

UM ESTUDO EM SÉRIES TEMPORAIS NA ANÁLISE DA RECEITA NOMINAL DE VENDAS DE VEÍCULOS E MOTOS

Samuel de Oliveira

Especialista em ensino de Física

Mestrando em Engenharia de Sistemas pela Universidade Federal de Lavras-UFLA

Professor do Centro Universitário de Formiga-UNIFOR

e-mail: professor.samuel@yahoo.com.br

Recebido em: 01/03/2012

Aprovado em: 06/11/2012

RESUMO

O propósito deste artigo é a análise da série de dados pesquisada pelo IBGE que produz indicadores que permitem acompanhar o comportamento do comércio varejista no país. O trabalho apresenta uma abordagem dos dados de índice mensal de comércio varejista ampliado, uma contextualização e conceitos da análise de séries temporais, bem como suas principais características. Para o conjunto de dados analisados fez-se o uso do software livre gretl 1.9.5 cvs e a partir dele, construiu-se o gráfico dos dados e também o gráfico da série estacionária, bem como o ajustamento dos dados para um modelo pertinente ao de séries temporais. O resultado obtido foi satisfatório, ainda que a abordagem utilizada foi apenas uma apresentação de ferramentas para esse tipo de ajuste em série de dados que possuem correlação temporal.

Palavras-chave: Séries Temporais. Vendas. Modelos Estatísticos.

A STUDY ON TIME SERIES IN ANALYSIS OF NOMINAL REVENUE OF VEHICLES AND MOTORCYCLES

ABSTRACT

The purpose of this paper is to analyze the data series investigated by IBGE that produces indicators to monitor the behavior of retail trade in the country. The paper presents an approach to data from the monthly retail trade expanded, a context and concepts of time series analysis and its main characteristics. For the analyzed data set was made using free software gretl 1.9.5 cvs, and it was constructed the graph of data and also the graph of the stationary series, as well adjustment of the data to a model appropriate to the time series. The result was satisfactory, though the approach was just a presentation tool for this kind of adjustment in the data series that have temporal correlation.

Keywords: Times Series. Sales. Statistical Models.

1 INTRODUÇÃO

O comércio pode estar relacionado com a economia formal que é firma registrada dentro da lei ou à economia informal que são firmas sem registros que não pagam impostos. O comércio informal traz prejuízos pois copiam vários tipos de produtos para uma venda mais barata, acarretando assim em graves perdas econômicas para o país.

O surgimento do mercado como um espaço físico ocorreu mesmo antes da invenção do dinheiro. Independentemente da existência do dinheiro, o fato de termos a oferta e a procura por mercadoria, mão de obra ou serviços gerais é o que permite a existência do comércio.

É de grande importância para as empresas que atuam no comércio conhecer a futura demanda de produtos comercializados por eles, pois isto seria uma grande ajuda em várias das atividades diárias efetuadas por estas empresas, especialmente se permitir otimizar o escalonamento da produção, além de permitir um agendamento de manutenção que não perturbe a sua produção.

2 METODOLOGIA

2.1 Séries temporais

Uma série temporal é um conjunto de observações ordenadas no tempo (não necessariamente igualmente espaçadas) e que apresentam dependência serial, isto é, dependência de instantes de tempo. (MORETTIN; TOLOI, 2006). Também denominada série histórica, é uma sequência de dados obtidos em intervalos regulares de tempo durante um período específico. Este conjunto pode ser obtido através de observações periódicas do evento de interesse.

Segundo Morettin e Toloi (2006), uma série temporal é um conjunto de observações compreendidas seqencialmente no tempo. Se a série histórica for denominada como Z , o valor da série no momento t pode ser escrito como Z_t ($t = 1, 2, \dots, n$).

Uma série temporal é um conjunto de dados que se distribuem de forma contínua no tempo, ou seja, as suas observações são ligadas diretamente umas às outras, sendo assim, não podemos trabalhar uma série temporal sem considerar a ordem cronológica das observações. O estudo de séries temporais requer o uso de técnicas diferenciadas daquelas trabalhadas em séries estatísticas, em que a ordem das observações é independente, ou seja, podemos agrupar as observações de uma maneira que não levamos em consideração o tempo.

As observações de séries temporais aparecem em campos diversos do conhecimento como, por exemplo, a Economia (preços diários de ações, taxa mensal de desemprego, produção industrial), Medicina (eletrocardiograma, eletroencefalograma), Epidemiologia (número mensal de

novos casos de meningite), Meteorologia (precipitação pluviométrica, temperatura diária, velocidade do vento) etc.

2.2 Objetivos da análise de séries temporais

Segundo Morettin e Toloí (2006), uma série temporal $Z(t_1), \dots, Z(t_n)$ observada nos instantes t_1, \dots, t_n pode ser utilizada para:

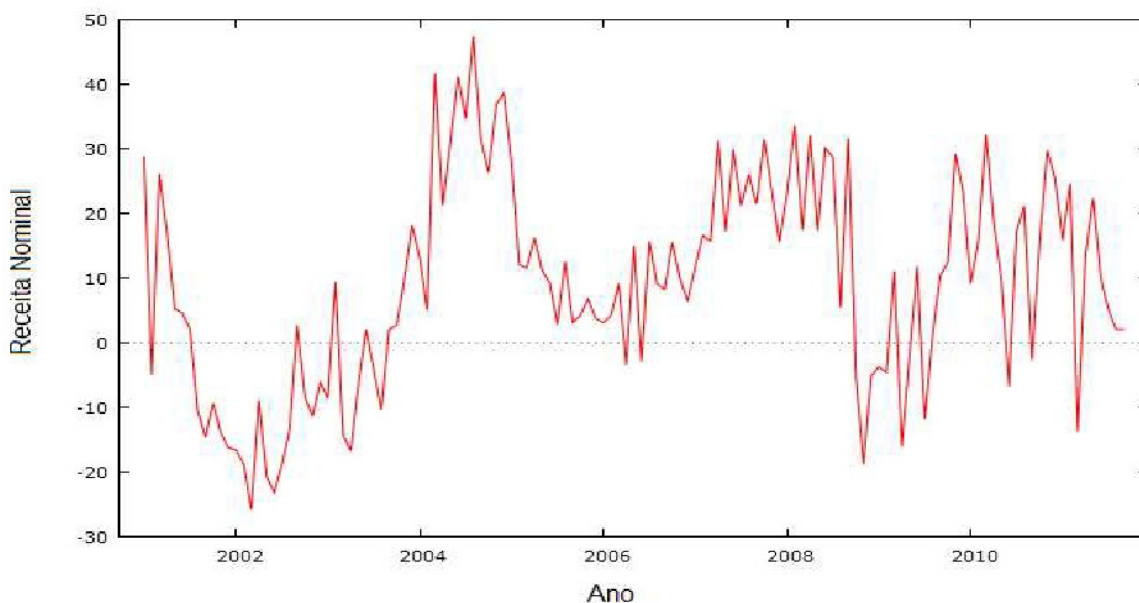
- Investigar o mecanismo gerador da série temporal;
- Fazer previsões de valores futuros da série, sendo que as previsões podem ser a curto e longo prazo;
- Descrever apenas o comportamento da série; neste caso a construção de histogramas e diagramas de dispersão, entre outros, podem ser ferramentas úteis;
- Verificar a existência de tendências, ciclos e variações sazonais;
- Procurar periodicidades relevantes nos dados; neste caso, a análise espectral pode ser de grande utilidade.

2.3 Gráfico

Em se tratando de séries temporais, o primeiro ponto de partida é fazer a representação gráfica da série, pois, através do gráfico, podemos fazer uma melhor análise e identificar as características que podem ser relevantes para o estudo da série em questão. O presente estudo mostra a análise da receita nominal de vendas de veículos e motos que compara os índices nominais e de volume da receita bruta de revenda do mês com os obtidos em igual mês do ano anterior.

Na FIG. 1, apresentam-se os valores de receita nominal de vendas de veículos e motos extraídos do Instituto Brasileiro de Estatística e Geografia, para o período de janeiro de 2001 a setembro de 2011.

Figura 1 - Receita nominal de vendas de veículos e motos (jan/01 a set/11)



2.4 Estacionariedade

De acordo com Morettin e Tolo (2006), uma das suposições mais frequentes que se faz a respeito de uma série temporal é a de que ela é estacionária, ou seja, ela se desenvolve no tempo aleatoriamente ao redor de uma média constante, refletindo alguma forma de equilíbrio estável.

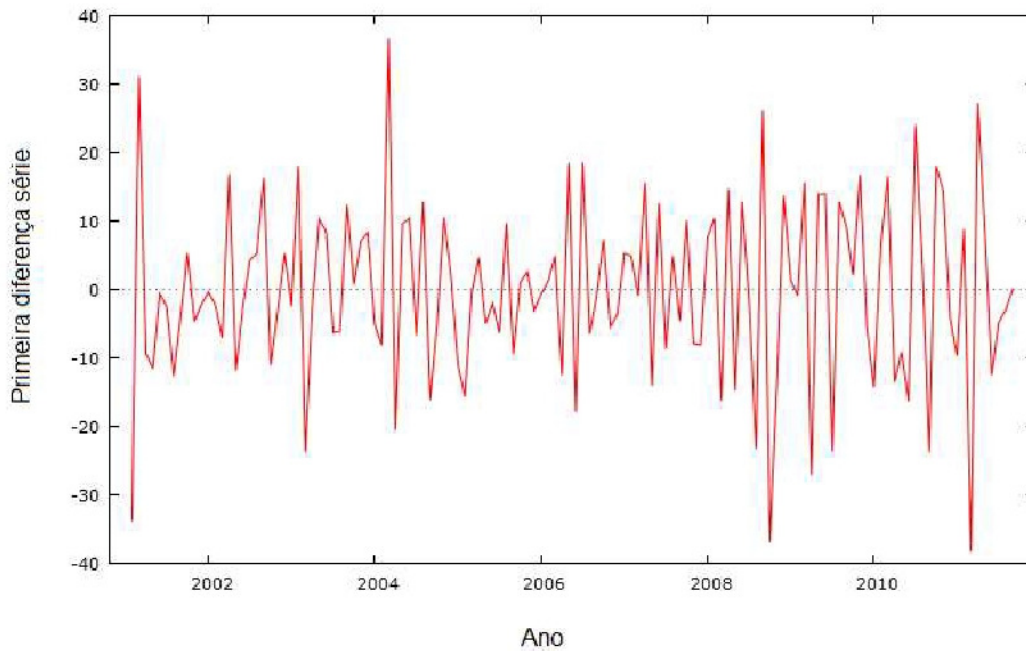
Na maioria dos procedimentos de análise estatística de séries temporais, supõe-se que estas sejam estacionárias. Com isso, caso a série não seja estacionária, será necessário transformar os dados originais. A transformação mais comum consiste em tomar diferenças sucessivas da série original, até se obter uma série estacionária. (MORETIN; TOLOI, 2006, p. 5).

$$\Delta Z(t) = Z(t) - Z(t - 1) \quad (1.0)$$

Segundo Stock e Watson (2004), para uma série de dados ser estacionária, suas variáveis não podem apresentar tendências e devem ser estáveis ao longo do tempo. Para as séries temporais, é importante que as variáveis sejam estacionárias ou passíveis de sua estacionariedade. Essa característica é fundamental para previsão do futuro com base na regressão de séries temporais, solidificando a premissa de que o futuro se comportará de acordo com o passado.

Para a série apresentada, foi realizada uma diferença para que, assim, possamos trabalhar a série temporal de forma invariante, ou seja, tornando-a estacionária. Quando a série é corretamente diferenciada, a variância da série transformada diminui. Por outro lado, um excesso de diferenças aumentará sua variância. A FIG. 2 mostra a primeira diferença tomada na série, tornando a série original em uma série estacionária.

Figura 2 - Primeira diferença tomada na série - $\Delta Z(t) = Z(t) - Z(t-1)$



2.5 Métodos para previsão de uma série temporal

A classe de modelos mais comumente utilizada na análise de séries temporais é a dos modelos auto-regressivos integrados e de médias móveis ou, simplesmente, ARIMA(p, d, q), que foi proposta por Box e Jenkins (1970).

Box e Jenkins (1970) formalizam a teoria da utilização de componentes auto-regressivos e de médias móveis na modelagem de séries temporais utilizando-se de duas ideias básicas na criação de sua metodologia de construção de modelos: a parcimônia que é a utilização do menor número possível de parâmetros para se obter uma representação adequada do fenômeno em estudo e a interatividade, ou seja, a informação empírica é analisada teoricamente e o resultado deste estágio é confrontado com a prática sucessivas vezes, até que o modelo obtido seja satisfatório.

2.5.1 Modelos auto-regressivos

A classe de modelos conhecida como modelos auto-regressivos-AR (no caso, de ordem p) é um processo iterativo em que há a identificação da ordem p através da função de auto-correlação; a partir daí, faz-se a estimativa de um modelo de previsão bem como a análise dos resíduos para a avaliação da existência de vieses e/ou grandes erros de estimativas. A dificuldade desta técnica é a identificação do modelo, pois é possível que pessoas diferentes identifiquem modelos de ordem diferentes para a mesma série temporal.

Suponha que $\{\epsilon_t\}$ seja um processo puramente aleatório com média zero e variância σ^2 .

Um processo Z_t é chamado de processo auto-regressivo de ordem p , ou AR(p), se

$$Z_t = \alpha_1 Z_{t-1} + \dots + \alpha_p Z_{t-p} + \epsilon_t \quad (1.1)$$

Sobre o modelo auto-regressivo:

Este modelo é clássico e foi, talvez, um dos primeiros a serem utilizados notadamente em Astronomia e Física. No primeiro caso, o interesse era determinar a posição de um planeta em dado momento do tempo e é claro que o erro obtido ao estimar a posição num instante não terá influência na posição do planeta em instantes posteriores. Por outro lado, o modelo (2.35)¹ é bastante utilizado em Física, quando, por exemplo, os a_t representam erros de mensuração de uma quantidade Q . O modelo reduz-se ao caso mais simples $Z_t = Q + a_t$, $t = 1, \dots, N$. (MORETTIN; TOLOI, 2006, p. 34).

Note a similaridade com um modelo de regressão múltipla, em que os valores passados de Z_t fazem o papel das regressoras. Assim, processos AR podem ser usados como modelo se for razoável assumir que o valor atual de uma série temporal depende do seu passado imediato.

2.5.2 Modelos médias-móveis

O modelo de média móvel de ordem q - MA(q) é usado quando há autocorrelação entre os resíduos, ou seja, existe uma relação de dependência entre o conjunto de erro em períodos passados (DELURGIO, 1998).

Este modelo em série temporal utiliza, como previsão para uma determinada observação no futuro, a média das observações passadas. As médias móveis podem ser simples, centradas ou ponderadas.

O termo média móvel é utilizado porque, à medida que a próxima observação está disponível, a média das observações é recalculada, incluindo esta observação no conjunto de observações e desprezando a observação mais antiga.

Seja $\{\epsilon_t\}$ um processo discreto puramente aleatório com média zero e variância $\sigma \in 2$.

Um processo Z_t é chamado de processo de médias móveis de ordem q , ou MA(q), se

$$Z_t = \epsilon_t + \beta_1 \epsilon_{t-1} + \dots + \beta_q \epsilon_{t-q} \quad (1.2).$$

Wheelwright e Makridaki (1985) colocam que, quanto maior o número de observações concluídas na média móvel, maior o efeito de alisamento na previsão. Assim, caso a série temporal

¹ O modelo 2.35 refere-se a : $Z = f(t) + a_t$, $t = 1, \dots, N$

apresente muita aleatoriedade ou pequenas mudanças nos padrões dessa série, um número maior de valores pode ser utilizado no cálculo da média móvel, obtendo-se uma previsão mais alisada.

Entretanto, se houver pouca flutuação aleatória nos dados ou mudança significativa no padrão da série, um número menor de observações deve ser incluído no conjunto de valores empregado na determinação da média para que se possa reagir a essas alterações mais rapidamente.

3 AJUSTAMENTO

Para uma identificação do modelo e ajustamento do próprio, DeLurgio (1998) sugere a utilização dos seguintes procedimentos: projeção dos dados em gráficos para verificar a existência de algum padrão; ajustes como deflacionar ou logaritmizar, estabilizando a variância; usar a função de auto-correlação (FAC) e a função parcial de auto-correlação (FACP) para verificar a existência de algum padrão na série; diferenciar os dados para obter a estacionariedade; examinar os FAC e FACP para identificar potenciais modelos, podendo ter ajuda de softwares especializados.

No estudo apresentado, utiliza-se o software gretl 1.9.5 cvs para atingir o objetivo deste trabalho que é a identificação de um modelo para a série de dados mostrada.

O objetivo da identificação é determinar os valores de p , d e q do modelo ARIMA(p,d,q), e também encontrar os parâmetros a serem usados no modelo.

O procedimento de identificação consiste em três partes:

- i) Verificar se existe a necessidade de uma transformação na série original.
- ii) Tomar diferenças da série, tantas vezes necessárias, para se obter uma série estacionária, de modo que o processo seja reduzido a um ARMA(p,q).
- iii) Identificar o processo ARMA(p,q) através da análise das autocorrelações e autocorrelações parciais estimadas.

4 DESENVOLVIMENTO

O correlograma é útil na identificação do tipo de modelo ARIMA que fornece a melhor representação de uma série observada. Um correlograma, em que os valores de r_k decaem para zero de forma relativamente lenta, indica não estacionariedade e a série precisa ser diferenciada.

Para séries estacionárias, o correlograma é comparado com as autocorrelações teóricas de vários processos ARMA para auxiliar na identificação daquele mais apropriado.

Por exemplo, se r_1 é significativamente diferente de zero e todos os valores subsequentes r_2, r_3, \dots são próximos de zero, então um modelo MA(1) é indicado já que sua função de

autocorrelação teórica se comporta assim. Por outro lado, se r_1, r_2, r_3, \dots parecem estar decaindo exponencialmente então um modelo AR(1) pode ser apropriado.

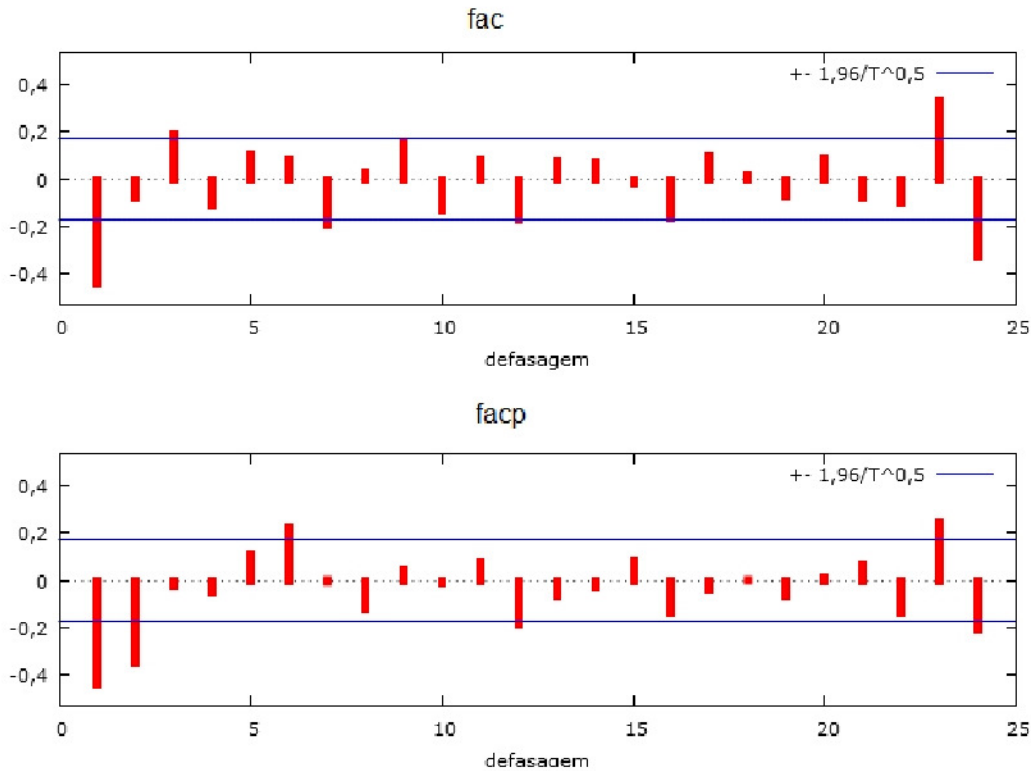
Vale notar, entretanto, que a interpretação de correlogramas é um dos aspectos mais difíceis da análise de séries temporais.

Alguns autores apresentam procedimentos para identificação das ordens p e q do modelo ARMA, a partir dos gráficos FAC e FACP, como no caso apresentado no QUADRO 1. Segundo a recomendação de Montgomery e Johnson (1976, p.469), a melhor tentativa é o ajuste de um modelo ARMA (1,0) ou simplesmente AR(1).

| MODELO | FAC | FACP |
|---------------|--|--|
| ARMA (p,0) | Decaimento gradativo | Decaimento brusco, após defasagem |
| ARMA (0,q) | Decaimento brusco, após defasagem | Decaimento gradativo |
| ARMA (p,q) | Decaimento gradativo, com onda senoidal amortecida, após a defasagem (p-q) | Decaimento gradativo, com onda senoidal amortecida, após a defasagem (q-p) |

A função de autocorrelação parcial é um importante coadjuvante nesta etapa de identificação se houver termos auto-regressivos no modelo, já que seus valores estimados tendem a ficar próximos de zero após a defasagem p .

O correlograma abaixo é referente à primeira diferença para a série em questão, lembrando que a função de autocorrelação nos mostra a ordem do modelo MA(q) que será ajustado, enquanto que a função de autocorrelação parcial nos mostra a ordem de um modelo AR(p).



Analisando os dois gráficos, vemos que um modelo ARIMA(2,1,0) será escolhido, pois a distribuição dos lags estão uniformes a partir do 2º lag, comparado à distribuição da função de autocorrelação. Tomamos assim o último lag fora da confiabilidade.

O modelo ajustado será então a partir da equação 1.1:

$$Z_t = -0,6159Z_{t-1} - 0,3587Z_{t-2} + \epsilon_t$$

Esse modelo foi ajustado utilizando o software gretl 1.9.5 cvs que obteve os parâmetros para a equação 1.1: $\alpha_1 = -0,6159$ e $\alpha_2 = -0,3587$

Após identificar a ordem e estimar eficientemente os parâmetros de um modelo, é necessário verificar sua adequação antes de utilizá-lo, por exemplo, para fazer previsões. Pode-se fazer testes de sobreajustamento que consistem em incluir parâmetros extras no modelo e verificar sua significância estatística.

Para os modelos vistos nesse estudo, o valor ajustado é a previsão um passo à frente, de modo que o resíduo bem ajustado esteja distribuído aleatoriamente em torno de zero com uma variância aproximadamente constante e que ainda não seja correlacionada.

5 CONCLUSÃO

O presente trabalho teve sua objetividade em descrever uma forma simples e suave na descrição e ajustamento de um modelo de séries temporais. Obtendo os dados a partir de pesquisa

feita pelo IBGE, podemos notar que uma simples comparação dos dados obtidos através do gráfico nos leva a uma efetividade concreta de valores futuros. Desse pressuposto, decorre a necessidade de avaliação de séries temporais através de modelos estatísticos diversos, cada um mais adequado a uma específica trajetória no tempo. Neste artigo, foi apresentada uma maneira sumária apenas dos modelos estatísticos mais utilizados em análise de séries temporais. Espera-se, com isso, mostrar as inúmeras aplicações da análise de séries temporais, estimulando-se a busca das técnicas estatísticas adequadas.

A disponibilidade de softwares permite aos pesquisadores a utilização da melhor técnica estatística, ainda que a importância técnica não deva ser esquecida para o desenvolvimento das hipóteses.

REFERÊNCIAS

BOX, G. E. P.; JENKINS, G. M. **Time Series Analysis: forecasting and control**. San Francisco: Holden-Day, 1970.

DELURGIO, S. A. **Forecasting principles and applications**. Singapore: McGraw-Hill, 1998.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Disponível em: <<http://www.ibge.gov.br/home/estatistica/indicadores/comercio/pmc/default.shtm>>. Acesso em: 02 nov. 2011.

MONTGOMERY, D. C.; JOHNSON, L. A. **Forecasting and Time Series Analysis**. New York: McGraw-Hill, 1976.

MORETTIN, P. A.; TOLOI, C. M. **Séries Temporais**. 2. ed. São Paulo: Atual, 2006.

MORETTIN, P. A. **Econometria financeira: um curso em Séries Temporais financeiras**. São Paulo: Departamento de Estatística-Instituto de Matemática e Estatística-USP, 2004.

STOCK, J. H.; WATSON, M. W. **Econometria**. São Paulo: Addison Wesley, 2004.

WHEELWRIGTH, S. C.; MAKRIDAKIS, S. **Forecasting Methods for Management**. 4. ed. New York: John Wiley & Sons Inc, 1985.