

## Previsão de Vendas do Varejo Brasileiro por meio de *Singular Spectrum Analysis* (SSA)

### Resumo

O setor varejista se mostra de grande importância no cenário econômico brasileiro, sendo responsável em 2013 por 25,8% do PIB do Brasil, com taxa de crescimento de 2,2% acumulada no ano de 2014, segundo a Pesquisa Mensal de Comércio (PMC) de dezembro de 2014 divulgada pelo IBGE. Prever o nível de vendas pode auxiliar no controle gerencial, em especial no controle interno de atividades operacionais, nas políticas de investimento e, conseqüentemente, na tomada de decisão. Sendo assim, este estudo trabalha com a série temporal de vendas do varejo brasileiro com dados de janeiro de 2000 a agosto de 2014. Uma forma eficiente de lidar com séries temporais que possuem sazonalidade e erros e imprecisões consiste no uso do *Singular Spectrum Analysis* (SSA). O SSA é um método eficiente na extração e reconstrução de componentes periódicas e não periódicas de séries temporais com elevados níveis de ruído e é usualmente aplicada em diversos outros campos do conhecimento. Assim, o objetivo deste trabalho é o de utilizar este método em séries da área contábil e financeira, especificamente, a série de vendas do varejo brasileiro. As previsões geradas pelo SSA para 12 meses à frente, foram comparadas a previsões geradas por um SARIMA. Quando levado em consideração a métrica dos Soma dos Quadrados dos Resíduos, o SSA obteve desempenho superior ao do SARIMA, porém, quando comparado o Erro Percentual Absoluto Médio, o SARIMA obteve melhor desempenho.

**Palavras-chave:** Vendas do Varejo, Séries Temporais, *Singular Spectrum Analysis*

### 1. INTRODUÇÃO

Na economia atual a tomada de decisão deve estar a par dos efeitos gerados pela variação das condições econômicas ao longo do tempo. Uma das formas de conciliar estes efeitos com o interesse do tomador de decisão é fazer previsões sobre eventos futuros relevantes (LEVINE *et al.*, 2000).

Vender constitui uma atividade indispensável para o funcionamento de toda sociedade. É por meio das vendas que as empresas ou organizações garantem sua própria sustentabilidade e que os consumidores atendem suas necessidades e seus desejos. Essa atividade, contudo, demanda capacidade de adaptação e bom trabalho preliminar.

No âmbito da contabilidade, por sua vez, as previsões auxiliam, principalmente, no controle gerencial, em especial no controle interno de atividades operacionais e nas políticas de investimento. Especificamente, este artigo tem como objeto de estudo o volume de vendas reais do varejo brasileiro, setor que é o maior gerador de empregos formais no país e foi responsável em 2013 por 25,8% do PIB do Brasil (IBGE, 2014).

Para execução destas previsões vários métodos podem ser empregados, sendo estes métodos divididos, basicamente, em dois grupos: qualitativos e quantitativos.

Os métodos qualitativos são vistos como técnicas subjetivas cujo objetivo é captar a percepção dos indivíduos sobre os fluxos futuros (ANGELO *et al.*, 2011). Por sua vez, os métodos quantitativos são vistos como técnicas de estruturas bem definidas que utilizam de dados históricos para prever fluxos futuros. Em outras palavras, os métodos quantitativos buscam explicar fluxos futuros com base nos fluxos passados.

Portanto, é possível afirmar que os métodos qualitativos são mais flexíveis, considerando diversas variáveis, contudo perde-se com a falta de objetividade. Por outro lado, ao contrário das

técnicas qualitativas, os métodos quantitativos são menos flexíveis, considerando menor número de situações, porém com ganho na objetividade (ANGELO *et al.*, 2011).

Este trabalho tem como objetivo principal prever a série de vendas do varejo brasileiro pelo método SSA. Este método permite que uma série temporal possa ser transformada, por meio de um mapa invertível, em uma matriz trajetória, uma matriz passível de ser expandida em termos de valores singulares. Cada componente nesta expansão concentra uma parcela de energia contida na matriz trajetória gerada a partir da série temporal e pode ser reconhecida como tendência, sazonalidade ou ruído (HASSANI, 2007).

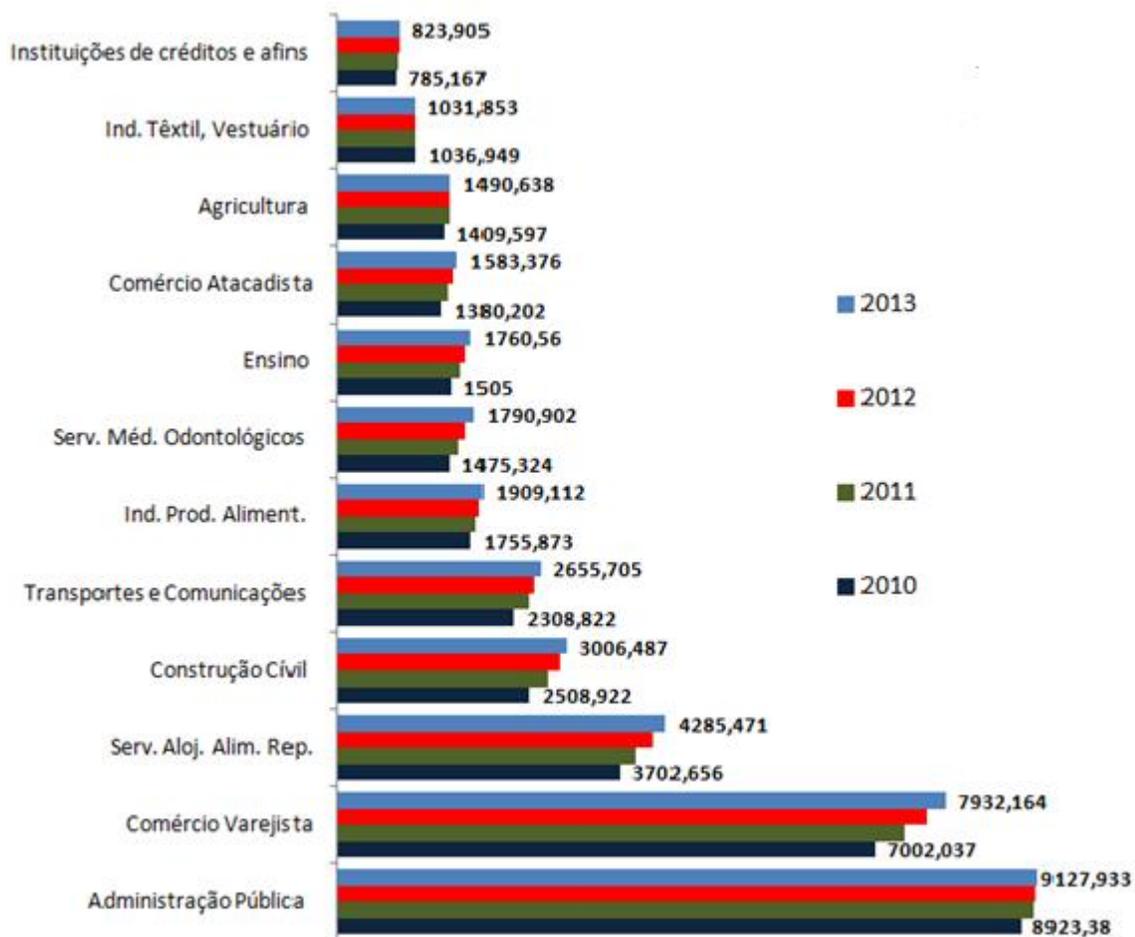
Além desta introdução, este artigo dividido da seguinte maneira: 1) na segunda seção, é realizada uma contextualização das vendas de varejo no Brasil e uma revisão das técnicas de séries temporais utilizadas em contabilidade e finanças em nível nacional; 2) na terceira seção, a metodologia a ser utilizada é apresentada; 3) em seguida, há a discussão dos resultados; e 4) por fim, as considerações finais desse trabalho.

## **2. CONTEXTUALIZAÇÃO**

Com o crescimento da renda real e o conseqüente aumento do poder aquisitivo da população, assim como o maior incentivo ao consumo, as empresas ou organizações devem estar preparadas para atender de forma mais eficiente as necessidades do consumidor. É necessário estar pronto para contornar as constantes atualizações de tecnologia e fornecer de forma rápida um produto ou determinado serviço. Para tal, o varejo se mostra uma ótima solução. Ele consiste na atividade de venda de produtos e serviços para o consumidor final, sendo o varejista qualquer organização ou empresa cuja atividade primordial se baseia no varejo (PARENTE, 2000).

O setor varejista demonstra importância crescente no cenário econômico brasileiro, sendo responsável em 2013 por 25,8% do PIB do Brasil, com taxa de crescimento de 2,2% acumulada no ano de 2014, segundo a Pesquisa Mensal de Comércio (PMC) de dezembro de 2014 divulgada pelo IBGE. Além disso, o setor é o maior gerador de empregos formais no País, sendo responsável por empregar, em 2013, cerca de 8 milhões de pessoas, ficando atrás apenas do setor de administração pública em número de funcionários (Figura 1).

Sua importância, contudo, não se limita ao seu próprio resultado. Este setor é responsável, frequentemente, pelo intermédio entre o setor industrial e o consumo das famílias, tendo influência direta em seus resultados. Isto se dá pelo fato de que, ao estipular o volume de vendas previsto, o setor estabelece a quantidade necessária de estoque, o que conseqüentemente influencia a quantidade produzida pelo setor industrial e limita o consumo das famílias. De tal modo, verifica-se a importância da estimativa coerente e bem ajustada da projeção de vendas.



Fonte: Instituto de Desenvolvimento do Varejo (2014).

**Figura 1 – Total de Funcionários por setor**

Para obter estimativas coerentes e bem ajustadas vários métodos de previsão podem ser utilizados. A variedade dos métodos, assim como sua área de aplicação, permite às empresas ou organizações a análise de diversos pontos, sendo a avaliação de vendas futuras uma das mais frequentes.

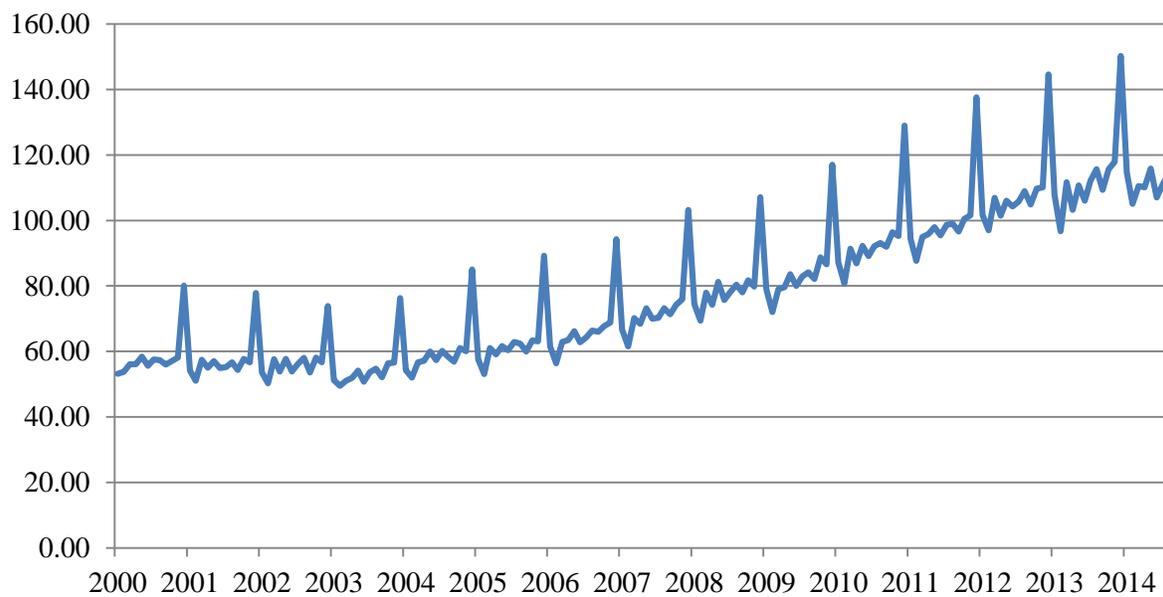
No âmbito da avaliação de desempenho econômico algumas técnicas se destacam dando realce para aquelas provenientes da análise de series temporais. Para as próprias áreas de contabilidade e finanças, é possível identificar na literatura a utilização e comparação de técnicas provenientes da análise de series temporais. Alguns exemplos apresentados na Figura 2 evidenciam a variedade destas técnicas e de suas utilidades, em especial, o trabalho publicado pela revista *Brazilian Business Review (BBR)* que compara técnicas na previsão de vendas do varejo (ANGELO *et al.*, 2011).

Referência	Título	Ano	País	Setor Analisado	Metodologia	Revista
Medeiros <i>et al.</i> , 2011	Modelando e estimando as demonstrações financeiras de uma empresa com o modelo VAR - VECM	2011	Brasil	Petrobras S/A.	VAR; VECM	Brazilian Business Review
Angelo <i>et al.</i> , 2011	Séries temporais e redes neurais: uma análise comparativa de técnicas na previsão de vendas do varejo brasileiro	2011	Brasil	Varejo	ARIMA; Redes Neurais	Brazilian Business Review
Moraes <i>et al.</i> , 2013	Estimativas de longo prazo para volatilidade de series temporais no mercado financeiro brasileiro	2013	Brasil	Mercado Financeiro	GARCH; EGARCH; TGARCH; ARLS	Revista Brasileira de Finanças
Bressan, 2004	Tomada de decisão em futuros agropecuários com modelos de previsão de séries temporais	2014	Brasil	Agropecuário	ARIMA; Redes Neurais; Modelos Lineares Dinâmicos	Revista de Administração de Empresas Electron
Lima <i>et al.</i> , 2010	Previsão de preços de commodities com modelos ARIMA-GARCH e redes neurais com ondaletas: velhas tecnologias - novos resultados	2010	Brasil	Soja	ARIMA-GARCH; Ondaletas em redes neurais	Revista de Administração
Farias; <i>et al.</i> , 2006	Estudo comparativo da aplicação de modelos clássicos de previsão de demanda no agronegócio	2006	Brasil	Agronegócio	Medias moveis;	Revista Organizações em Contexto
Mól <i>et al.</i> , 2014	Volatilidade dos Índices de Ações Mid-Large Cap e Small Cap: Uma Investigação a partir de Modelos Arima/Garch	2014	Brasil	Mid-Large and Small CAP	ARIMA-GARCH	Revista de Gestão, Finanças e Contabilidade

Fonte: Elaborado pelos autores

**Figura 2 – Literatura nacional de métodos de previsão em contabilidade**

Este trabalho, por sua vez, visa introduzir no estudo da contabilidade e finanças, a técnica de Análise Espectral Singular (*Singular Spectrum Analysis – SSA*). Para comparar sua performance, suas previsões são comparadas às previsões resultantes de uma modelagem de Box & Jenkins. Para isso, será utilizada a série de janeiro de 2000 a agosto de 2014 de importância previamente citada, das vendas reais do varejo brasileiro – Índice (Figura 3).



Fonte: Elaborado pelos autores com base no IPEADATA.

**Figura 3 – Vendas reais do varejo brasileiro – Índice (média 2011= 100)**

### 3. METODOLOGIA

Nesta seção está disposta a metodologia utilizada neste artigo. Inicialmente, é apresentada a metodologia de SSA, destacando seus passos de decomposição e reconstrução. A previsão por meio do SSA é apresentada em seguida. Neste artigo, utiliza-se o modelo ARIMA como comparação, assim, ele é brevemente descrito na subseção seguinte. Por fim, são apresentados os testes estatísticos utilizados na modelagem.

#### 3.1. Análise Espectral Singular (*Singular Spectrum Analysis – SSA*)

O SSA consiste na transformação de uma série temporal  $Y_T = (y_1, \dots, y_T)$ , de valores reais não nulos, em uma soma de séries, de modo que cada componente dessa soma possa ser reconhecido como tendência, sazonalidade ou ruído (HASSANI, 2007). Para tanto, o algoritmo dessa técnica resume-se em duas etapas: decomposição e reconstrução. A primeira etapa é feita pela transformação dos dados em uma matriz trajetória e sua decomposição em valores singulares (SVD). A segunda etapa, por sua vez, trata do agrupamento de elementos da matriz trajetória decomposta para formação de subgrupos a partir dos quais uma nova série de tempo é obtida.

##### 3.1.1. Decomposição

O estágio de decomposição de uma série temporal via SSA está subdividido em duas fases: transformação e decomposição em valores singulares (*singular value decomposition – SVD*).

A transformação se dá pelo mapeamento da série unidimensional  $Y_T = (y_1, \dots, y_T)$  em uma matriz  $X_{k \times l}$ , em que  $K = T - L + 1$  e  $2 < L < T$ . A matriz trajetória, então, será composta por  $X_1, \dots, X_k$  com  $X_i = (y_i, \dots, y_{i+L-1}) \in \mathbb{R}^L$ .  $L$  representa o tamanho da janela de decomposição e é o único parâmetro a ser definido. A matriz trajetória resultante desse passo é uma matriz de Hankel, e, assim sendo,  $a_{i,j} \in X \rightarrow a_{i,j} = a_{i-1,j+1}$ .

A decomposição em valores singulares transforma a matriz trajetória  $X$  simétrica em uma soma de matrizes elementares. Seja  $S = XX^T$ ,  $\lambda_1 \geq \dots \geq \lambda_L$  seus autovalores ordenados em ordem decrescentes de magnitude e  $U_1, \dots, U_L$  o sistema ortogonal de seus autovetores correspondentes aos autovetores.

Seja  $d = \text{posto}(X) = \max(i \mid \lambda_i > 0)$ . Então, se  $V_i = \frac{X^T U_i}{\sqrt{\lambda_i}}$  ( $i = 1, \dots, d$ ), a SVD da matriz trajetória pode ser escrita como

$$X = X_1 + \dots + X_d \quad (1)$$

em que  $X_i = \sqrt{\lambda_i} U_i V_i^T$ . Cada matriz  $X_i$  é chamada de matriz elementar e o conjunto  $(\sqrt{\lambda_i}, U_i, V_i^T)$  é denominado autotripla.

### 3.1.2. Reconstrução

Inicialmente, realiza-se o agrupamento, que consiste em separar em grupos as matrizes elementares obtidas no passo anterior. Essa separação é feita com base nas correlações entre cada uma delas: utiliza-se a *Weighted Correlation Matrix* como uma matriz de distâncias. Assim, é possível obter uma definição prévia dos grupos de matrizes elementares (ou seja, os conjuntos de autotriplas) que formam a tendência, sazonalidade e ruído da série original.

### 3.2. Previsão

Na análise de séries temporais, podem-se fazer previsões de dados futuros se a estrutura da série for bem definida e se houver algum modelo que seja derivado dos dados e descreva essa estrutura. No SSA tal modelo é descrito por Equações de Recorrência Linear (ERL). Uma determinada série pode ser descrita por ERL se for combinação linear de produtos exponenciais, polinômiais ou de séries harmônicas.

Seja  $d$  a dimensão mínima de todas as ERL que descrevem uma série temporal  $Y_T$ . Pode-se provar que caso a janela de decomposição  $L$  seja maior que  $d$ , então a matriz trajetória ótima da série  $Y_T$  tem espaço vetorial cuja dimensão é  $d$ . Esse espaço determina uma ERL de dimensão  $L-1$  que descreve a série. Assim, pode-se obter a continuação da série ao se aplicar essa ERL a seus últimos elementos (GOLYANDINA *et al.*, 2001).

### 3.3 ARIMA

Modelos ARIMA são usados para ajustar estruturas lineares de dependência temporal em séries temporais. Segundo Souza e Camargo (2004), a forma genérica de um modelo SARIMA (*Seasonal Autoregressive Integrated Moving Average*) é dado por:

$$y_t \nabla^d \nabla_S^D \Phi(B) \Phi(B^S) \theta^{-1}(B) \Theta^{-1}(B^S) = \varepsilon_t \quad (2)$$

Em que:  $\Phi(B^S) = [1 - \Phi_1 B^S - \dots - \varphi \Phi_p B^{SP}]$  é o fator auto-regressivo sazonal;  $\Phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$  é o fator auto-regressivo;  $\nabla^d = (1 - B)^d$  é o operador de diferença;  $\nabla_S^D =$

$(1 - B^S)^D$  é o operador de diferença sazonal;  $\Theta(B^S) = [1 - \Theta_1 B^S - \dots - \Theta_Q B^{SQ}]$  é o fator de médias móveis sazonal; e  $\theta(B) = [1 - \theta_1 - \dots - \theta_q B^q]$  é o fator de médias móveis.

### 3.4. Testes Estatísticos

Os testes estatísticos são mecanismos usados para a tomada de decisões. Tais testes baseiam-se na observação de dados quantitativos modelados por variáveis aleatórias e ajudam a determinar se há evidência o suficiente para rejeitar alguma conjectura ou hipótese estabelecida sobre o processo. Tal conjectura é denominada *Hipótese Nula*.

#### 3.4.1. Teste de Jarque-Bera

O teste de Jarque-Bera computa a assimetria e a curtose de uma distribuição através da seguinte estatística:

$$JB = n \left[ \frac{S^2}{6} + \frac{(K-3)^2}{24} \right] \quad (3)$$

em que  $n$  é o tamanho da amostra,  $S$  é o coeficiente de assimetria e  $K$  o coeficiente de curtose. Para uma variável normalmente distribuída,  $S = 0$  e  $K = 3$ . Esse teste estabelece a hipótese nula de que os resíduos são normalmente distribuídos.

#### 3.4.2. Teste de Ljung-Box

O teste de Ljung-Box verifica se existe autocorrelação entre os valores observados de uma determinada série através da seguinte estatística:

$$Q = n(n+2) \sum_{k=1}^h \frac{\rho'_k{}^2}{n-k} \quad (4)$$

em que  $n$  é o tamanho da amostra,  $\rho'_k$  é a autocorrelação da amostra no *lag*  $k$ , e  $h$  é o número de *lags* testados. Esse teste estabelece a hipótese nula de que os dados são independentemente distribuídos.

## 4. ESTIMAÇÕES E RESULTADOS

Para a estimação dos modelos utilizou-se do logaritmo da série de vendas de varejo (Figura 3), sendo a série dividida em dois períodos. O primeiro, com os dados de janeiro de 2000 até agosto de 2013, foi utilizado para a definição dos modelos. Os dados referentes aos meses de setembro de 2013 a agosto de 2014 foram empregados na comparação dos ajustes destes modelos. Os resultados apresentados foram obtidos por meio dos softwares R 3.1.2 e Eviews® 6.

### 4.1. Decomposição

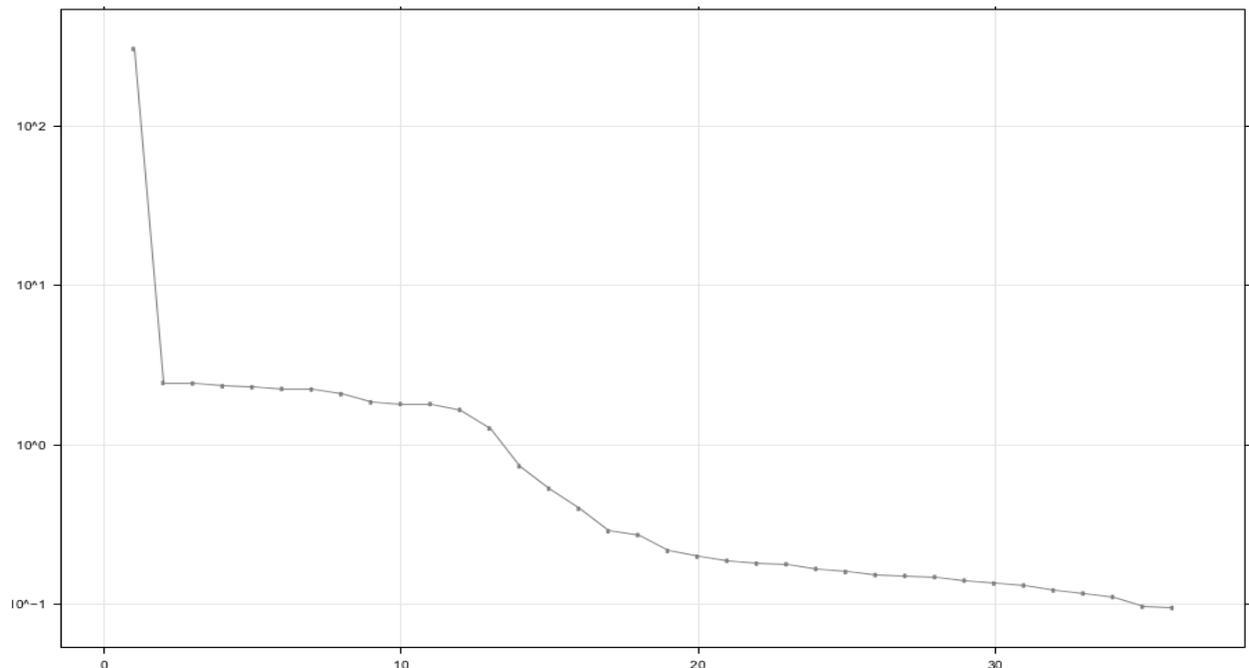
Como mencionado na seção 3.1, o tamanho  $L$  da janela é o único parâmetro a ser definido. A escolha ótima desse parâmetro é fundamental para a qualidade da decomposição (HASSANI, 2007). Por esse motivo há na literatura um amplo debate sobre qual deve ser o tamanho ideal da janela. Há um consenso de que o valor de  $L$  não deve ser maior que metade do tamanho da série. Alguns autores como Elsner e Tsonis (1996) e Hassani (2007) sugerem que a janela deve ter um valor próximo a  $T/4$  e também proporcional ao período de sazonalidade da série, caso tal componente periódico exista. Assim, pode-se obter uma melhor separação dos elementos aditivos.

Seguindo essas recomendações utilizou-se neste trabalho,  $L = 48$ , uma vez que  $T = 164$  e o período sazonal da série estudada é igual a 12. Sendo assim, por meio da decomposição em valores singulares da matriz trajetória obteve-se 48 autotriplas, das quais se extraiu os componentes da série. Para agrupar adequadamente as autotriplas em tendência, sazonalidade ou ruído, três ferramentas auxiliares foram utilizadas: o espectro de autovalores, os gráficos de pares de autovetores e a matriz-w de correlação das autotriplas.

#### 4.1.1. Espectro de Autovalores

Segundo Hassani (2007), os autovalores dos componentes harmônicos são quase idênticos, ao passo que as séries de ruídos produzem uma sequência de autovalores que decrescem lentamente. Assim, pode-se fazer a identificação visual dos componentes que formam a série observando-se variações bruscas, quedas suaves e valores quase constantes.

De acordo com a Figura 4 a grande disparidade entre o logaritmo do autovalor 1 e o autovalor 2, seguida de um período de estabilidade até o autovalor 12, é um indício de que o primeiro componente se mostra um componente de tendência enquanto os componentes 2 a 12 formam um grupo de sazonalidade. É também possível identificar uma sequência de autovalores, 19 a 48, que decrescem lentamente, indicando que este conjunto representa um grupo de ruído. O conjunto formado pelos autovalores 13 a 18 representam um grupo transitório, no qual sua classificação como ruído, sazonalidade ou tendência, dependerá do resultado das demais ferramentas auxiliares.



Fonte: Elaborado pelos autores com base no software R 3.1.2.

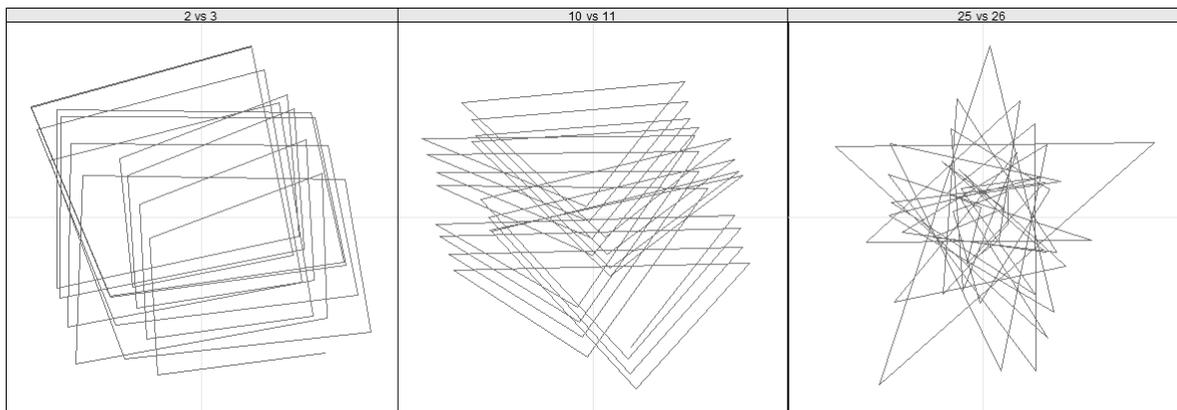
**Figura 4 – Logaritmos dos autovalores**

#### 4.1.2. Pares de Autovetores

A análise dos pares de autovetores se baseia na sua combinação gráfica, com o objetivo de identificar períodos harmônicos e a periodicidade destes períodos. Desse modo, se esta combinação

produzir figuras geométricas com vértices bem definidos, é possível afirmar que estes autovetores representam um período harmônico, devendo ser considerados como sazonalidade.

Seguindo esta lógica as conclusões da análise feita na seção 4.1.1 foram confirmadas, em parte, pelos gráficos de pares de autovetores. Eles mostraram que os pares de autovetores considerados como sazonais na seção 4.1.1 formam polígonos de  $n$  vértices, como por exemplo, o par dos autovetores 2 e 3, polígono de 4 vértices, e o par 10 e 11, polígono de 3 vértices, representados na Figura 5, respectivamente. Além disso, a análise dos pares de autovetores mostra que há um período harmônico, de frequência igual a 2.5, na combinação dos autovetores 25 e 26 (Figura 5), vetores considerados anteriormente como componentes do ruído.



Fonte: Elaborado pelos autores com base no software R 3.1.2.

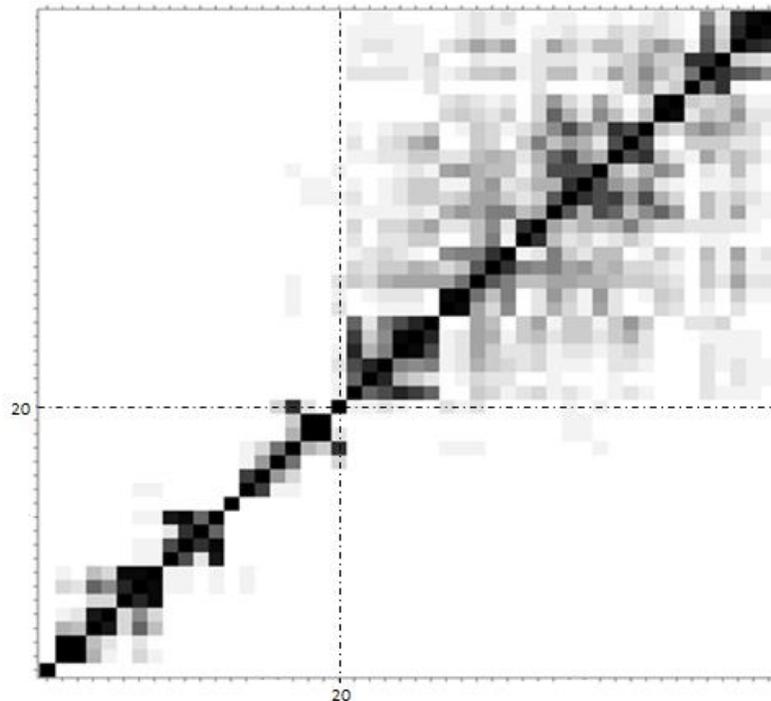
**Figura 5 – Pares dos autovetores 2-3, 10-11 e 25-26**

#### 4.1.3. Matriz de Correlação Ponderada

O conceito de separabilidade é, segundo Hassani (2007), o principal ponto no estudo da SSA, de modo que, a decomposição dos valores singulares só será bem sucedida se os componentes gerados forem separáveis uns dos outros. O método para identificar, portanto, a separabilidade das autotriplas é chamado de *w-correlation* ou correlação ponderada (*weighted correlation*), que mensura a dependência entre duas series.

Após identificar os períodos harmônicos é necessário checar há separabilidade entre as autotriplas. Este método tem por objetivo confirmar que os grupos formados não apresentam correlação entre si, podendo então ser reconstruídos corretamente. Ele ainda permite uma última análise, mostrando indícios de quais autotriplas devem ser combinadas.

A Figura 6 mostra a matriz de correlação ponderada das 48 autotriplas. Ele permite a confirmação da análise feita nas seções 4.1.1 e 4.1.2, estabelecendo três grupos bem definidos e um quarto grupo classificado como transitório. O primeiro, composto pela autotripla 1, representa a parte relacionada a tendência da série, o segundo, composto pelas autotriplas 2 a 16, representa o grupo relacionado a sazonalidade e o terceiro, composto pelas autotriplas 21 a 48, representa o grupo relacionado ao ruído. O quarto grupo, classificado como um grupo transitório deve ser analisado com maior atenção; nele buscar-se-á indícios de qual a melhor classificação de cada autotripla separadamente.



Fonte: Elaborado pelos autores com base no software R 3.1.2.

**Figura 6 – Matriz de correlação ponderada**

## 4.2. Reconstrução

A reconstrução é o segundo passo da técnica SSA. Nela se faz o agrupamento dos componentes da série, reconstruindo separadamente os componentes da tendência, sazonalidade e ruído. Para determinar qual a melhor forma de agrupamento utiliza-se a análise das secções 4.1.1, 4.1.2 e 4.1.3.

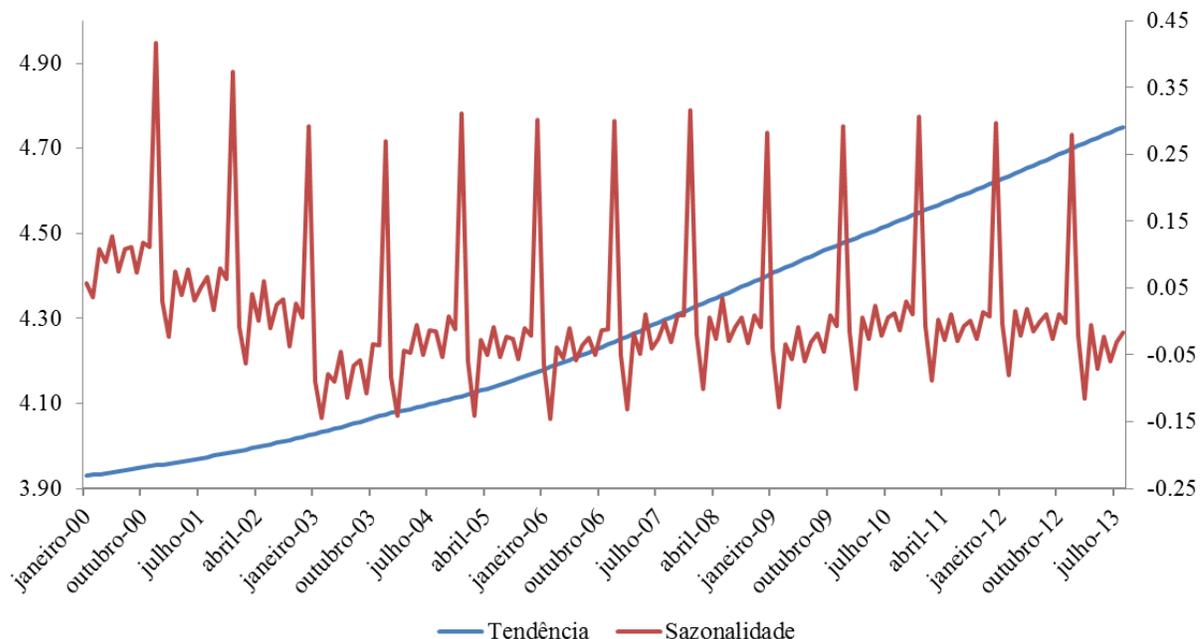
### 4.2.1. Tendência

A tendência, em geral, é expressa pelas primeiras autotriplas, sendo ela um componente que não contém elementos de oscilação (HASSANI, 2007). Como demonstrado anteriormente na Figura 6, a primeira autotripla estabelece um grupo único, que não se correlaciona com as demais, o que é um sinal de que a tendência será composta unicamente pela primeira autotripla. Este indício é reforçado ao analisar a combinação de autovetores e a Figura 4, do logaritmo dos autovalores. Desse modo, a tendência da série, reconstruída a partir da primeira autotripla, é representada na Figura 7.

### 4.2.2. Sazonalidade

A sazonalidade é expressa pelos períodos considerados como harmônicos, sendo ela um componente de oscilações constantes em dado período de tempo. Como demonstrado anteriormente, em especial na secção 4.1.2, as autotriplas consideradas como harmônicas são aquelas cujos pares de autovetores compõem polígonos com vértices bem definidos, de modo que o número de vértices representa a frequência destas oscilações. Portanto, este componente seria composto pelas autotriplas 2 -16 e 25-26.

Contudo, após a análise de cada autotripla do grupo considerado como transitório na secção 4.1.3, identificou-se períodos harmônicos nas autotriplas 19-20. Desse modo, a composição final da sazonalidade, reconstruída na Figura 7, será pelas autotriplas 2-16, 19-20 e 25-26.



Fonte: Resultados da pesquisa.

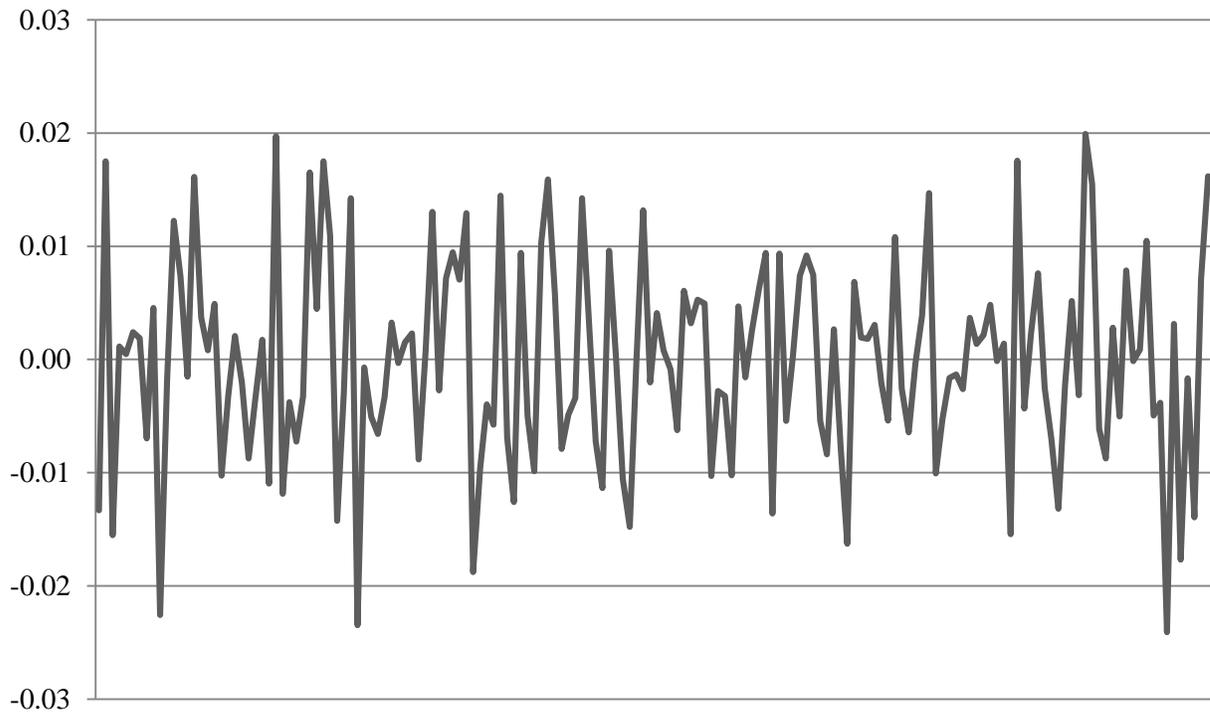
Obs: Tendência como Eixo da esquerda e Sazonalidade como Eixo da direita

**Figura 7 – Tendência e Sazonalidade**

#### 4.2.3. Ruído

O ruído embora composto pelas autotriplas restantes, não incorporadas aos componentes da tendência ou sazonalidade, deverá seguir alguns pressupostos básicos. Estes pressupostos são: i) os resíduos devem seguir uma distribuição normal, com média igual a 0 e variância constante; ii) os resíduos não devem ser autocorrelacionados. Tendo atendido a tais pressupostos, pode-se então classificar o ruído como ruído branco (*White Noise*) o que permite sua retirada da composição da série, sem perda de informação e possibilitando a melhor modelagem da mesma.

Tendo isto em vista o ruído foi reconstruído a partir das autotriplas 17-18, 20-24 e 27-48, com o resultado representado na Figura 8. Após sua reconstrução a série de ruídos passou pelos testes de Ljung-Box e Jarque-Bera (seção 3.4), com o objetivo de averiguar se atendia aos pressupostos descritos.



Fonte: Resultados da pesquisa.

**Figura 8 – Série de resíduos**

Em primeiro lugar buscou identificar se a série possui distribuição normal, com média igual a 0 e variância constante. Para isso utilizou-se da análise do histograma da série e do teste de Jarque-Bera. O resultado desse teste não forneceu evidências estatísticas suficientes para refutar a hipótese nula, indicando que não é possível rejeitar a hipótese de que a série possui distribuição normal. A análise do histograma, por sua vez, indicou uma curtose próxima a 3 e simetria próxima a 0, retificando o resultado encontrado pelo teste de Jarque-Bera.

Em seguida foi verificada a autocorrelação desses resíduos, com base na análise da função de autocorrelação e do teste de Ljung-Box. Os resultados desses procedimentos indicaram que não há correlação entre os componentes do ruído, dado um nível de significância de 5%. Desse modo, pode-se classificar a série de ruído como *White Noise*, indicando a separação correta dos componentes da série original.

### **4.3. SARIMA**

Utilizando da metodologia de Box & Jenkins, o melhor modelo encontrado foi um SARIMA(2,1,0)x(2,1,0) com os parâmetros descritos na Tabela 1. É importante destacar que nesta modelagem também foram utilizados os testes de Jarque-Bera e de Ljung-Box para a validação do modelo.

**Tabela 1 – Parâmetros do modelo SARIMA**

<b>Termo</b>	<b>Coefficiente</b>	<b>Erro Padrão</b>	<b>Estatística <i>t</i></b>	<b>p-valor</b>
<b>AR(1)</b>	-0.5943	0.0857	-6.9389	0.0000
<b>AR(2)</b>	-0.2916	0.0892	-3.2681	0.0014
<b>SAR(12)</b>	-0.5075	0.0896	-5.6634	0.0000
<b>SAR(24)</b>	-0.2433	0.0880	-2.7667	0.0066

Fonte: Resultados da pesquisa.

#### **4.4. Previsões**

Após a estimação dos modelos foram feitas as previsões relativas aos meses de setembro de 2013 a agosto de 2014 e calculados os índices relativos aos erros de previsão destes modelos. Os índices utilizados foram a Soma dos Quadrados dos Erros (SQE) e o Erro Percentual Absoluto Médio (MAPE). Por meio da Tabela 2, é possível observar estas estatísticas para ambos os modelos. Em relação ao SQE, o modelo SSA se mostrou com melhor capacidade preditiva, contudo, em relação ao MAPE, o modelo SARIMA (2,1,0) (2,1,0) apresentou melhor capacidade preditiva.

**Tabela 2 – Comparação entre os modelos**

<b>Modelo</b>	<b>SQE</b>	<b>MAPE</b>
<b>SSA</b>	0.0228	0.74%
<b>ARIMA</b>	0.0236	0.72%

Fonte: Resultados da pesquisa.

### **5. CONSIDERAÇÕES FINAIS**

A estimativa coerente e bem ajustada da projeção de vendas pode ser destacada no âmbito da contabilidade, para o controle gerencial, em especial no controle interno de atividades operacionais e nas políticas de investimento, sendo assim um importante instrumento para a tomada de decisão.

Para se obter estimativas coerentes e bem ajustadas vários métodos de previsão podem ser utilizados. Este trabalho utilizou da metodologia de SSA – que tem sido utilizada com sucesso em diversos campos do conhecimento – na análise da série temporal de vendas do varejo brasileiro. Foi possível decompor a série original em tendência, sazonalidade e ruído. As componentes foram detectadas por meio da análise do espectro de autovalores, dos pares de autovetores e da matriz de correlação ponderada. Além disso, os testes de Jarque-Bera e de Ljung-Box foram utilizados, respectivamente, para comprovar que os resíduos seguiam uma distribuição normal e eram desconcorrelatados.

Na comparação com a modelagem de Box & Jenkins não foi possível verificar ganhos preditivos. Daí surge a primeira extensão natural para este trabalho: o uso de janelas de decomposição diferentes, já que não foram testadas janelas que não fossem proporcionais ao período sazonal. A escolha ótima desse parâmetro é fundamental para a qualidade da decomposição e existe na literatura um amplo debate sobre qual deve ser o tamanho ideal da janela. Há um consenso de que o valor de *L* não deve ser maior que metade do tamanho da série, porém alguns autores citam que ele não precisa ser proporcional ao período de sazonalidade da série. Assim, pode-se obter uma melhor separação dos elementos aditivos.

Outro trabalho futuro é a utilização da metodologia de SSA em outras séries diferentes do campo da contabilidade e finanças, *e.g.* séries de retornos de ações. Além disso, em trabalhos futuros, objetiva-se utilizar a metodologia de SSA para o pré-processamento das séries de duas formas distintas: i) por meio da remoção de ruídos existentes nas séries originais, e ii) por meio da decomposição das séries originais em subséries. O passo posterior ao pré-processamento é a aplicação dos modelos de séries temporais (*e.g.* ARIMA e Redes Neurais) às séries transformadas. Assim, esperam-se obter uma modelagem mais adequada para explicar o comportamento das séries em análise, dado que i) no caso do encolhimento de ruídos, as influências desproporcionais, de um ou mais pares de observações contaminados por resíduos podem ser amenizados e ii) no caso da decomposição das séries originais em subséries, melhores padrões de comportamento podem ser obtidos.

## 6. REFERÊNCIAS

ANGELO, C.; ZWICKER, R.; FOUTO, N.; LUPPE, M. Séries temporais e redes neurais: uma análise comparativa de técnicas na previsão de vendas do varejo brasileiro. **Brazilian Business Review**, v.8, n.2, p. 01-21, 2011.

BRESSAN, A. Tomada de decisão em futuros agropecuários com modelos de previsão de séries temporais. **RAE Eletrônica**, v.3, n.1, 2004.

ELSNER, J. B.; TSONIS, A. A. **Singular Spectral Analysis: A New Tool in Time Series Analysis**. New York: Plenum Press, 1996.

FARIAS, O.; NOGUEIRA NETO, M.; NOGUEIRA, C. C. Estudo comparativo da aplicação de modelos clássicos de previsão de demanda no agronegócio. **Revista Organizações em Contexto**, v.2, n.3, 2006

GOLYANDINA, N.; NEKRUTKIN, V.; ZHIGLJAVSKY A. A. **Analysis of Time Series Structure: SSA and Related Techniques**. Boca Raton: Chapman & Hall/CRC, 2001.

HASSANI, H. Singular Spectrum Analysis: Methodology and Comparison. **Journal of Data Science**, v.5, p. 239-257, 2007.

IBGE (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA). **Pesquisa Mensal de Comércio Dezembro** 2014. Instituto Brasileiro de Geografia e Estatística, Fevereiro, 2015. Disponível em: <[ftp://ftp.ibge.gov.br/Comercio\\_e\\_Servicos/Pesquisa\\_Mensal\\_de\\_Comercio/Fasciculo\\_Indicadores\\_IBGE/pmc\\_201412caderno.pdf](ftp://ftp.ibge.gov.br/Comercio_e_Servicos/Pesquisa_Mensal_de_Comercio/Fasciculo_Indicadores_IBGE/pmc_201412caderno.pdf)>. Acesso em Jan. 2015.

INSTITUTO DE DESENVOLVIMENTO DO VAREJO. Disponível em: <http://www.idv.org.br/>. Acesso em Jan. 2015.

IPEADATA. Disponível em: <[www.ipeadata.gov.br](http://www.ipeadata.gov.br)>. Acesso em Jan. 2015.

MEDEIROS, O.R.; VAN DOORNIK, B.F.N.; OLIVEIRA, G.R., Modelando e estimando as demonstrações financeiras de uma empresa com o modelo VAR – VECM. **Brazilian Business Review**, v.8, n.3, 2011.

MÓL, A.; FELIPE, I.; GALVÃO JÚNIOR, F. Volatilidade dos Índices de Ações Mid-Large Cap e Small Cap: Uma Investigação a partir de Modelos Arima/Garch. **Revista de Gestão, Finanças e Contabilidade**, v.4, n. 1, 2014

MORAES, A.; PINTO, A. C.; KLOTZLE, M. Estimativas de longo prazo para volatilidade de series temporais no mercado financeiro brasileiro. **Revista Brasileira de Finanças**, v.11, n.4, 2013.

LEVINE, D.; BERENSON, M.; STEPHAN, D. **Estatística: Teoria e Aplicações**. Rio de Janeiro: LTC Editora S.A., 2000.

LIMA, F. G.; KIMURA, H.; ASSAF NETO, A.; PERERA, L. C. Previsão de preços de commodities com modelos ARIMA-GARCH e redes neurais com ondaletas: velhas tecnologias - novos resultados. **Revista de Administração**, v.45, n.2, p.188-202, 2010.

PARENTE, J. **Varejo no Brasil: Gestão e Estratégia**. São Paulo: Editora Atlas, 2000.

SOUZA, R. C.; CAMARGO, M. E. **Análise e Previsão de Séries Temporais: Os Modelos ARIMA**. 2ed. Rio de Janeiro: Gráfica e Editora Regional, 2004.