$ is the weighting of this period. Ching et al. (2004) presents in Equations 12, 13, 14 and 15 the matrix form.

$$\mathbf{X}_{r+1} = \begin{pmatrix} \mathbf{X}_{r+1}^{(1)} \\ \mathbf{X}_{r+1}^{(2)} \\ \vdots \\ \mathbf{X}_{r+1}^{(s)} \end{pmatrix} = \begin{pmatrix} B^{(11)} & B^{(12)} & \dots & B^{(1s)} \\ B^{(21)} & B^{(22)} & \dots & B^{(2s)} \\ \vdots & \vdots & \vdots & \vdots \\ B^{(s1)} & B^{(s2)} & \dots & B^{(ss)} \end{pmatrix} \begin{pmatrix} \mathbf{X}_r^{(1)} \\ \mathbf{X}_r^{(2)} \\ \vdots \\ \mathbf{X}_r^{(s)} \end{pmatrix} = Q\mathbf{X}_r \tag{12}$$

which:

$$B^{(ii)} = \begin{pmatrix} \lambda_{ii}^{(1)} P_1^{(ii)} & \lambda_{ii}^{(2)} P_2^{(ii)} & \dots & \lambda_{ii}^{(n-1)} P_{n-1}^{(ii)} & \lambda_{ii}^{(n)} P_n^{(ii)} \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & I & 0 \end{pmatrix} \tag{13}$$

If $i \neq j$, so:

$$B^{(ij)} = \begin{pmatrix} \lambda_{ij}^{(1)} P_1^{(ij)} & \lambda_{ij}^{(2)} P_2^{(ij)} & \dots & \lambda_{ij}^{(n-1)} P_{n-1}^{(ij)} & \lambda_{ij}^{(n)} P_n^{(ij)} \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix} \tag{14}$$

$$Q = \begin{pmatrix} B^{(11)} & B^{(12)} & \dots & B^{(1s)} \\ B^{(21)} & B^{(22)} & \dots & B^{(2s)} \\ \vdots & \vdots & \vdots & \vdots \\ B^{(s1)} & B^{(s2)} & \dots & B^{(ss)} \end{pmatrix} \tag{15}$$

Where Q is the one-step transition probability matrix of the multivariate model, which determines the probability of making a transition depending on the current state.

3 METHODOLOGY AND RESULTS

In order to achieve the proposed goal of this study, the elaborated methodology was divided into four stages, as shown in Figure 1.

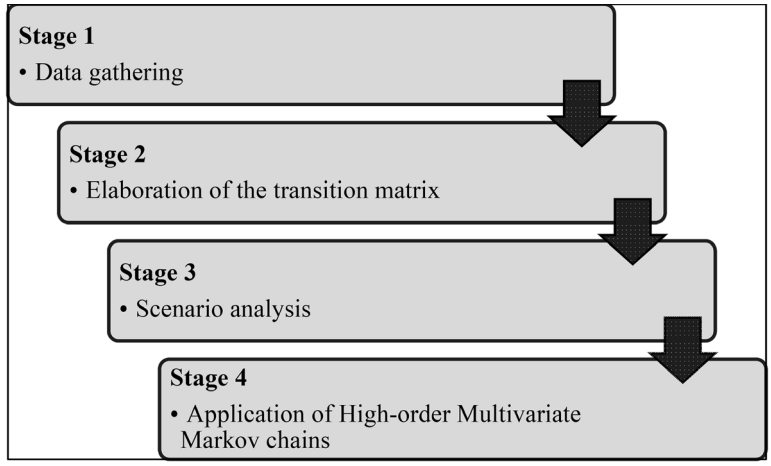

Figure 1 – Methodology of research.

3.1 Data gathering

For this study, the values of the percentage variation of the Ibovespa index and the Dow Jones index were collected. The Ibovespa index is a theoretical portfolio of companies, which serves to indicate the average performance of asset prices of greater tradability and representativeness of

the Brazilian stock market, while the Dow Jones is one of the main indicators of the movements of the American market. This data was taken from the InfoMoney electronic address, the time period was commercial day-to-day trading of derivatives on the São Paulo and New York stock exchanges, and index values refer to January 2, 2008 to March 21, 2018. Figure 2 shows the Ibovespa data and Figure 3 the Dow Jones index.

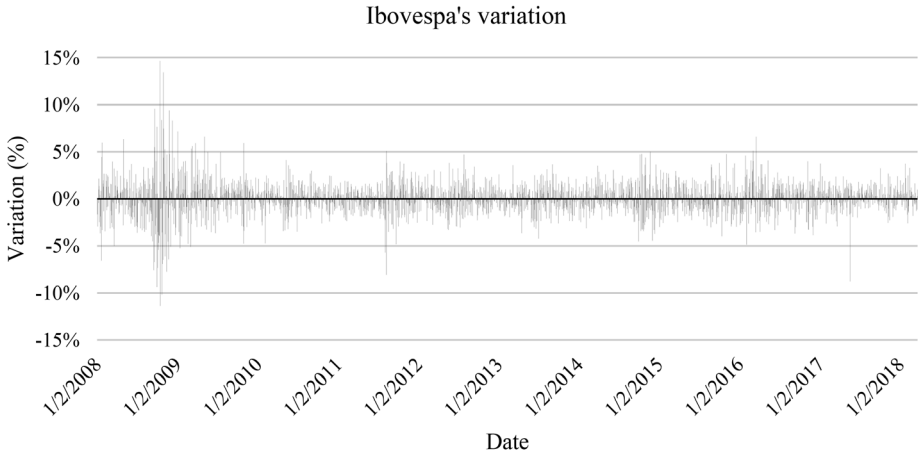

Figure 2 – Percentage variation of the Ibovespa index.

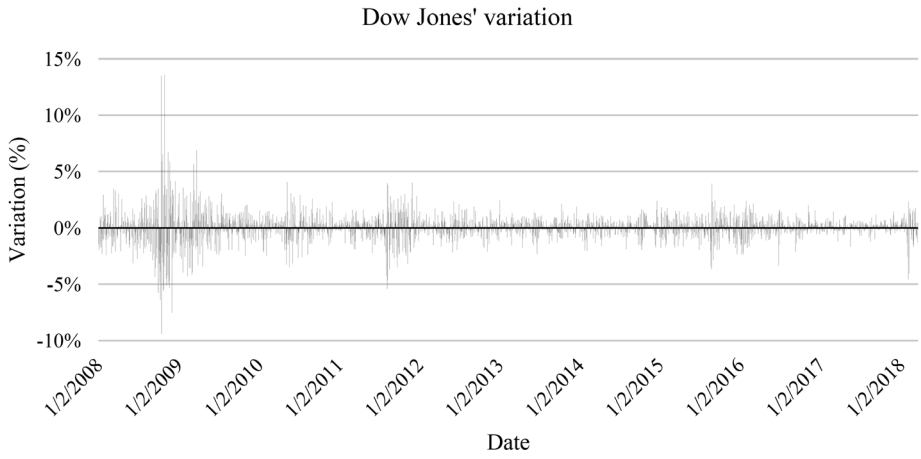

Figure 3 – Percentage variation of the Dow Jones index.

3.2 Elaboration of the transition matrix

In order to calculate the transition matrices, it was necessary to define intervals for the percentage changes of the two indexes, as shown in Table 1.

Table 1 – Range of variation.

	$< -2.5\%$		
from	-2.49%	to	-1.5%
from	-1.49%	to	-0.5%
from	-0.49%	to	0.5%
from	0.51%	to	1.5%
from	1.51%	to	2.5%
	$2.51\% <$		

The development of the transition matrix, it was separated into two subsections, one for the Ibovespa index and another for Dow Jones.

3.2.1 Transition matrix for Ibovespa

In this step the transition matrix of the Ibovespa index was created. For this, the first stage was the creation of the frequency matrix, which represents the total number of times the transition occurred, shown in Figure 4.

	from / to	< –2.5%	–2.49% to –1.5%	–1.49% to –0.5%	–0.49% to 0.5%	0.51% to 1.5%	1.51% to 2.5%	2.51% <
	< –2.5%	18	15	10	36	27	13	24
	–2.49% to –1.5%	19	16	40	69	56	17	9
	–1.49% to –0.5%	29	47	120	131	83	46	28
F =	–0.49% to 0.5%	33	63	154	226	147	88	25
	0.51% to 1.5%	19	49	91	157	95	48	26
	1.51% to 2.5%	13	21	48	79	57	20	13
	2.51% <	12	14	21	38	21	19	13

Figure 4 – Transition frequency matrix for Ibovespa.

Based on this matrix the cumulative percentage that occurred in the variations of each of the intervals was observed, as shown in Figure 5.

By means of this chart it is noticed that the highest frequency of variation occurred in the range of –0.49 to 0.5%. Following the calculation procedure and considering this information, it was possible to calculate the matrix with the transition probabilities. The value P_{ij} is the probability of price variation from state i to state j in one step of time, which is presented in Figure 6.

3.2.2 Transition matrix for Dow Jones

Based on data from the Dow Jones daily percentage change, it was possible to analyze the frequency of variations, shown in Figure 7.

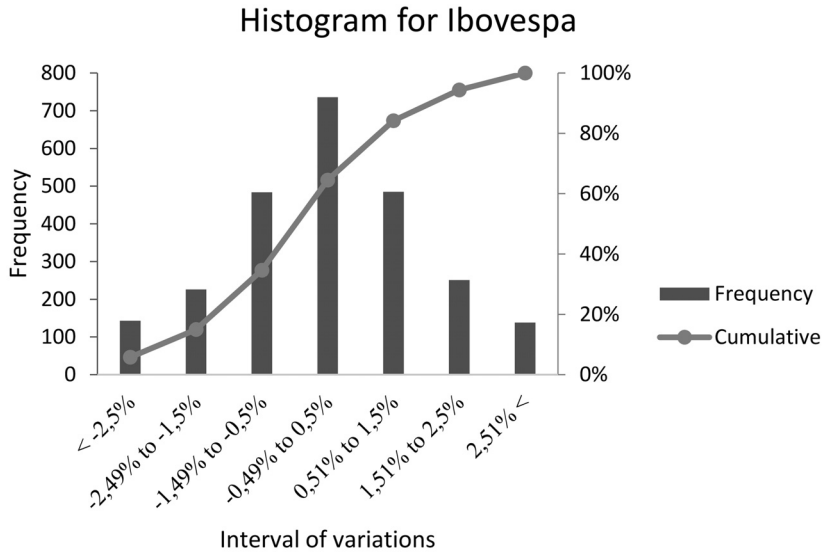

Figure 5 – Histogram of variations interval for Ibovespa.

	from / to	< –2.5%	–2.49% to –1.5%	–1.49% to –0.5%	–0.49% to 0.5%	0.51% to 1.5%	1.51% to 2.5%	2.51% <
	< –2.5%	0.126	0.105	0.070	0.252	0.189	0.091	0.168
	–2.49% to –1.5%	0.084	0.071	0.177	0.305	0.248	0.075	0.040
	–1.49% to –0.5%	0.060	0.097	0.248	0.271	0.171	0.095	0.058
P =	–0.49% to 0.5%	0.045	0.086	0.209	0.307	0.200	0.120	0.034
	0.51% to 1.5%	0.039	0.101	0.188	0.324	0.196	0.099	0.054
	1.51% to 2.5%	0.052	0.084	0.191	0.315	0.227	0.080	0.052
	2.51% <	0.087	0.101	0.152	0.275	0.152	0.138	0.094

Figure 6 – Transition matrix for Ibovespa.

	from / to	< –2.5%	–2.49% to –1.5%	–1.49% to –0.5%	–0.49% to 0.5%	0.51% to 1.5%	1.51% to 2.5%	2.51% <
	< –2.5%	9	6	7	12	9	3	13
	–2.49% to –1.5%	5	4	19	39	35	6	5
	–1.49% to –0.5%	7	26	73	152	93	17	9
F =	–0.49% to 0.5%	17	44	177	757	234	36	17
	0.51% to 1.5%	8	21	78	266	89	19	7
	1.51% to 2.5%	5	8	14	39	16	5	2
	2.51% <	8	4	8	18	12	3	2

Figure 7 – Transition frequency matrix for Dow Jones.

As has been created previously, the cumulative percentage that occurred in the variations of each of the intervals was developed, as shown in Figure 8.

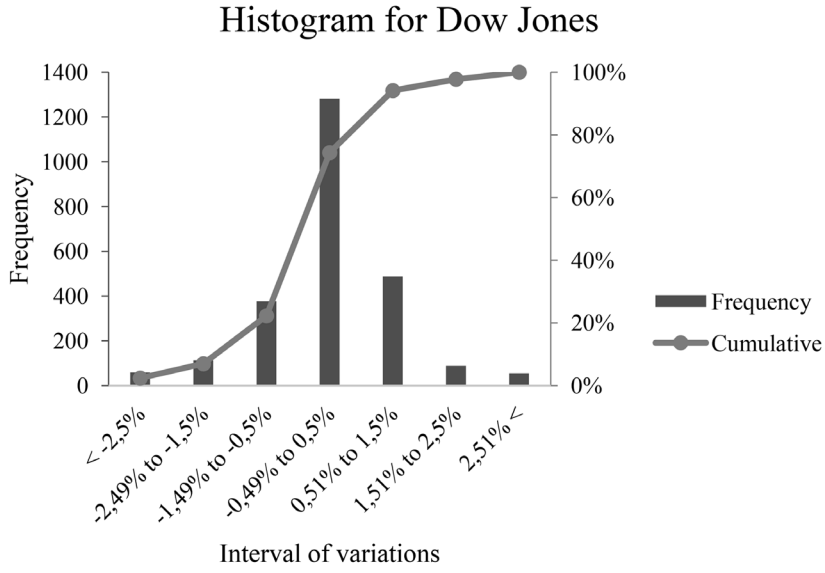

Figure 8 – Histogram of variations interval for Dow Jones.

In this specific case by analyzing the graph, it is noted that the higher frequency of variation also occurred in the range of –0.49 to 0.5%. The Markov transition probabilities is shown in Figure 9.

	from / to	< –2.5%	–2.49% to –1.5%	–1.49% to –0.5%	–0.49% to 0.5%	0.51% to 1.5%	1.51% to 2.5%	2.51% <
	< –2.5%	0.153	0.102	0.119	0.203	0.153	0.051	0.220
	–2.49% to –1.5%	0.044	0.035	0.168	0.345	0.310	0.053	0.044
	–1.49% to –0.5%	0.019	0.069	0.194	0.403	0.247	0.045	0.024
P =	–0.49% to 0.5%	0.013	0.034	0.138	0.590	0.183	0.028	0.013
	0.51% to 1.5%	0.016	0.043	0.160	0.545	0.182	0.039	0.014
	1.51% to 2.5%	0.056	0.090	0.157	0.438	0.180	0.056	0.022
	2.51% <	0.145	0.073	0.145	0.327	0.218	0.055	0.036

Figure 9 – Transition matrix of transition for Dow Jones.

3.3 Scenario analysis

The analysis of the scenarios of the two indices was also separated into two subsections.

3.3.1 Scenario analysis for Ibovespa

Scenario analysis requires the calculation of steady state, as presented in Equations 16 and 17.

$$(\pi_0\pi_1\pi_2\pi_3\pi_4\pi_5\pi_6) = (\pi_0\pi_1\pi_2\pi_3\pi_4\pi_5\pi_6)\begin{pmatrix} 0.126 & 0.105 & 0.070 & 0.252 & 0.189 & 0.091 & 0.168 \\ 0.084 & 0.071 & 0.177 & 0.305 & 0.248 & 0.075 & 0.040 \\ 0.060 & 0.097 & 0.248 & 0.271 & 0.171 & 0.095 & 0.058 \\ 0.045 & 0.086 & 0.209 & 0.307 & 0.200 & 0.120 & 0.034 \\ 0.039 & 0.101 & 0.188 & 0.324 & 0.196 & 0.099 & 0.054 \\ 0.052 & 0.084 & 0.191 & 0.315 & 0.227 & 0.080 & 0.052 \\ 0.087 & 0.101 & 0.152 & 0.275 & 0.152 & 0.138 & 0.094 \end{pmatrix} \tag{16}$$

$$(\pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 + \pi_6) = 1 \tag{17}$$

Resulting in the following system of equations, shown by Equations 18 to 25.

$$\pi_0 = 0.126\pi_0 + 0.084\pi_1 + 0.060\pi_2 + 0.045\pi_3 + 0.039\pi_4 + 0.052\pi_5 + 0.087\pi_6 \tag{18}$$
$$\pi_1 = 0.105\pi_0 + 0.071\pi_1 + 0.097\pi_2 + 0.086\pi_3 + 0.086\pi_4 + 0.084\pi_5 + 0.101\pi_6 \tag{19}$$
$$\pi_2 = 0.070\pi_0 + 0.177\pi_1 + 0.248\pi_2 + 0.209\pi_3 + 0.209\pi_4 + 0.191\pi_5 + 0.152\pi_6 \tag{20}$$
$$\pi_3 = 0.252\pi_0 + 0.305\pi_1 + 0.271\pi_2 + 0.307\pi_3 + 0.307\pi_4 + 0.315\pi_5 + 0.275\pi_6 \tag{21}$$
$$\pi_4 = 0.189\pi_0 + 0.248\pi_1 + 0.171\pi_2 + 0.200\pi_3 + 0.200\pi_4 + 0.227\pi_5 + 0.152\pi_6 \tag{22}$$
$$\pi_5 = 0.091\pi_0 + 0.075\pi_1 + 0.095\pi_2 + 0.120\pi_3 + 0.120\pi_4 + 0.080\pi_5 + 0.138\pi_6 \tag{23}$$
$$\pi_6 = 0.168\pi_0 + 0.040\pi_1 + 0.058\pi_2 + 0.034\pi_3 + 0.034\pi_4 + 0.052\pi_5 + 0.094\pi_6 \tag{24}$$
$$1 = \pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 + \pi_6 \tag{25}$$

Because there are eight equations and seven unknowns, it is necessary to disregard one of the equations. Solving this system of equations, we obtain the following results:

$$\begin{array}{ll} \pi_0 = 0.058 & \pi_1 = 0.091 \\ \pi_2 = 0.197 & \pi_3 = 0.299 \\ \pi_4 = 0.197 & \pi_5 = 0.102 \\ \pi_6 = 0.056 & \end{array}$$

These values represent the possibility of being in certain states, according to Table 2.

Table 2 shows that the Ibovespa is most likely to range from –0.49 to 0.50%, with a probability of 29.9%. Moreover, it is possible to analyze the expected recurrence time for each stable state probability, according to Table 3.

In Table 3 it is observed that, when the range of variation is –0.49 to 0.5%, for example, the expected recurrence time for this variation is 3.4 days, this means that in 3.4 days a value ranging from –0.49 to 0.5% occurs again.

Table 2 – Probability of steady-states for Ibovespa.

Range of variation	Probability
< –2.5%	5.8%
from –2.49% to –1.5%	9.1%
from –1.49% to –0.5%	19.7%
from –0.49% to 0.5%	29.9%
from 0.51% to 1.5%	19.7%
from 1.51% to 2.5%	10.2%
2.51% <	5.6%

Table 3 – Expected recurrence time for Ibovespa.

Probability of steady-state π_i	Expected recurrence time μ_{ii}
π_0	17.2 days
π_1	11 days
π_2	5.1 days
π_3	3.4 days
π_4	5.1 days
π_5	9.8 days
π_6	17.9 days

3.3.2 Scenario analysis for Dow Jones

Scenario analysis requires the calculation of steady state, as shown in Equations 26 and 27.

$$(\pi_0\pi_1\pi_2\pi_3\pi_4\pi_5\pi_6) = (\pi_0\pi_1\pi_2\pi_3\pi_4\pi_5\pi_6) \begin{pmatrix} 0.153 & 0.102 & 0.119 & 0.203 & 0.153 & 0.051 & 0.220 \\ 0.044 & 0.035 & 0.168 & 0.345 & 0.310 & 0.053 & 0.044 \\ 0.019 & 0.069 & 0.194 & 0.403 & 0.247 & 0.045 & 0.024 \\ 0.013 & 0.034 & 0.138 & 0.590 & 0.183 & 0.028 & 0.013 \\ 0.016 & 0.043 & 0.160 & 0.545 & 0.182 & 0.039 & 0.014 \\ 0.056 & 0.090 & 0.157 & 0.438 & 0.180 & 0.056 & 0.022 \\ 0.145 & 0.073 & 0.145 & 0.327 & 0.218 & 0.055 & 0.036 \end{pmatrix} \tag{26}$$

$$(\pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 + \pi_6) = 1 \tag{27}$$

Resulting in the following system of equations, shown by Equations 28 to 35.

$$\pi_0 = 0.153\pi_0 + 0.044\pi_1 + 0.019\pi_2 + 0.013\pi_3 + 0.016\pi_4 + 0.056\pi_5 + 0.145\pi_6 \quad (28)$$

$$\pi_1 = 0.102\pi_0 + 0.035\pi_1 + 0.069\pi_2 + 0.034\pi_3 + 0.043\pi_4 + 0.090\pi_5 + 0.073\pi_6 \quad (29)$$

$$\pi_2 = 0.119\pi_0 + 0.168\pi_1 + 0.194\pi_2 + 0.138\pi_3 + 0.160\pi_4 + 0.157\pi_5 + 0.145\pi_6 \quad (30)$$

$$\pi_3 = 0.203\pi_0 + 0.345\pi_1 + 0.403\pi_2 + 0.590\pi_3 + 0.545\pi_4 + 0.438\pi_5 + 0.327\pi_6 \quad (31)$$

$$\pi_4 = 0.153\pi_0 + 0.310\pi_1 + 0.247\pi_2 + 0.183\pi_3 + 0.182\pi_4 + 0.180\pi_5 + 0.218\pi_6 \quad (32)$$

$$\pi_5 = 0.051\pi_0 + 0.053\pi_1 + 0.045\pi_2 + 0.028\pi_3 + 0.039\pi_4 + 0.056\pi_5 + 0.055\pi_6 \quad (33)$$

$$\pi_6 = 0.220\pi_0 + 0.044\pi_1 + 0.024\pi_2 + 0.013\pi_3 + 0.014\pi_4 + 0.022\pi_5 + 0.036\pi_6 \quad (34)$$

$$1 = \pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 + \pi_6 \quad (35)$$

Because there are eight equations and seven unknowns, it is necessary to disregard one of the equations. Solving this system of equations, we obtain the following results:

$$\pi_0 = 0.024 \qquad \pi_1 = 0.046$$
$$\pi_2 = 0.153 \qquad \pi_3 = 0.522$$
$$\pi_4 = 0.198 \qquad \pi_5 = 0.036$$
$$\pi_6 = 0.022$$

These values represent the possibility of being in certain states, according to Table 4.

Table 4 – Probability of steady-states for Dow Jones.

Interval	Probability
< –2.5%	2.3%
from –2.49% to –1.5%	4.4%
from –1.49% to –0.5%	15%
from –0.49% to 0.5%	52.4%
from 0.51% to 1.5%	20.1%
from 1.51% to 2.5%	3.7%
2.51% <	2.1%

With Table 4, it is observed that the Dow Jones is most likely to range from –0.49 to 0.50%, with a probability of 52.4%. Furthermore, it is possible to analyze the expected recurrence time for each stable state probability, as shown in Table 5.

With Table 5, it is observed that, when analyzing a variation of –0.49 to 0.5%, for example, the expected recurrence time for this variation is 1.9 days, that is, it is expected that in 1.9 days a value ranging from –0.49 to 0.5% occurs again.

3.4 Application of High-order Multivariate Markov Chains

Considering 1 as the Dow Jones data set and 2 as the Ibovespa data set for both with the same intervals stipulated previously, then $F_1^{(12)}$ represents the transition frequency matrix by analyzing

Table 5 – Expected recurrence time for Dow Jones.

Probability of steady-state π_i	Expected recurrence time μ_{ii}
π_0	47.8 days
π_1	21.8 days
π_2	6.6 days
π_3	1.9 days
π_4	5.1 days
π_5	27.7 days
π_6	44.8 days

the Dow Jones for Ibovespa with one step of time, that is, Dow Jones at time t and Ibovespa at $t+1$. So, $F_2^{(12)}$ portrays for two steps of time, that is, Dow Jones at time t and Ibovespa at $t+2$. These matrices are shown in Figures 10 and 11.

	from Dow Jones / to Ibovespa	< –2.5%	–2.49% to –1.5%	–1.49% to –0.5%	–0.49% to 0.5%	0.51% to 1.5%	1.51% to 2.5%	2.51% <
	< –2.5%	14	3	9	8	10	4	11
	–2.49% to –1.5%	10	14	18	30	22	13	6
	–1.49% to –0.5%	27	37	66	108	72	39	28
$F_1^{(12)} =$	–0.49% to 0.5%	56	112	266	398	256	138	56
	0.51% to 1.5%	26	44	104	147	100	43	24
	1.51% to 2.5%	5	11	15	28	13	10	7
	2.51% <	5	4	6	17	13	4	6

Figure 10 – Transition frequency matrix for $F_1^{(12)}$.

	from Dow Jones / to Ibovespa	< –2.5%	–2.49% to –1.5%	–1.49% to –0.5%	–0.49% to 0.5%	0.51% to 1.5%	1.51% to 2.5%	2.51% <
	< –2.5%	9	5	7	13	7	6	12
	–2.49% to –1.5%	15	16	15	22	17	14	14
	–1.49% to –0.5%	25	36	68	102	78	40	28
$F_2^{(12)} =$	–0.49% to 0.5%	51	116	268	422	259	122	43
	0.51% to 1.5%	25	42	97	143	101	52	28
	1.51% to 2.5%	10	7	18	21	14	13	6
	2.51% <	8	3	11	12	10	4	7

Figure 11 – Transition frequency matrix for $F_2^{(12)}$.

Using this information, it was possible to calculate the transition matrices, where $P_1^{(12)}$ represents the probability of the transition from the Dow Jones data to the Ibovespa data with one step in time, that is, Dow Jones at time t and Ibovespa at $t+1$. So, $P_2^{(12)}$ portrays for two

steps in time, that is, Dow Jones at time t and Ibovespa at $t + 2$. These matrices are shown in Figures 12 and 13.

	from Dow Jones / to Ibovespa	< –2.5%	–2.49% to –1.5%	–1.49% to –0.5%	–0.49% to 0.5%	0.51% to 1.5%	1.51% to 2.5%	2.51% <
	< –2.5%	0.237	0.051	0.153	0.136	0.169	0.068	0.186
	–2.49% to –1.5%	0.088	0.124	0.159	0.265	0.195	0.115	0.053
	–1.49% to –0.5%	0.072	0.098	0.175	0.286	0.191	0.103	0.074
$P_1^{(12)} =$	–0.49% to 0.5%	0.044	0.087	0.207	0.310	0.200	0.108	0.044
	0.51% to 1.5%	0.053	0.090	0.213	0.301	0.205	0.088	0.049
	1.51% to 2.5%	0.056	0.124	0.169	0.315	0.146	0.112	0.079
	2.51% <	0.091	0.073	0.109	0.309	0.236	0.073	0.109

Figure 12 – Transition matrix for $P_1^{(12)}$.

	from Dow Jones / to Ibovespa	< –2.5%	–2.49% to –1.5%	–1.49% to –0.5%	–0.49% to 0.5%	0.51% to 1.5%	1.51% to 2.5%	2.51% <
	< –2.5%	0.153	0.085	0.119	0.220	0.119	0.102	0.203
	–2.49% to –1.5%	0.133	0.142	0.133	0.195	0.150	0.124	0.124
	–1.49% to –0.5%	0.066	0.095	0.180	0.271	0.207	0.106	0.074
$P_2^{(12)} =$	–0.49% to 0.5%	0.040	0.091	0.209	0.329	0.202	0.095	0.034
	0.51% to 1.5%	0.051	0.086	0.199	0.293	0.207	0.107	0.057
	1.51% to 2.5%	0.112	0.079	0.202	0.236	0.157	0.146	0.067
	2.51% <	0.145	0.055	0.200	0.218	0.182	0.073	0.127

Figure 13 – Transition matrix for $P_2^{(12)}$.

Then, the probability of steady state was calculated, as presented in Equations 4 and 5. The results for the probabilities with one and two steps in time are presented in Table 6.

Table 6 – Probability of steady-states.

Range of variation	π_j for $P_1^{(12)}$	π_j for $P_2^{(12)}$
< –2.5%	7.36%	8.03%
from –2.49% to –1.5%	9.32%	9.07%
from –1.49% to –0.5%	18.34%	18.62%
from –0.49% to 0.5%	28.75%	27.21%
from 0.51% to 1.5%	19.37%	18.62%
from 1.51% to 2.5%	9.89%	10.61%
2.51% <	6.98%	7.84%

4 DISCUSSION OF RESULTS

For this analysis, the data from January 2, 2008 to March 21, 2018, it was observed that the highest probability of variation of the two indexes is within the range of –0.49 to 0.5% and for Ibovespa, this variation is likely to occur again in 3.4 days, while for Dow Jones, it will be in 1.9 days. This information can be important, since repeating patterns can help in choosing the time to buy or sell stocks.

By the application it was studied the behavior of probabilities by analyzing the Dow Jones data scenario for time t, while Ibovespa was for $t + 1$ and $t + 2$. It was analyzed that the variation of these probabilities was small. It was also observed that there is a 27.21% chance that the Ibovespa variation, with two steps in time, is in the range of –0.49 to 0.5%, when given the Dow Jones also vary within these limits in time t.

In relation to the steady state, its behavior was similar for the two indexes, that is, a low probability for the first and last data intervals, and a high probability for the intervals close to 0% of variation. This behavior was repeated for the expected recurrence time, but the number of days to occur again the highest probability is, for Ibovespa, practically twice when compared to Dow Jones. This indicates that the Brazilian stock exchange is more volatile than the American stock market, a fact that can be verified with the tables of the historical variations of these two indexes, since the one of the Ibovespa has a greater intensity in the oscillations of the values.

Information such as this can be advantageous for investors who negotiate on the stock exchange, since the knowledge of the behavior of the markets allows to assist in the decision making regarding the purchase and sale of shares. For practical applications, it is suggested that the database be fed in the desired time step, and that the analysis be done from one period to the next, maximizing the possibility of gain through the transition probabilities.

5 CONCLUSIONS

The applications of High-order Multivariate Markov Chains are still little explored in the literature. However, it can be a tool of great potential. This paper studied an application to analyze the variation of the main indexes of the Brazilian and American stock exchange, which was efficient in obtaining results.

Thus, this work aimed to instruct Markov Chains and their possible practical applications, providing information to investors and, consequently, enabling a better targeting of their efforts and investments, as well as better risk management. It is known that other techniques may be associated with valuation for investors, the technique presented here may be considered as one more option.

From this application, it was possible and feasible to forecast the range of variation of the Ibovespa index, using its data and the Dow Jones data, because the Brazilian economy is related to the American stock market, as mentioned in the literature.

This work had the purpose of presenting the application of the High-order Multivariate Markov Chains for two stock indexes. As a continuation hint more comparisons with other tools can be made from this implementation.

REFERENCES

[1] ACHCAR JA, CUERVO EC & BAROSSI-FILHO M. 2012. Multivariate volatility models: an application to Ibovespa and Dow Jones Industrial. *Cuadernos de Economía*, **31**(56): 301–320.

[2] ALMEIDA D. 2013. Assimetrias na volatilidade e nas perturbações nos modelos de volatilidade. Dissertação (Mestrado em Estatística) – Universidade Estadual de Campinas.

[3] ANDERSEN AR, NIELSEN BF & REINHARDT LB. 2017. Optimization of hospital ward resources with patient relocation using Markov chain modeling. *European Journal of Operational Research*, **260**(3): 1152–1163.

[4] BOLLERSLEV T, CHOU RY & KRONER KF. 1992. ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of econometrics*, **52**(1-2): 5–59.

[5] CHING WK, FUNG ES & NG MK. 2004. Higher-order Markov chain models for categorical data sequences. *Naval Research Logistics (NRL)*, **51**(4): 557–574.

[6] CHING WK, NG MK & FUNG ES. 2008. Higher-order multivariate Markov chains and their applications. *Linear Algebra and its Applications*, **428**(2-3): 492–507.

[7] CHIU CWJ, HARRIS R, STOJA E & CHIN M. 2018. Financial market volatility, macroeconomic fundamentals and investor sentiment. *Journal of Banking & Finance*, **92**: 130–145.

[8] CHRISTIANSEN C, SCHMELING M & SCHRIMPF A. 2012. A comprehensive look at financial volatility prediction by economic variables. *Journal of Applied Econometrics*, **27**(6): 956–977.

[9] DOUBLEDAY KJ & ENSUGE JN. 2011. Application of Markov Chains to Stock Trends. *Journal of Mathematics and Statistics*, **7**(2): 103–106.

[10] FÁVERO LPL & CONFORTINI D. 2010. Modelos multinível de coeficientes aleatórios e os efeitos firma, setor e tempo no mercado acionário Brasileiro. *Pesquisa Operacional*, **30**(3): 703–727.

[11] FITRIYANTO A & LESTARI TE. 2018. Application of Markov Chain to stock trend: A study of PT HM Sampoerna, tbk. In. *IOP Conference Series: Materials Science and Engineering*, **434**(1): 012007.

[12] FORNARI F & MELE A. 2013. Financial volatility and economic activity. *Journal of Financial Management, Markets and Institutions*, **1**(2): 155–198.

[13] HILLIER F & LIEBERMAN G. 2005. Introduction to Operations Research. 8. ed. Nova Iorque: McGraw-Hill.

[14] HO LL & QUININO R. 2012. An analysis of Taguchi's on-line quality monitoring procedure for variables based on the results of a sequence of inspections. *Pesquisa Operacional*, **32**(1): 197–212.

[15] HUNTER J. 2018. The computation of the mean first passage times for Markov chains. *Linear Algebra and its Applications*, **549**: 100–122.

[16] INFOMONEY. Site institucional. Disponível em: `http://www.infomoney.com.br/`.

[17] ISHIZAWA DK. 2008. *Modelos de volatilidade estatística*. Dissertação (Mestrado em Estatística) – Universidade Federal de São Carlos.

[18] KÁRNÝ M. 2016. Recursive estimation of high-order Markov chains: approximation by finite mixtures. *Information Sciences*, **326**: 188–201.

[19] KY DX & TUYEN L. 2018. A Higher order Markov model for time series forecasting. *International Journal of Applied Mathematics and Statistics*, **57**(3): 1–18.

[20] MELLO ARAF. 2009. Volatilidade implícita das opções de ações: uma análise sobre a volatilidade futura. Dissertação (Mestrado em Finanças e Economia) – Fundação Getúlio Vargas.

[21] RIBEIRO BPS. 2009. Modelagem e previsão de volatilidade para o setor siderúrgico brasileiro: volatilidade estocástica versus determinística. Dissertação (Mestrado em Economia) – Universidade Federal do Rio Grande do Sul.

[22] ROSHAN G & NASTOS PT. 2018. Assessment of extreme heat stress probabilities in Iran's urban settlements, using first order Markov chain model. *Sustainable Cities and Society*, **36**: 302–310.

[23] SCHWERT GW. 1989. Why does stock market volatility change over time?. *The Journal of Finance*, **44**(5): 1115–1153.

[24] SHEPHARD N. 2008. Stochastic volatility. New Palgrave Dictionary of Economics: MacMillan.

[25] SOUZA GCUI, SAMANEZ CP, RAPOSO GS & GONÇALVES AC. 2011. Application of numerical methods, derivatives theory and Monte Carlo simulation in evaluating BM&F BOVESPA's POP. *Pesquisa Operacional*, **31**(2): 195–215.

[26] STAUDT FH, COELHO AS & GONÇALVES MB. 2011. Determinação da capacidade real necessária de um processo produtivo utilizando cadeia de Markov. *Production*, **21**(4): 634–644.

[27] TAHA H. 2007. *Operations Research: an introduction*. 8. ed. New Jersey: Pearson Prentice Hall.

[28] YANG H, LI Y, LU L & QI R. 2011. First order multivariate Markov chain model for generating annual weather data for Hong Kong. *Energy and Buildings*, **43**(9): 2371–2377.

[29] YU J & MEYER R. 2006. Multivariate Stochastic Volatility Models: Bayesian Estimation and Model Comparison. *Research Collection School of Economics*, **25**(2-3): 361–384.

[30] ZHU Q & SINGH G. 2016. The impacts of oil price volatility on strategic investment of oil companies in North America, Asia, and Europe. *Pesquisa Operacional*, **36**(1): 1–21.