

THIAGO RAFAEL CORRÊA DE ALMEIDA

MODELOS PARA PREVISÃO DO ÍNDICE DE VOLUME TRIMESTRAL DO PIB DO
ESTADO DE MINAS GERAIS

BELO HORIZONTE

2013

THIAGO RAFAEL CORRÊA DE ALMEIDA

MODELOS PARA PREVISÃO DO ÍNDICE DE VOLUME TRIMESTRAL DO PIB DO
ESTADO DE MINAS GERAIS

Monografia de conclusão apresentada ao Curso de Especialização em Estatística da Universidade Federal de Minas Gerais (UFMG), como requisito para a obtenção do título de pós-graduado em Estatística com Ênfase em Indústria e Mercado.

Orientação: Edna Afonso Reis.

BELO HORIZONTE

2013

Thiago Rafael Corrêa de Almeida

Modelos para Previsão do Índice de Volume Trimestral do PIB
do Estado de Minas Gerais.

Monografia de conclusão apresentada ao Curso de Especialização em Estatística da Universidade Federal de Minas Gerais (UFMG), como requisito para a obtenção do título de pós-graduado em Estatística com Ênfase em Indústria e Mercado.

Aprovada na Banca Examinadora

Profa. Dra. Edna Afonso Reis (orientadora)
Departamento de Estatística
Universidade Federal de Minas Gerais (UFMG)

Profa. Dra. Ela Mercedes Medrano de Toscano (avaliadora)
Departamento de Estatística
Universidade Federal de Minas Gerais (UFMG)

Profa. Dra. Ilka Afonso Reis (avaliadora)
Departamento de Estatística
Universidade Federal de Minas Gerais (UFMG)

Belo Horizonte, 16 de dezembro de 2013

AGRADECIMENTOS

A Deus em primeiro lugar e sobre todas as coisas;

A minha orientadora Edna, pela paciência, pelo conhecimento transmitido e por ter apostado no tema deste trabalho;

Ao meu amigo e colega de trabalho Raimundo pelo aprendizado e incentivo;

Aos amigos e companheiros de trabalho que me deram a oportunidade de trabalhar no Centro de Estatística e Informações (CEI) e contribuíram para o meu aprendizado: Ronaldo, Reinaldo, Maria Aparecida, Frederico Poley, Nira, Marilene, João, Claudinéia, Mauro, Juliana, Mirela, Adriana, Raquel, Carla, Glauber, Léo, Mariana, Alexandre e muitas outras pessoas que me auxiliaram na Fundação João Pinheiro;

Aos meus amigos do Curso de Especialização em Estatística que me apoiaram durante todo o curso, em especial ao meu amigo Marccone;

A todos os professores que, de alguma maneira, contribuíram para a construção do meu conhecimento;

Aos meus pais, Antônio e Joana, por me ensinarem o respeito e a ética e por terem me ajudado em toda a minha trajetória acadêmica, muito obrigado;

Aos familiares, amigos e a minha namorada Raquel pelas palavras de incentivo e apoio.

“A melhor maneira de prever o futuro é criá-lo”
(Peter Drucker).

RESUMO

Este trabalho tem como objetivo principal a construção de modelos estatísticos de acompanhamento e previsão da série trimestral de índice de volume do Produto Interno Bruto (PIB) de Minas Gerais, tradicionalmente divulgado pela Fundação João Pinheiro. A estratégia adotada foi a de utilizar a própria série (modelos univariados) ou outras variáveis concomitantes que expliquem a variável de interesse (modelos multivariados). No primeiro caso, o propósito foi alcançado com a utilização da metodologia típica para séries temporais de Box-Jenkins (modelos ARIMA ou o caso mais geral – modelos SARIMA) e através da técnica de Alisamento Exponencial de Holt-Winters. No segundo caso, foi feita uma análise via construção de um modelo de Regressão Dinâmica. O trabalho também justifica a importância em se modelar (e tentar prever) esta variável tão complexa e tenta preencher uma lacuna no que se refere ao comportamento da economia mineira. Finalmente, ele apresenta as principais conclusões identificando qual dos modelos obtidos teve melhor desempenho no período amostral e no período de validação.

Palavras-chave: Produto Interno Bruto (PIB), Economia de Minas Gerais, Séries temporais e modelos de Box-Jenkins, Alisamento Exponencial, Regressão Dinâmica.

ABSTRACT

This monograph mainly aims at the construction of statistical models for monitoring and forecasting the quarterly series of volume index of gross domestic product (GDP) of Minas Gerais, traditionally published by Fundação João Pinheiro. The strategy adopted was to use its own series (univariate) or other concomitant variables that explain the variable of interest (multivariate models). In the first case, the purpose has been achieved with the use of typical approach for time-series Box-Jenkins (ARIMA models or the more general case - SARIMA models) and through exponential smoothing technique Holt-Winters. In the second case, an analysis was done via the construction of a dynamic regression model. The monograph also justifies the importance of modeling (and try to predict) such a complex and variable and attempts to fill a gap in relation to the behavior of the State economy. Finally, it presents the main conclusions obtained identifying which model performed best in the sample period and the validation period.

Keywords: Gross Domestic Product (GDP), Economy of Minas Gerais, Time Series and Box-Jenkins models, Exponential Smoothing, Dynamic Regression.

LISTA DE ILUSTRAÇÕES

Gráfico 1 – Gráfico de dispersão dos resultados trimestrais da série de base fixa (média 2002 = 100) em volume da arrecadação de Impostos e do PIB de Minas Gerais – 1º Trimestre de 2002-2º Trimestre de 2013.....	26
Gráfico 2 – Função de Autocorrelação (FAC) de uma série qualquer (Y(t)) com a representação dos limites com 5% de significância.....	37
Gráfico 3 – Função de Autocorrelação Parcial (FACP) de uma série qualquer (Y(t)) com a representação dos limites com 5% de significância.....	38
Figura 1 – Simulação de um processo Ruído Branco Gaussiano.....	39
Figura 2 – Ciclo Iterativo sugerido por Box-Jenkins para identificação e escolha do modelo ARIMA (p, d, q) apropriado.....	41
Figura 3 – Correlogramas representativos da FAC e FACP dos processos AR (1).....	43
Figura 4 – Correlogramas representativos da FAC e FACP dos processos AR (2).....	44
Figura 5 – Correlogramas representativos da FAC e FACP dos processos MA (1).....	45
Figura 6 – Correlogramas representativos da FAC e FACP dos processos MA (2).....	46
Figura 7 – Características da série em relação à sazonalidade, nível e tendência e o método de alisamento exponencial recomendado.....	52
Figura 8 – Construção de um modelo de Regressão Dinâmica.....	58
Gráfico 4 – Comportamento via gráfico de linhas da série Índice de Volume do PIB de Minas Gerais – 1º Trimestre de 2002-4º Trimestre de 2011.....	62
Gráfico 5 – Boxplot do Índice de Volume do PIB de Minas Gerais – 1º Trimestre de 2002-4º Trimestre de 2011.....	63
Gráfico 6 – Histograma do Índice de Volume do PIB de Minas Gerais – 1º Trimestre de 2002-4º Trimestre de 2011.....	64
Gráfico 7 – Histograma do logaritmo natural do Índice de Volume do PIB de Minas Gerais – 1º Trimestre de 2002-4º Trimestre de 2011.....	65
Gráfico 8 – Função de Autocorrelação (FAC) e Função de Autocorrelação Parcial (FACP) da série de índice de volume do PIB de Minas Gerais e após a aplicação de uma diferença não sazonal na série.....	66
Gráfico 9 – Série de Índice de Volume do PIB de Minas Gerais (Y _t) e Valores ajustados pelo modelo SARIMA (0, 1, 1)(1, 1, 0) ₄ na série original (dentro da amostra).....	68
Figura 9 – Resultados da simulação do modelo SARIMA (0, 1, 1)(1, 1, 0) ₄ na série original (dentro da amostra) para o comportamento do vetor de resíduos.....	69

Gráfico 10 – Resíduos regulares e padronizados versus valores ajustados pelo modelo SARIMA (0, 1, 1)(1, 1, 0) ₄ na série original (dentro da amostra).....	70
Gráfico 11 – Previsões do modelo SARIMA (0, 1, 1)(1, 1, 0) ₄ na série original seis passos a frente e medidas de desempenho.....	71
Gráfico 12 – Série de Índice de Volume do PIB de Minas Gerais (Y _t) e Valores ajustados pelo modelo SARIMA (2, 0, 0)(0, 1, 1) ₄ na série original (dentro da amostra).....	72
Figura 10 – Resultados da simulação do modelo SARIMA (2, 0, 0)(0, 1, 1) ₄ na série original (dentro da amostra) para o comportamento do vetor de resíduos.....	73
Gráfico 13 – Resíduos regulares e padronizados versus valores ajustados pelo modelo SARIMA (2, 0, 0)(0, 1, 1) ₄ na série original (dentro da amostra).....	74
Gráfico 14 – Previsões do modelo SARIMA (2, 0, 0)(0, 1, 1) ₄ na série original seis passos a frente e medidas de desempenho.....	75
Gráfico 15 – Série de Índice de Volume do PIB de Minas Gerais (Y _t) e Valores ajustados pelo modelo de Alisamento Exponencial Multiplicativo (α = 0,9; γ = 0,2; δ = 0,2) e medidas de desempenho dentro da amostra.....	76
Figura 11 – Resultados da simulação do modelo de Alisamento Exponencial Multiplicativo (α = 0,9; γ = 0,2; δ = 0,2) para o comportamento do vetor de resíduos (dentro da amostra).....	77
Gráfico 16 – Resíduos regulares e padronizados versus valores ajustados pelo modelo de Alisamento Exponencial Multiplicativo (α = 0,9; γ = 0,2; δ = 0,2) dentro da amostra.....	78
Gráfico 17 – Previsões do modelo de Alisamento Exponencial Multiplicativo (α = 0,9; γ = 0,2; δ = 0,2) seis passos a frente e medidas de desempenho.....	78
Gráfico 18 – Série de Índice de Volume do PIB de Minas Gerais (Y _t) e Valores ajustados pelo modelo de Alisamento Exponencial Aditivo (α = 0,9; γ = 0,2; δ = 0,2) e medidas de desempenho dentro da amostra.....	80
Gráfico 19 – Resíduos regulares e padronizados versus valores ajustados pelo modelo de Alisamento Exponencial Aditivo (α = 0,9; γ = 0,2; δ = 0,2) dentro da amostra.....	80
Figura 12 – Resultados da simulação do modelo de Alisamento Exponencial Aditivo (α = 0,9; γ = 0,2; δ = 0,2) para o comportamento do vetor de resíduos (dentro da amostra).....	81
Gráfico 20 – Previsões do modelo de Alisamento Exponencial Aditivo (α = 0,9; γ = 0,2; δ = 0,2) seis passos a frente e medidas de desempenho.....	82
Gráfico 21 – Comportamento via gráfico de linhas do Valor Adicionado da Agropecuária de Minas Gerais e da variável proxy (A _t) utilizada para representar o mesmo – média de 2002 = 100 – 1º Trimestre de 2002-4º Trimestre de 2011.....	84
Gráfico 22 – Comportamento via gráfico de linhas do Valor Adicionado da Indústria de Minas Gerais e da variável proxy (I _t) utilizada para representar o mesmo – média de 2002 = 100 – 1º Trimestre de 2002-4º Trimestre de 2011.....	86

Gráfico 23 – Comportamento via gráfico de linhas do Valor Adicionado do setor de Serviços de Minas Gerais e da variável proxy (S_t) utilizada para representar o mesmo – média de 2002 = 100 – 1º Trimestre de 2002-4º Trimestre de 2011.....	87
Gráfico 24 – Série de Índice de Volume do PIB de Minas Gerais (Y_t) e Valores ajustados pelo modelo de Regressão Dinâmica na série original (dentro da amostra).....	90
Figura 13 – Resultados da simulação do modelo de regressão dinâmica (na amostra) para o comportamento do vetor de resíduos (ϵ_t).....	91
Gráfico 25 – Resíduos regulares e padronizados versus valores ajustados pelo modelo de Regressão Dinâmica (dentro da amostra).....	92
Gráfico 26 – Previsões do modelo de Regressão Dinâmica seis passos a frente e medidas de desempenho.....	94

LISTA DE QUADROS E TABELAS

Quadro 1 – Valores do parâmetro (λ) e Transformação de Box-Cox na série original correspondente ao parâmetro.....	35
Quadro 2 – Pressupostos do Modelo de Regressão Linear.....	54
Quadro 3 – Critérios de desempenho dos modelos através dos erros associados – MAD, MAPE e MSD.....	60
Quadro 4 – Medidas de desempenho (MAD, MAPE e MSD) dos cinco modelos construídos tanto no período amostral quanto no período de validação.....	96
Quadro 5 – Taxas de crescimento real (%) da economia mineira segundo o desempenho dos dois melhores modelos no período de validação em diferentes períodos analisados comparativamente a divulgada pela Fundação João Pinheiro.....	98

SUMÁRIO

CAPÍTULO 1: INTRODUÇÃO.....	13
CAPÍTULO 2: A RELEVÂNCIA DO ÍNDICE DE VOLUME DO PIB.....	18
CAPÍTULO 3: FUNDAMENTAÇÃO TEÓRICA.....	33
3.1) Aspectos gerais da análise de séries Temporais.....	33
3.2) Função de Autocorrelação (FAC) e Função de Autocorrelação parcial (FACP).....	36
3.3) O conceito de “Ruído Branco”.....	39
3.4) Utilização de Operadores.....	40
3.5) Modelos de Box-Jenkins.....	40
3.5.1) Modelos Autorregressivos de Ordem p – AR (p).....	42
3.5.1.1) Modelos Autorregressivos de Ordem 1 – AR (1).....	42
3.5.1.2) Modelos Autorregressivos de Ordem 2 – AR (2).....	43
3.5.2) Modelos de Médias Móveis de Ordem q – MA (q).....	44
3.5.2.1) Modelos de Médias Móveis de Ordem 1 – MA (1).....	44
3.5.2.2) Modelos de Médias Móveis de Ordem 2 – MA (2).....	45
3.5.3) Modelos Autorregressivos e de Médias Móveis de Ordem (p,q) – ARMA (p,q).....	46
3.5.4) Modelos Autorregressivos Integrados e de Médias Móveis de Ordem (p, d, q) – ARIMA (p, d, q).....	47
3.6) Modelo ARIMA Sazonal – SARIMA (p, d, q) (P, D, Q)_s.....	48
3.7) Escolha e verificação dos modelos construídos.....	48
3.7.1) Verificação dos modelos.....	48
3.7.2) Escolha entre modelos.....	51
3.8) Alisamento Exponencial de Holt-Winters.....	52
3.9) Regressão Dinâmica.....	54
3.10) Comparação do desempenho dos modelos construídos.....	60
CAPÍTULO 4: CARACTERÍSTICAS DA SÉRIE DE ÍNDICE DE VOLUME DO PIB DE MINAS GERAIS.....	61
CAPÍTULO 5: APRESENTAÇÃO DOS MODELOS CONSTRUIDOS.....	67
5.1) Modelos Univariados.....	67
5.1.1) Modelos SARIMA.....	67
5.1.1.1) Modelo SARIMA (0, 1, 1)(1, 1, 0) ₄ da série Y _t	67
5.1.1.2) Modelo SARIMA (2, 0, 0)(0, 1, 1) ₄ da série Y _t	71
5.1.2) Modelos de Alisamento Exponencial de Holt-Winters.....	75
5.1.2.1) Alisamento Exponencial Multiplicativo (α = 0,9; γ = 0,2; δ = 0,2).....	76
5.1.2.2) Alisamento Exponencial Aditivo (α = 0,9; γ = 0,2; δ = 0,2).....	79
5.2) Modelo Multivariado.....	82

5.2.1) Escolha das variáveis explicativas.....	82
5.2.1.1) Proxy para o comportamento do VA da Agropecuária (A _t).....	83
5.2.1.2) Proxy para o comportamento do VA da Indústria (I _t).....	85
5.2.1.3) Proxy para o comportamento do VA de Serviços (S _t).....	86
5.2.2) O modelo de Regressão Dinâmica.....	88
CAPÍTULO 6: CONCLUSÃO.....	96
REFERÊNCIAS.....	100
APÊNDICE.....	103

CAPÍTULO 1: INTRODUÇÃO

Este trabalho tem como propósito a construção de um ou mais modelos estatísticos de acompanhamento e previsão no curto prazo da série encadeada de volume do Produto Interno Bruto (PIB) de Minas Gerais. A ideia é utilizar a própria série (modelos univariados) ou outras variáveis intrinsecamente relacionadas à variável de interesse (modelos multivariados) para alcançar este objetivo.

Outra finalidade desse trabalho é demonstrar a dificuldade de construção de modelos que sejam satisfatórios, sobretudo os univariados, visto que a série encadeada de volume do PIB é bastante complexa e depende ou se relaciona, direta ou indiretamente, com uma infinidade de fatores: política macroeconômica do governo federal; conjuntura econômica interna e externa (por exemplo, em situações de crise econômica espera-se uma diminuição no nível de atividade econômica); grau de endividamento das famílias e do comprometimento da renda; nível de investimentos e de utilização da capacidade instalada por parte dos empresários; nível de inflação e do comportamento dos preços; taxa de câmbio; taxa de juros; estabilidade do sistema financeiro e do mercado de crédito; etc.

Portanto, é um desafio muito grande prever e acompanhar uma série que depende de um conjunto de fatores com comportamentos tão difusos. Essa dificuldade fica evidenciada nos modelos univariados, normalmente construídos com base no passado da série original – seja através da metodologia das séries temporais ou através da técnica de alisamento exponencial. O principal problema aqui envolvido é que o passado da série, embora contenha um conjunto de “informações” importantes deste arcabouço de fatores difusos, ele sozinho não é capaz de lidar de forma eficiente com situações extremas típicas de séries econômicas. Por exemplo, o passado de uma série como a do PIB pode apresentar um comportamento padrão aparentemente “conhecido”, mas que pouco servirá para a realização de uma previsão futura quando já se sabe da existência de um período de turbulência no ciclo econômico vigente (como acontece nos períodos de crise econômica). O raciocínio é basicamente o seguinte: como prever uma queda brusca no nível de atividade econômica se o passado da série não traz nenhuma informação a respeito desta inflexão?

Possivelmente, espera-se que os modelos univariados se mostrem mais ineficientes na tentativa de fornecer uma resposta adequada para esta pergunta.

Mesmo assim, será dispendido um esforço de construção de modelos univariados, que descrevam com alguma “qualidade” estatística a série encadeada de volume do PIB de Minas Gerais, com todas as limitações que rodeiam esse processo de elaboração.

Procurando identificar o melhor modelo de acompanhamento e previsão da série, também se buscou a construção de algum tipo de modelo multivariado. O raciocínio por trás deste procedimento foi buscar soluções para os problemas típicos dos modelos autoprojativos. A estratégia foi utilizar não apenas o passado da série de interesse, mas os valores do passado e do presente de variáveis relacionadas à série de volume do PIB de Minas Gerais através dos modelos de regressão dinâmica. De fato, nesses modelos, a variável dependente pode ser explicada tanto pelos seus valores defasados quanto pelos valores atuais e passados de variáveis causais ou exógenas (ZANINI, 2000).

Portanto, as previsões obtidas a partir de um modelo de regressão dinâmica dependem não apenas dos valores passados da série, mas também dos valores previstos para as variáveis causais. Assim, faz-se necessário fornecer ao modelo os valores “futuros” do vetor de variáveis causais, pois se as previsões destas variáveis exógenas não forem adequadas, o modelo de regressão dinâmica também irá gerar previsões inapropriadas (ZANINI, 2000). Todavia, é interessante observar que o conjunto de variáveis que se pretende utilizar como independentes no modelo de acompanhamento do PIB de Minas Gerais são variáveis que antecedem o nível e o resultado da atividade econômica, isto é, estão disponibilizadas antes do resultado oficial do volume encadeado do PIB. Assim, com a obtenção das informações das variáveis causais no tempo oportuno, não será necessária a realização de previsões para o vetor de variáveis explicativas.

Apesar da tempestividade das variáveis que se pretende utilizar na modelagem como intervenientes (explicativas) do nível de atividade econômica, algumas modificações foram necessárias nas séries originais das variáveis explicativas. A justificativa para estas alterações está no fato de que a variável dependente (a série encadeada de volume do PIB de Minas Gerais) é divulgada trimestralmente, ao passo que as variáveis explicativas que se pretende utilizar na construção do modelo são divulgadas mensalmente ou anualmente. O tratamento destas adaptações e da escolha das variáveis explicativas será descrito em um capítulo próprio com os aspectos metodológicos.

Em consonância com a discussão da utilização de modelos estatísticos, (univariados ou multivariados) para acompanhamento e previsão de séries

econômicas como as do PIB, é importante salientar que existe certa divergência teórica entre economistas clássicos e adeptos do keynesianismo a respeito da percepção de previsão do “futuro” com base no presente e no passado de séries econômicas via métodos estatísticos. Para os clássicos, um resultado futuro pode, em algum grau de medida, ser previsível e calculado probabilisticamente a partir dos dados de mercado do presente e do passado. A visão keynesiana caminha em sentido contrário: devido à existência de um universo incerto em que as decisões dos agentes econômicos são ambíguas, o comportamento da atividade econômica não poderia ser previsto de forma confiável através da modelagem estatística. A esse arcabouço de expectativas incertas e imprevisíveis, no curto e no longo prazo, foi o que Keynes denominou de “animal spirits” (DAVIDSON, 1999).

Levando-se em conta esta discussão de cunho ideológico, esta monografia converge para uma análise que, de certa forma, compreende os dois pontos de vista do pensamento econômico. Ela concorda com o raciocínio keynesiano, na medida em que qualquer modelagem estatística de uma série como a do PIB de Minas Gerais, será uma simplificação da realidade e como tal estará sujeita a críticas e a flutuações “inesperadas” diante da complexidade dos fatores envolvidos. Por outro lado, ela caminha para uma solução através do pensamento clássico, já que o processo de modelagem estatística tende a examinar apenas os aspectos considerados “essenciais” do problema, o que por si só já fornece um conjunto de informações importantes para o entendimento da série, mesmo que essa compreensão seja limitada e parcial.

Portanto, este capítulo introdutório tem o objetivo de trazer a ideia geral do trabalho, que é a de construir um ou mais modelos univariados (autoprojetivos), isto é, modelos que não fazem uso de outras variáveis que não a própria série de interesse (a série encadeada de volume do PIB de Minas Gerais) e, a partir daí, verificar as limitações desses modelos inclusive no que se refere à utilização dos mesmos para fazer projeções. Posteriormente, pretende-se elaborar um modelo de regressão dinâmica com o intuito de explicar a série do PIB através de outras variáveis e, com isso, verificar se a inserção das “novas causalidades” melhoraram a performance das previsões.

Evidentemente que, para a consecução desses objetivos, será necessário a utilização de algum sistema computacional voltado para o cálculo estatístico. Segundo Cordeiro Júnior (2007), existe uma variedade de softwares estatísticos que trazem em seus pacotes funções relacionadas à modelagem estatística: SAS, Eviews, Stata,

Minitab, R, SACS, etc. são apenas alguns destes programas. O autor argumenta que critérios ligados ao preço do software (quando o mesmo não é gratuito), a portabilidade, a facilidade de uso e de instalação e a disponibilidade de manuais são alguns dos fatores utilizados no processo de escolha por parte dos pesquisadores. O programa adotado no presente trabalho foi o Minitab por ser o software estatístico mais utilizado no curso de Especialização em Estatística na Universidade Federal de Minas Gerais (UFMG), além da sua facilidade de manuseio e interface com o usuário.

Apenas como forma de contextualizar a estrutura da monografia para além desta primeira parte introdutória, onde foram discutidos os objetivos e os mecanismos para alcançá-los, o próximo capítulo justifica a realização desta monografia respondendo a seguinte pergunta: afinal, por que é tão importante buscar mecanismos para acompanhar e projetar o resultado da série encadeada de volume do PIB? O capítulo fornecerá a resposta a essa pergunta dando a definição de Produto Interno Bruto, explicando como ele é calculado e mostrando a sua interface com os diversos agentes econômicos interessados no seu resultado.

Por sua vez, o Capítulo 3 apresenta uma fundamentação teórica através de uma revisão bibliográfica dos assuntos atinentes para concretização dos propósitos do trabalho. Nele serão discutidos os fundamentos que permeiam os conceitos, as definições e todo o arcabouço que norteia a utilização da metodologia das séries temporais, da técnica de alisamento exponencial e dos modelos de regressão dinâmica.

O Capítulo 4 apresenta de maneira sucinta as características da série encadeada de volume do PIB de Minas Gerais. Além das peculiaridades mais específicas da série, como o período em que ela se inicia e o número de observações, que serão de suma importância para o entendimento dos modelos construídos, o capítulo traz também uma análise exploratória dos dados em que se identifica ou não a presença de variações sazonais, de tendência ao longo do tempo, de estacionariedade, de outliers, etc. Assim, para além da descrição dessas características, busca-se, na medida do possível, interpretar e encontrar explicações razoáveis que auxiliem no entendimento da série, de suas peculiaridades e, conseqüentemente, dos modelos propostos.

Finalmente, o Capítulo 5 traz a apresentação dos modelos construídos tanto via séries temporais e alisamento exponencial quanto através do modelo de regressão dinâmica proposto. Ademais, nesse capítulo são discutidos os aspectos metodológicos e as adaptações efetuadas nas variáveis causais escolhidas para construção do

modelo multivariado. A justificativa para a escolha das variáveis independentes para explicar a série encadeada de volume do PIB de Minas Gerais também é demonstrada neste capítulo.

O último capítulo apresenta as principais conclusões a respeito da realização deste trabalho. Nele pretende-se comparar os modelos construídos e apontar o “melhor” modelo estatístico para acompanhar a série no período amostral (passado da série) e, principalmente, para realizar as previsões de curto prazo (período de validação), sempre levando em consideração as limitações da modelagem estatística proposta (dado o tamanho da série) e reconhecendo a dificuldade em se modelar séries econômicas como as do PIB.

CAPÍTULO 2: A RELEVÂNCIA DO ÍNDICE DE VOLUME DO PIB

Este capítulo tem o intuito de justificar a realização desta monografia. Para a concretização deste propósito, serão discutidos os fatores que motivaram a escolha do tema e da problemática principal que circunda o assunto envolvido, que é a de acompanhar e projetar o resultado do índice de volume do PIB para tomada de decisão dos agentes econômicos interessados. Todavia, para que a relevância do PIB seja mais bem compreendida, convém entender melhor o conceito deste agregado econômico, o que ele aborda, como é calculado e quais são suas limitações.

O Produto Interno Bruto, ou PIB, corresponde ao valor total de todos os bens e serviços finais produzidos em uma região durante um intervalo de tempo (normalmente um ano). Nesse ponto, é importante entender a distinção entre bens e serviços finais dos chamados bens intermediários. Um bem intermediário é aquele produzido por uma empresa e vendido à outra firma que vai utilizá-lo na produção de outro bem. Já os bens e serviços finais são aqueles vendidos aos usuários da “ponta” do processo produtivo e que não mais serão transformados. A não inclusão dos bens intermediários no cálculo do PIB ocorre pelo simples mecanismo de se evitar dupla contagem no cômputo da apuração dos bens produzidos. De fato, se os bens intermediários fossem incluídos diretamente no PIB, seriam contabilizados duas vezes, por que fazem parte do produto final em que foram utilizados (CLARK; DOWNING, 1998).

Tradicionalmente, o PIB pode ser calculado em relação a três óticas ou pontos de vistas diferenciados: a ótica da produção, da demanda e da renda. Na ótica da produção, o raciocínio de cálculo é direto: basta pesquisar todas as firmas e empresas envolvidas e descobrir o valor da produção de bens e serviços finais, excluindo o valor dos bens e serviços intermediários (KRUGMAN; WELLS, 2007).

A maneira mais tradicional de evitar o processo de dupla contagem via ótica da produção é através do conceito de valor adicionado (ou valor agregado), que corresponde à diferença do valor das vendas dos produtos e o valor da compra de insumos de outras empresas. Em outras palavras, o Valor Adicionado (VA) traduz a diferença entre o Valor Bruto de Produção (VBP) de cada empresa (ou setor) e o Consumo Intermediário (CI) ali envolvido, isto é, $VA = VBP - CI$ (IBGE, 2008). O somatório dos valores adicionados dos diferentes setores presentes na economia

corresponde exatamente ao PIB avaliado a preços básicos ou a custos dos fatores¹ (FUNDAÇÃO JOÃO PINHEIRO, 2010). A equação abaixo resume esta afirmativa:

$$\text{PIB (preços básicos)} = \sum (\text{VA setoriais})$$

Essa primeira forma de calcular o PIB será de suma importância para o entendimento do modelo de regressão dinâmica que será construído mais adiante, visto que a modelagem proposta será embasada na análise do comportamento de algumas variáveis que representem, com algum grau de fidedignidade, os valores adicionados setoriais.

A segunda maneira de calcular o PIB é através da ótica da demanda (ou da despesa). O raciocínio por trás desse ponto de vista é que toda a produção gerada tem uma destinação econômica. Nesse sentido, o gasto agregado em bens e serviços finais produzidos internamente corresponderia à soma dos gastos dos consumidores – ou do Consumo das famílias (C) –, com os gastos relativos à Formação Bruta de Capital Fixo – gastos de Investimento pelas empresas (I), com as compras governamentais de bens e serviços ou gastos do Governo (G) e da diferença entre as exportações (X) e as importações (IM). Assim, essa segunda forma de mensurar o PIB simplesmente calcula o gasto agregado em bens e serviços finais internamente produzidos na economia (KRUGMAN; WELLS, 2007). Todavia, como as importações não correspondem a bens e serviços produzidos em âmbito interno, elas são subtraídas da equação de estimação do PIB pela ótica da demanda. Constata-se ainda que, nesse caso, o PIB é mensurado a preços de mercado, pois representa o valor efetivamente pago pelos agentes econômicos na aquisição dos bens e serviços finais transacionados.

$$\text{PIB (preços de mercado)} = C + I + G + X - \text{IM}$$

A terceira forma de mensurar o PIB é por meio da ótica da renda. A ideia por trás desse método de cálculo é que toda venda realizada na economia precisa representar uma renda para alguém independentemente da forma como a mesma é auferida. Assim, os salários ganhos pelo trabalho, os aluguéis recebidos por aqueles que alugam suas casas, os juros ganhos por aqueles que emprestam suas poupanças para as empresas e ao governo, e os lucros recebidos pelos proprietários do capital

¹ O conceito de PIB avaliado a preços básicos ou a custos dos fatores difere do conceito de PIB avaliado a preços de mercado. O conceito de PIB a preços básicos (custos dos fatores) corresponde ao produto agregado avaliado na “porta da fábrica” ou a preços ao produtor. Já o PIB a preços de mercado leva em consideração a incidência dos impostos líquidos de subsídios sobre a produção. Por isso, o PIB a preços de mercado também é conhecido como PIB a preços ao consumidor, pois neste caso o produto agregado está sendo mensurado pelo valor efetivamente pago (despesa) dos agentes econômicos.

físico das firmas representam o somatório total das rendas ganhas pelos fatores de produção e oriundas das empresas presentes na economia (KRUGMAN; WELLS, 2007).

Tradicionalmente, costuma-se separar na ótica da renda a remuneração do fator capital do fator trabalho. De fato, o fator trabalho é mensurado pela remuneração dos empregados incluindo os funcionários não residentes no país (W) e pelo rendimento misto – ou rendimento dos autônomos (WA), enquanto o fator capital é representado pelos lucros, juros e aluguéis auferidos pelos proprietários das empresas – também chamados de excedente operacional bruto (EOB) (ARAÚJO NETO; COSTA, 2005). Considerando a incidência dos impostos sobre a produção e a importação (T) líquidos de subsídios (S), tem-se a seguinte equação para o cálculo do PIB pela ótica da renda e a preços de mercado:

$$\text{PIB (preços de mercado)} = W + WA + \text{EOB} + (T - S)$$

Na prática, o cálculo do PIB através das três óticas tem a sua utilização difundida como medida do tamanho de determinada economia. Além disso, ele fornece uma escala para mensurar o desempenho econômico no tempo (comparações entre anos) e no espaço (comparação entre regiões). Todavia, o PIB ainda é um indicador que apresenta algumas inconsistências no que tange a mensuração do desempenho e do desenvolvimento econômico de uma região.

Essas inconsistências ocorrem por que o cálculo do PIB não tem a pretensão de medir o bem-estar social, muito embora o seu comportamento possa dar pistas do grau de desenvolvimento econômico e social de uma região. De fato, o PIB não leva em consideração (ou subestima) os serviços executados pelas donas de casa e as atividades de produção voltadas para o autoconsumo ou atividades “extramercado” (CLARK; DOWNING, 1998). Esse mesmo problema ocorre em relação à produção de bens e serviços gerados de forma ilegal, isto é, a presença de estabelecimentos não registrados que produzem algum tipo de produto ou serviço e que, além de não registrarem seus funcionários, também não pagam impostos.²

Além disso, o PIB poderá expandir sem que isso necessariamente signifique melhora no padrão de vida desfrutado pela população. O crescimento desenfreado

² Ressalta-se aqui a presença também de uma quantidade significativa de atividades transacionadas de forma ilegal e que funcionam na chamada economia “subterrânea” ou “invisível”, formada por diversas operações irregulares como: jogo do bicho, cassinos clandestinos, bingos, comércio de drogas e de animais silvestres, contrabando, prostituição, etc. Esse conjunto de atividades gera “renda” de difícil quantificação e que não é desejável do ponto de vista do bem-estar social dado os efeitos danosos que circundam esse conjunto de atividades.

das indústrias pode vir acompanhado do aumento da emissão de poluentes no meio ambiente e, com isso, faz-se necessário a ampliação na produção de serviços e mercadorias para combater os efeitos danosos da poluição. Tem-se assim um exemplo típico de expansão na taxa do PIB e, ao mesmo tempo, de decréscimo na qualidade de vida. O aumento da produção agrícola através da agricultura expansiva é outro exemplo notório, já que o desmatamento resultante da atividade gera impactos negativos à natureza e no padrão de vida do homem, que não são mensurados pelo indicador oficial (SILVA, 2011).

Essa limitação relacionada ao conceito de PIB acontece por que esse agregado macroeconômico é uma medida quantitativa de desempenho econômico. De fato, informações mais qualitativas a respeito da realidade econômica e social de uma região não devem ser buscadas estritamente através deste indicador. Por isso, muitas das limitações que circundam essa variável econômica vêm contribuindo para o surgimento de propostas alternativas para a metodologia de mensuração deste indicador, na tentativa de incorporar variáveis qualitativas que reflitam o padrão de vida social. Recentemente, o governo chinês expressou essa preocupação em se buscar um crescimento do PIB mais sustentável com a construção de algum modelo que possa considerar aspectos mais qualitativos:

“ [...] a intenção de calcular um “Índice de qualidade do PIB”, por meio da metodologia desenvolvida pelo economista Niu Wenyuan que é assessor do Conselho de Estado da China, e inclusive já havia criado um índice para o PIB “Verde”, ganha força. O índice é composto das seguintes variáveis: qualidade econômica [que estabelece uma relação inversa entre o gasto de energia e unidade monetária do PIB]; a qualidade social [que tem como foco a distribuição de renda]; a qualidade ambiental; a qualidade de vida; qualidade de serviços e qualidade de gestão [que mede a eficiência do setor público]” (SILVA, 2011).

Assim, a mensuração do PIB com um viés qualitativo e que contabilize possíveis efeitos nocivos da expansão do nível de atividade econômica sobre a biodiversidade, é uma exigência para o futuro, mas que ainda permanece como um “sonho” distante nas metodologias tradicionais de cálculo deste agregado macroeconômico, muito embora a preocupação com a sustentabilidade venha ganhando força nos últimos anos. De outra forma, o crescimento econômico desenfreado no presente através da exploração excessiva dos recursos naturais poderá representar em enormes prejuízos para gerações futuras (sobretudo se ocorrer o esgotamento dessas reservas naturais).

Para além dessas limitações que o indicador que mensura o nível de atividade econômica não consegue captar de forma efetiva, existem outros cuidados

necessários acerca da utilização correta das informações do PIB, sobretudo quando a mesma é utilizada para fazer comparações ao longo do tempo. Krugman e Wells (2007) argumentam que uma parcela do aumento do valor do PIB no tempo representa apenas aumentos decorrentes da elevação nos preços de bens e serviços produzidos. Para medir a efetiva mudança na quantidade de bens e serviços gerados faz-se necessário a utilização de outra variável conhecida como PIB real. Essa variável representa “o valor total dos bens e serviços finais produzidos na economia durante um ano, calculada como se os preços tivessem permanecido constantes ao nível de algum dado ano base” (KRUGMAN; WELLS, 2007, p. 494). Assim, o PIB real difere do PIB nominal (ou a preços correntes) no sentido de que o primeiro capta apenas o incremento na quantidade produzida, ao passo que o segundo capta tanto a elevação em volume quanto a elevação no nível dos preços.³

A série encadeada de volume do Produto Interno Bruto (PIB) de Minas Gerais, foco da realização desta monografia, é uma série de PIB real do estado. Isto significa dizer que ela mensura apenas o incremento de quantidade nos bens e serviços finais produzidos ou, em outras palavras, que ela capta o “verdadeiro” crescimento econômico da economia mineira ao longo do tempo.

Outro cuidado importante em relação ao uso das informações do PIB deve ocorrer quando os dados são utilizados para comparações entre regiões. Em relação a esse aspecto, Clark e Downing (1998) argumentam que se o interesse for à quantidade de bens e serviços médios à disposição de cada habitante deve-se, indubitavelmente, apelar para a informação contida no PIB per capita real ao invés do PIB real total. Na prática, a divisão pelo número de habitantes elimina o efeito de diferenças entre as regiões em virtude do tamanho populacional, já que um país com população maior tem um PIB maior simplesmente pelo fato de existir mais pessoas trabalhando. Todavia, a informação contida no PIB per capita real contém as mesmas limitações destacadas anteriormente para o PIB total. Além de não computar às atividades voltadas para o autoconsumo e a ausência de informações acerca dos aspectos mais qualitativos daquilo que é produzido, o PIB per capita real representa um potencial expansionista no padrão de vida que não necessariamente se confirmará. O raciocínio é basicamente o seguinte:

“[...] um aumento no PIB real per capita significa uma expansão da fronteira das possibilidades de produção da economia. Se a economia aumentou sua capacidade produtiva, há mais coisas que a sociedade pode alcançar. Mas

³ A elevação no nível dos preços no universo da metodologia de cálculo do PIB é informada pelo Deflator implícito do PIB, que é obtido dividindo-se o PIB nominal (PIB a preços correntes) pelo PIB Real de determinado ano.

saber se a sociedade de fato faz bom uso desse potencial expandido para melhorar seu padrão de vida é uma outra questão. Em palavras um pouco diferentes: sua renda pode ser mais alta este ano do que foi no ano passado, mas se você de fato usa sua renda mais alta para melhorar sua qualidade de vida é uma escolha sua. [Além disso, existe o problema da concentração da renda gerada nas mãos de poucas pessoas, o que significa dizer que esse potencial de melhora está restrito a um número limitado de pessoas]” (KRUGMAN; WELLS, 2007, p. 495-496).

Outra limitação relacionada ao PIB diz respeito à mensuração do valor agregado pelas atividades da administração pública. De fato, os serviços prestados pelo governo a população são, de maneira geral, avaliados pelo seu custo de produção. Isto acontece porque a maior parte dos serviços realizados pelo governo não são transacionados no mercado (vendidos) como ocorre com os serviços dos demais setores produtivos e, com isso, é difícil atribuir-lhes valores (CLARK; DOWNING, 1998). O PIB também não capta, como mencionado anteriormente, alguns aspectos fundamentais para a obtenção de um padrão de bem-estar social mais elevado – como a capacidade de distribuir a renda gerada, o valor do tempo de lazer e descanso ou mesmo a questão da sustentabilidade vinculada à exploração dos recursos naturais – muito embora ele possa fornecer indícios positivos de elevação potencial no padrão de vida – como quando acontece a expansão do PIB per capita real mencionada na citação anterior.

Apesar de todas as limitações que envolvem o conceito do PIB, ainda assim ele permanece como um indicador importante que influencia as diversas decisões tomadas pelos agentes econômicos. Todavia, todas as inconsistências que envolvem o cálculo do produto agregado e que foram mencionadas nos parágrafos anteriores deixam algumas perguntas no ar: por que os agentes econômicos continuam interessados no índice de volume do PIB? Por que as atenções se voltam para o resultado da taxa de crescimento do PIB real? Qual a expectativa de crescimento econômico⁴ para o período subsequente?

A resposta a esse conjunto de perguntas justifica o propósito dessa monografia, que é encontrar modelos estatísticos de acompanhamento e previsão da série de índice de volume do PIB de Minas Gerais. Entretanto, é importante lembrar que a resposta a essas indagações já foi iniciada nos parágrafos anteriores. De fato, os diversos agentes econômicos estão interessados na taxa de crescimento real da

⁴ O resultado do crescimento econômico de determinada região constantemente aparece na mídia e nos noticiários e o seu significado muitas vezes não é totalmente compreendido. A expressão “crescimento econômico” é equivalente ao resultado da expansão que ocorre no PIB real, ou seja, corresponde ao crescimento em volume (ou na quantidade) dos bens e serviços finais produzidos de um período para o outro supondo a não alteração do nível de preços.

economia porque, em primeiro lugar, isto representará em uma expansão do PIB per capita real e, conseqüentemente, de uma melhora “potencial” no padrão de vida.

Em segundo lugar, o resultado da expansão do índice de volume do PIB carrega em si mesmo um conjunto de informações que expressa em um único número o desempenho da economia de determinada região em um período do ano. A divulgação do resultado faz com que especialistas analisem o número obtido e emitam opiniões no sentido de tentar entender a realidade econômica vivenciada no momento. Portanto, a importância deste indicador é que ele é um índice resumo do nível de atividade econômica e o seu resultado afeta, de maneira indiscriminada, diversos atores e instituições da sociedade. A citação abaixo confirma essa afirmação:

“A utilidade desse indicador macroeconômico não se restringe apenas ao âmbito do setor público, onde o mesmo é demandado regularmente como referencial na formulação e acompanhamento dos planos e programas governamentais e/ou previsão de efeitos de políticas econômicas globais e setoriais. O PIB é também amplamente utilizado por entidades privadas, estudiosos da realidade econômica e/ou elaboradores de projetos, análises e cenários prospectivos” (SEI, 2013, p. 3).

Em consonância com a citação acima, percebe-se que um dos atores interessados no resultado da taxa de expansão real do PIB é o governo e sua equipe técnica. Esse interesse está associado à capacidade do mesmo de intervir na economia suavizando as turbulências no ciclo econômico. De fato, em períodos de crise econômica (ou de expectativa de crise), é comum a intervenção do governo para minimizar as incertezas e redirecionar a economia para uma trajetória mais compreensível aos olhos dos empresários e tomadores de decisão.⁵ Assim, em virtude do seu poder de mobilização e de interferência na demanda agregada, o governo pode ser considerado um agente “construtor” do futuro, uma vez que as decisões tomadas pelo aparato governamental influenciam a sociedade como um todo – seja por meio de políticas públicas e programas de governo ou através de intervenções mais pontuais (CARVALHO, 1999).

Essas intervenções ocorrem, muitas vezes, de maneira expansionista quando o governo aumenta os seus gastos afetando diretamente a demanda agregada – lembre-se de que os gastos governamentais constituem uma variável importante no

⁵ Essa intervenção estatal na suavização dos efeitos de possíveis períodos conturbados e de instabilidade econômica (como os períodos de crise) é tratada pela teoria keynesiana. O pensamento keynesiano se fundamenta no princípio de que o ciclo econômico não é auto-regulado dado o “espírito animal” dos empresários. Por isso, Keynes defende a intervenção do governo para “inflar” a demanda agregada e reduzir as incertezas, mas deixando a cargo dos agentes privados às decisões de alocação dos recursos disponíveis (CARVALHO, 1999).

cálculo do PIB pela ótica da demanda –, ou quando toma medidas para reativar a economia – na tentativa de influenciar positivamente a taxa de crescimento real do PIB. Um exemplo típico da utilização dessas medidas aconteceu no ano de 2012 quando o governo brasileiro adotou um pacote de medidas econômicas para ampliar a produção e evitar uma queda mais acentuada na taxa de expansão real do PIB nacional. Nesse período, houve desonerações fiscais para diversos setores; redução do IPI (Imposto sobre Produtos Industrializados) para a “linha branca” (móveis e eletrodomésticos), para compra de automóveis e de bens de capital; além de cortes na taxa de juros (RESENDE; SIMÃO, 2012).

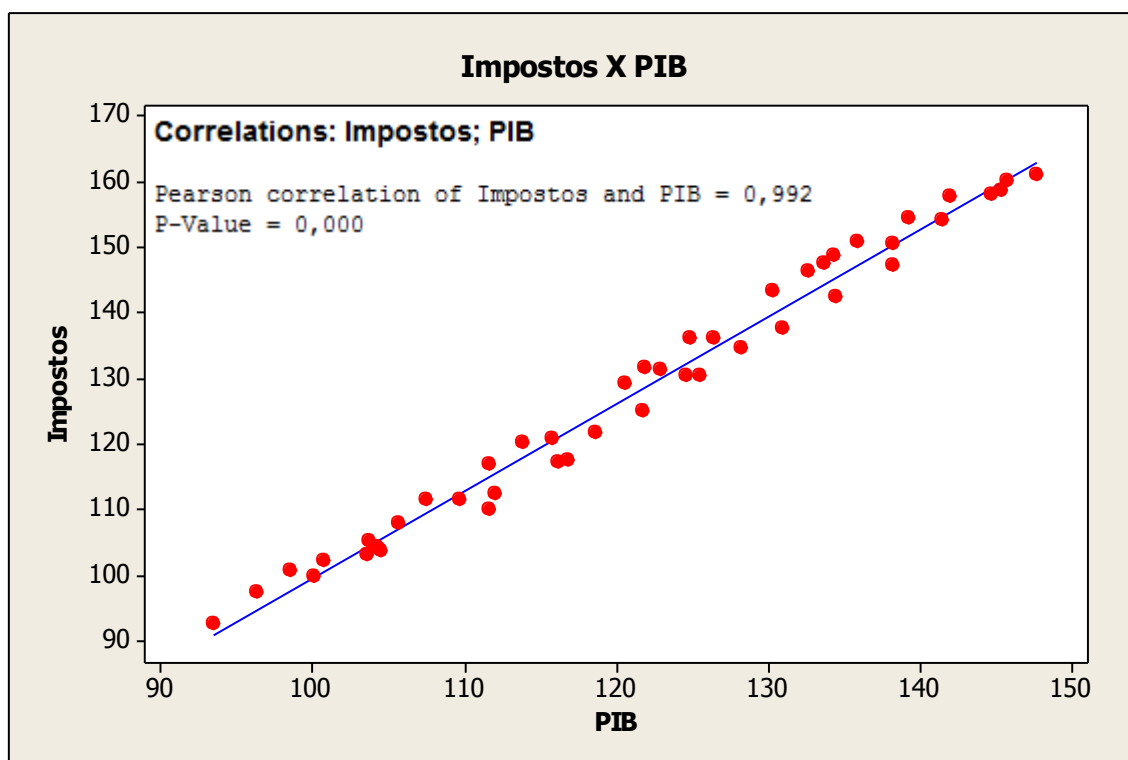
O governo também se preocupa com o resultado do índice de volume do PIB porque esse indicador é, em alguma medida, utilizado pelos eleitores na sua tomada de decisão para escolha de um governante. Evidentemente, que a taxa de crescimento do PIB não é a única variável observada pelo eleitorado, já que uma inflexão nem sempre está associada às decisões políticas tomadas anteriormente. Ela pode ser fruto, por exemplo, de um ambiente econômico externo desfavorável. Um exemplo dessa situação ocorre quando a geração do produto agregado cai a nível global. Nesse caso, quando o ritmo de crescimento econômico cai em escala mundial é de se esperar, naturalmente, uma redução no ritmo de crescimento na maior parte das economias abertas, já que o comércio bilateral entre os diferentes países sofrerá uma redução significativa – lembre-se de que o PIB pela ótica da demanda é afetado pelo saldo das exportações e importações.

Mesmo assim, existe uma preocupação latente por parte dos governantes com a taxa de crescimento real da economia ao longo do seu mandato e, sobretudo, quando se aproxima o período eleitoral. Por isso, é comum o aumento dos gastos públicos com obras em infraestrutura no período final do seu governo visando à obtenção de votos. Essa estratégia é frequentemente corroborada pela Teoria da Escolha Racional, em que os agentes tentam maximizar o seu retorno e, no caso do político, a sua probabilidade de ser reeleito. A citação abaixo confirma essa constatação:

“Levantamento feito pelo Instituto de Pesquisa Econômica e Aplicada (Ipea), com base em dados de 1995 a 2011, demonstra que as eleições têm impacto direto no investimento público feito por prefeituras e governos estaduais e federal. A pesquisa, divulgada em dezembro do ano passado, verificou um aumento expressivo nos gastos públicos em anos pares, em que há eleições; nos anos subsequentes, por sua vez, há forte contenção das despesas”. (IPEA, 2012 *apud* BECKER, 2012).

Apesar do uso político da taxa de expansão real do PIB ser generalizado, talvez a principal preocupação dos governantes e administradores públicos em relação ao índice de volume do produto agregado seja em relação aos condicionantes que afetam a arrecadação tributária. Para o correto dimensionamento e planejamento de políticas públicas e programas, o setor público precisa saber o quanto arrecadou de impostos, taxas e contribuições sociais das demais unidades produtivas e instituições. Todavia, essa arrecadação cai substancialmente quando o nível de atividade econômica é baixo, ou em outras palavras, quando o crescimento do índice de volume do PIB é reduzido. Assim, o fato de existir uma correlação positiva entre a expansão do produto agregado e a arrecadação de impostos é mais um dos motivos pelo qual o governo se preocupa tanto com o resultado do índice de volume do PIB. O Gráfico 1 abaixo demonstra exatamente essa correlação: cada ponto representa o resultado trimestral do índice de volume do PIB e do nível de impostos em Minas Gerais no período 1º Trimestre de 2002-2º Trimestre de 2013, sendo o ano de 2002 o ano de referência (FUNDAÇÃO JOÃO PINHEIRO, 2013).

Gráfico 1 – Gráfico de dispersão dos resultados trimestrais da série de base fixa (média 2002 = 100) em volume da arrecadação de Impostos e do PIB de Minas Gerais – 1º Trimestre de 2002-2º Trimestre de 2013



Fonte: Elaboração própria com base em informações divulgadas pela FUNDAÇÃO JOÃO PINHEIRO (2013).

Outro agente importante e interessado na taxa de crescimento real da economia é o Banco Central. De fato, como essa instituição é responsável pela condução da política monetária do país, frequentemente ela precisa tomar decisões que “sacrifiquem” uma parcela do crescimento da economia, mas que assegurem um sistema financeiro sólido com a manutenção da estabilidade do poder de compra da moeda. Desse modo, em situações de elevação no nível geral dos preços praticados no mercado, o Banco Central precisa tomar medidas restritivas que contenham o avanço da inflação: o aperto monetário através da elevação da taxa de juros (a SELIC⁶) costuma ser o principal “remédio” para o avanço inflacionário.

Todavia, a política de contenção e restrição do crédito (via elevação dos juros) afeta negativamente a taxa de crescimento do PIB, pois aumenta o custo da obtenção de empréstimos e, com isso, desmobiliza empresários no que tange à realização de novos investimentos, além de diminuir o consumo das famílias. O investimento e consumo das famílias são variáveis com um peso importante no que se refere ao cálculo do PIB pela ótica da demanda. Portanto, o Banco Central tem a difícil missão de manter a inflação sob controle com a adoção de medidas restritivas, porém, na proporção correta. Caso contrário, se a elevação na taxa básica de juros for acima do estritamente necessário, o impacto negativo no nível de atividade econômica poderá ser bastante significativo.

Um exemplo deste tipo de intervenção por parte do Banco Central ocorreu no ano de 2013, no qual a preocupação com a inflação ganhou espaço midiático desde o início do ano, sobretudo diante do impacto esperado com o repasse da desvalorização do câmbio para a inflação, já que os produtos importados acabam entrando no país a um preço mais elevado com o Real desvalorizado. A citação abaixo explicita a estratégia de interferência por parte do Banco Central na taxa de juros no que se refere ao direcionamento da política monetária:

“A ata do Comitê de Política Monetária (Copom) divulgada ontem trouxe a sinalização de que o Banco Central dará continuidade ao ciclo de aperto monetário, reforçando as apostas refletidas nos contratos futuros de juros de uma nova alta de 0,5 ponto percentual da taxa Selic, hoje em 8,5%, na reunião do Copom em agosto. A ata divulgada pelo BC destacou que a depreciação cambial constitui uma “fonte de pressão inflacionária” em prazos mais curtos, reforçando a necessidade de “adequada condução da política monetária” para limitar os efeitos secundários do choque de câmbio ao longo do tempo” (PEREZ; ROSA; CASTRO, 2013).

⁶ A SELIC corresponde a uma taxa média ponderada pelo volume das operações de financiamento por um dia, lastreadas em títulos públicos federais e calculada com base no Sistema Especial de Liquidação e de Custódia (SELIC).

Além do Banco Central, outras instituições privadas bancárias também se preocupam com o resultado do índice de volume do PIB. Isto acontece porque, em períodos conturbados e de baixo crescimento econômico, é comum a redução no nível de atividade do setor de intermediação financeira, já que a desconfiança que impregna a ação de empresários e consumidores poderá reduzir o ritmo das operações e concessões de crédito na economia. Esse problema ainda pode ser acentuado em um contexto inflacionário, tanto pela perda de parte do poder de compra da moeda quanto pela adoção de possíveis políticas restritivas por parte do Banco Central (como a elevação dos juros), que aumenta o custo do dinheiro para realização de empréstimos bancários.

Os empresários, acionistas e os industriais vinculados as grandes corporações constituem outro conjunto de atores interessados no resultado da taxa de crescimento real da economia. De fato, esses *stakeholders* constantemente balizam suas decisões com base no comportamento do índice de volume do PIB. Se a expectativa de expansão real da economia é baixa, eles planejam suas ações de maneira a minimizar os riscos e possíveis prejuízos advindos de uma tomada de decisão equivocada, ou dito de outra forma, são mais cautelosos e menos confiantes. Esse cuidado referente ao acompanhamento da economia por parte dos empresários frequentemente é acompanhado por decisões acerca de quão deva ser elevada a sua própria produção empresarial e também no que se refere ao correto dimensionamento do nível de estoques. Ademais, muitas das decisões acerca da realização de novos investimentos (Formação Bruta de Capital Fixo) são tomadas ou proteladas com base na taxa de crescimento real do PIB.

Neste ponto, é importante destacar outro aspecto fundamental, que é a relação ambígua entre empresários e o crescimento econômico. Ao mesmo tempo que as ações tomadas pela direção das empresas são, em alguma medida, norteadas pelo resultado do índice de volume do PIB, os empresários são, na prática, a força motriz do crescimento econômico sustentado no longo prazo, porque são eles que implementam ou financiam a inovação da economia, mesmo que motivados por fatores diversos – o aumento do lucro, a ampliação do *market share* (fatia do mercado) e, principalmente, a concorrência (MIGLIOLI, 1986).

Como a inovação, o progresso tecnológico e as ações em Pesquisa & Desenvolvimento (P&D) das empresas estão no cerne da teoria do crescimento

econômico, sobretudo, a “teoria do crescimento endógeno⁷”, pode-se dizer que os empresários possuem papel decisivo na expansão do produto agregado, mesmo que, em virtude disso, ocorra a extinção de empresas que não se adaptem à prática fundamental de inovar. É este o fator essencial do capitalismo destacado por Joseph Schumpeter como “o processo de destruição criadora”. A importância do empresário inovador para o crescimento econômico persistente no longo prazo pode ser melhor compreendida pela citação abaixo que destaca a lógica da teoria schumpeteriana:

“O sistema teórico schumpeteriano [...] pode ser interpretado como [uma] teoria do ciclo da capacidade empresarial, devido ao encadeamento lógico explicativo do *modus operandi* do desenvolvimento capitalista: inovação, novas plantas industriais, novas firmas, novo homem (empresário) e, por fim, a necessidade de crédito *ad hoc*. Em outros termos, o surgimento do desenvolvimento econômico [e o crescimento econômico no longo prazo] é motivado pela existência de grupos empresariais de qualificação especial (*ability*), distinguidos pela liderança social, os quais encaminham com êxito uma nova combinação dos fatores de produção. A lógica desse discurso teórico evidencia a importância da hipótese do empresário inovador para a teoria schumpeteriana [e também para a teoria do crescimento econômico endógeno]” (CALAZANS, 1992, p. 1).

Portanto, percebe-se que a ocorrência de inovação é fundamental para a expansão do índice de volume do PIB no longo prazo, uma vez que dela decorre um novo patamar de acumulação de capital e de geração de valor adicionado. No entanto, a relação intrínseca entre inovação e produtividade – a capacidade de criar valor adicionado com uma quantidade cada vez menor de funcionários e de horas trabalhadas, decorrente do processo inovativo – vem ganhando contornos dramáticos na economia brasileira e de muitas unidades da federação. Isto ocorre porque a produtividade em nosso país (e também no estado de Minas Gerais) ainda é bastante baixa (FUNDAÇÃO JOÃO PINHEIRO, 2012).

Ademais, Lima (2013) e Giambiagi (2013) argumentam que esse problema ficará ainda mais nítido com a passagem do período intitulado como de “bônus demográfico”. De fato, o primeiro autor destaca três mudanças demográficas importantes na sociedade brasileira: a estimativa dos órgãos oficiais de estatística de que a população brasileira deixará de crescer por volta do ano 2040; a consequente diminuição da População em Idade Ativa (PIA) e o declínio também da População Economicamente Ativa (PEA) (LIMA, 2013). Já o segundo autor chama a atenção para o fato de que uma parte não desprezível do crescimento em volume do PIB nacional até então sempre esteve associado ao dinamismo da população (o crescimento

⁷ Endógeno por que o progresso e o avanço tecnológico são explicados pela teoria econômica atinente aos modelos, sobretudo o modelo desenvolvido por Paul e Romer, uma vez que nos modelos anteriores (o modelo de crescimento econômico de Solow) a tecnologia era uma variável exógena ao modelo, o que significa dizer que as ações das empresas em P&D não afetariam o crescimento (NASCIMENTO, 2006).

populacional), ao que o autor denominou de “exército industrial de reserva” (GIAMBIAGI, 2013).

No entanto, se as taxas de crescimento da população e das pessoas em idade de trabalhar vem sendo declinantes ao longo do tempo, isto significa dizer que cada vez mais o país (e Minas Gerais) terão dificuldade de crescer economicamente por meio de mais ocupação e cada vez mais dependerá dos ganhos de produtividade. Assim, o índice de volume do PIB no cenário mineiro e brasileiro ganha em importância, visto que o resultado da taxa de crescimento real da economia revelará cada vez mais a capacidade da mesma em implementar processos inovadores.

A população de forma geral e, sobretudo, aquela de baixa renda também possui interesse particular no resultado referente à taxa de crescimento real da economia. Esse interesse se relaciona exatamente com a questão da ocupação e a inserção (ou manutenção) no mercado de trabalho. Isto ocorre porque existe uma relação imperfeita, porém significativa, entre a taxa de desemprego e a taxa de crescimento do PIB real. De maneira geral, essa forte relação é negativa e sugere que “quando o crescimento é superior à média, a taxa de desemprego cai; e quando o crescimento é inferior à média, a taxa de desemprego aumenta” (KRUGMAN; WELLS, 2007, p. 499).

Evidentemente que essa relação não é sempre direta, já que é perfeitamente possível a combinação de um crescimento econômico mais lento com taxas de desemprego crescentes, o que é comumente tratado na literatura como recuperação sem emprego e se relaciona, possivelmente, com a ampliação da utilização da capacidade instalada. Também pode acontecer uma desaceleração econômica sem que isso necessariamente resulte em ampliação da taxa de desemprego, sobretudo quando há indícios de que a diminuição no ritmo de atividade econômica se manifesta como transitória e reversível (KRUGMAN; WELLS, 2007).

Todavia, a preocupação com o índice de volume do PIB se acentua quando ocorre um declínio exagerado no produto agregado e o baixo nível de atividade produtiva permanece por um período de tempo significativo. As recessões econômicas são um exemplo típico dessa situação e as depressões simbolizam os casos mais críticos. O grande problema de uma queda acentuada no produto agregado e que perdure por um período de tempo prolongado é que, mais cedo ou mais tarde, isso rebaterá no desempenho do mercado de trabalho. Nesses casos, a elevação do desemprego surge como uma resposta quase que natural à inflexão duradoura no nível de atividade econômica.

Além da preocupação com a manutenção no mercado de trabalho, a população brasileira mais carente também possui a atenção voltada para o resultado do índice de volume do PIB porque a taxa de crescimento real da economia, juntamente com a inflação observada, são critérios utilizados no reajuste do salário de referência da economia. De fato, a regra do governo federal estabelece o aumento no salário mínimo com base em dois critérios: a variação do INPC (Índice Nacional de Preços ao Consumidor) do ano anterior e o crescimento real do PIB nos últimos 24 meses (DIEESE, 2013 *apud* LAPORTA, 2013).

Portanto, nota-se que existe um interesse generalizado e diversificado em relação ao resultado do índice de volume do PIB por parte dos diversos agentes econômicos e tomadores de decisão. O governo e os órgãos da administração pública; o Banco Central e os bancos privados; os empresários e investidores; e a população de forma geral – constituem apenas em uma parcela dos atores vinculados, direta ou indiretamente, ao desempenho da economia registrado pela taxa de expansão real do produto agregado. Outro conjunto de atores envolvidos com o crescimento econômico são os grupos de análise de cenários prospectivos.

Esta monografia vem preencher uma lacuna exatamente no que concerne a análise de cenários e prognósticos futuros em relação ao comportamento da economia de Minas Gerais, informando assim os *stakeholders* interessados no seu resultado. De fato, em âmbito nacional, existe um conjunto de grupos que fornecem pistas acerca do comportamento “futuro” da economia brasileira: o próprio Banco Central (através do Boletim Focus e da divulgação do IBC-BR – Índice de Atividade Econômica do Banco Central), e diversas consultorias como a LCA Soluções Estratégicas em Economia fazem este tipo de trabalho.

Porém, em âmbito estadual existem ainda poucas organizações que sinalizem o comportamento futuro da economia mineira no curto prazo ⁸. Tanto é verdade que a Fundação João Pinheiro divulga tradicionalmente os resultados trimestrais oficiais preliminares do índice de volume do PIB mineiro e uma das maiores demandas atinentes a divulgação se refere a realização de previsões e análise de prognósticos. Assim, a construção dos modelos estatísticos propostos no trabalho vem apontar algumas sugestões para o preenchimento desta lacuna através dos modelos de

⁸ Embora também aqui o próprio Banco Central (através IBCR – Índice de Atividade Econômica Regional) e a LCA Soluções Estratégicas em Economia tentem precisar o comportamento futuro de algumas economias estaduais. Porém, o desconhecimento por parte da realização deste tipo de trabalho destes órgãos em nível regional e a própria dificuldade em encontrar indicadores satisfatórios para a realização dos prognósticos motivem a realização da monografia.

acompanhamento e previsão do índice de volume do PIB de Minas Gerais construídos.

CAPÍTULO 3: FUNDAMENTAÇÃO TEÓRICA

Este capítulo tem o objetivo de apresentar uma breve revisão bibliográfica atinente à utilização de séries temporais, da técnica de alisamento exponencial e dos modelos de regressão dinâmica. Nele também serão analisados os aspectos para a escolha e verificação dos modelos propostos.

3.1) Aspectos gerais da análise de séries Temporais

Uma série temporal é um registro de observações de algum fenômeno medido de forma sequencial no tempo. Nesse sentido, os resultados trimestrais do índice de volume do PIB de Minas Gerais, foco da realização deste trabalho, constituem em um exemplo típico de série temporal. Os modelos utilizados para descrever as séries temporais constituem em processos estocásticos, isto é, processos controlados por leis probabilísticas e representados por uma família de variáveis aleatórias indexadas no tempo (MORETTIN; TOLOI, 1987).

Os objetivos que se pretendem alcançar através da análise de séries temporais podem ser os mais diversos possíveis como investigar o mecanismo gerador da série temporal, fazer previsões de valores futuros ou, simplesmente, descrever o comportamento geral e dos componentes da série (MORETTIN; TOLOI, 1987).

Uma das suposições mais necessárias que se faz acerca de uma série temporal é a de que ela seja estacionária, isto é, que ela se desenvolva no tempo de forma aleatória ao redor de uma média constante, refletindo alguma forma de equilíbrio estável. Entretanto, a maior parte das séries encontradas apresenta alguma forma de não estacionariedade. Nos casos mais simples, a série flutua ao redor de uma reta, com inclinação positiva ou negativa, indicando a presença de uma tendência linear. Há também casos mais extremos, onde a não estacionariedade é explosiva, como acontece com o crescimento populacional de uma colônia de bactérias (MORETTIN; TOLOI, 1987).

O conjunto que permeia os modelos ARIMA (Autoregressivos integrados e de Médias móveis), vinculados à metodologia de Box-Jenkins e que serão abordados mais adiante, é capaz de representar de maneira eficaz séries estacionárias e não estacionárias, mas que não apresentem comportamento explosivo. Nesses casos, afirma-se que a estacionariedade é homogênea:

“[...] a série pode ser estacionária, flutuando ao redor de um nível, por certo tempo, depois mudar de nível e flutuar ao redor de um novo nível e assim por diante, ou então mudar de inclinação, ou ambas as coisas” (MORETTIN; TOLOI, 1987, p. 6-7).

A presença ou a ausência de estacionariedade na série original é um fator relevante, tendo em vista que a maior parte dos procedimentos de análise estatística pressupõe a presença de estacionariedade. Assim, muitas vezes é necessária a transformação dos dados originais caso esses não formem uma série estacionária. Costuma-se, por exemplo, tomar diferenças sucessivas da série original, até que se obtenha uma série estacionária. Assim, a primeira diferença de uma série $Y(t)$ pode ser definida como: $\Delta Y(t) = Y(t) - Y(t-1)$. Já a segunda diferença pode ser representada por: $\Delta^2 Y(t) = \Delta[\Delta Y(t)] = \Delta[Y(t) - Y(t-1)]$, que escrita de outro modo configura a seguinte relação: $\Delta^2 Y(t) = Y(t) - 2Y(t-1) + Y(t-2)$. Todavia, Morettin e ToloI (1987) argumentam que, na maior parte dos casos, será suficiente tomar uma ou duas diferenças para que a série se transforme em uma série estacionária.

Outras transformações podem também ser efetuadas nos dados originais visando a atingir objetivos diversos, seja para estabilizar a variância e tornar o efeito sazonal aditivo ou para obter uma distribuição mais simétrica e próxima da normal. Em séries financeiras e econômicas, por exemplo, é comum a ocorrência de acréscimos na variância que justificam a aplicação de transformações não lineares antes de aplicar diferenças na série (MORETTIN; TOLOI, 1987). Um conjunto de transformações que podem ser utilizada foi proposta por Box-Cox, sendo definida como:

$$Y(t)^{(\lambda)} = \begin{cases} \frac{Y(t)^\lambda - c}{\lambda} & \lambda \neq 0 \\ \log Y(t) & \lambda = 0 \end{cases}$$

Aqui λ e c são parâmetros a serem estimados. O Quadro 1 abaixo traz alguns dos possíveis valores de λ e o tipo de transformação relacionada. É importante salientar que a série transformada não será necessariamente estacionária. Além disso, vale lembrar também que, em muitas séries econômicas, a transformação logarítmica ($\lambda = 0$) se mostra adequada em muitos casos, sobretudo quando essa transformação é acompanhada pela aplicação de uma diferença (MORETTIN; TOLOI, 1987).

Quadro 1 – Valores do parâmetro (λ) e Transformação de Box-Cox na série original correspondente ao parâmetro

Parâmetro (λ)	Transformação (T) na série original Y(t)
-1,0	$T [Y(t)] = 1/Y(t)$
-0,5	$T [Y(t)] = 1/\sqrt{Y(t)}$
0,0	$T [Y(t)] = \text{Log}Y(t)$
0,5	$T [Y(t)] = \sqrt{Y(t)}$
1,0	$T [Y(t)] = Y(t)$

Fonte: Elaboração própria com base em informações de TOSCANO (2012).

O grande problema em relação à aplicação das transformações é que, muitas vezes, elas não melhoram a qualidade das previsões, podendo inclusive conduzir a resultados inapropriados. Além disso, alguns autores recomendam a utilização direta do modelo sazonal como instrumento de realização das previsões ao invés da utilização de um modelo não sazonal já com a série transformada. Morettin e Tolo (1987) citam diversos autores que exemplificam este tipo de argumentação:

“[...] Nelson (1976) conclui que transformações não melhoram a qualidade da previsão; Makridakis e Hibon (1979) verificaram que os dados transformados têm pouco efeito na melhoria da previsão e, sob bases mais teóricas, Granger e Newbold (1976) mostram que as previsões dos antilogaritmos dos dados transformados são estimadores viciados e deveriam, portanto, ser ajustados, mais isto não é feito nos principais programas de computador, o que significa que, depois que os dados são transformados, um vício é introduzido nas previsões, decorrentes de tal transformação. Além disso, Granger e Newbold observam que a heterocedasticidade não afeta a adequação da previsão, como no caso de regressão múltipla. [...] Plosser (1979) conclui que parece ser preferível fazer a previsão usando diretamente o modelo sazonal ao invés de ajustar sazonalmente a série e depois utilizar um modelo não sazonal” (MORETTIN; TOLOI, 1987, p. 10-11).

Independente da aplicação de diferenças ou transformações na série original, os diversos modelos construídos são, muitas vezes, representados pelos modelos da metodologia de Box-Jenkins. Todavia, o termo “método” ou “metodologia” de previsão ARIMA não configuram a expressão mais adequada a ser utilizada. De fato, “não há algo chamado “método” de previsão ou algo chamado de previsão de Box-Jenkins. Há algo chamado método de previsão de “mínimos quadrados”” (PRIESTLEY, 1979 *apud* MORETTIN; TOLOI, 1987, p. 8). É esse método que fornece a base para todo o embasamento teórico posterior em que os parâmetros são estimados de forma a minimizar a soma dos quadrados dos resíduos.

Com relação à decomposição de uma série temporal Y(t) pode-se dizer que a mesma pode ser escrita conforme a soma de três componentes: uma componente tendencial – T(t), uma componente sazonal – S(t), e um termo aleatório – a(t),

conforme a equação: $Y(t) = T(t) + S(t) + a(t)$. De maneira geral, a tendência em séries econômicas está relacionada a fatores medidos durante períodos extensos de tempo. Já a componente sazonal vincula-se a existência de observações com comportamentos “similares” dentro de um período do ano, como acontece com os dados registrados mensalmente ou trimestralmente. Retirando as componentes de tendência e de sazonalidade o que resta é a componente aleatória ou residual. A suposição usual é que o termo $a(t)$ seja um processo estocástico puramente aleatório e, em alguns casos, pode-se considerá-lo um processo estacionário com média zero e variância constante. A representação da série temporal acima como a soma das três componentes configura um modelo aditivo, indicado quando a componente $S(t)$ não depende de outras componentes como $T(t)$. Todavia, quando as amplitudes sazonais mudam com a tendência, um modelo multiplicativo do tipo: $Y(t) = T(t) \cdot S(t) \cdot a(t)$, se mostra mais adequado (MORETTIN; TOLOI, 1987).

A partir desta pequena introdução dos aspectos mais gerais relacionados ao entendimento das diversas séries temporais, convém entender agora o conceito de Função de Autocorrelação (FAC) e Função de Autocorrelação parcial (FACP); o conceito de “ruído branco” e de utilização de operadores. O entendimento destes três aspectos será de suma importância para a interpretação dos modelos descritos mais adiante.

3.2) Função de Autocorrelação (FAC) e Função de Autocorrelação parcial (FACP)

A Função de Autocorrelação (FAC) e a Função de Autocorrelação parcial (FACP) são instrumentos fundamentais para identificação do tipo de processo estacionário que descreve uma série temporal. A autocorrelação entre dois instantes de tempo (t e $t+k$) pode ser definida pela seguinte relação:

$$\rho_k = \frac{\text{covariância}(Y_t, T_{t+k})}{\sqrt{\text{Variância}(Y_t)} \cdot \sqrt{\text{Variância}(T_{t+k})}} = \frac{Y_k}{Y_0}$$

Através do cálculo acima, pode-se representar todas as autocorrelações ρ_k para $k = 0, \pm 1, \pm 2, \dots$; em um tipo de gráfico conhecido como correlograma. Neste tipo de gráfico fica evidenciado algumas propriedades da Função de Autocorrelação (FAC):

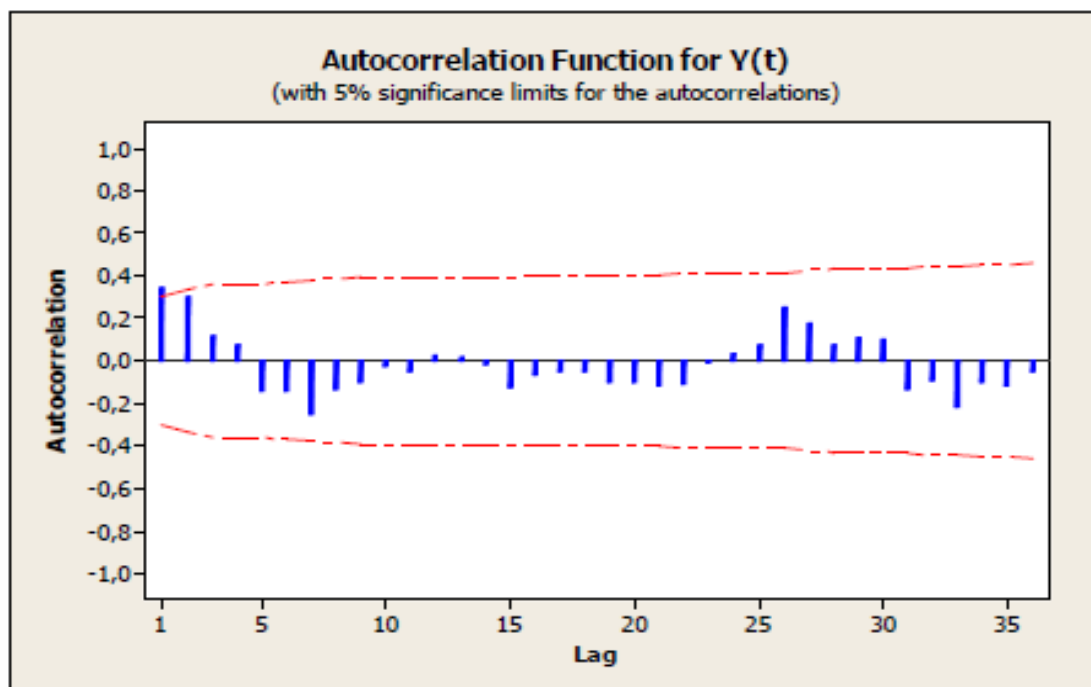
$$\begin{aligned} \rho_0 &= 1 \\ |\rho_k| &\leq 1 \quad \text{para todo } k, \\ \rho_{-k} &= \rho_k \quad \text{para todo } k. \end{aligned}$$

Ademais, considerando que as correlações amostrais seguem aproximadamente uma distribuição normal com desvio padrão S_{ρ_k} , pode-se calcular e representar no correlograma da FAC os limites com um nível de 5% de significância e então testar a hipótese se as autocorrelações são estatisticamente diferentes de zero (TOSCANO, 2012). A fórmula para o cálculo do desvio padrão S_{ρ_k} é mostrada abaixo:

$$S_{\hat{\rho}_k} \cong \sqrt{\frac{1}{n} (1 + 2\hat{\rho}_1^2 + \dots + 2\hat{\rho}_{k-1}^2)}$$

Na prática, se a correlação situar-se acima da linha “pontilhada” traçada pelo software estatístico quando da construção do correlograma, conclui-se que a correlação é estatisticamente diferente de zero. O gráfico 2 traz um exemplo de correlograma da FAC em que a primeira correlação é estatisticamente diferente de zero (maior que o limite de confiança superior):

Gráfico 2 – Função de Autocorrelação (FAC) de uma série qualquer (Y(t)) com a representação dos limites com 5% de significância



Fonte: TOSCANO (2012).

Já a Função de Autocorrelação Parcial (FACP) explicita uma espécie de correlação condicional que investiga a correlação entre Y_t e Y_{t+K} após a remoção das dependências lineares das variáveis intermediárias $Y_{t+1}, Y_{t+2}, \dots, Y_{t+K-1}$ (TOSCANO, 2012). Nesse sentido, a autocorrelação parcial no lag K é definida como o coeficiente da seguinte regressão:

$$Y_t = \phi_{k1}Y_{t-1} + \dots + \phi_{kk}Y_{t-k} + a_t$$

$$\phi_{kk} = \text{Corr}(Y_t, Y_{t+k} / Y_{t+1}, \dots, Y_{t+k-1})$$

Exemplo:

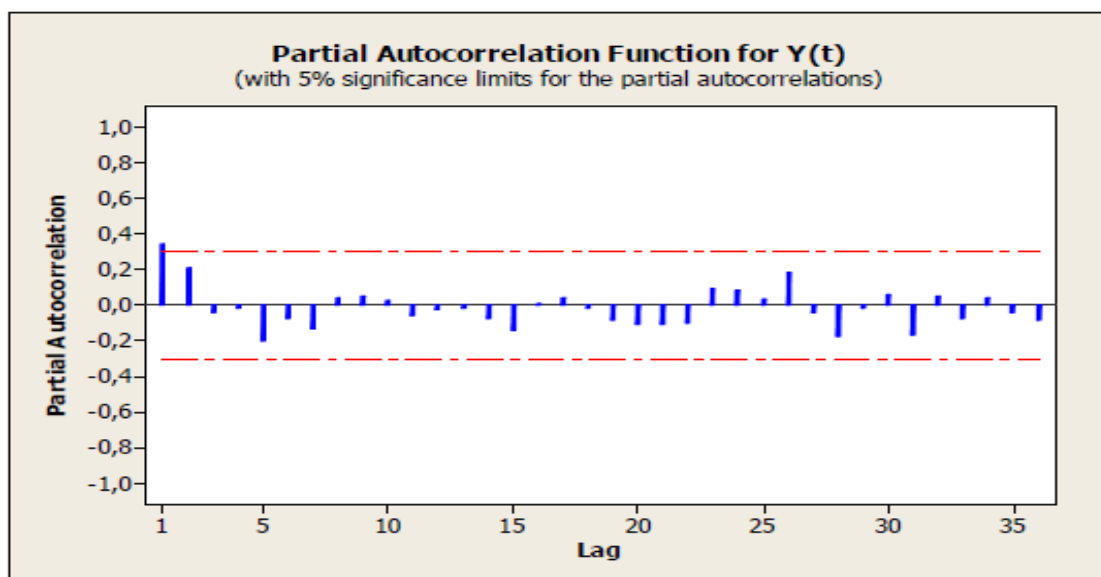
$$Y_{\varepsilon} = \phi_{11}Y_{\varepsilon-1} + a_t$$

$$Y_{\varepsilon} = \phi_{21}Y_{\varepsilon-1} + \phi_{22}Y_{\varepsilon-2} + a_{\varepsilon}$$

$$\phi_{11} = \rho_1 \quad \phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} \quad \phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \dots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \dots & \rho_1 & 1 \end{vmatrix}}$$

Além disso, da mesma forma que no caso anterior, as correlações amostrais parciais também seguem uma distribuição normal com desvio padrão $S_{\phi_{kk}} \cong \sqrt{1/n}$. Assim, também neste caso pode-se calcular e representar os limites com um nível de 5% de significância e testar a hipótese se as autocorrelações parciais são estatisticamente diferentes de zero. O gráfico 3 traz um exemplo de correlograma da FACP em que a primeira correlação parcial é estatisticamente diferente de zero (maior que o limite de confiança superior):

Gráfico 3 – Função de Autocorrelação Parcial (FACP) de uma série qualquer (Y(t)) com a representação dos limites com 5% de significância



Fonte: TOSCANO (2012).

3.3) O conceito de “Ruído Branco”

O termo ruído branco (a_t) refere-se a um processo com média zero, $E(a_t) = 0$, variância constante, $Var(a_t) = \sigma^2$, e não correlacionados, isto é, com $E(a_r, a_s) = 0$ para $r \neq s$. Na prática, um ruído branco é um processo estacionário com Função de Autocorrelação (FAC) e Função de Autocorrelação Parcial (FACP) com características específicas, conforme se pode notar abaixo:

função de autocorrelação

$$\rho_k = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases}$$

função de autocorrelação parcial

$$\phi_{kk} = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases}$$

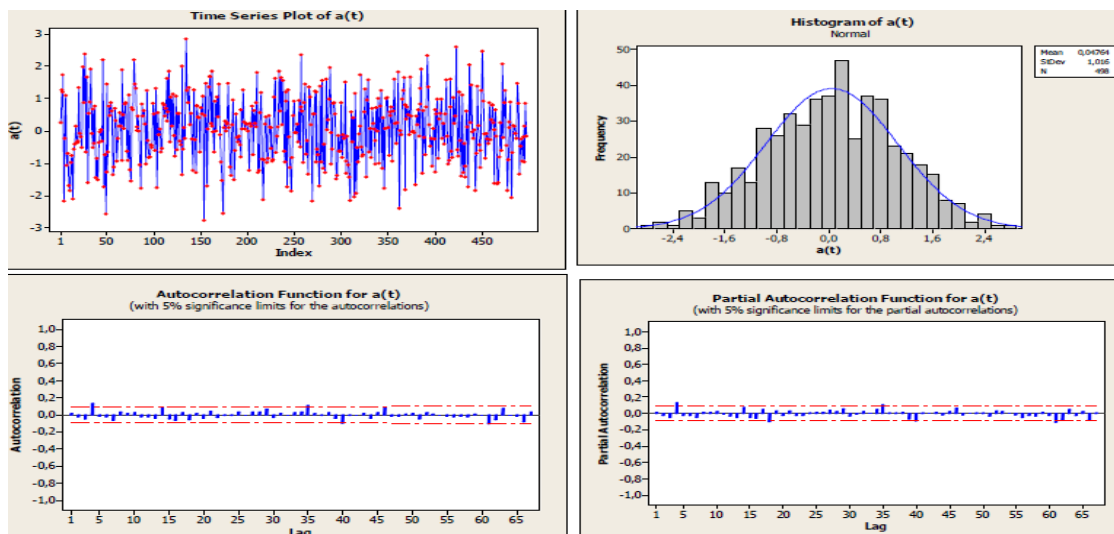
A sequência de choques aleatórios (a_t) será um ruído branco gaussiano se o processo apresentar distribuição normal com média zero e variância constante, possuindo as seguintes características:

$$a_t \sim N(\mu, \sigma_a^2) \\ t = 1, 2, \dots$$

$$E(a_t) = \mu = 0 \\ V(a_t) = \sigma_a^2 \\ Cov(a_t, a_{t+k}) = 0$$

A Figura 1 traz um exemplo simulado da ocorrência de um processo de ruído branco gaussiano retirado da Apostila de Análise de Séries Temporais da professora Ela Mercedes Medrano de Toscano. Nota-se que o processo possui média zero, é simétrico conforme a distribuição normal, e não apresenta autocorrelações significativas para $K \neq 0$, conforme os correlogramas da FAC e da FACP.

Figura 1 – Simulação de um processo Ruído Branco Gaussiano



Fonte: TOSCANO (2012).

3.4) Utilização de Operadores

Antes de concentrar os estudos nos modelos de Box-Jenkins, convém introduzir a noção de utilização de operadores no que se refere ao entendimento e compreensão das equações e fórmulas dos modelos construídos. Na prática, três tipos de operadores são suficientes:

- a) Operador de retardo ou de translação para o passado, denotado por B e definido por: $B^K Y_t = Y_{t-K}$.
- b) Operador de diferença (Δ): $Z_t = \Delta Y_t = Y_t - Y_{t-1} = (1-B)Y_t$, onde $\Delta = 1 - B$.
- c) Operador de diferenciação sazonal (S): $Z_t = S Y_t = Y_t - Y_{t-S} = (1-B^S)Y_t$, onde $S = 1 - B^S$.

A seguir temos alguns exemplos da aplicação dos operadores propostos acima:

$$B Y_t = Y_{t-1}.$$

$$B^2 Y_t = Y_{t-2}.$$

$$(1-B)^2 Y_t = (1 - 2B + B^2)Y_t = Y_t - 2B Y_t + B^2 Y_t = Y_t - 2Y_{t-1} + Y_{t-2}.$$

$$(1-B^{12})Y_t = Y_t - B^{12}Y_t = Y_t - Y_{t-12}.$$

A familiaridade com as notações apresentadas facilitará também a manipulação matemática dos modelos bem como sua interpretação (MORETTIN; TOLOI, 1987).

3.5) Modelos de Box-Jenkins

A modelagem de Box-Jenkins para séries temporais consiste em ajustar modelos Autorregressivos integrados e de Médias móveis - ARIMA (p, d, q) a um conjunto de dados. A simbologia (p, d, q) informa, respectivamente, a ocorrência de um processo autorregressivo (p), a aplicação de diferenças na série (d), e a ocorrência de um processo de média móvel (q). Os modelos em que não é necessária a aplicação de diferenças pelo fato da série já ser estacionária são conhecidos como modelos ARMA (p, q), que combina características dos modelos autorregressivos e de médias móveis. Existem ainda os modelos puramente Autorregressivos – conhecidos como AR (p) –, e os modelos com características exclusivas da ocorrência de Médias Móveis – conhecidos como MA (q).

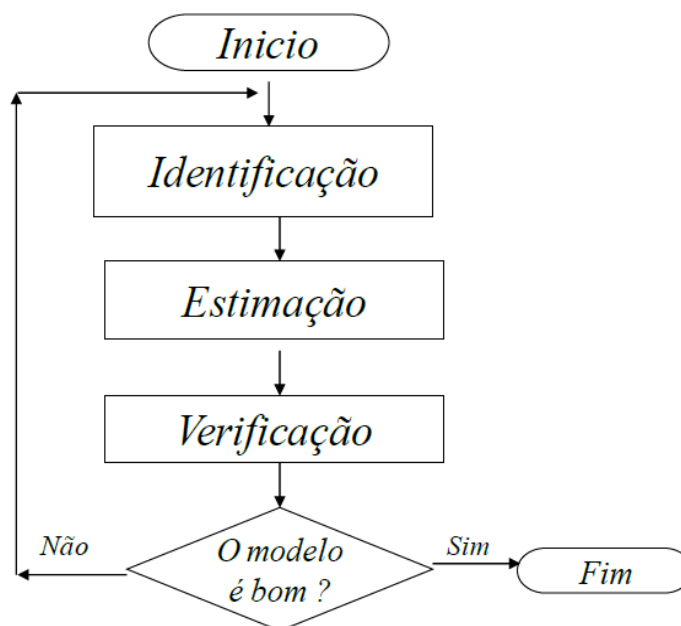
A estratégia de construção e identificação do modelo ARIMA (p, d, q) apropriado é baseado em um ciclo iterativo em que a escolha da estrutura do modelo

é direcionada pela própria característica dos dados. Os estágios do ciclo iterativo sugerido por Box-Jenkins foi resumido por Morettin e Tolo (1987):

“a) uma classe geral de modelos é considerada para a análise (especificação); b) há a identificação de um modelo, com base na análise de autocorrelações, autocorrelações parciais e outros critérios; c) a seguir, vem a fase de estimação, na qual os parâmetros do modelo identificado são estimados; d) finalmente, há a verificação do modelo ajustado, através de uma análise de resíduos, para se saber se este é adequado para os fins em vista” (MORETTIN; TOLOI, 1987, p. 154).

Esses estágios do procedimento proposto por Box-Jenkins para identificação do modelo podem ser resumidos pela Figura 2. Nota-se que, quando o modelo construído não é apropriado, o ciclo se repete voltando a fase de identificação (fase b). Na prática, costuma-se identificar mais de um modelo para estimação e verificação. Se o objetivo é fazer previsões, por exemplo, pode-se optar pelo modelo que apresentar o menor erro quadrático médio de previsão.

Figura 2 – Ciclo Iterativo sugerido por Box-Jenkins para identificação e escolha do modelo ARIMA (p, d, q) apropriado



Fonte: TOSCANO (2012).

Como se pode notar pela Figura 2, a fase de identificação do modelo é complicada, visto que existe a possibilidade de vários pesquisadores identificarem modelos diferentes para uma mesma série temporal. Por isso, a técnica de Box-Jenkins, assim como qualquer procedimento de modelagem, pressupõe por parte do

usuário conhecimento e experiência para além da simples aplicação de pacotes estatísticos (MORETTIN; TOLOI, 1987).

3.5.1) Modelos Autorregressivos de Ordem p – AR (p)

Um modelo Autorregressivo AR (p) – modelo ARMA (p, 0) ou ARIMA (p, 0, 0) – pode ser representado da seguinte forma:

$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t$$

O processo é estacionário com média (μ), tal que $\theta_0 = \mu(1 - \phi_1 - \phi_2 - \dots - \phi_p)$ e a_t é um processo Ruído Branco com média zero, $E(a_t) = 0$, e variância constante, $\text{Var}(a_t) = \sigma^2$. Existe ainda uma restrição para os parâmetros para que o modelo autorregressivo satisfaça a condição de estacionariedade:

Para $p = 1$, $-1 < \phi_1 < 1$;

Para $p = 2$, $\phi_1 + \phi_2 < 1$, $\phi_2 - \phi_1 < 1$ e $-1 < \phi_2 < 1$;

Para $p > 2$, as condições de estacionariedade tornam-se mais complicadas, levando a necessidade de análises mais refinadas.

3.5.1.1) Modelos Autorregressivos de Ordem 1 – AR (1)

Um modelo Autorregressivo AR (1) – modelo ARMA (1,0) ou ARIMA (1,0,0) – pode ser representado da seguinte forma através da utilização do conceito de operadores:

$$(1 - \phi_1 B)Y_t = \theta_0 + a_t$$

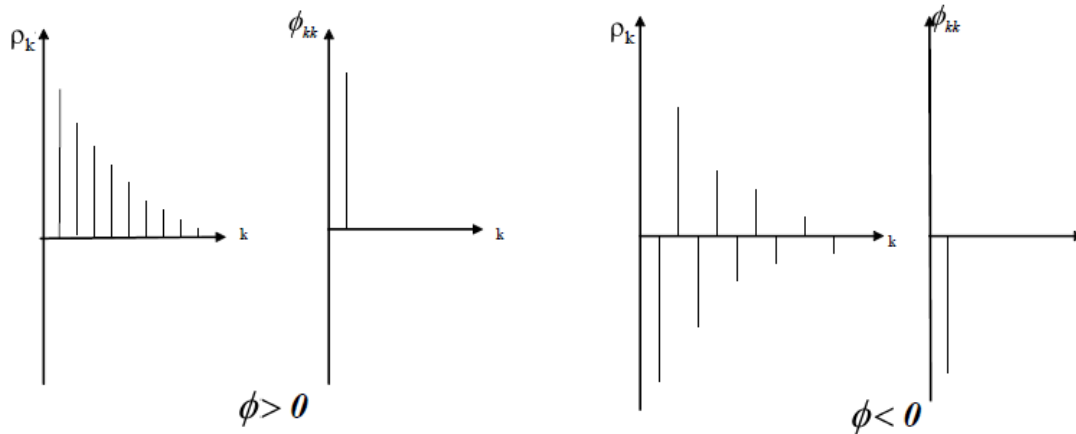
$$Y_t = \theta_0 + \phi_1 Y_{t-1} + a_t$$

O processo é estacionário com média (μ), tal que $\theta_0 = \mu(1 - \phi_1)$ e a_t é um processo Ruído Branco com média zero, $E(a_t) = 0$, e variância constante, $\text{Var}(a_t) = \sigma^2$. Além disso, o coeficiente autorregressivo ϕ_1 deve variar entre -1 e 1 para satisfazer a condição de estacionariedade.

Os correlogramas da Função de Autocorrelação (FAC) e da Função de Autocorrelação Parcial (FACP) de um processo AR (1) possuem configurações especiais (Figura 3). Na FAC de um processo AR (1), as autocorrelações decaem “exponencialmente” para zero e as primeiras autocorrelações decaem de forma um pouco mais devagar. Além disso, esse decaimento pode ocorrer de maneira alternada. Já a FACP de um processo AR (1) indica que a primeira autocorrelação parcial foi

significativa e apenas ela. As demais autocorrelações parciais decaem para zero rapidamente após a primeira defasagem.

Figura 3 – Correlogramas representativos da FAC e FACP dos processos AR (1)



Fonte: TOSCANO (2012).

3.5.1.2) Modelos Autorregressivos de Ordem 2 – AR (2)

Um modelo Autorregressivo AR (2) – modelo ARMA (2,0) ou ARIMA (2,0,0) – pode ser representado da seguinte forma através da utilização do conceito de operadores:

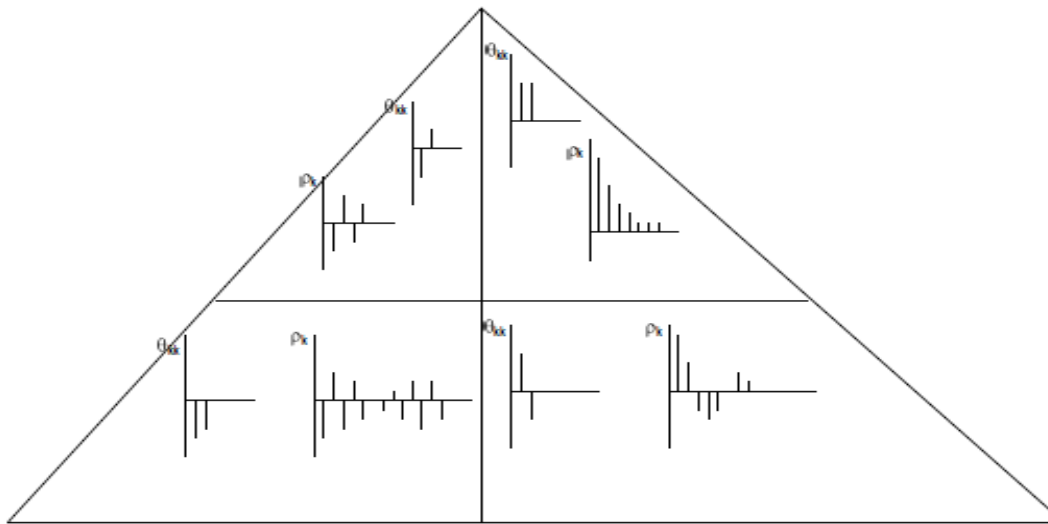
$$(1 - \phi_1 B - \phi_2 B^2)Y_t = \theta_0 + a_t$$

$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t$$

O processo é estacionário com média (μ), tal que $\theta_0 = \mu(1 - \phi_1 - \phi_2)$ e a_t é um processo Ruído Branco com média zero, $E(a_t) = 0$, e variância constante, $\text{Var}(a_t) = \sigma^2$. Além disso, a restrição para os parâmetros ϕ_1 e ϕ_2 para que a condição de estacionariedade seja observada é a seguinte: $\phi_1 + \phi_2 < 1$, $\phi_2 - \phi_1 < 1$ e $-1 < \phi_2 < 1$.

Os correlogramas da Função de Autocorrelação (FAC) e da Função de Autocorrelação Parcial (FACP) de um processo AR (2) possuem configurações especiais tal como acontece com o processo AR (1) (Figura 4). Na FAC de um processo AR (2) as autocorrelações também decaem “exponencialmente” para zero. Esse decaimento pode ocorrer de maneira alternada ou as autocorrelações podem apresentar uma mistura de decaimento “exponencial senoidal” dependendo do sinal dos coeficientes. Já a FACP de um processo AR (2) apresenta a primeira e a segunda autocorrelação parcial significativas. As demais autocorrelações parciais decaem para zero rapidamente após a segunda defasagem – não são significativas.

Figura 4 – Correlogramas representativos da FAC e FACP dos processos AR (2)



Fonte: TOSCANO (2012).

3.5.2) Modelos de Médias Móveis de Ordem q – MA (q)

Um modelo de Médias Móveis MA (q) – modelo ARMA (0, q) ou ARIMA (0, 0, q) – pode ser representado da seguinte forma:

$$Y_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

O processo é estacionário com média (μ) e a_t é um processo Ruído Branco com média zero, $E(a_t) = 0$, e variância constante, $\text{Var}(a_t) = \sigma^2$. Existe ainda uma restrição para os parâmetros para que o modelo de Médias Móveis satisfaça a condição de invertibilidade:

Para $q = 1$, $-1 < \theta_1 < 1$;

Para $q = 2$, $\theta_1 + \theta_2 < 1$, $\theta_2 - \theta_1 < 1$ e $-1 < \theta_2 < 1$;

Para $q > 2$, as condições de invertibilidade tornam-se mais complicadas, levando a necessidade de análises mais refinadas.

3.5.2.1) Modelos de Médias Móveis de Ordem 1 – MA (1)

Um modelo de Médias Móveis MA (1) – modelo ARMA (0,1) ou ARIMA (0,0,1) – pode ser representado da seguinte forma através da utilização da notação de operadores:

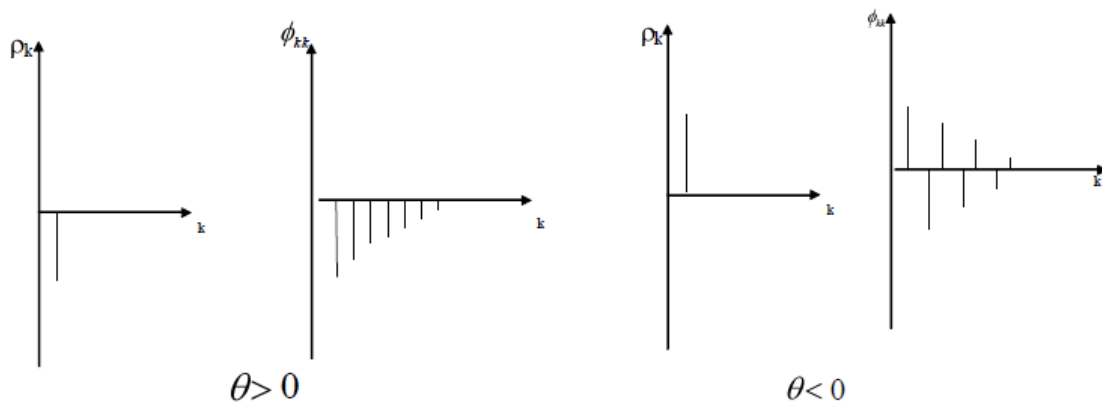
$$Y_t = \mu + (1 - \theta_1 B)a_t$$

$$Y_t = \mu + a_t - \theta_1 a_{t-1}$$

O processo é estacionário com média (μ) e a_t é um processo Ruído Branco com média zero, $E(a_t) = 0$, e variância constante, $\text{Var}(a_t) = \sigma^2$. Além disso, o parâmetro θ_1 deve variar entre -1 e 1 para satisfazer a condição de invertibilidade.

Os correlogramas da Função de Autocorrelação (FAC) e da Função de Autocorrelação Parcial (FACP) de um processo MA (1) possuem configurações especiais diferentes daquelas observadas nos processos autorregressivos (Figura 5). Na FAC de um processo MA (1), observa-se uma única autocorrelação estatisticamente diferente de zero na primeira defasagem. Todas as outras se situam próximas de zero. Já na FACP de um processo MA (1) as autocorrelações parciais decaem “exponencialmente” para zero e, dependendo do sinal do coeficiente do termo de Média Móvel (θ_1), esse decaimento poderá ocorrer de forma alternada.

Figura 5 – Correlogramas representativos da FAC e FACP dos processos MA (1)



Fonte: TOSCANO (2012).

3.5.2.2) Modelos de Médias Móveis de Ordem 2 – MA (2)

Um modelo de Média Móveis MA (2) – modelo ARMA (0,2) ou ARIMA (0,0,2) – pode ser representado da seguinte forma através da utilização da notação de operadores:

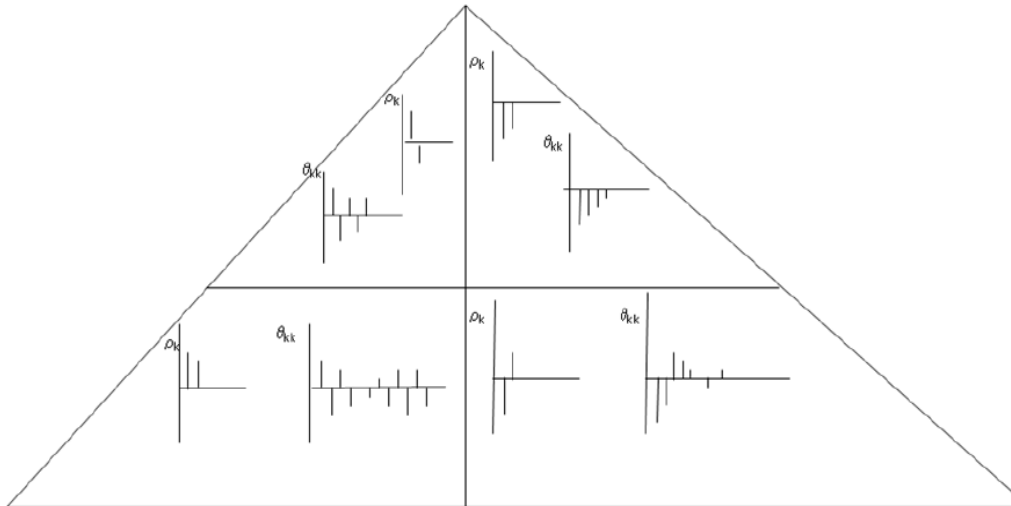
$$Y_t = \mu + (1 - \theta_1 B - \theta_2 B^2)a_t$$

$$Y_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$

O processo é estacionário com média (μ) e a_t é um processo Ruído Branco com média zero, $E(a_t) = 0$, e variância constante, $\text{Var}(a_t) = \sigma^2$. Além disso, a restrição para os parâmetros θ_1 e θ_2 para que a condição de invertibilidade seja observada é a seguinte: $\theta_1 + \theta_2 < 1$, $\theta_2 - \theta_1 < 1$ e $-1 < \theta_2 < 1$.

A FAC de um processo MA (2) apresenta as duas primeiras autocorrelações estatisticamente diferentes de zero. Todas as outras se situam próximas de zero não sendo, portanto, significativas. Já a FACP de um processo MA (2) pode apresentar um padrão de decaimento “exponencial senoidal”, decaimento “exponencial” e “senoidal alternado”, dependendo dos sinais dos coeficientes dos termos de Média Móvel (θ_1 e θ_2). A Figura 6 representa as características da FAC e FACP de um processo MA (2):

Figura 6 – Correlogramas representativos da FAC e FACP dos processos MA (2)



Fonte: TOSCANO (2012).

3.5.3) Modelos Autorregressivos e de Médias Móveis de Ordem (p,q) – ARMA (p,q)

Um modelo Autorregressivo e de Médias móveis de ordem (p, q) – ARMA (p, q) ou ARIMA (p, 0, q) – pode ser representado da seguinte forma:

$$Y_t - \phi_1 Y_{t-1} - \phi_p Y_{t-p} = \theta_0 + a_t - \theta_1 a_{t-1} - \theta_q a_{t-q}$$

O modelo ARMA (p, q) combina, portanto, características dos modelos autorregressivos ao utilizar a variável Y_t defasada no tempo e características dos modelos de médias móveis onde explicita os sucessivos erros defasados no tempo.

Na prática, este tipo de modelo é utilizado quando os modelos autorregressivos se mostram inadequados para descrever a série e quando a utilização de um modelo de médias móveis puro parece não ser natural ou intuitiva (MORETTIN; TOLOI, 1987).

O modelo ARMA (p, q) também é estacionário com média (μ), tal que $\theta_0 = \mu(1 - \phi_1 - \phi_2 - \dots - \phi_p)$ e a_t é um processo Ruído Branco com média zero, $E(a_t) = 0$,

e variância constante, $\text{Var}(a_t) = \sigma^2$. A seguir temos a representação de um modelo ARMA (1,1) utilizando a notação de operadores para facilitar o entendimento deste tipo de modelo:

$$(1 - \phi_1 B)Y_t = \theta_0 + (1 - \theta_1 B)a_t$$

$$Y_t = \theta_0 + \phi_1 Y_{t-1} + a_t - \theta_1 a_{t-1}$$

3.5.4) Modelos Autorregressivos Integrados e de Médias Móveis de Ordem (p, d, q) – ARIMA (p, d, q)

Um modelo Autorregressivo Integrado e de Médias móveis de ordem (p, d, q) – ARIMA (p, d, q) – representa, de maneira geral, os modelos anteriores AR, MA e ARMA quando não há necessidade de aplicar diferenças na série para que a mesma se torne estacionária. Nesses casos tem-se que $d = 0$. Todavia, existem processos conhecidos como “lineares não-estacionários homogêneos”, que supõem que “as séries são não-estacionárias em nível e/ou em inclinação. Estas séries podem tornar-se estacionárias através de um número finito (geralmente um ou dois) de diferenças” (MORETTIN; TOLOI, 1987, p. 30). Portanto, na maioria dos casos têm-se $d = 1$ ou $d = 2$ que representam os dois casos mais comuns de não-estacionariedade homogênea:

“a) Séries não-estacionárias quanto ao nível: oscilam ao redor de um nível médio durante algum tempo e depois saltam para outro nível temporário. Para torna-las estacionárias é suficiente tomar uma diferença; este é o caso típico de séries econômicas; b) Séries não-estacionárias quanto à inclinação: oscilam numa direção por algum tempo e depois mudam para outra direção temporária. Para torna-las estacionárias é necessário tomar a segunda diferença” (MORETTIN; TOLOI, 1987, p. 205).

A título de curiosidade, representaremos abaixo alguns modelos ARIMA com $d \neq 0$, já que os modelos em que não há necessidade de aplicar diferenças ($d = 0$) já foram apresentados anteriormente. De novo, o entendimento da notação dos operadores é fundamental para interpretar as equações dos modelos:

$$\text{ARIMA (0,1,1): } (1 - B)Y_t = \theta_0 + (1 - \theta_1 B)a_t$$

$$\text{ARIMA (1,1,0): } (1 - \phi_1 B)(1 - B)Y_t = \theta_0 + a_t$$

$$\text{ARIMA (1,1,1): } (1 - \phi_1 B)(1 - B)Y_t = \theta_0 + (1 - \theta_1 B)a_t$$

$$\text{ARIMA (3,1,0): } (1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3)(1 - B)Y_t = \theta_0 + a_t$$

$$\text{ARIMA (0,2,1): } (1 - B)^2 Y_t = \theta_0 + (1 - \theta_1 B)a_t$$

3.6) Modelo ARIMA Sazonal – SARIMA (p, d, q) (P, D, Q)_s

O modelo ARIMA mostrado anteriormente pode ser generalizado por um modelo SARIMA (p, d, q) (P, D, Q)_s com a consideração da sazonalidade (s) observada na série. A simbologia (P, D, Q)_s informa, respectivamente, a ocorrência de um processo autorregressivo sazonal (P), a aplicação de diferenças sazonais na série (D), e a ocorrência de um processo de média móvel sazonal (Q).

Não existe, em princípio, nenhuma dificuldade adicional para a construção dos modelos sazonais em comparação aos tradicionais modelos ARIMA. A diferença é que temos que diferenciar a série em relação a ela mesma (1 – B) e considerar também a diferença sazonal (1 – B^s), para produzir estacionariedade e assim encontrar os valores de d e D.

Obtendo a série estacionária, basta inspecionar a FAC e a FACP amostrais da série nos “lags” 1, 2, 3, ... para obter os valores iniciais de p e q; e nos “lags” sazonais s, 2s, 3s, ... para obter os valores de P e Q iniciais. Tem-se assim a formulação de um primeiro modelo tentativo SARIMA para descrever a série (MORETTIN; TOLOI, 1987).

A título de curiosidade, representaremos abaixo alguns modelos SARIMA considerando que os dados foram observados mensalmente e a sazonalidade é indicada, portanto, com um período s = 12:

$$\text{SARIMA } (0,1,0)(0,1,1)_{12}: (1 - B)(1 - B^{12})Y_t = \theta_0 + (1 - \theta_{12}B^{12})a_t$$

$$\text{SARIMA } (0,1,0)(1,1,0)_{12}: (1 - \phi_{12}B^{12})(1 - B)(1 - B^{12})Y_t = \theta_0 + a_t$$

$$\text{SARIMA } (0,1,1)(0,1,1)_{12}: (1 - B)(1 - B^{12})Y_t = \theta_0 + (1 - \theta_1B)(1 - \theta_{12}B^{12})a_t$$

$$\text{SARIMA } (0,1,0)(1,1,1)_{12}: (1 - B)(1 - B^{12})(1 - \phi_{12}B^{12})Y_t = \theta_0 + (1 - \theta_{12}B^{12})a_t$$

3.7) Escolha e verificação dos modelos construídos

3.7.1) Verificação dos modelos

Como podemos saber se um modelo construído se ajusta razoavelmente aos dados? Um diagnóstico simples é obter os resíduos e inspecionar as funções de autocorrelação e de autocorrelação parcial desses resíduos. Se nenhuma das autocorrelações, geral e parcial, são significativas do ponto de vista estatístico, há indícios fortes de que o vetor de resíduos seja puramente aleatório e, portanto, um ruído branco.

Todavia, ao invés de verificar (ou testar) a significância estatística de qualquer coeficiente de autocorrelação de forma individual, pode-se testar a hipótese conjunta de que todos os ρ_k até uma certa defasagem são simultaneamente iguais a zero (GUJARATI, 2006). De fato, a estatística Q de Box e Pierce permite este tipo de análise, em que n = tamanho da amostra e m = tamanho da defasagem:

$$Q = n \cdot \sum_{k=1}^m (\rho_k)^2$$

Essa estatística é utilizada frequentemente para testar se a série temporal representa um ruído branco e, em amostras grandes, ela se aproxima de uma distribuição qui-quadrado com m graus de liberdade. Quanto ao uso da estatística, a análise deve ser feita comparando o Q calculado com o Q crítico. Se o Q calculado excede o valor crítico de Q na distribuição qui-quadrado ao nível de significância escolhido, pode-se rejeitar a hipótese nula de que todos os (verdadeiros) ρ_k são iguais a zero e, neste caso, têm-se que pelo menos um deles deverá ser diferente de zero.

Uma alternativa à estatística Q de Box e Pierce é a estatística de Ljung e Box (LB), também conhecida como estatística de Box e Pierce modificada:

$$LB = n \cdot (n + 2) \cdot \sum_{k=1}^m \left(\frac{(\rho_k)^2}{n - k} \right) \sim \chi^2 m$$

De acordo com Gujarati (2006), em amostras grandes, tanto a estatística Q quanto a estatística LB seguem a distribuição qui-quadrado com m graus de liberdade. Todavia, o autor argumenta que a estatística LB tem propriedades melhores para pequenas amostras comparativamente a estatística Q. A interpretação da estatística LB é a mesma da estatística Q: quando o p-valor obtido for maior que um determinado α , não se rejeita a hipótese nula ($H_0: \rho_1 = \rho_2 = \rho_3 = \dots = \rho_k = 0$) e, portanto, é um indicativo da ocorrência de um ruído branco.

Outra estatística utilizada para a verificação de autocorrelação nos resíduos é a estatística de Durbin-Watson (DW). Na prática, o teste de Durbin-Watson testa se a primeira autocorrelação é nula verificando se ρ é igual a zero na seguinte equação do vetor de erros defasados: $\varepsilon_t = \rho \cdot \varepsilon_{t-1} + a_t$.

Assim, se $\rho = 0$, tem-se que $a_t = \varepsilon_t$ e nenhum problema existiria. Porém, como não se conhece a distribuição de probabilidade de seu estimador ρ , cria-se a seguinte estatística com o intuito de verificar o problema em questão, onde e_t corresponde aos

valores dos resíduos estimados pelo modelo em determinado instante de tempo e n é o número de observações:

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n (e_t)^2}$$

É possível demonstrar por meio de uma manipulação matemática que a relação existente entre a estatística DW e o valor de ρ é a seguinte: $DW \cong 2(1 - \rho)$. Assim, caso não ocorra autocorrelação dos resíduos, a estatística DW possui valor próximo de “dois” (FÁVERO, 2009). Vale lembrar também a representação das hipóteses do teste de Durbin-Watson: $H_0: \rho = 0$; $H_1: \rho \neq 0$.

Além das estatísticas testes de Box e Pierce (Q), Ljung e Box (LB) e de Durbin-Watson (DW), outras análises sobre o comportamento dos resíduos podem ser efetuadas para além da presença (ou ausência) de autocorrelação e, portanto, da ocorrência ou não de ruído branco. Pode-se, por exemplo, verificar se os resíduos padronizados situam dentro de um intervalo “apropriado”: -3,1 e 3,1. Essa análise pode ser importante para detecção de elementos discrepantes (como outliers). Além disso, no próprio gráfico de resíduos pode-se verificar a validade da suposição de que os erros são ou não correlacionados: a presença de configurações especiais neste gráfico pode indicar que os erros são correlacionados. Pode-se também verificar a suposição de normalidade dos resíduos através do seguinte teste: H_0 : os erros seguem distribuição Normal; H_1 : os erros não seguem distribuição Normal. Se, no teste de normalidade aplicado, o p-valor for maior que um determinado α , não se rejeita a hipótese nula. A importância da hipótese de normalidade será explicada de forma mais detalhada mais adiante, mas em todo caso, ela é importante para as conclusões acerca dos parâmetros estimados pelo modelo.

Além da análise dos resíduos, deve-se também verificar se os parâmetros estimados pelo modelo (ϕ 's e θ 's) são estatisticamente diferentes de zero. Para isso, pode-se efetuar os seguintes testes de hipótese:

$$H_0: \phi's = 0 ; H_1: \phi's \neq 0$$

$$H_0: \theta's = 0 ; H_1: \theta's \neq 0$$

Se, nos testes efetuados, os p-valores forem menores que um determinado α (que determina o nível de significância), rejeita-se a hipótese nula e, portanto, conclui-

se que os parâmetros estimados são estatisticamente diferentes de zero. Ademais, lembrando a construção dos modelos autorregressivos e de medias móveis, os parâmetros obtidos devem satisfazer também as condições de estacionariedade e de invertibilidade.

3.7.2) Escolha entre modelos

Entre os critérios utilizados para a escolha de modelos concorrentes e/ou para comparar modelos para fins de previsão dois critérios de decisão se destacam: o Critério de Informação de Akaike (CIA) e o Critério de Informação de Schwarz (CIS).

O Critério de Informação de Akaike impõe uma punição pelo acréscimo de regressores ao modelo ($2k/n$ é o fator punitivo). A sua fórmula de cálculo escrita em sua versão logarítmica é demonstrada abaixo, onde k corresponde ao número de regressores (incluindo o intercepto), n é o número de observações e SQR representa a soma dos quadrados dos resíduos (GUJARATI, 2006):

$$CIA = \left(\frac{2k}{n} \right) + \ln \left(\frac{SQR}{n} \right)$$

Já o Critério de Informação de Schwarz impõe uma punição ainda mais dura comparativamente ao Critério de Akaike. O fator de punição como se pode notar pela fórmula abaixo escrita em sua versão logarítmica é de $[(k/n) \cdot \ln(n)]$:

$$CIS = \left(\frac{k}{n} \right) \cdot \ln(n) + \ln \left(\frac{SQR}{n} \right)$$



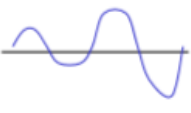



Ao comparar dois ou mais modelos utilizando o CIA ou o CIS, o preferido será aquele que apresentar o menor valor para o critério de informação. Uma vantagem do cálculo desses critérios é que os mesmos podem ser utilizados para avaliar o desempenho dos modelos dentro e fora da amostra⁹ (GUJARATI, 2006). Todavia, nesta monografia, o seu uso se dará preferencialmente para comparar modelos dentro da amostra. Para comparar os modelos fora da amostra, serão utilizados os erros de previsão para a construção de algum tipo de “função perda”, como por exemplo, o erro quadrático médio (MORETTIN; TOLOI, 1987).

⁹ A análise “dentro” da amostra se refere ao conjunto de observações utilizadas para estimar os modelos, ou seja, corresponde a construção no período amostral. A análise “fora” da amostra refere-se à utilização do modelo para fazer previsões e também é conhecida como análise no período de validação.

3.8) Alisamento Exponencial de Holt-Winters

A técnica de alisamento exponencial é utilizada frequentemente por ser simples, precisa e de fácil entendimento conceitual por parte do usuário. De forma geral, existem dois tipos de métodos (multiplicativo e aditivo) de aplicação da técnica de alisamento exponencial de Holt-Winters, que depende das características da série. O método de alisamento aditivo é recomendado quando a sazonalidade da série também é aditiva. Já o método de alisamento multiplicativo deve ser utilizado quando as amplitudes sazonais variam com o tempo. A Figura 7 abaixo ilustra as características da série e o método recomendado para a realização da análise:

Figura 7 – Características da série em relação à sazonalidade, nível e tendência e o método de alisamento exponencial recomendado

	<i>Sem Sazonalidade</i>	<i>Sazonalidade Aditiva</i>	<i>Sazonalidade Multiplicativa</i>
<i>Nível Constante</i>	 <p>Alisamento Exponencial Simples</p>	 <p>Alisamento Exponencial Winter</p>	
<i>Tendência Linear</i>	 <p>Alisamento Exponencial Duplo Holt</p>		

Fonte: TOSCANO (2012).

A utilização do método baseia-se em três equações com constantes de alisamento diferentes e que são associadas a cada uma das componentes da série: nível, tendência e sazonalidade (MORETTIN; TOLOI, 1987). As fórmulas a seguir trazem as três equações que resumem o método multiplicativo e o método aditivo. Note que L_t representa o alisamento do nível até o final do instante t e α corresponde a constante (peso) de alisamento do nível; T_t simboliza o alisamento da tendência até o final do instante t , sendo γ a constante (peso) de alisamento da tendência; S_t representa o alisamento do fator sazonal até o instante t e δ é a constante (peso) da componente sazonal; Y_t corresponde a observação da série no instante t ; \hat{Y}_t simboliza a equação de previsão e p representa o período sazonal.

Método Multiplicativo

$$L_t = \alpha (Y_t / S_{t-p}) + (1-\alpha) [L_{t-1} + T_{t-1}]$$

$$T_t = \gamma [L_t - L_{t-1}] + (1 - \gamma)T_{t-1}$$

$$S_t = \delta (Y_t / L_t) + (1 - \delta) S_{t-p}$$

$$\hat{Y}_t = (L_{t-1} + T_{t-1}) S_{t-p}$$

Método Aditivo

$$L_t = \alpha (Y_t - S_{t-p}) + (1 - \alpha) [L_{t-1} + T_{t-1}]$$

$$T_t = \gamma [L_t - L_{t-1}] + (1 - \gamma)T_{t-1}$$

$$S_t = \delta (Y_t - L_t) + (1 - \delta) S_{t-p}$$

$$\hat{Y}_t = L_{t-1} + T_{t-1} + S_{t-p}$$

Tendo em vista a facilidade no que se refere ao uso da ferramenta de análise aqui apresentada, convém acrescentar algumas recomendações e sugestões quanto à utilização do método de alisamento exponencial. Em primeiro lugar, a técnica é recomendada quando o passado de uma sequência temporal é muito pequeno para encontrar o seu padrão de comportamento, já que um menor número de observações passadas da série é requerido para ajustar o modelo comparativamente a outros modelos. Também se recomenda a utilização da técnica nos seguintes casos:

“a) não existe um padrão definido do passado da série; b) os modelos não se ajustam bem ao passado da série; c) os modelos que se ajustam bem ao passado da série, podem não ajustar bem para prever os valores futuros da série; d) a série não apresenta um comportamento padrão; e) quando não se deseja depender de suposições fortes de outros métodos” (TOSCANO, 2012, p. 1-2).

Todavia, existem algumas restrições em relação às constantes (pesos) de alisamento, já que dependendo do valor de cada um dos pesos a série alisada pode ficar instável ou fora de controle. Por isso, adota-se normalmente o intervalo de [0,1] para cada uma das constantes de alisamento (α , γ , δ). De forma geral, os softwares estatísticos restringem os valores das constantes para que estes permaneçam em intervalos ainda menores (0,1 e 0,3). Quando α , por exemplo, tem valor bem próximo de “1” isto é um indicativo de que as informações do passado (nível da série) estão sendo descontadas rapidamente pelo modelo. Se α possui um valor maior do que “1”, o efeito de uma previsão errada poderá crescer com o passar do tempo.

Apesar dessas restrições, as constantes de alisamento podem exceder o valor da unidade e, mesmo assim, o processo permanecer estável. Por isso, alguns autores defendem o argumento de que o valor dos pesos deve ser extraído dos dados de forma a minimizar os erros (GARDNER, 1985 *apud* TOSCANO, 2012). Em todo caso, nesta monografia, será dispendido um esforço de extrair os valores das constantes por meio dos próprios dados, mas adotando o intervalo [0,1] como referência.

A principal desvantagem da técnica de alisamento exponencial é que a mesma não faz nenhuma suposição acerca da distribuição dos dados. Assim, não se pode ter

certeza de que o modelo correto foi encontrado e também só é possível a construção de intervalos de confiança aproximados para as previsões. Mesmo assim, a necessidade de testes permanece pequena, já que prevalecem evidências de que os métodos de alisamento exponencial são incrivelmente robustos (MAKRIDAKIS; HIBON, 1989 *apud* TOSCANO, 2012).

3.9) Regressão Dinâmica:

Os modelos de regressão linear múltipla tradicionalmente estudados supõe que os erros do modelo possuem algumas características típicas como: média zero, variância constante, distribuição normal e independência, pressupondo, portanto, a inexistência de correlação serial (ZANINI, 2000). Estes e outros pressupostos do modelo de regressão linear podem ser sintetizados no Quadro 2 abaixo:

Quadro 2 – Pressupostos do Modelo de Regressão Linear

	Pressupostos	Violações
1	A variável dependente é uma função linear de um conjunto específico de variáveis e do erro	Regressores inadequados e não-linearidade
2	O valor esperado do termo de erro é zero	Estimadores viesados
3	O erro tem distribuição normal e não apresenta autocorrelação ou correlação com qualquer variável X	Heterocedasticidade (o erro não possui a mesma variância) e autocorrelação dos resíduos
4	As observações das variáveis explicativas podem ser consideradas fixas em amostras repetidas	Erros de levantamento ou medida das variáveis
5	Não existe relação linear exata entre as variáveis explicativas e existem mais observações do que variáveis explicativas	Multicolinearidade (presença de variáveis explicativas com comportamento semelhantes ou apresentando relação linear entre si)

Fonte: Adaptado de FÁVERO (2009) *apud* KENNEDY (2003).

A hipótese de normalidade dos resíduos pode ser explicada por diferentes aspectos. Em primeiro lugar, espera-se que a influência de variáveis omitidas ou esquecidas na modelagem seja pequena e, quando muito, aleatória. Pelo Teorema do Limite Central, pode-se demonstrar que, se houver uma grande quantidade de variáveis aleatórias independentes e identicamente distribuídas, então, com raras exceções, a distribuição de sua soma tenderá para uma distribuição normal à medida que o número de variáveis aumentarem indefinidamente. Além disso, uma variante do Teorema do Limite Central ainda acrescenta que, mesmo que o número das variáveis não seja grande ou que essas variáveis não sejam absolutamente independentes, ainda assim a sua soma pode-se distribuir normalmente (GUJARATI, 2000).

Além disso, com a hipótese de normalidade, as distribuições de probabilidade dos estimadores (parâmetros do modelo) obtidos pelo método dos mínimos quadrados

ordinários podem ser facilmente derivadas, já que uma propriedade da distribuição normal é que qualquer função linear de variáveis distribuídas normalmente é ela própria distribuída normalmente. Assim, sob a hipótese de normalidade, os parâmetros do modelo (por exemplo, β_1 e β_2) também se distribuem normalmente. Em função disso, esses parâmetros são não-viesados, possuem variância mínima e apresentam consistência (com o aumento do tamanho da amostra convergem para os verdadeiros valores na população) (GUJARATI, 2000).

Por isso, a hipótese de normalidade é importante para a realização de inferências estatísticas (testes de hipóteses) e para a construção de intervalos de confiança no que diz respeito aos estimadores obtidos. Como no caso dos parâmetros encontrados quando do tratamento dos modelos ARIMA, também no caso da regressão linear múltipla deve-se verificar a existência do modelo e se os parâmetros obtidos através da modelagem foram significativos.

Para testar a existência do modelo o seguinte teste pode ser efetuado: $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$; H_1 : existe pelo menos um $\beta_i \neq 0$. Esse teste é conhecido como teste F (refere-se a distribuição F) e avalia a significância conjunta das variáveis explicativas. Também pode-se testar individualmente a significância dos parâmetros através de um teste como: $H_0: \beta_1 = 0$; $H_1: \beta_1 \neq 0$. Neste caso, a distribuição de referência para a realização do teste é a distribuição t de student. Na maior parte dos casos não é necessário calcular as estatísticas (F e t) a que os testes estão relacionados, pois os pacotes estatísticos já trazem a informação do p-valor em sua saída.¹⁰ Assim, se o p-valor mostrado for menor do que um determinado α (que traduz o nível de significância), rejeita-se a hipótese nula e os coeficientes obtidos podem ser considerados como parâmetros significativos (FÁVERO, 2009). Nesta monografia, o valor para o nível de significância adotado será sempre de 5% ($\alpha = 0,05$).

Outra informação utilizada como grau de ajuste do modelo de regressão linear múltipla é o R-quadrado (R^2), definido como a fração da variância da amostra de Y_i explicada pelos regressores. Em outras palavras, o R^2 é traduzido como a proporção da variação amostral da variável dependente explicada pelas variáveis explicativas. A fórmula abaixo representa o cálculo do R-quadrado, onde SQR representa a soma dos quadrados dos resíduos; SQE corresponde a soma dos quadrados da regressão e STQ é a soma total dos quadrados:

¹⁰ Caso seja necessário calcular as estatísticas F e t, ver o livro *Análise de dados: modelagem multivariada para tomada de decisões* de FÁVERO (2009) pág. 354 e 355.

$$R^2 = \frac{SQE}{STQ} = 1 - \frac{SQR}{STQ}$$

Segundo Gujarati (2006), o R^2 mede a qualidade do ajustamento do modelo dentro da amostra, ou seja, a proximidade entre os valores estimados de Y e seu valor observado na amostra analisada. Todavia, não existe garantia de que o R-quadrado elevado garanta uma boa previsão das observações fora da amostra. Na prática, o R^2 pode variar entre 0 e 1 (0 a 100%). Valores elevados do R-quadrado (próximo de 1) indica um grau de ajustamento do modelo satisfatório dentro da amostra.

Entretanto, quando se tem a intenção de comparar o coeficiente de ajuste entre dois modelos quaisquer ou entre um mesmo modelo com tamanho de amostra diferente, faz-se necessário a utilização do R-quadrado ajustado, que é uma medida do R^2 da regressão ajustado pelo número de graus de liberdade (FÁVERO, 2009). A expressão do R^2 -ajustado é demonstrada a seguir, onde n representa o tamanho da amostra e k corresponde ao número de parâmetros do modelo de regressão (número de variáveis explicativas mais o intercepto):

$$R^2_{\text{ajust}} = 1 - (1 - R^2) \left(\frac{n-1}{n-k} \right)$$

Após essa breve descrição acerca dos modelos tradicionais de regressão linear múltipla, convém focalizar a análise nos modelos de regressão dinâmica. Um dos principais problemas ao se tentar modelar uma série temporal através de um modelo de regressão tradicional é que a hipótese de independência dos ruídos não é realista e, com isso, os testes e resultados oriundos do modelo deixam de ser válidos (BARROS; SOUZA, 1995 *apud* DIAS, 2008). Algumas das consequências da presença de autocorrelação dos resíduos são:

“1)Os estimadores usuais por mínimos quadrados são ainda não tendenciosos, mas não têm variância mínima; 2)os estimadores da variância e dos erros padrões dos coeficientes da regressão são subestimados, o que levaria à conclusão de que os estimadores são mais precisos do que na realidade; 3)os intervalos de confiança para os parâmetros da regressão e os testes de hipóteses relacionados a estes intervalos perdem a validade (como uma consequência direta de 2)” (ZANINI, 2000, p. 21).

Assim, pode-se chegar a conclusões equivocadas se ignoradas as consequências citadas acima e devem-se procurar alternativas para que os problemas sejam superados. Uma das possíveis soluções é exatamente a utilização dos modelos de regressão dinâmica (DIAS, 2008).

Os modelos de regressão dinâmica consideram, além da variável de interesse e os seus valores defasados, também o efeito de variáveis explicativas (causais ou exógenas) e seus valores defasados. Atenta-se ainda que a expressão “regressão dinâmica” não indica que os parâmetros do modelo evoluem no tempo. Na verdade, a palavra “dinâmica” significa um modelo de regressão em que foi incluída a estrutura de dependência de uma série temporal (ZANINI, 2000).

A estimação dos parâmetros de um modelo de regressão dinâmica é obtida através do método dos mínimos quadrados ordinários tal como acontece nos modelos de regressão usuais. Porém, a estimação para um modelo de regressão dinâmica é mais complexa e envolve um processo iterativo com várias etapas (ZANINI, 2000).

Com relação à estrutura dos modelos de regressão dinâmica percebe-se que a variável endógena Y_t no modelo é explicada pelas suas defasagens e por uma ou mais variáveis causais (exógenas), como, por exemplo, a variável X_t e suas defasagens. A equação abaixo traduz a representação de um modelo de regressão dinâmica com essas características:

$$Y_t = \beta_0 X_t + \beta_1 X_{t-1} + \dots + \beta_k X_{t-k} + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_m Y_{t-m} + \varepsilon_t$$

O modelo pode ser reescrito utilizando a notação de vetores e de operadores de defasagem, onde Y_t é a variável dependente (endógena) no instante t ; β é o vetor de coeficientes das variáveis causais que será estimado por mínimos quadrados; X_t é o vetor de variáveis causais e exógenas incluindo as suas possíveis defasagens; ε_t é o ruído aleatório associado ao modelo, onde se supõe que estes sejam independentes e identicamente distribuído com distribuição normal $(0, \sigma^2)$ e $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_m B^m$ é um polinômio autorregressivo, com B sendo o operador de defasagem (DIAS, 2008):

$$\varphi(B)Y_t = \beta X_t + \varepsilon_t$$

Este modelo de regressão dinâmica ainda pode ser considerado como um caso particular do que é comumente conhecido na literatura como modelos de Cochrane e Orcutt generalizados, dado por:

$$\varphi(B)Y_t = \beta X_t + W_t$$

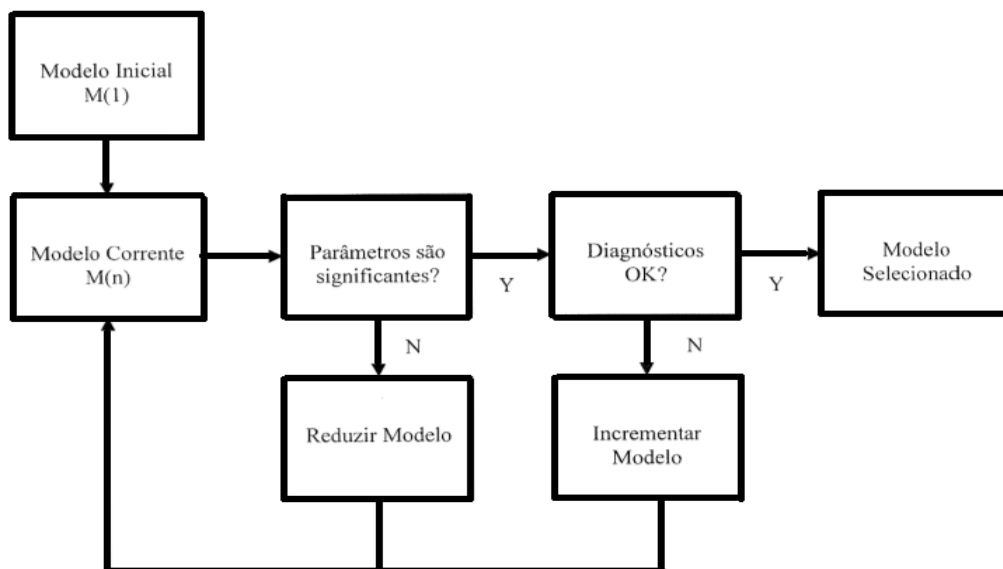
$$R(B)W_t = \varepsilon_t$$

Nota-se que este modelo tem a mesma estrutura que o caso anterior (primeira equação). Entretanto, os erros W_t apresentam uma estrutura autorregressiva dada pela segunda equação e representada pelo polinômio autorregressivo $R(B)$ (DIAS, 2008).

Percebe-se, portanto, que a diferença básica entre os modelos de regressão dinâmica e os modelos ARIMA tratados anteriormente é justamente a inclusão do vetor de variáveis causais βX_t . De fato, nos modelos ARIMA, apenas os valores passados da série Y_t e os valores defasados da série de erros são usados na modelagem e na previsão de Y_t (DIAS, 2008).

Com relação à construção de modelos de regressão dinâmica, a estratégia normalmente utilizada é uma estratégia *bottom-up*, isto é, parte-se de um modelo simplificado incrementando-o até que seja encontrado um modelo adequado e satisfatório. Esse processo é, muitas vezes, custoso e demorado, pois envolve a escolha de variáveis explicativas e suas defasagens. Além disso, não basta levar em consideração apenas a significância dos parâmetros encontrados. É preciso avaliar se o modelo encontrado é consistente com a teoria econômica, ou seja, se os coeficientes obtidos são coerentes (ZANINI, 2000). A Figura 8 abaixo descreve, de maneira genérica, os passos recomendados para a construção de um modelo de regressão dinâmica:

Figura 8 – Construção de um modelo de Regressão Dinâmica



Fonte: ZANINI (2000).

Para responder as perguntas presentes na Figura 8 sobre a significância dos parâmetros e diagnósticos, até que se alcance um modelo apropriado, são utilizados

testes com o objetivo de definir a especificação do modelo explicativo, encontrar a dinâmica do modelo (com a inclusão ou não de variáveis defasadas) e verificar o ajuste do modelo (DIAS, 2008).

Na prática, com a obtenção de um modelo dito satisfatório (parâmetros significativos e coerentes e também com a inclusão de variáveis explicativas e suas defasagens na medida necessária), a verificação do modelo também pode ser efetuada através de testes ou gráficos que identifiquem a presença ou não de autocorrelações nos resíduos. A presença de autocorrelações significativas nos resíduos pode indicar uma das seguintes situações, já que algum tipo de estrutura presente na série Y_t não foi captada pelo modelo:

“1)Deve-se incluir mais “lags” da variável dependente; 2)deve-se incluir “lags” adicionais das variáveis exógenas já presentes no modelo ou incluir novas variáveis causais” (BARROS; SOUZA, 1995 apud ZANINI, 2000, p. 34).

Por fim, um último aspecto que merece ser abordado no que se refere à construção de modelos de regressão (múltipla ou dinâmica) diz respeito à presença de multicolinearidade, isto é, a presença de regressores correlacionados no modelo. Em relação a esse aspecto Gujarati (2006) argumenta que, em casos de alta colinearidade, as principais consequências para a modelagem são:

“1)Embora sejam melhores estimadores lineares não tendenciosos, os estimadores do método dos Mínimos Quadrados Ordinários (MQO) têm grande variância e covariância, tornando difícil uma estimação exata; 2)em decorrência da consequência 1, os intervalos de confiança tendem a ser muito mais amplos, facilitando a aceitação da “hipótese nula igual a zero” (isto é, que o coeficiente populacional verdadeiro seja igual a zero); 3)também como efeito de 1, a razão t de um ou mais coeficientes tende a ser estatisticamente insignificante; 4)Embora a razão t de um ou mais coeficientes seja estatisticamente insignificante, R^2 , a medida geral da qualidade do ajustamento, pode ser muito alto; 5)Os estimadores de MQO e seus erros-padrão podem ser sensíveis a pequenas alterações nos dados” (GUJARATI, 2006, p. 282).

Uma medida utilizada como indicador de multicolinearidade é o Fator de Inflação da Variância (VIF), em que se observa a velocidade com que as variâncias e covariâncias aumentam. O VIF é definido como:

$$VIF = \frac{1}{(1 - R_j^2)}$$

Assim, a medida que o R_j^2 , o coeficiente de determinação da regressão do regressor X_j em relação aos regressores restantes do modelo, aumenta no sentido de se aproximar da unidade, isto é, a medida que a colinearidade de X_j aumenta em

relação aos demais regressores, o VIF também cresce e, no limite, pode ser infinito. Assim, quanto maior o valor do VIF mais problemática ou colinear será a variável X_j . Na maior parte das análises, se o VIF de uma variável for maior que “10”, considera-se que a variável em questão é altamente colinear. No entanto, esse número de referência poderá ser relativizado dependendo do tipo de análise (GUJARATI, 2006).

3.10) Comparação do desempenho dos modelos construídos

Além dos critérios de Akaike, Schwarz e o R-quadrado (sobretudo o ajustado) utilizados como critérios para seleções e comparações de modelos (GUJARATI, 2006), existem outras medidas de avaliação usadas para determinar a qualidade da modelagem dentro e fora da amostra que levam em consideração, respectivamente, o grau de precisão dos valores ajustados e das previsões (MARTÍNEZ; ZAMPROGNO, 2003). Estas medidas, também consideradas como “funções-perda” (MORETTIN; TOLOI, 1987), são o Erro Médio Absoluto (MAD), o Erro Percentual Médio Absoluto (MAPE) e o Erro Quadrático Médio (MSD).

O Quadro 3 sintetiza as três medidas de desempenho dos modelos construídos (MAD, MAPE e MSD), onde Y_t representa o valor verdadeiro (atual ou real) da observação no instante t , \hat{Y}_t corresponde ao valor ajustado ou de previsão no instante t e n corresponde ao número de observações obtidas. É importante lembrar que o uso dessas medidas como critérios de comparação da capacidade de ajuste ou preditiva dos modelos só fazem sentido se os modelos construídos satisfizerem as demais condições já detalhadas anteriormente nesta monografia.

Quadro 3 – Critérios de desempenho dos modelos através dos erros associados – MAD, MAPE e MSD

Critério	Notação	Expressão
Erro Médio Absoluto	MAD	$\frac{\sum_{t=1}^n y_t - \hat{y}_t }{n}$
Erro Percentual Médio Absoluto	MAPE	$\frac{\sum (y_t - \hat{y}_t)/y_t }{n} \times 100 \quad (y_t \neq 0)$
Erro Quadrático Médio	MSD	$\frac{\sum_{t=1}^n y_t - \hat{y}_t ^2}{n}$

Fonte: Elaboração Própria com base em MARTÍNEZ e ZAMPROGNO (2003).

CAPÍTULO 4: CARACTERÍSTICAS DA SÉRIE DE ÍNDICE DE VOLUME DO PIB DE MINAS GERAIS

A série encadeada do índice de volume do PIB de Minas Gerais é uma série trimestral iniciada no ano de 2002 e divulgada pela Fundação João Pinheiro. Essa instituição é, portanto, a responsável pelo cálculo e pela validação dos resultados da produção agregada no âmbito da economia mineira em um trabalho que é de parceria com o Instituto Brasileiro de Geografia e Estatística (IBGE), instituição que coordena o cálculo dos PIB's regionais junto aos órgãos estaduais de estatísticas (FUNDAÇÃO JOÃO PINHEIRO, 2013).

Os últimos resultados desta série trimestral podem sofrer pequenas revisões ao longo do tempo, sobretudo dentro de um intervalo que corresponde aos dois últimos anos, em razão de atualizações nas bases de dados e na estrutura de ponderação das atividades econômicas. Mesmo assim, as revisões não costumam ser significativas a ponto de alterar o diagnóstico econômico divulgado anteriormente (FUNDAÇÃO JOÃO PINHEIRO, 2013).

A última divulgação pela Fundação João Pinheiro da série do índice de volume trimestral de Minas Gerais faz referência ao 2º trimestre de 2013 e o resultado do 3º trimestre só será divulgado no final deste ano. Assim, a série de interesse a ser modelada neste trabalho se inicia no 1º trimestre de 2002 e se encerra no 2º trimestre de 2013. Vale lembrar também que a série de índice de volume do PIB de Minas Gerais é construída tendo o ano de 2002 como ano base, isto é, como a série se inicia neste ano nada mais natural do que ele ser o ano de referência (média de 2002 = 100).¹¹

Portanto, contando o número de observações do 1º trimestre de 2002 até o 2º trimestre de 2013, têm-se um total de 46 observações. Como forma de estruturação do trabalho, as seis últimas observações serão utilizadas para a validação dos modelos construídos no que se refere a sua capacidade de realizar previsões (análise fora da amostra). Assim, serão utilizadas 40 observações ao todo para a construção dos modelos e que fazem referência ao período amostral (análise dentro da amostra).

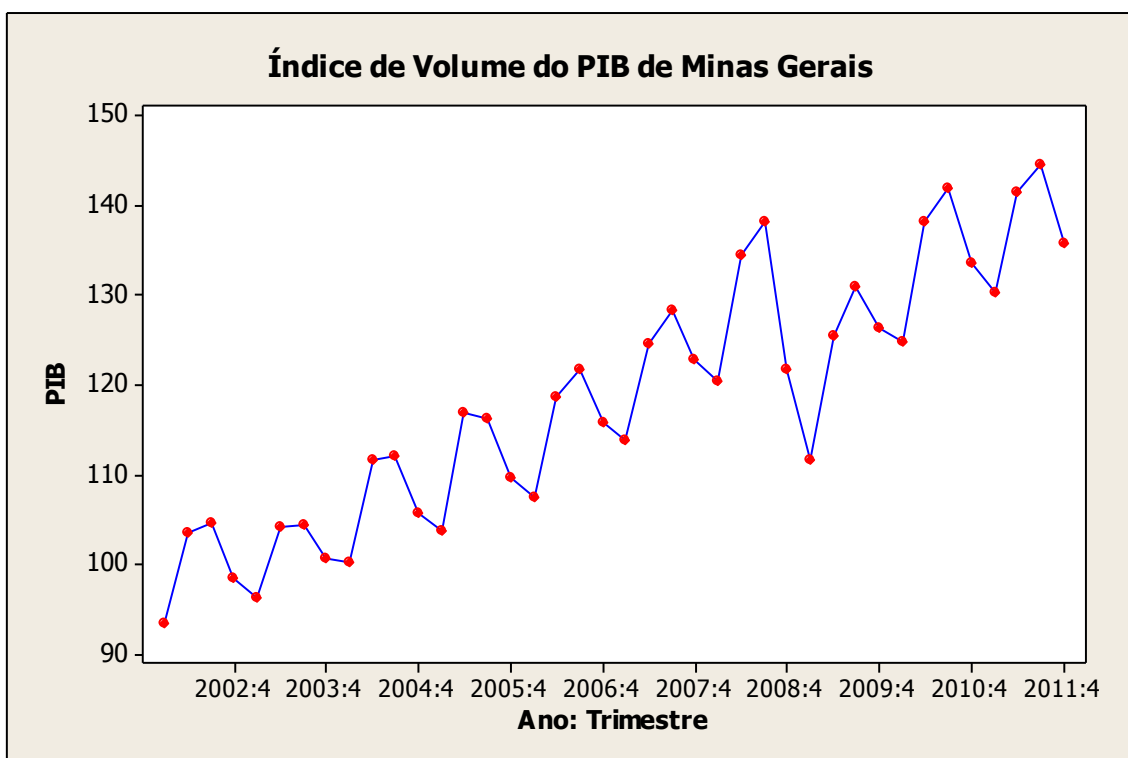
Embora a recomendação para a construção de modelos, sobretudo aqueles construídos com base na metodologia de Box-Jenkins, seja a utilização de, no mínimo,

¹¹ Ver Apêndice 1 ao final desta monografia que traz a série completa até então divulgada do índice de volume do PIB de Minas Gerais.

50 observações (SIQUEIRA, 2002), será realizado um esforço de modelagem utilizando apenas as 40 observações, já sabendo previamente que o tamanho reduzido da amostra poderá ser um obstáculo adicional para obtenção de modelos ditos satisfatórios. Por isso, serão utilizadas outras técnicas para além da metodologia de Box-Jenkins, como a técnica de alisamento exponencial e os modelos de regressão dinâmica.

Utilizando as 40 observações iniciais, que vão do 1º trimestre de 2002 até o 4º trimestre de 2011, nota-se que a série de índice de volume do PIB de Minas Gerais apresenta uma tendência crescente aparentemente linear ao longo do tempo. A série é não estacionária, tendo em vista que as observações não variam aleatoriamente em torno de um valor médio. O Gráfico 4 traduz essas características:

Gráfico 4 – Comportamento via gráfico de linhas da série Índice de Volume do PIB de Minas Gerais – 1º Trimestre de 2002-4º Trimestre de 2011



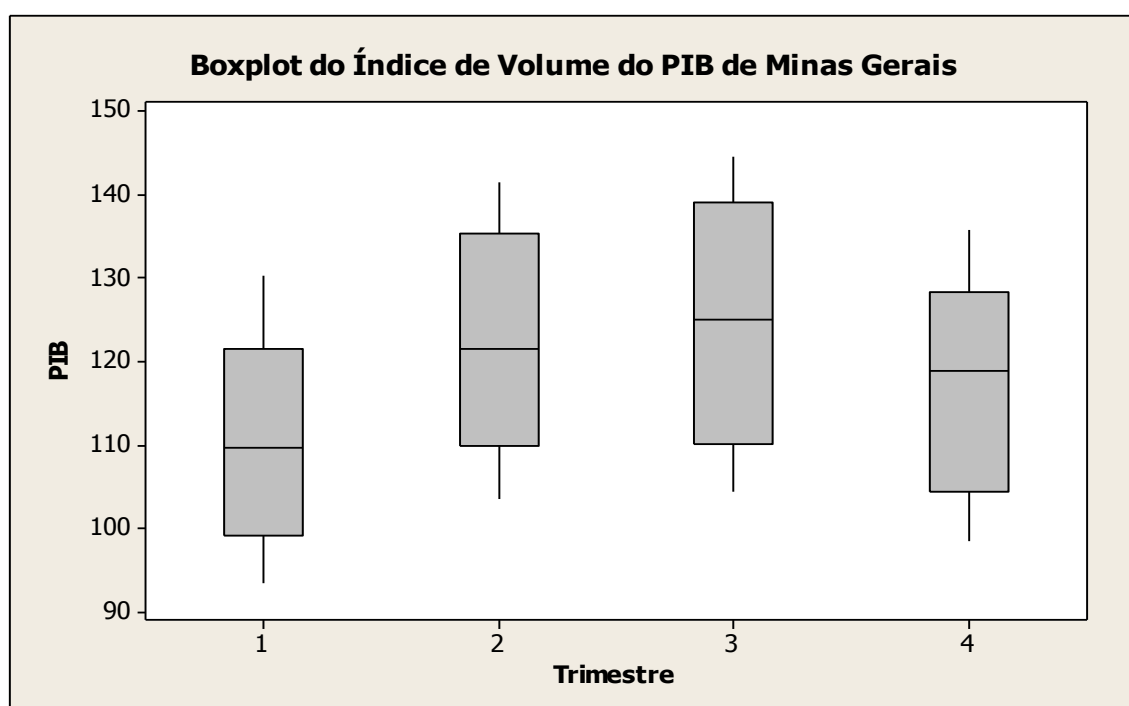
Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Como se pode observar no Gráfico 4, o ano de 2008 foi caracterizado por uma inflexão abrupta do nível de atividade econômica internacional e que foi considerada uma das crises econômicas mais graves desde a Grande Depressão de 1929, atingindo inclusive economias em desenvolvimento como o Brasil e, conseqüentemente, Minas Gerais.

Originada ainda em 2007 no mercado imobiliário norte-americano, a crise financeira se aprofundou e adquiriu contornos dramáticos no último trimestre de 2008, quando problemas de insolvência envolvendo algumas das mais importantes instituições financeiras dos Estados Unidos e da Europa desencadearam uma crise de confiança de escala global contra os sistemas financeiros. A falência do Lehman Brothers, um dos bancos de investimento mais importantes da economia norte-americana, representou a quebra de confiança e o estopim para que a crise tomasse proporções sistêmicas (IBGE, 2010).

O fluxo líquido de capitais privados destinados às economias emergentes a partir do último trimestre de 2008 e ao longo de 2009 sofreu uma queda acentuada, bem como o fluxo de transações comerciais. Foi neste contexto que a economia mineira foi significativamente afetada, sobretudo, por sua especificidade produtiva em *commodities* minerais e agrícolas. Portanto, a ocorrência da crise econômica de 2008-2009 será um fator de dificuldade adicional para a modelagem da série econômica do índice de volume do PIB de Minas Gerais.

Gráfico 5 – Boxplot do Índice de Volume do PIB de Minas Gerais – 1º Trimestre de 2002-4º Trimestre de 2011



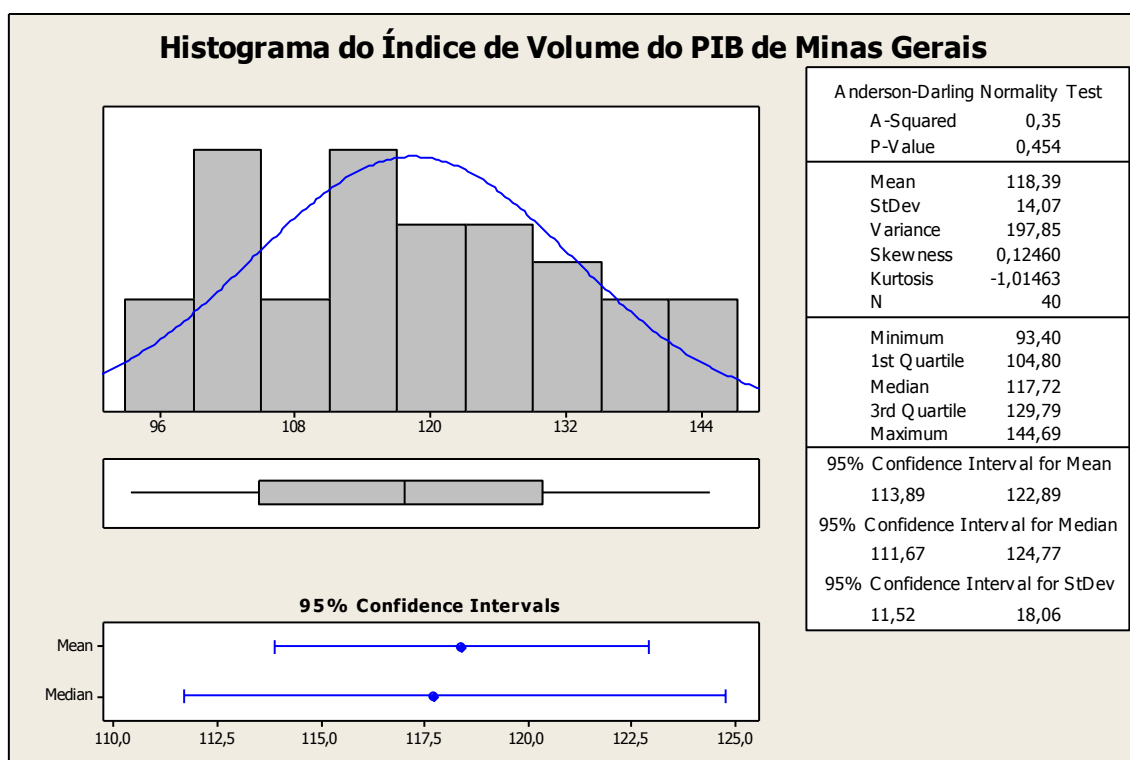
Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Analisando o Gráfico 5, nota-se que a série encadeada do índice de volume do PIB de Minas Gerais também possui uma marcada variação sazonal trimestral, uma vez que a mediana foi diferente para cada um dos trimestres do ano. Percebe-se que

a produção é mais acentuada no segundo e no terceiro trimestre de cada ano e mais baixa no primeiro e quarto trimestre.

Existem duas explicações principais para este fato: 1) a agropecuária mineira possui uma marcada variação sazonal correspondente à observada para o índice de volume do PIB de Minas Gerais, porque o principal produto da agricultura mineira (o café) é colhido, exatamente, no segundo e terceiro trimestre de cada ano; 2) o primeiro e o quarto trimestre de cada ano são, de forma geral, o período de chuvas no estado de Minas Gerais (sobretudo na região central do estado). Isto prejudica outro segmento com marcada variação sazonal na economia de Minas Gerais: a extração do minério de ferro. As chuvas prejudicam também outros segmentos como a construção civil e o comércio. Embora no caso do comércio este efeito seja suavizado com as compras de natal no quarto trimestre do ano.

Gráfico 6 – Histograma do Índice de Volume do PIB de Minas Gerais – 1º Trimestre de 2002-4º Trimestre de 2011

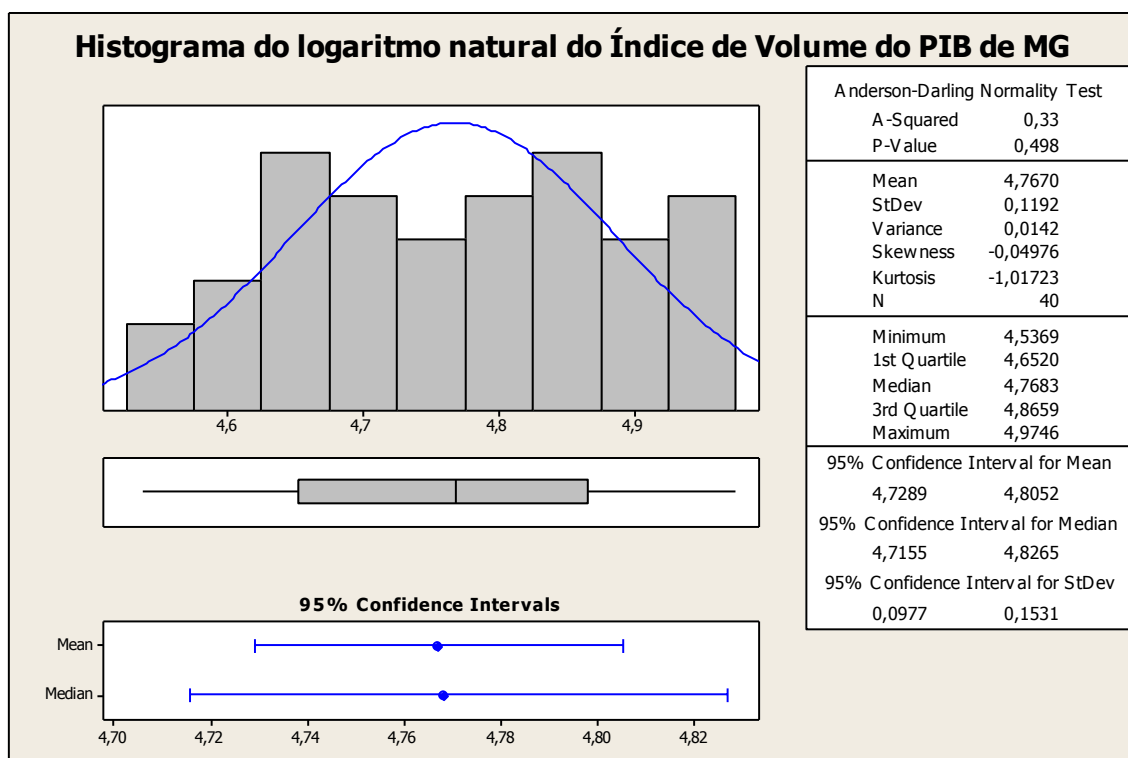


Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

O Gráfico 6 traz o histograma da série de índice de volume do PIB de Minas Gerais e outras características relacionadas ao comportamento dos dados. Percebe-se que a hipótese de normalidade, quando se analisa as 40 observações, não foi rejeitada ($p\text{-valor} = 0,454 > \alpha = 0,05$). Porém, a distribuição dos dados apresentou uma leve assimetria positiva (Skewness = 0,12460). Aplicando uma transformação

logarítmica nos dados percebe-se que a distribuição dos dados tornou-se mais simétrica (Skewness = - 0,04976) e a hipótese de normalidade dos dados continuou não sendo rejeitada (p-valor = 0,498 > $\alpha = 0,05$). O Gráfico 7 traz o histograma com as características do logaritmo natural da série de índice de volume do PIB de Minas Gerais.

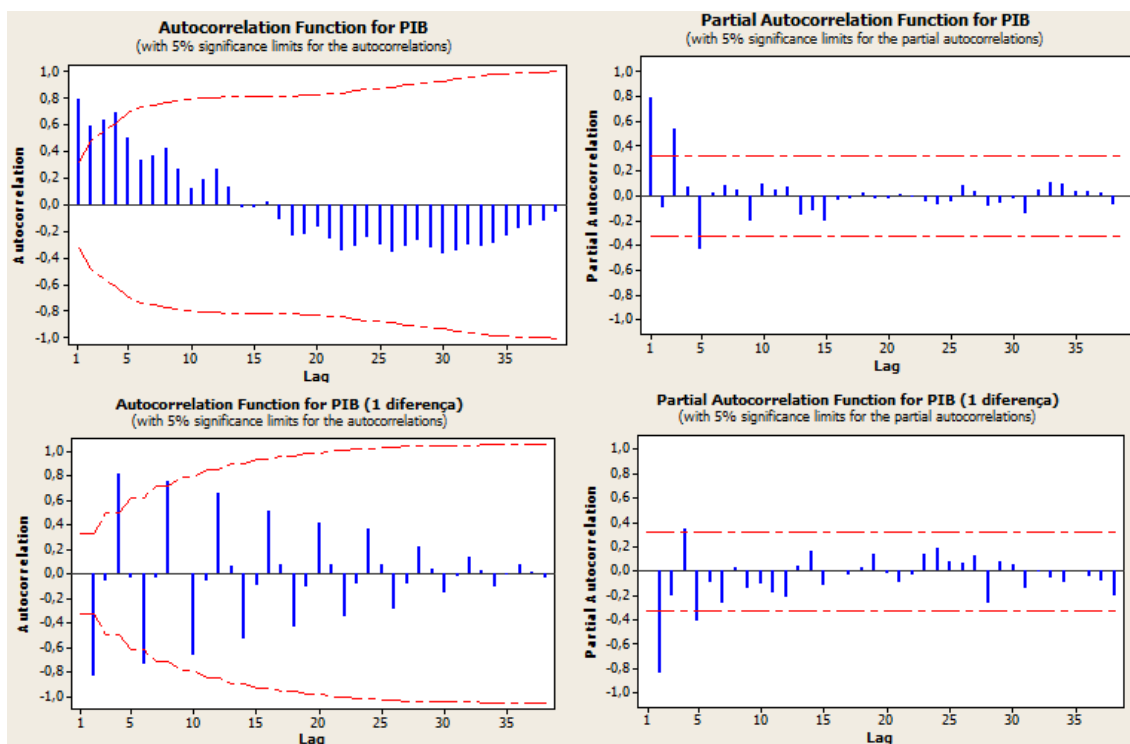
Gráfico 7 – Histograma do logaritmo natural do Índice de Volume do PIB de Minas Gerais – 1º Trimestre de 2002-4º Trimestre de 2011



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

O Gráfico 8 traz os correlogramas – Função de Autocorrelação (FAC) e Função de Autocorrelação Parcial (FACP) da série de índice de volume do PIB de Minas Gerais e após a aplicação de uma diferença não sazonal na série. Nota-se que apenas observando os correlogramas da série não diferenciada já se percebe que a mesma não é estacionária. Vale lembrar que a interpretação dos correlogramas não é uma tarefa fácil. Na verdade, entendê-los requer bastante experiência por parte do usuário. Mesmo assim, eles terão a sua utilidade, junto com a análise descritiva aqui efetuada, para a construção dos modelos.

Gráfico 8 – Função de Autocorrelação (FAC) e Função de Autocorrelação Parcial (FACP) da série de índice de volume do PIB de Minas Gerais e após a aplicação de uma diferença não sazonal na série



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

CAPÍTULO 5: APRESENTAÇÃO DOS MODELOS CONSTRUÍDOS

Após a análise exploratória dos dados efetuada no capítulo anterior, este capítulo concentrará a atenção nos modelos propostos para analisar a série. Inicialmente, serão apresentados os “melhores” modelos univariados obtidos (com todas as limitações e dificuldades que esses modelos possuem), partindo da modelagem ARIMA Sazonal até o uso da técnica de Alisamento Exponencial. Em seguida, a análise focará o modelo de Regressão Dinâmica construído, as variáveis explicativas utilizadas bem como suas adaptações e o modelo propriamente dito.

5.1) Modelos Univariados:

Com base na análise descritiva da série, os dois melhores modelos SARIMA¹² propostos para a série de índice de volume do PIB de Minas Gerais (Y_t), mesmo com a limitação do tamanho da amostra (apenas 40 observações) foram:

5.1.1) Modelos SARIMA:

5.1.1.1) Modelo SARIMA (0, 1, 1)(1, 1, 0)₄ da série Y_t :

A equação do modelo, considerando as observações do período amostral, é a seguinte:

$$(1 - B)(1 - B^4)(1 - \phi B^4)Y_t = \theta_0 + (1 - \theta B)a_t$$

$$(1 - B^4 - B + B^5)(1 - \phi B^4)Y_t = \theta_0 + (1 - \theta B)a_t$$

$$(1 - \phi B^4 - B^4 + \phi B^8 - B + \phi B^5 + B^5 - \phi B^9)Y_t = \theta_0 + (1 - \theta B)a_t$$

Logo:

$$Y_t = \phi Y_{t-4} + Y_{t-4} - \phi Y_{t-8} + Y_{t-1} - \phi Y_{t-5} - Y_{t-5} + \phi Y_{t-9} + \theta_0 - \theta a_{t-1} + a_t; \text{ onde } \theta_0 = 0 \text{ (modelo sem constante); } \phi = -0,4949 \text{ e } \theta = -0,6147.$$

Os resultados dos parâmetros estimados e de outras características e informações do modelo (dentro da amostra) são mostrados abaixo:

¹² Além dos dois modelos SARIMA aplicados sobre a série original (Y_t), testou-se os mesmos modelos na série com transformação logarítmica ($Z_t = \ln(Y_t)$). Na prática, a transformação logarítmica não melhorou de forma significativa a performance na análise “dentro” da amostra. Ela apenas deixou a distribuição um pouco mais simétrica e mais próxima da distribuição normal tal como foi observado na análise exploratória realizada no Capítulo 4. Na análise “fora” da amostra, a transformação logarítmica piorou os resultados das previsões. As observações de Granger e Newbold (1976) citada por Morettin e Tolo (1987), de que as previsões dos antilogaritmos dos dados transformados são estimadores viciados e que deveriam sofrer algum processo de ajuste foi confirmada na análise aqui efetuada.

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
SAR 4	-0,4949	0,1558	-3,18	0,003
MA 1	-0,6147	0,1442	-4,26	0,000

Differencing: 1 regular, 1 seasonal of order 4

Number of observations: Original series 40, after differencing 35

Residuals: SS = 256,234 (backforecasts excluded)
MS = 7,765 DF = 33

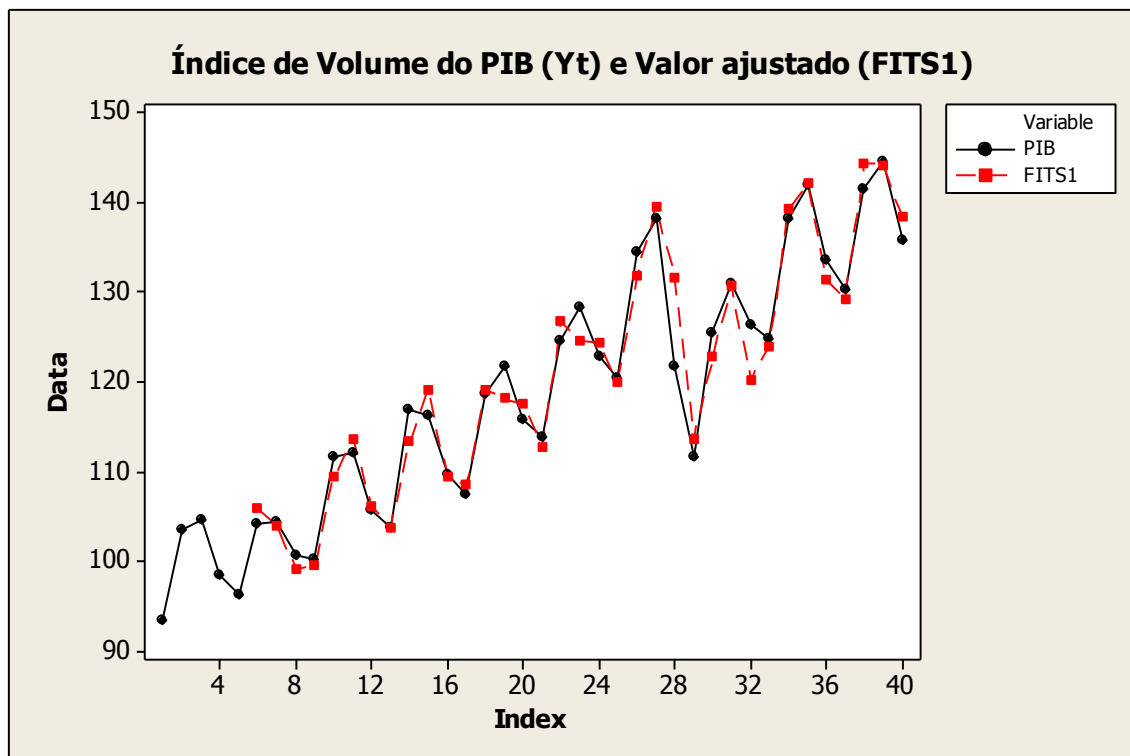
Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	10,8	24,1	*	*
DF	10	22	*	*
P-Value	0,372	0,341	*	*

MAD	1,956
MAPE	1,603
MSD	7,321
SQR	256,234
R-quadrado	0,954
R-quadrado ajustado	0,952
Durbin-Watson (DW)	2,047
Akaike (CIA)	1,957
Schwarz (CIS)	2,042

O Gráfico 9 traz a série original de índice de volume do PIB (Y_t) e os valores ajustados pelo modelo SARIMA $(0, 1, 1)(1, 1, 0)_4$ na série original considerando as 40 observações do período amostral:

Gráfico 9 – Série de Índice de Volume do PIB de Minas Gerais (Y_t) e Valores ajustados pelo modelo SARIMA $(0, 1, 1)(1, 1, 0)_4$ na série original (dentro da amostra)

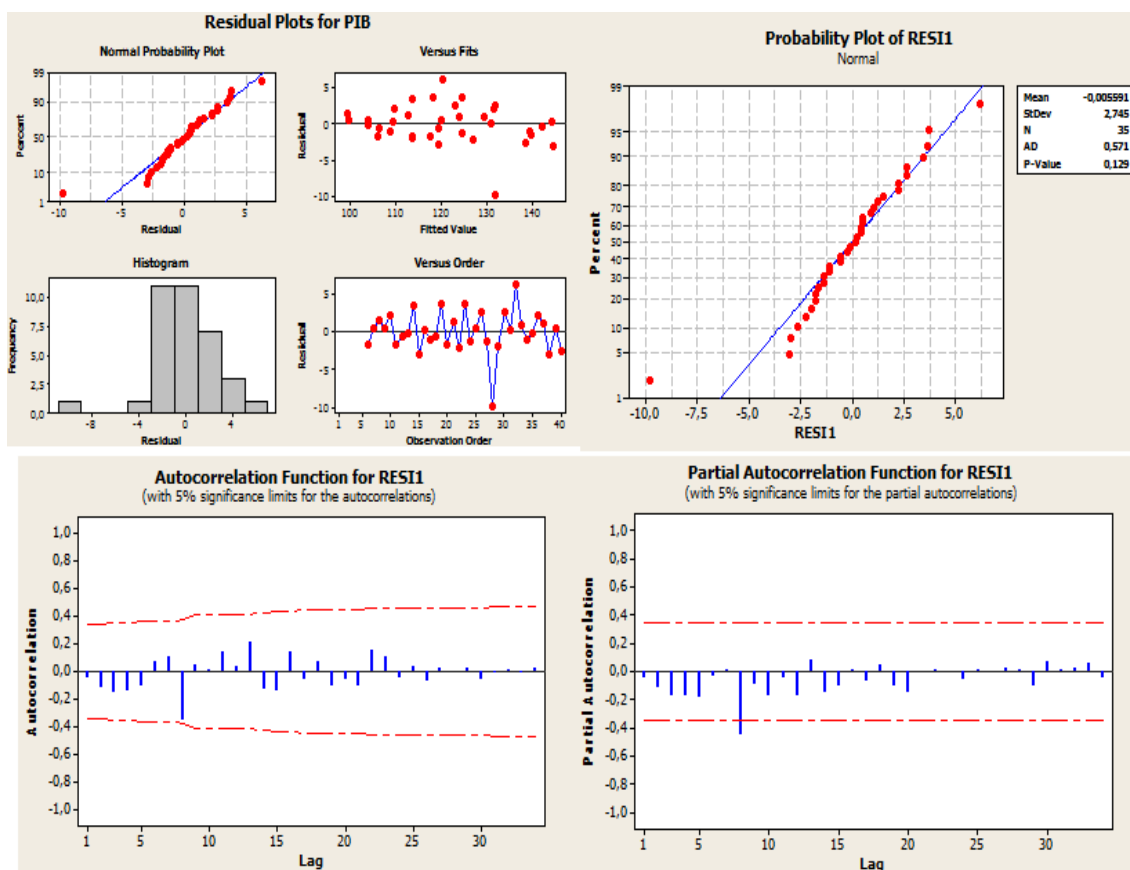


Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

O modelo ajustado foi bastante satisfatório, tendo em vista o tamanho pequeno da amostra (apenas 40 observações). Percebe-se que os coeficientes obtidos foram estatisticamente diferentes de zero (p-valores obtidos $< \alpha = 0,05$). Além disso, os coeficientes obtidos foram menores que “1” e maiores que “-1”.

Com relação ao comportamento dos resíduos, a informação de Ljung-Box apresentou valores para o p-valor maiores que $\alpha = 0,05$, indicando que o resíduo (a_t) obtido deva ser um ruído branco. O resultado da estatística de Durbin-Watson confirma que o modelo não apresenta correlação significativa de ordem 1 (DW = 2,047). De fato, os correlogramas da FAC e da FACP dos resíduos praticamente não apontaram correlações estatisticamente diferentes de “zero”, uma vez que as correlações situaram, de forma geral, “dentro” do intervalo pontilhado em vermelho. Ademais, há indícios de que os resíduos siga uma distribuição normal (p-valor do teste de normalidade = $0,129 > \alpha = 0,05$). O vetor de resíduos também apresentou média aproximadamente igual a “zero”.

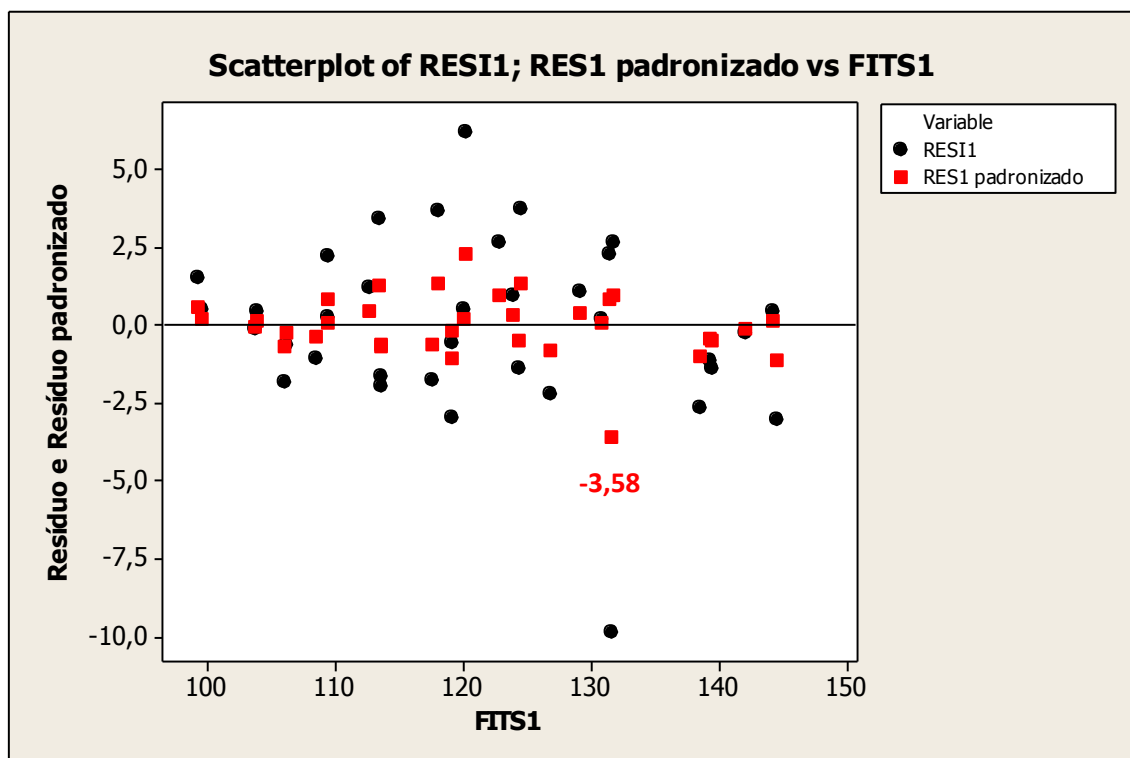
Figura 9 – Resultados da simulação do modelo SARIMA (0, 1, 1)(1, 1, 0)₄ na série original (dentro da amostra) para o comportamento do vetor de resíduos



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

O principal problema desse modelo foi a presença de um ponto outlier, que se refere ao quarto trimestre de 2008 e coincide com o início da crise econômica mundial. De fato, calculando, além do resíduo regular, o resíduo padronizado obtém-se um valor para o último trimestre de 2008 de -3,58 para o resíduo padronizado, valor menor do que -3,1 e, portanto, um indicativo da presença de um outlier. A presença desse ponto discrepante já era esperada, visto que como o modelo é construído com base apenas no passado da série, não haveria como ele prever de maneira eficaz um acontecimento tão diferenciado como a ocorrência de uma grave crise econômica. O Gráfico 10 representa os resíduos regulares e padronizados versus os valores ajustados pelo modelo:

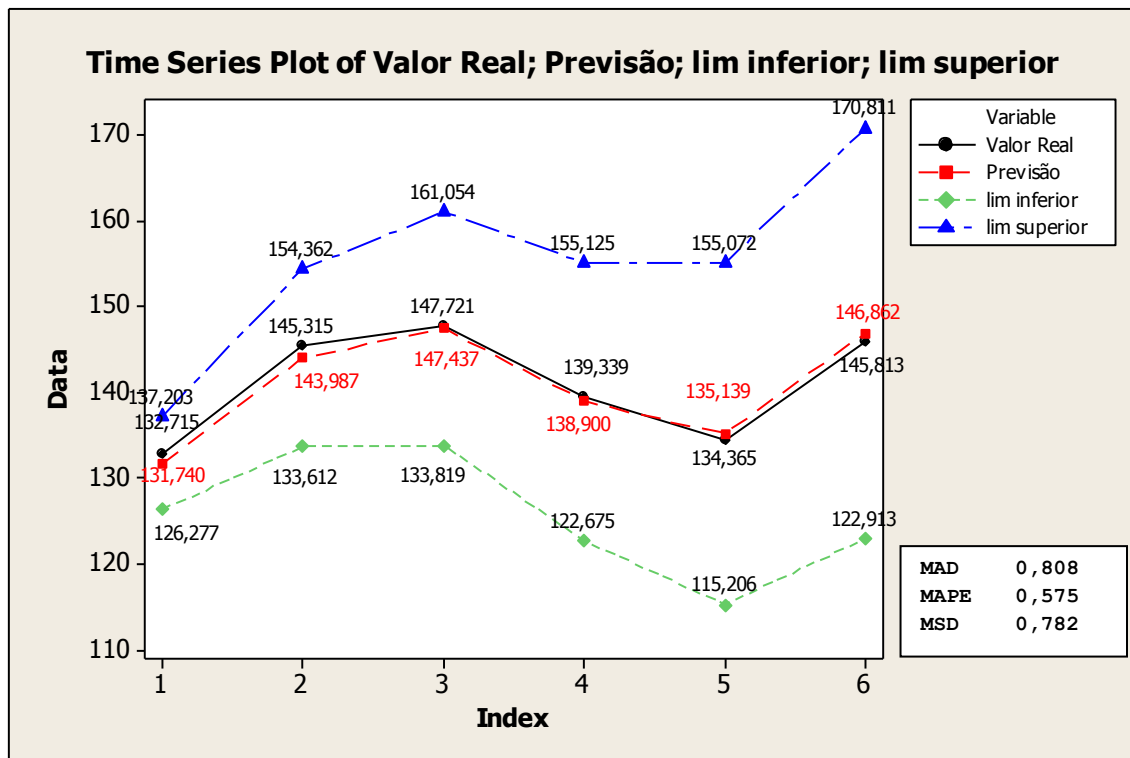
Gráfico 10 – Resíduos regulares e padronizados versus valores ajustados pelo modelo SARIMA (0, 1, 1)(1, 1, 0)₄ na série original (dentro da amostra)



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Ao realizar a análise no período de validação do modelo (fora da amostra), nota-se que o modelo foi bastante satisfatório. Os valores efetivamente observados situaram dentro do intervalo de confiança. Além disso, o comportamento das previsões acompanhou exatamente o comportamento dos valores reais observados. As estatísticas de desempenho do modelo (MAD, MAPE e MSD) no período de validação também se encontram no Gráfico 11:

Gráfico 11 – Previsões do modelo SARIMA (0, 1, 1)(1, 1, 0)₄ na série original seis passos a frente e medidas de desempenho



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

5.1.1.2) Modelo SARIMA (2, 0, 0)(0, 1, 1)₄ da série Y_t :

A equação do modelo, considerando as observações do período amostral, é a seguinte:

$$(1 - \phi_1 B - \phi_2 B^2)(1 - B^4)Y_t = \theta_0 + (1 - \theta B^4)a_t$$

$$(1 - B^4 - \phi_1 B + \phi_1 B^5 - \phi_2 B^2 + \phi_2 B^6)Y_t = \theta_0 + (1 - \theta B^4)a_t$$

Logo:

$$Y_t = Y_{t-4} + \phi_1 Y_{t-1} - \phi_1 Y_{t-5} + \phi_2 Y_{t-2} - \phi_2 Y_{t-6} + \theta_0 - \theta a_{t-4} + a_t; \text{ onde } \phi_1 = 1,0422; \phi_2 = -0,5651; \theta = 0,8376 \text{ e } \theta_0 = 2,24865 \text{ (modelo com constante).}$$

Os resultados dos parâmetros estimados e de outras características e informações do modelo (dentro da amostra) são mostrados abaixo:

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
AR	1	1,0422	0,1510	6,90	0,000
AR	2	-0,5651	0,1498	-3,77	0,001

SMA 4 0,8376 0,1377 6,08 0,000
 Constant 2,24865 0,06947 32,37 0,000

Differencing: 0 regular, 1 seasonal of order 4
 Number of observations: Original series 40, after differencing 36
 Residuals: SS = 154,042 (backforecasts excluded)
 MS = 4,814 DF = 32

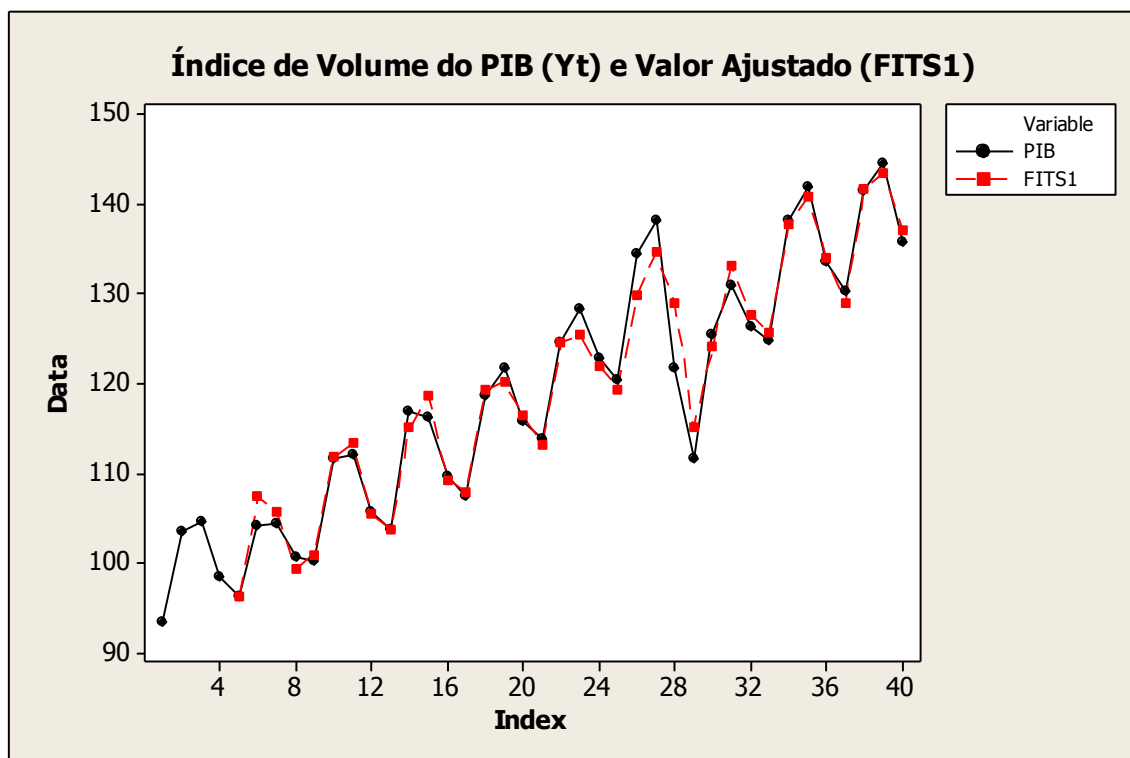
Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	6,4	15,3	*	*
DF	8	20	*	*
P-Value	0,600	0,757	*	*

MAD	1,451
MAPE	1,197
MSD	4,279
SQR	154,042
R-quadrado	0,975
R-quadrado ajustado	0,973
Durbin-Watson (DW)	1,818
Akaike (CIA)	1,548
Schwarz (CIS)	1,717

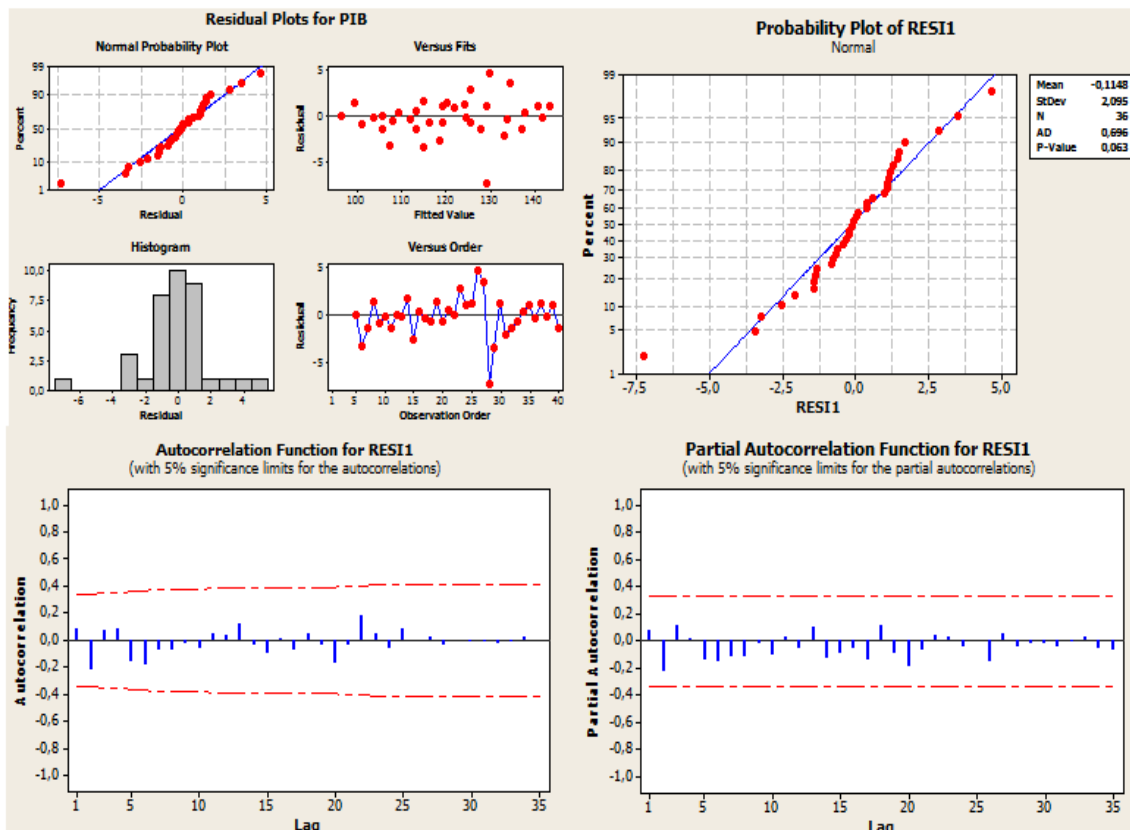
O Gráfico 12 traz a série original de índice de volume do PIB (Y_t) e os valores ajustados pelo modelo SARIMA $(2, 0, 0)(0, 1, 1)_4$ na série original considerando as 40 observações do período amostral:

Gráfico 12 – Série de Índice de Volume do PIB de Minas Gerais (Y_t) e Valores ajustados pelo modelo SARIMA $(2, 0, 0)(0, 1, 1)_4$ na série original (dentro da amostra)



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Figura 10 – Resultados da simulação do modelo SARIMA (2, 0, 0)(0, 1, 1)₄ na série original (dentro da amostra) para o comportamento do vetor de resíduos



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

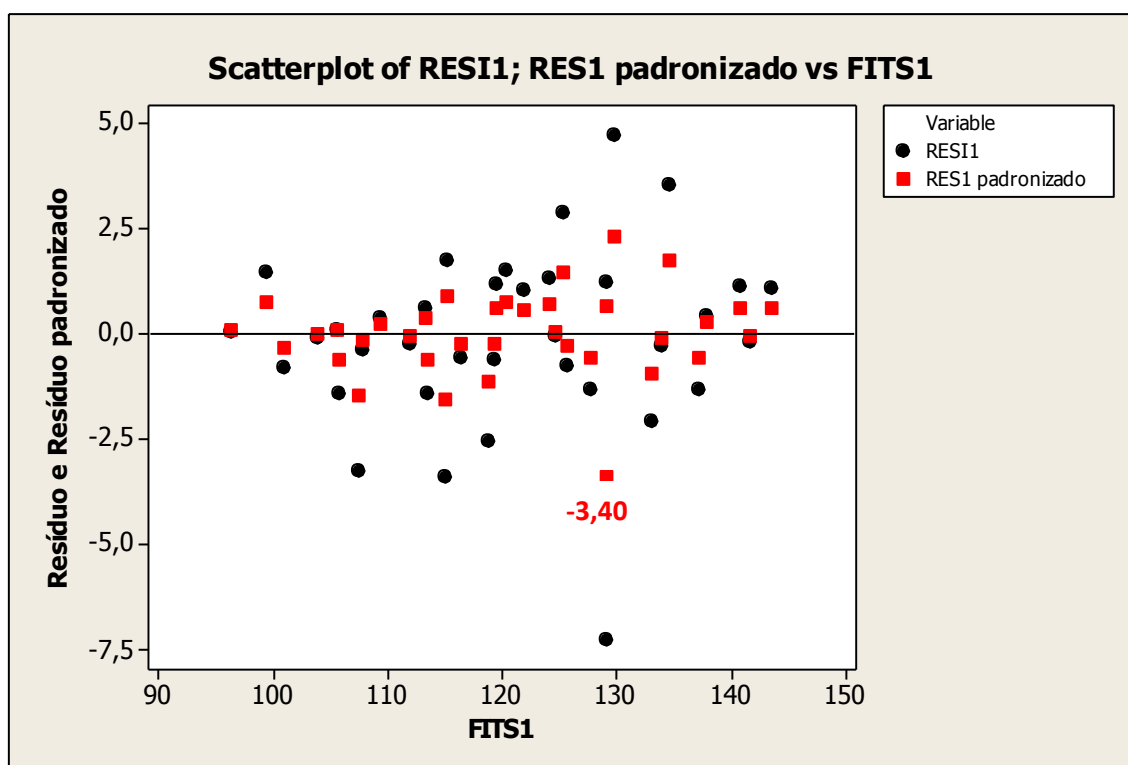
O modelo ajustado também foi bastante satisfatório, tendo em vista o tamanho pequeno da amostra (apenas 40 observações). Percebe-se que os coeficientes obtidos foram estatisticamente diferentes de zero (p-valores obtidos $< \alpha = 0,05$).

Com relação ao comportamento dos resíduos, a informação de Ljung-Box apresentou valores para o p-valor maiores que $\alpha = 0,05$, indicando que o resíduo (a_t) obtido deva ser um ruído branco. O resultado da estatística de Durbin-Watson apresentou valor próximo de “dois”, porém não tão próximo quanto o observado no modelo anterior (DW = 1,818). Os correlogramas da FAC e da FACP dos resíduos confirmaram os indícios da ocorrência de ruído branco, pois não apontaram correlações estatisticamente diferentes de “zero”. As correlações situaram “dentro” do intervalo pontilhado em vermelho (Figura 10). Ademais, há indícios não tão significativos de que os resíduos tenham uma distribuição normal, já que a hipótese nula de normalidade dos resíduos não foi rejeitada no limite (p-valor do teste de normalidade = 0,063 $> \alpha = 0,05$). O vetor de resíduos também apresentou média próxima de “zero”.

Um aspecto favorável deste modelo em relação ao anterior foi que o mesmo apresentou melhores resultados em relação às informações de escolha e de desempenho dos modelos. De fato, o modelo SARIMA (2, 0, 0)(0, 1, 1)₄ aplicado sobre a série original (Y_t) apresentou menores valores para os critérios de Akaike e Schwarz e também menores valores para as medidas da performance do modelo (MAD, MAPE, MSD e SQR). Conseqüentemente, o grau de ajuste deste modelo também foi mais acentuado do que o observado no modelo anterior (maior valor para o R-quadrado e o R-quadrado ajustado).

Por outro lado, o problema da presença de um ponto outlier referente ao quarto trimestre de 2008 (início da crise econômica mundial) permaneceu neste modelo. De fato, calculando o resíduo padronizado obtém-se um valor de -3,40 para o último trimestre de 2008, valor menor do que -3,1 e, portanto, um indicativo da presença de um outlier. O Gráfico 13 traz os resíduos regulares e padronizados versus os valores ajustados pelo modelo. Nele é notória a presença de um ponto em vermelho (resíduo padronizado) mais afastado do que os demais em relação à linha horizontal de média “zero”:

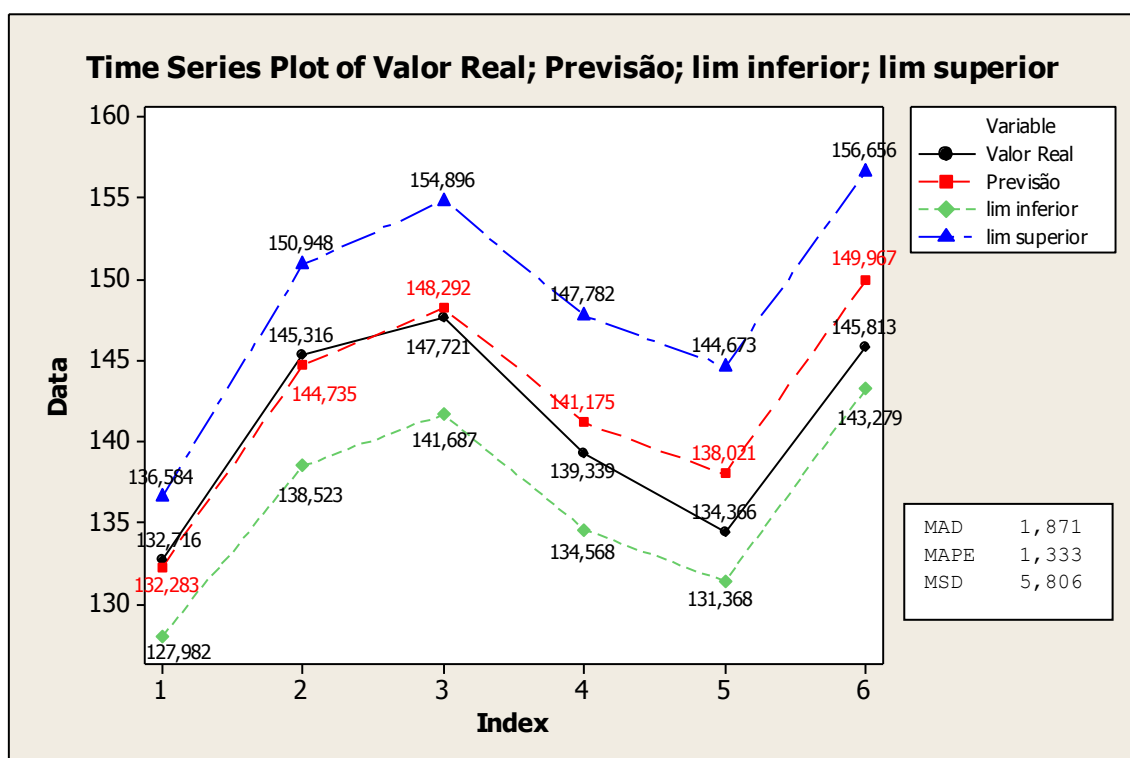
Gráfico 13 – Resíduos regulares e padronizados versus valores ajustados pelo modelo SARIMA (2, 0, 0)(0, 1, 1)₄ na série original (dentro da amostra)



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Ao realizar a análise fora da amostra, percebe-se que o modelo encontrado acompanhou o comportamento dos valores reais observados. Porém, percebe-se que as previsões mais distantes começaram a se afastar dos valores efetivamente observados. Com isso, as estatísticas de desempenho do modelo (MAD, MAPE e MSD) no período de validação foram piores nesse modelo comparativamente ao modelo anterior. O Gráfico 14 explicita este resultado:

Gráfico 14 – Previsões do modelo SARIMA (2, 0, 0)(0, 1, 1)₄ na série original seis passos a frente e medidas de desempenho



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

5.1.2) Modelos de Alisamento Exponencial de Holt-Winters:

A técnica de Alisamento Exponencial é recomendada em casos em que o passado da série é relativamente pequeno para encontrar o padrão de comportamento da série. Isto acontece com o índice de volume do PIB de Minas Gerais (Y_t) que possui apenas 40 observações no período amostral. Assim, decidiu-se aplicar a técnica nos seus dois métodos tradicionais (aditivo e multiplicativo) e escolher os dois melhores modelos que se ajustam a cada um deles através da observação das medidas de desempenho (MAD, MAPE e MSD). Os dois melhores modelos ajustados encontram-se abaixo.

5.1.2.1) Alisamento Exponencial Multiplicativo ($\alpha = 0,9$; $\gamma = 0,2$; $\delta = 0,2$)

A equação do melhor modelo de Alisamento Exponencial Multiplicativo obtido é mostrada abaixo com os respectivos valores das constantes de alisamento:

Método Multiplicativo

$$L_t = \alpha (Y_t / S_{t-p}) + (1-\alpha) [L_{t-1} + T_{t-1}]$$

$$T_t = \gamma [L_t - L_{t-1}] + (1-\gamma)T_{t-1}$$

$$S_t = \delta (Y_t / L_t) + (1-\delta) S_{t-p}$$

$$\hat{Y}_t = (L_{t-1} + T_{t-1}) S_{t-p}$$

Onde:

$$\alpha = 0,9$$

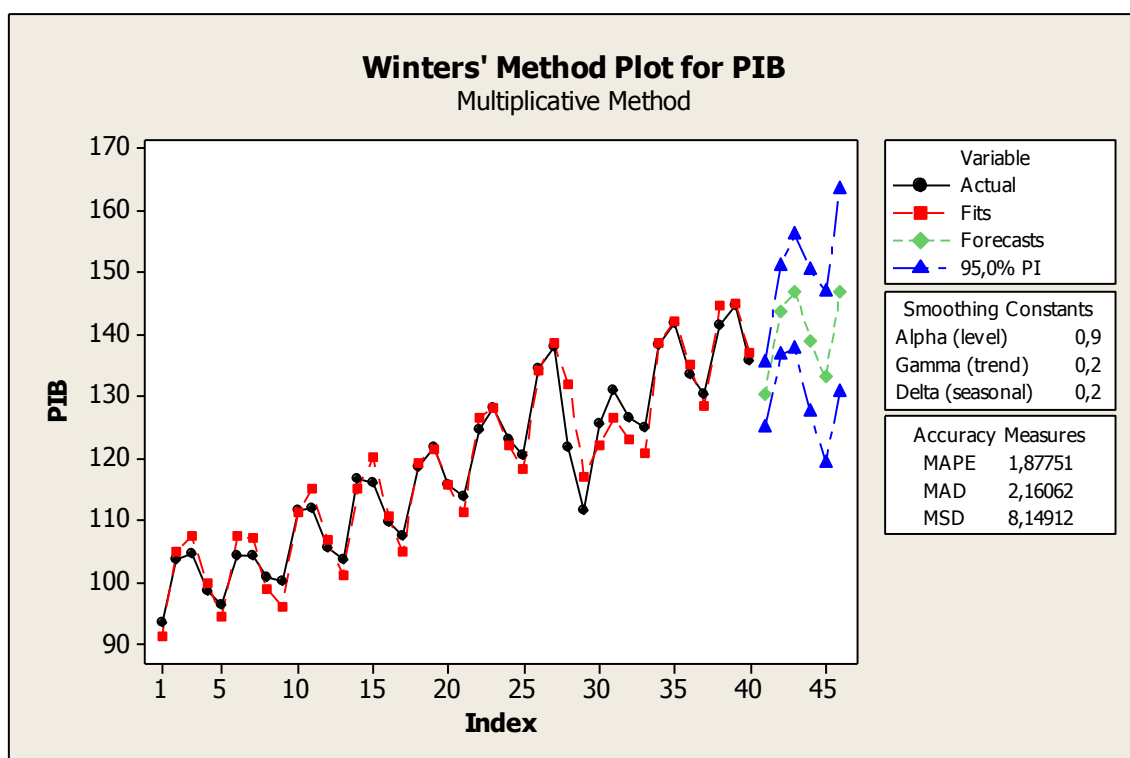
$$\gamma = 0,2$$

$$\delta = 0,2$$

$$p = 4$$

Nota-se que as constantes (pesos) situaram dentro do intervalo recomendado $[0,1]$, porém, a constante de nível “alfa” apresentou um valor bastante elevado e bem próximo de “1” ($\alpha = 0,9$).

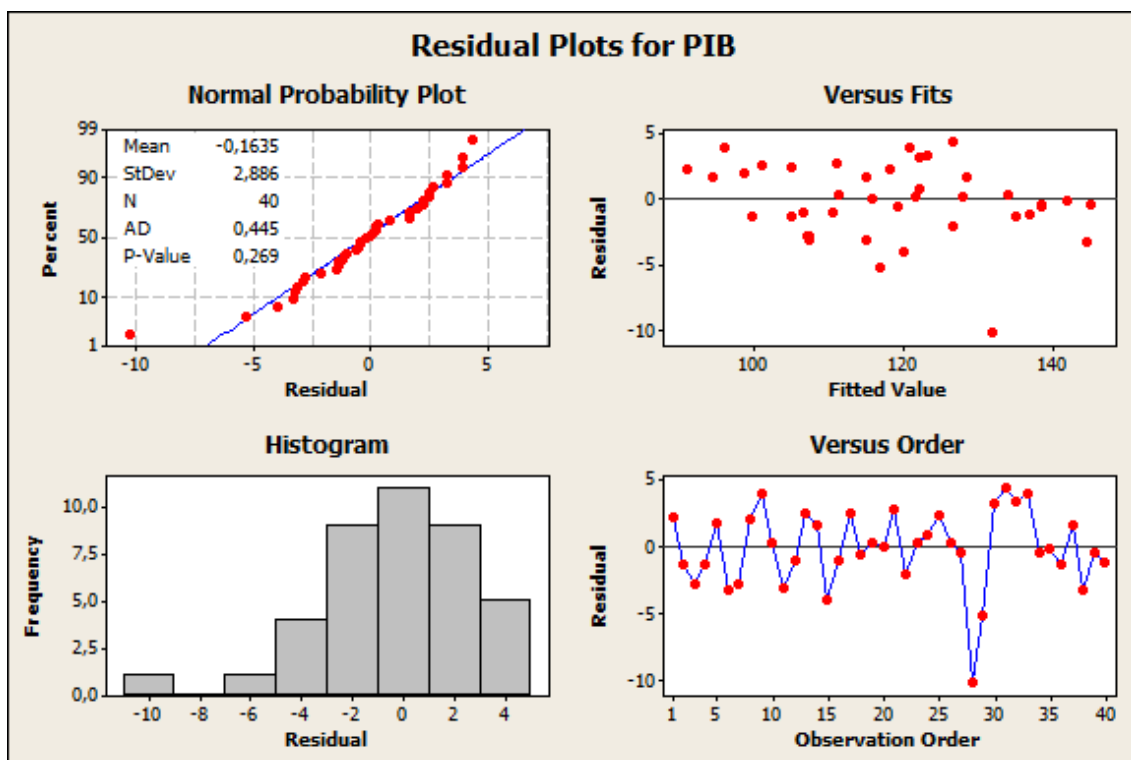
Gráfico 15 – Série de Índice de Volume do PIB de Minas Gerais (Y_t) e Valores ajustados pelo modelo de Alisamento Exponencial Multiplicativo ($\alpha = 0,9$; $\gamma = 0,2$; $\delta = 0,2$) e medidas de desempenho dentro da amostra



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Com relação ao comportamento do modelo dentro da amostra, percebe-se que, comparativamente aos modelos SARIMA construídos anteriormente, o modelo de alisamento multiplicativo foi o que apresentou os piores resultados nas medidas de desempenho ($MAD = 2,161$, $MAPE = 1,878$ e $MSD = 8,149$). De fato, o resultado deste modelo no período amostral foi ligeiramente pior do que o observado no modelo SARIMA $(0, 1, 1)(1, 1, 0)_4$ ($MAD = 1,956$, $MAPE = 1,603$ e $MSD = 7,321$), modelo com a menor performance no período amostral. O Gráfico 15 acima traduz esses resultados.

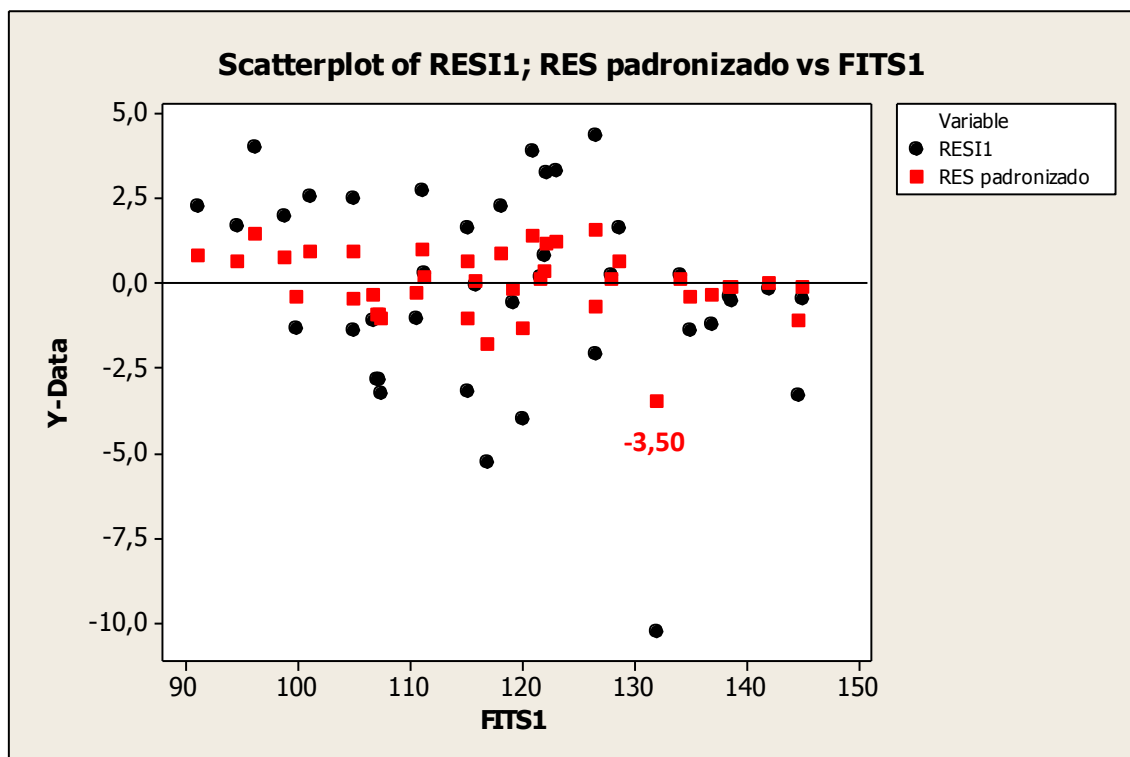
Figura 11 – Resultados da simulação do modelo de Alisamento Exponencial Multiplicativo ($\alpha = 0,9$; $\gamma = 0,2$; $\delta = 0,2$) para o comportamento do vetor de resíduos (dentro da amostra)



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

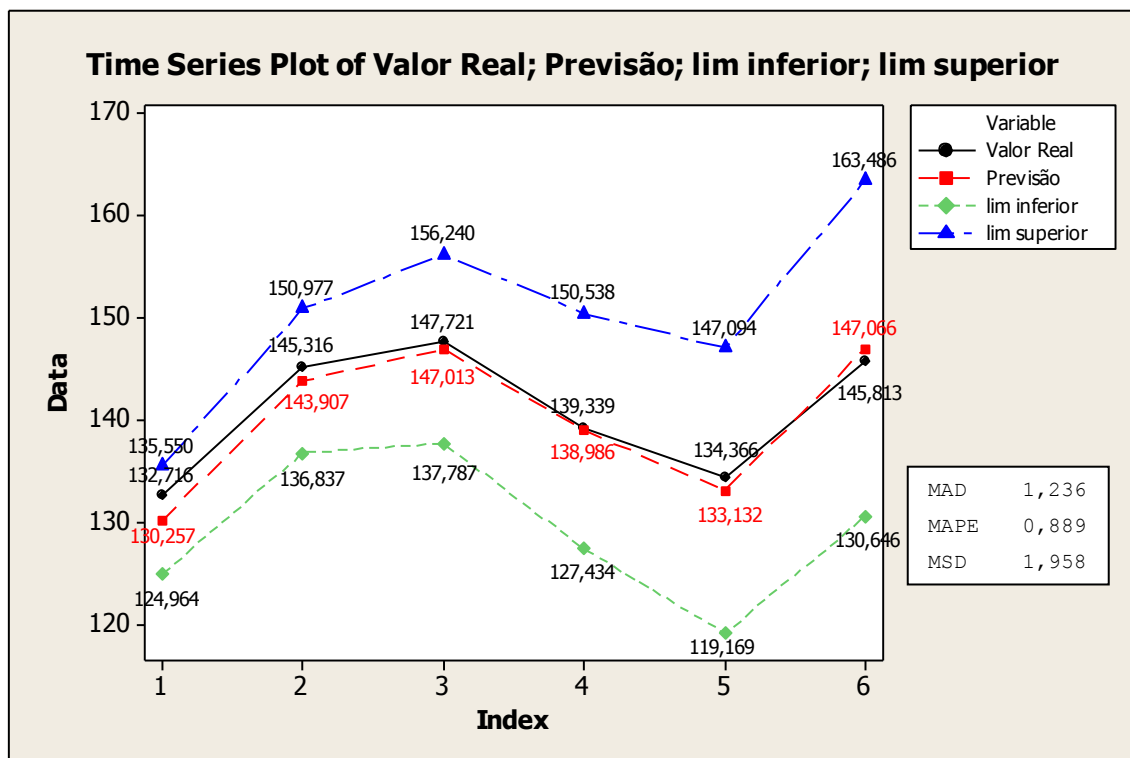
Já o comportamento do vetor de resíduos apresentou problemas similares aos já observados nos modelos SARIMA. Apesar da hipótese de normalidade dos resíduos não ter sido rejeitada ($p\text{-valor} = 0,269 > \alpha = 0,05$), o problema da presença de um ponto outlier relativo ao quarto trimestre de 2008 (início da crise econômica) permaneceu (Figura 11). De fato, calculando o resíduo padronizado do modelo encontra-se um valor para o último trimestre de 2008 de $-3,50$. O Gráfico 16 mostra o comportamento do resíduo regular e do resíduo padronizado do modelo de alisamento multiplicativo:

Gráfico 16 – Resíduos regulares e padronizados versus valores ajustados pelo modelo de Alisamento Exponencial Multiplicativo ($\alpha = 0,9$; $\gamma = 0,2$; $\delta = 0,2$) dentro da amostra



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Gráfico 17 – Previsões do modelo de Alisamento Exponencial Multiplicativo ($\alpha = 0,9$; $\gamma = 0,2$; $\delta = 0,2$) seis passos a frente e medidas de desempenho



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Com relação ao comportamento do modelo de alisamento multiplicativo fora da amostra, percebe-se que os valores previstos acompanharam a trajetória dos valores efetivamente observados (Gráfico 17). De maneira geral, os valores previstos também foram ligeiramente inferiores aos valores reais, com exceção da última previsão. As medidas de desempenho do modelo no período de validação (MAD, MAPE e MSD) também foram bastante razoáveis. Todavia, o desempenho deste modelo de alisamento fora da amostra ainda foi inferior ao do modelo SARIMA (0, 1, 1)(1, 1, 0)₄ aplicado sobre a série original (Y_t).

5.1.2.2) Alisamento Exponencial Aditivo ($\alpha = 0,9$; $\gamma = 0,2$; $\delta = 0,2$)

Também no caso da aplicação do método de alisamento aditivo, encontrou-se um valor elevado para a constante (peso) referente ao nível ($\alpha = 0,9$) e os mesmos valores para as constantes de alisamento da tendência ($\gamma = 0,2$) e para o peso da componente sazonal ($\delta = 0,2$). Em todo caso, as constantes situaram dentro do intervalo recomendado [0,1]. A equação do melhor modelo de alisamento exponencial aditivo ajustado à série (Y_t) é demonstrada abaixo:

Método Aditivo

$$L_t = \alpha (Y_t - S_{t-p}) + (1 - \alpha) [L_{t-1} + T_{t-1}]$$

$$T_t = \gamma [L_t - L_{t-1}] + (1 - \gamma) T_{t-1}$$

$$S_t = \delta (Y_t - L_t) + (1 - \delta) S_{t-p}$$

$$\hat{Y}_t = L_{t-1} + T_{t-1} + S_{t-p}$$

Onde:

$$\alpha = 0,9$$

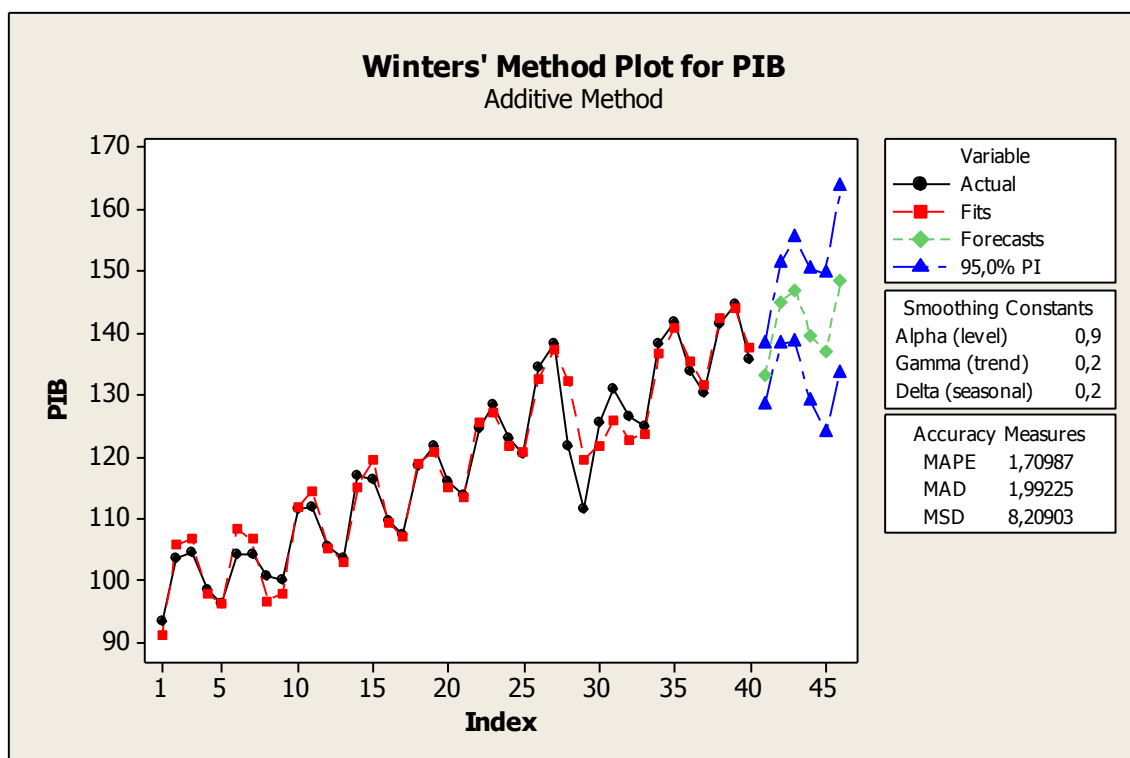
$$\gamma = 0,2$$

$$\delta = 0,2$$

$$p = 4$$

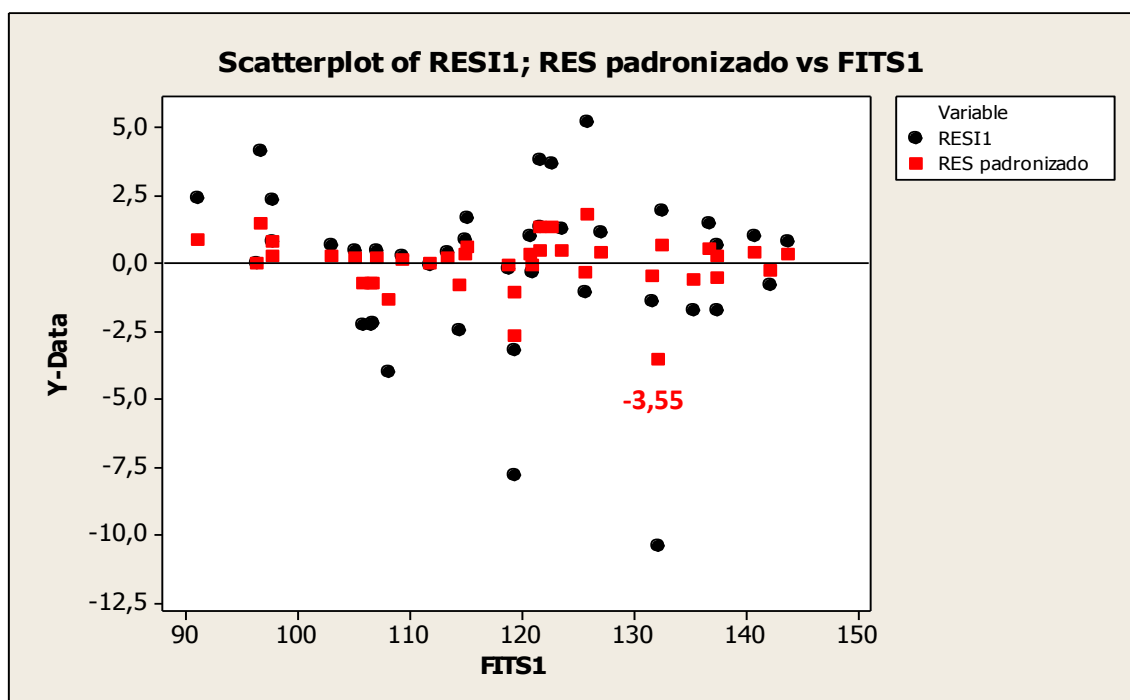
O comportamento do modelo de alisamento exponencial aditivo no período amostral foi bastante parecido com o observado no modelo multiplicativo, pelo menos no que se refere aos resultados das medidas de desempenho. De fato, o modelo de alisamento aditivo apresentou os seguintes resultados no que se refere à capacidade de ajuste dos valores preditos em relação aos verdadeiros valores: MAD = 1,992 (contra 2,161 do modelo multiplicativo), MAPE = 1,710 (contra 1,878) e MSD = 8,209 (contra 8,149). Portanto, em comparação aos modelos SARIMA construídos anteriormente, os modelos de alisamento apresentaram resultados desfavoráveis dentro da amostra. O Gráfico 18 abaixo traz o ajuste do modelo de alisamento aditivo com as medidas de desempenho e os pesos das constantes de alisamento utilizados na construção do modelo.

Gráfico 18 – Série de Índice de Volume do PIB de Minas Gerais (Y_t) e Valores ajustados pelo modelo de Alisamento Exponencial Aditivo ($\alpha = 0,9$; $\gamma = 0,2$; $\delta = 0,2$) e medidas de desempenho dentro da amostra



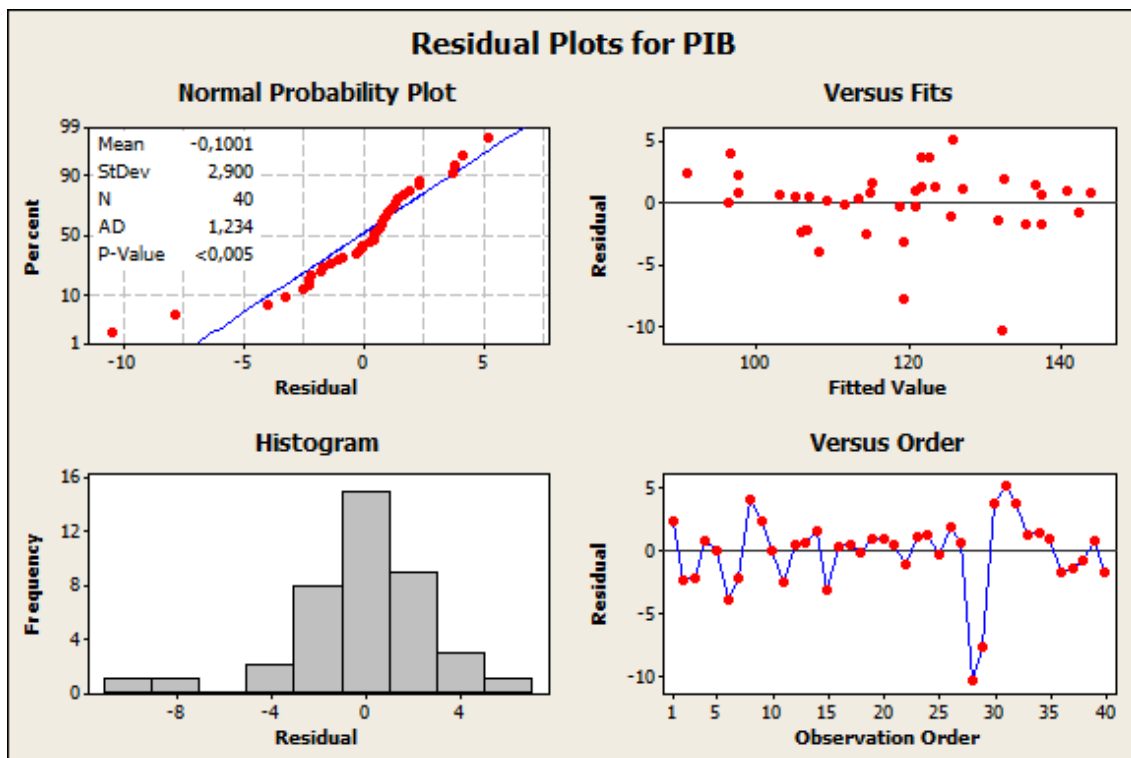
Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Gráfico 19 – Resíduos regulares e padronizados versus valores ajustados pelo modelo de Alisamento Exponencial Aditivo ($\alpha = 0,9$; $\gamma = 0,2$; $\delta = 0,2$) dentro da amostra



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Figura 12 – Resultados da simulação do modelo de Alisamento Exponencial Aditivo ($\alpha = 0,9$; $\gamma = 0,2$; $\delta = 0,2$) para o comportamento do vetor de resíduos (dentro da amostra)



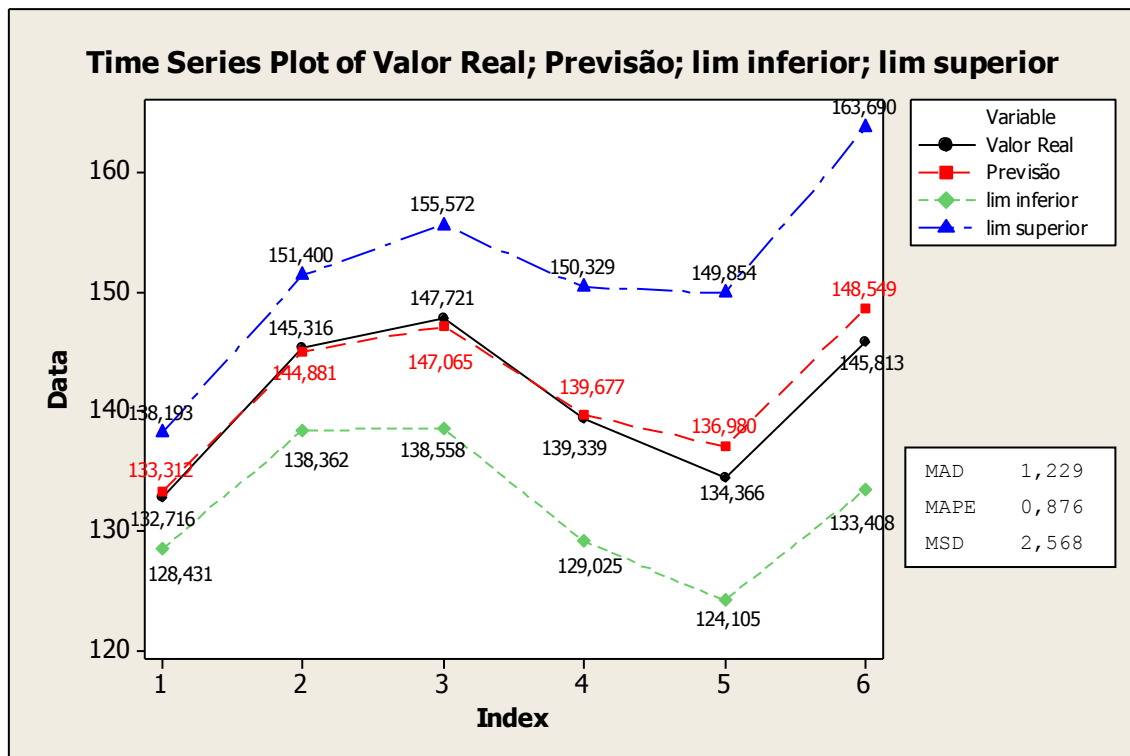
Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

O comportamento do vetor de resíduos também continuou apresentando problemas similares aos já observados nos modelos SARIMA. O problema da presença de um ponto outlier relativo ao último trimestre de 2008 permaneceu no modelo aditivo. De fato, o resíduo padronizado registrou um valor de -3,55 para o quarto trimestre de 2008 (Gráfico 19). Além disso, de todos os modelos anteriores construídos, o modelo de alisamento aditivo foi o único em que a hipótese de normalidade dos resíduos foi rejeitada (p -valor < 0,005), conforme a Figura 12.

Com relação à análise do modelo de alisamento aditivo fora da amostra, percebe-se que os valores previstos acompanharam a trajetória dos valores efetivamente observados. Porém, os últimos valores previstos se afastaram dos verdadeiros valores reais. As medidas de desempenho do modelo no período de validação (MAD, MAPE e MSD) também foram razoáveis. Todavia, o desempenho deste modelo de alisamento no período de validação ainda foi inferior ao do modelo SARIMA $(0, 1, 1)(1, 1, 0)_4$ aplicado sobre a série original (Y_t) e também teve performance ligeiramente menor do que o modelo de alisamento multiplicativo. O

comportamento do modelo de alisamento aditivo fora da amostra é mostrado no Gráfico 20:

Gráfico 20 – Previsões do modelo de Alisamento Exponencial Aditivo ($\alpha = 0,9$; $\gamma = 0,2$; $\delta = 0,2$) seis passos a frente e medidas de desempenho



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

5.2) Modelo Multivariado:

5.2.1) Escolha das variáveis explicativas:

As estratégias de construção do modelo de regressão dinâmica e escolha das variáveis explicativas intervenientes foram baseadas nos conceitos acerca da metodologia de cálculo do PIB apresentada no Capítulo 2. De fato, nesse capítulo, foi mencionado o procedimento de cálculo do PIB a preços básicos pela ótica da produção, ou seja, o PIB entendido como o somatório dos valores adicionados dos diferentes setores presentes na economia:

$$\text{PIB (preços básicos)} = \sum (\text{VA setoriais})$$

Considere que a economia seja formada pelos três grandes setores de atividade econômica: o setor agropecuário, o setor industrial e setor de serviços.

Nesse sentido, o PIB avaliado a preços básicos corresponderia, exatamente, ao somatório dos valores adicionados desses três setores:

$$\text{PIB (preços básicos)} = \text{VA Agropecuária} + \text{VA Indústria} + \text{VA Serviços}$$

A busca pelas variáveis explicativas que afetam o PIB (ou o índice de volume) foi pensada com base nesses três grandes setores de atividade econômica. A ideia era encontrar variáveis que representassem, com algum grau de fidedignidade, cada um dos valores adicionados setoriais e que estivessem disponíveis antes da divulgação do PIB. Assim, buscou-se uma variável *proxy* para cada um dos três grandes setores. Entretanto, ao encontrar essas variáveis, foi necessária a realização de algumas adaptações no banco de dados. Isto porque a variável a se modelar (o índice de volume do PIB de Minas Gerais) é uma variável divulgada trimestralmente, ao passo que as variáveis explicativas são divulgadas mensalmente ou anualmente. Portanto, foi necessário a “trimestralização” dessas variáveis antes da construção do modelo de regressão dinâmica.

Gujarati (2000) argumenta que a “manipulação” de dados ou a interpolação tal como se pretende fazer neste trabalho poderá resultar em um padrão sistemático nas perturbações, introduzindo assim o problema da autocorrelação. Todavia, pretende-se contornar este problema na monografia com a introdução das defasagens da variável de interesse e das variáveis explicativas: a ideia é a de que existe uma influência na produção corrente que depende da produção do passado e dos fatores intervenientes no presente e os mesmos defasados. As variáveis *proxy* para cada um dos valores adicionados setoriais são descritas abaixo.

5.2.1.1) *Proxy* para o comportamento do VA da Agropecuária (A_t):

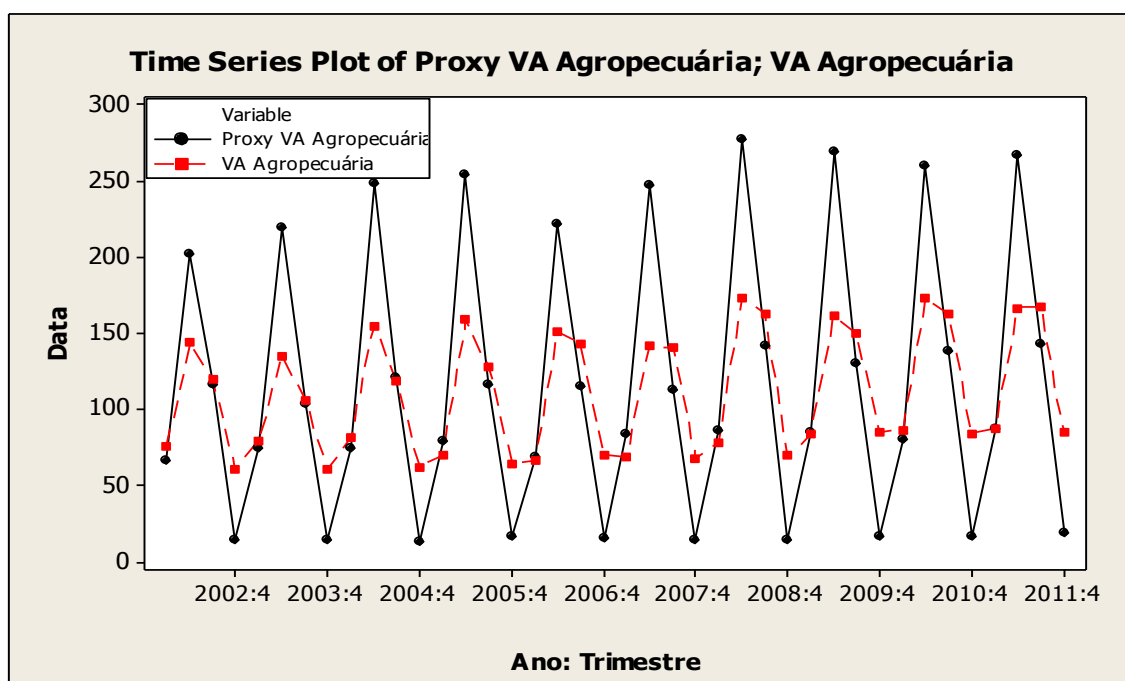
A variável utilizada como *proxy* para o valor adicionado do setor agropecuário foi a produção agrícola (em toneladas) de quatro dos principais produtos da agricultura de Minas Gerais e com impacto na produção agrícola nacional: o café, a batata, o feijão e o milho. O valor em toneladas para cada um dos produtos foi obtido por meio da Pesquisa Agrícola Municipal (PAM) e pelo Levantamento Sistemático da Produção Agrícola (LSPA) do IBGE. Todavia, esse dado é uma informação anual e precisa ser “trimestralizado”. Esse processo foi feito utilizando a informação da proporção de safra colhida por trimestre do último ano completo disponibilizada pelo Grupo de Coordenação de Estatísticas Agropecuárias de Minas Gerais (GCEA-MG). Na prática, a utilização deste processo não inviabiliza o método, visto que a proporção da safra

colhida por trimestre para cada produto não muda muito de ano para ano devido às condições climáticas e as características de cada cultura.¹³

Com a “trimestralização” de cada um dos produtos agrícolas, os resultados por trimestre foram somados em uma coluna única com o total da produção agrícola (em toneladas) desses quatro produtos analisados concomitantemente. Em seguida, a série obtida foi escrita no formato de número-índice com a média de 2002 = 100. Este procedimento foi feito visando à comparação com a série oficial de índice de volume do VA da agropecuária, que tradicionalmente é divulgada pela Fundação João Pinheiro na forma de número-índice com o mesmo ano base de referência.¹⁴

O Gráfico 21 traz o comportamento da série oficial do VA da agropecuária e da série construída e utilizada como *proxy*. Nota-se que a variável A_t utilizada como *proxy* captou a questão da sazonalidade do setor agropecuário com um ritmo mais intenso de produção no segundo e no terceiro trimestre de cada ano, sobretudo pela influência do café nesses trimestres (principal produto da agricultura mineira). Porém, a amplitude desta sazonalidade foi mais intensa na variável *proxy*.

Gráfico 21 – Comportamento via gráfico de linhas do Valor Adicionado da Agropecuária de Minas Gerais e da variável *proxy* (A_t) utilizada para representar o mesmo – média de 2002 = 100 – 1º Trimestre de 2002-4º Trimestre de 2011



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

¹³ Ver Apêndice 2 que traz os valores anuais em toneladas das quatro culturas (café, batata, feijão e milho) e a proporção de safra colhida por trimestre de cada um dos produtos utilizados na análise.

¹⁴ A série final *Proxy* para o comportamento do VA da Agropecuária (A_t) escrita no formato de número-índice e com a média de 2002 = 100 pode ser visualizada no Apêndice 4.

Também se percebe que a série construída diferencia-se da série oficial por dois motivos básicos. Em primeiro lugar, a série oficial é uma série de valor adicionado ao passo que a série construída e utilizada como *proxy* é uma série de produção (lembre-se que o valor adicionado corresponde ao valor de produção menos o consumo intermediário). Em segundo lugar, a série construída não abrange todos os setores da agricultura mineira, apenas quatro. Tampouco faz alguma referência ao comportamento da pecuária mineira, que tem uma importância significativa para o estado de Minas Gerais, seja pelo efetivo dos rebanhos ou pela produção de leite. Todos esses aspectos são captados pela série oficial. Mesmo assim, optou-se por utilizar essa série construída como variável *proxy* para compor o modelo de regressão dinâmica.

5.2.1.2) *Proxy* para o comportamento do VA da Indústria (I_t):

A variável utilizada como *proxy* para o valor adicionado do setor industrial foi o índice de produção física industrial de Minas Gerais divulgado pelo IBGE através da Pesquisa Industrial Mensal de Produção Física (PIM-PF). Entretanto, como o próprio nome da pesquisa informa, os dados são divulgados a cada mês e precisam ser “trimestralizados” para compor o modelo de regressão dinâmica. Este processo foi feito realizando a média aritmética dos meses que compõe cada trimestre em cada um dos anos. Além disso, como a série da PIM-PF já é divulgada no formato de número-índice (com a média de 2002 = 100), não foi necessário à adoção do procedimento feito no caso anterior da agropecuária.¹⁵

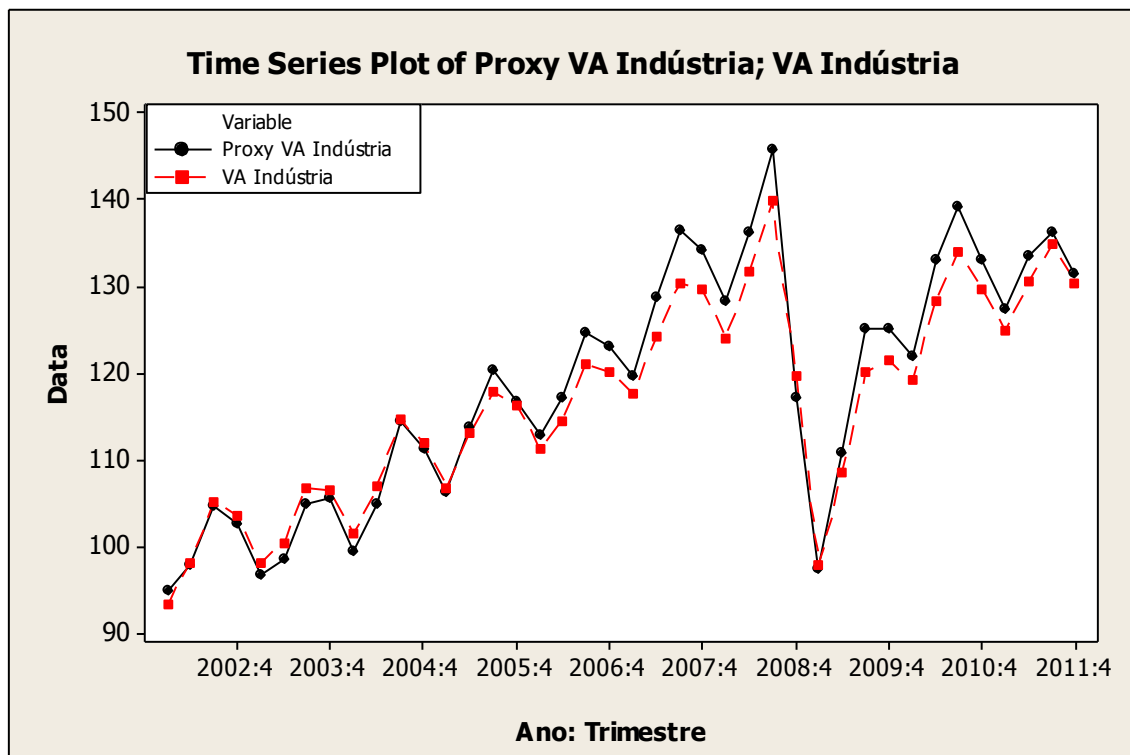
Também neste caso percebe-se que a série construída diferencia-se da série oficial por dois motivos básicos.¹⁶ Em primeiro lugar, a série oficial é uma série de valor adicionado ao passo que a série construída e utilizada como *proxy* é uma série de produção física. Em segundo lugar, o índice da PIM-PF capta o comportamento de apenas dois grandes setores industriais: a extrativa mineral e a indústria de transformação. Já o indicador oficial de VA da indústria mineira e divulgado pela Fundação João Pinheiro aborda todos os setores industriais: a extrativa mineral, a indústria de transformação, a construção civil e a produção e distribuição de energia elétrica e saneamento básico. Mesmo assim, apesar dessas peculiaridades, a série da

¹⁵ Ver Apêndice 3 que traz os valores mensais divulgados pela Pesquisa Industrial Mensal de Produção Física de Minas Gerais, desde janeiro de 2002 até junho de 2013.

¹⁶ A série final *Proxy* para o comportamento do VA da Indústria (I_t) e com a média de 2002 = 100 pode ser visualizada no Apêndice 4.

PIM-PF “trimestralizada” foi bastante aderente à série oficial, como se pode notar pelo Gráfico 22.

Gráfico 22 – Comportamento via gráfico de linhas do Valor Adicionado da Indústria de Minas Gerais e da variável *proxy* (I_t) utilizada para representar o mesmo – média de 2002 = 100 – 1º Trimestre de 2002-4º Trimestre de 2011



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

5.2.1.3) *Proxy* para o comportamento do VA de Serviços (S_t):

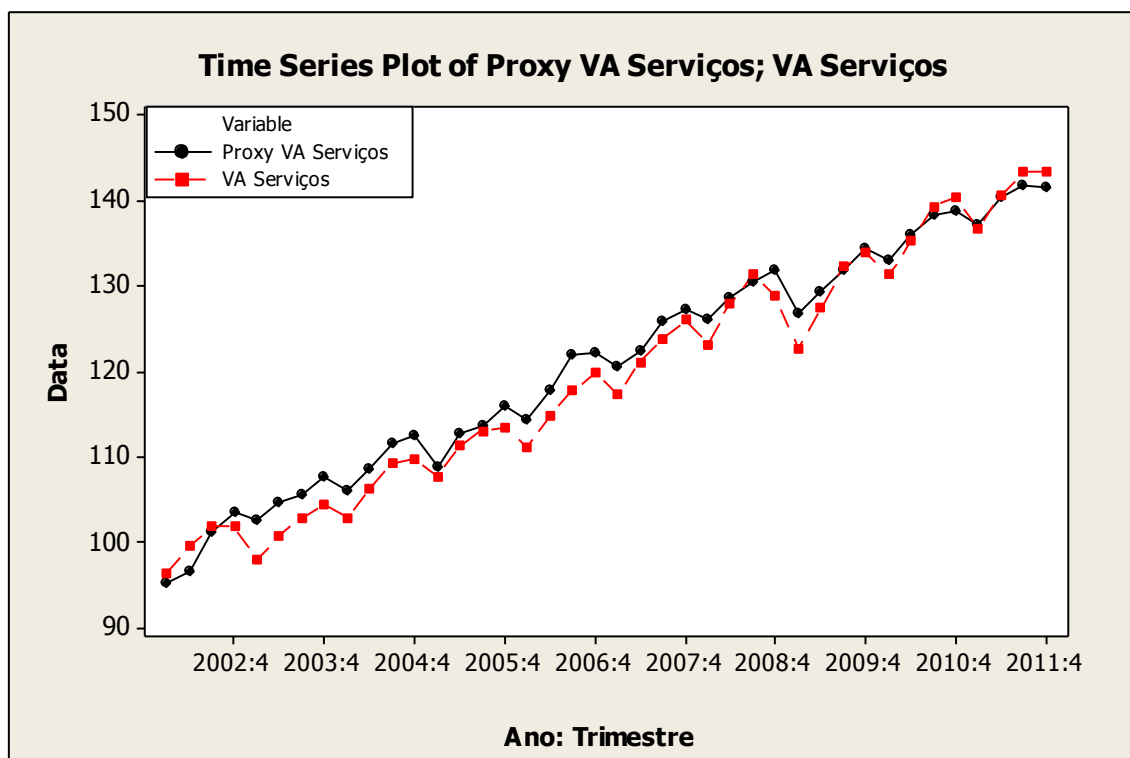
A variável utilizada como *proxy* para o valor adicionado do setor de serviços no estado foi o número de pessoas com 10 anos ou mais de idade ocupadas (mil pessoas) da Região Metropolitana de Belo Horizonte (RMBH) e divulgado pelo IBGE através da Pesquisa Mensal de Emprego (PME). Também neste caso, os dados divulgados mensalmente foram “trimestralizados” para serem utilizados no modelo de regressão dinâmica. Este processo foi feito realizando a média aritmética dos meses que compõe cada trimestre em cada um dos anos. Todavia, como os dados da pesquisa se iniciam em março de 2002, o resultado do primeiro trimestre do respectivo ano foi feito considerando apenas a informação do mês de março disponibilizada.¹⁷

¹⁷ Ver Apêndice 3 que traz os valores mensais divulgados pela Pesquisa Mensal de Emprego da Região Metropolitana de Belo Horizonte (RMBH), desde março de 2002 até junho de 2013.

Além disso, a série obtida foi escrita no formato de número-índice com a média de 2002 = 100. Este procedimento foi feito visando à comparação com a série oficial de índice de volume do VA de serviços, que tradicionalmente é divulgada pela Fundação João Pinheiro na forma de número-índice com o mesmo ano base de referência.¹⁸

A justificativa para utilização de uma série de emprego como *proxy* para o valor adicionado do setor de serviços é que esse setor depende substancialmente do comportamento do consumo das famílias e a capacidade destas consumirem está altamente atrelada ao ritmo de geração e distribuição de renda e do nível de emprego da população. A relação entre o PIB e o emprego é complexa no sentido de que quedas acentuadas no produto agregado podem afetar o mercado de trabalho e vice-versa, porém, o impacto dessa relação se manifesta, em grande medida, no setor terciário da economia (setor de serviços).

Gráfico 23 – Comportamento via gráfico de linhas do Valor Adicionado do setor de Serviços de Minas Gerais e da variável *proxy* (S_t) utilizada para representar o mesmo – média de 2002 = 100 – 1º Trimestre de 2002-4º Trimestre de 2011



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

¹⁸ A série final *Proxy* para o comportamento do VA de Serviços (S_t) e com a média de 2002 = 100 pode ser visualizada no Apêndice 4.

Evidentemente que a série construída não correspondeu exatamente ao comportamento do VA de serviços. Primeiro, por que se utilizou uma série de emprego como *proxy* para o índice de volume do valor adicionado. Segundo, por que a série de emprego faz referência apenas a Região Metropolitana de Belo Horizonte (RMBH) e não ao conjunto do estado de Minas Gerais. Porém, como o setor de serviços é mais representativo no estado nesta região, acredita-se que seja uma boa variável para o comportamento do que acontece no setor de serviços no estado como um todo. Tanto é verdade que a série construída foi bastante aderente à série oficial divulgada pela Fundação João Pinheiro, como se pode notar pelo Gráfico 23.

5.2.2) O modelo de Regressão Dinâmica:

A equação do melhor modelo de regressão dinâmica encontrado, considerando as 40 observações do período amostral, é a seguinte:

$$Y_t = -6,77620 + 0,371278 Y_{t-1} - 0,131681 Y_{t-2} + 0,349576 Y_{t-4} + 0,279648 I_t - 0,163884 I_{t-4} + 0,036674 A_t + 0,316937 S_t + \varepsilon_t$$

Os resultados dos parâmetros estimados e de outras características e informações do modelo (dentro da amostra) são mostrados abaixo:

Regression Equation

$$Y_t = -6,7762 + 0,371278 Y_{t-1} - 0,131681 Y_{t-2} + 0,349576 Y_{t-4} + 0,279648 I_t - 0,163884 I_{t-4} + 0,036674 A_t + 0,316937 S_t$$

Coefficients

Term	Coef	SE Coef	T	P	95% CI	VIF
Constant	-6,77620	1,60566	-4,22020	0,000	(-10,0652; -3,48715)	
Yt-1	0,37128	0,05473	6,78414	0,000	(0,2592; 0,48338)	27,8136
Yt-2	-0,13168	0,03902	-3,37454	0,002	(-0,2116; -0,05175)	13,1060
Yt-4	0,34958	0,08719	4,00924	0,000	(0,1710; 0,52818)	65,4010
It	0,27965	0,02828	9,88868	0,000	(0,2217; 0,33758)	7,1413
It-4	-0,16388	0,03856	-4,24988	0,000	(-0,2429; -0,08489)	14,6952
At	0,03667	0,00426	8,61818	0,000	(0,0280; 0,04539)	7,2668
St	0,31694	0,09750	3,25070	0,003	(0,1172; 0,51665)	69,3748

Summary of Model

S = 0,826079 R-Sq = 99,69% R-Sq(adj) = 99,61%
 PRESS = 32,3994 R-Sq(pred) = 99,47%

Analysis of Variance

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	7	6115,70	6115,70	873,672	1280,28	0,0000000
Yt-1	1	4254,95	31,41	31,407	46,02	0,0000002
Yt-2	1	28,56	7,77	7,771	11,39	0,0021810
Yt-4	1	1189,39	10,97	10,969	16,07	0,0004099
It	1	429,95	66,73	66,730	97,79	0,0000000
It-4	1	161,75	12,33	12,325	18,06	0,0002143
At	1	43,90	50,68	50,684	74,27	0,0000000
St	1	7,21	7,21	7,211	10,57	0,0029941
Error	28	19,11	19,11	0,682		
Total	35	6134,81				

Durbin-Watson statistic = 1,75660

		Correlations: Yt; At; It; St		
		Yt	At	It
MAD	0,586	At 0,297		
MAPE	0,492	0,063		
MSD	0,531	It 0,907	0,102	
SQR	19,107	0,000	0,530	
R-quadrado	0,997	St 0,930	0,045	0,830
R-quadrado ajustado	0,996	0,000	0,782	0,000
Durbin-Watson (DW)	1,757	Cell Contents: Pearson correlation		
Akaike (CIA)	-0,339	P-Value		
Schwarz (CIS)	-0,001			

O modelo de regressão dinâmica ajustado foi o melhor modelo construído para descrever o comportamento da série de índice de volume do PIB de Minas Gerais no período amostral. De fato, comparativamente aos modelos SARIMA e de alisamento exponencial construídos anteriormente, este modelo apresentou melhores resultados em relação às informações de escolha e de desempenho dos modelos. Esta afirmação é corroborada pelos menores valores obtidos para os critérios de Akaike (-0,339) e Schwarz (-0,001) e para as medidas de performance do modelo na amostra: MAD = 0,586; MAPE = 0,492 e MSD = 0,531. O valor extremamente baixo da soma dos quadrados dos resíduos confirma, indiscutivelmente, o excelente desempenho dentro da amostra do modelo de regressão dinâmica (SQR = 19,107).

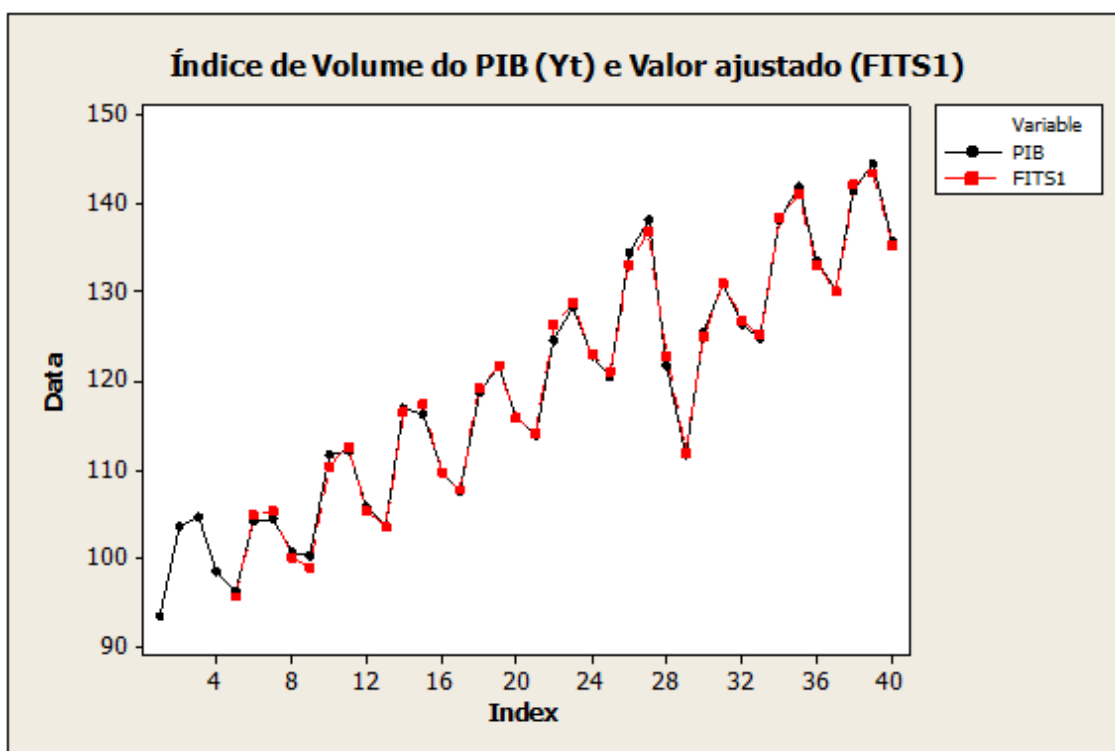
Com relação aos coeficientes obtidos pelo modelo, percebe-se que foram estatisticamente diferentes de zero (p-valores obtidos $< \alpha = 0,05$). Ademais, nota-se que o índice de volume do PIB de Minas Gerais (Y_t) foi influenciado tanto pelo passado da própria série de interesse quanto pela defasagem da variável *proxy* de valor adicionado para o setor industrial (I_{t-4}). O sinal negativo do coeficiente (I_{t-4}) afetando o índice de volume do PIB para baixo pode estar indicando que o crescimento em volume da indústria um ano antes atrapalha o crescimento no presente do PIB porque, possivelmente, a indústria ainda pode estar desovando uma quantidade do seu estoque. Percebe-se também que o índice de volume do PIB de Minas Gerais no presente é influenciado positivamente pelo passado “próximo” da série (Y_{t-1}), indicando, possivelmente, que um crescimento no trimestre imediatamente anterior afeta a confiança dos empresários de forma favorável; e positivamente pelo passado “distante” da série (Y_{t-4}), indicando a influência da sazonalidade da série – o produto agregado tende a ser maior que o observado no mesmo trimestre do ano anterior. Porém, é influenciado negativamente pelo passado “intermediário” da série (Y_{t-2}).

Os coeficientes das variáveis *proxys* dos valores adicionados setoriais (S_t , I_t e A_t) também foram coerentes com a estrutura da economia de Minas Gerais. De fato, de acordo com a Fundação João Pinheiro, o peso dos setores na geração do valor

adicionado da economia mineira é, aproximadamente, de: 60% para o setor de serviços, 30% para o setor industrial e em torno de 10% para o setor agropecuário (FUNDAÇÃO JOÃO PINHEIRO, 2010). Assim, nada mais natural que o coeficiente da variável *proxy* do setor de serviços (S_t) no modelo de regressão dinâmica tenha sido maior do que o coeficiente da *proxy* para a indústria (I_t), uma vez que o setor terciário tem uma importância no que se refere a composição do PIB mineiro muito mais elevada. Da mesma forma, era de se esperar que a influência da *proxy* para o setor agropecuário (A_t) no índice de volume do PIB tenha sido pequena, tendo em vista o peso relativamente reduzido do setor agropecuário na geração do produto agregado.

O Gráfico 24 traz a série original de índice de volume do PIB (Y_t) e os valores ajustados pelo modelo de regressão dinâmica considerando as observações do período amostral. Fica notória a proximidade dos valores ajustados dos verdadeiros valores da série de índice de volume do PIB de Minas Gerais.

Gráfico 24 – Série de Índice de Volume do PIB de Minas Gerais (Y_t) e Valores ajustados pelo modelo de Regressão Dinâmica na série original (dentro da amostra)

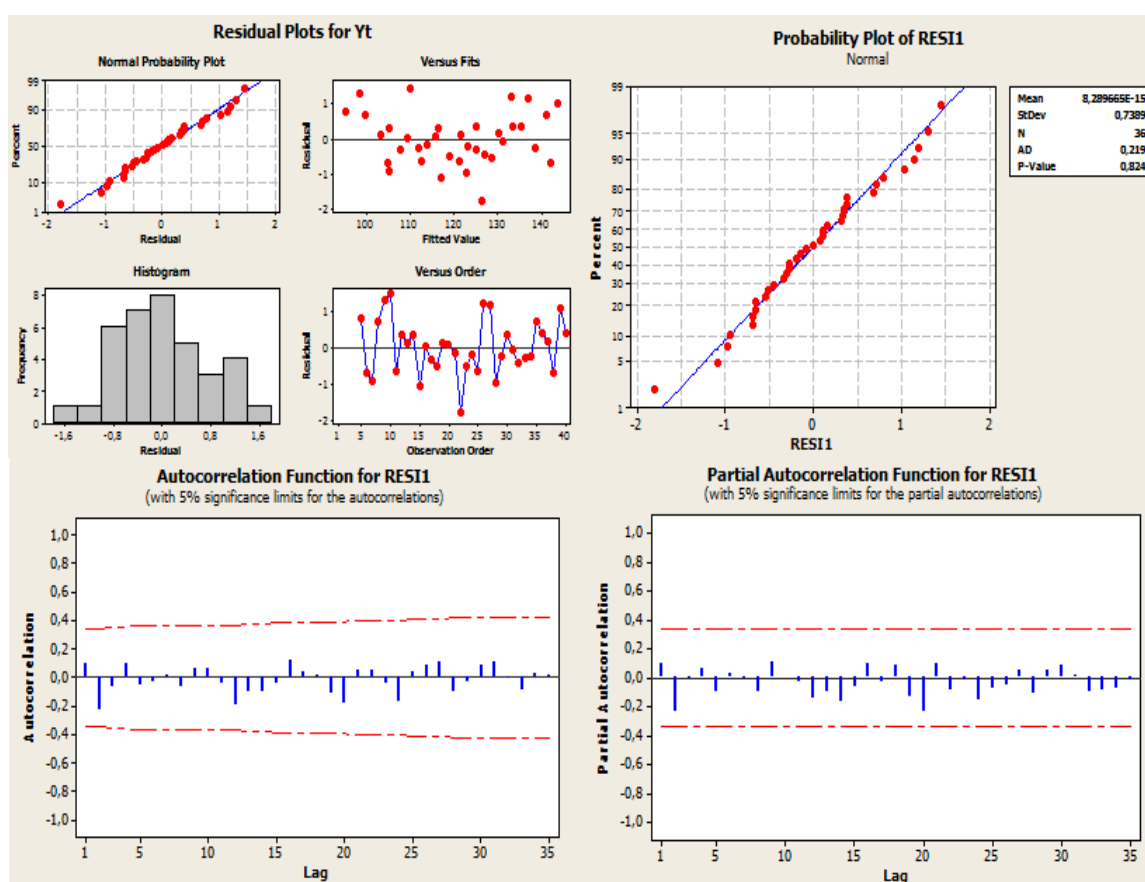


Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Com relação ao comportamento dos resíduos no modelo de regressão dinâmica, o resultado da estatística de Durbin-Watson ($DW = 1,757$) fornece pistas de que o vetor de resíduos (ε_t) deva ser um ruído branco, embora o seu valor não tenha

sendo tão próximo de “dois”. De fato, os correlogramas da FAC e da FACP dos resíduos não apontaram correlações estatisticamente diferentes de “zero”, uma vez que as correlações situaram “dentro” do intervalo pontilhado em vermelho. Ademais, há fortes indícios de que os resíduos siga uma distribuição normal, já que a hipótese nula de normalidade dos resíduos não foi rejeitada (p -valor do teste de normalidade = $0,824 > \alpha = 0,05$). O vetor de resíduos também apresentou média praticamente igual a “zero” (mean = $8,289665 \times 10^{-15}$). A Figura 13 abaixo traz os resultados relativos ao comportamento do vetor de resíduos do modelo de regressão dinâmica.

Figura 13 – Resultados da simulação do modelo de regressão dinâmica (na amostra) para o comportamento do vetor de resíduos (ϵ_t)

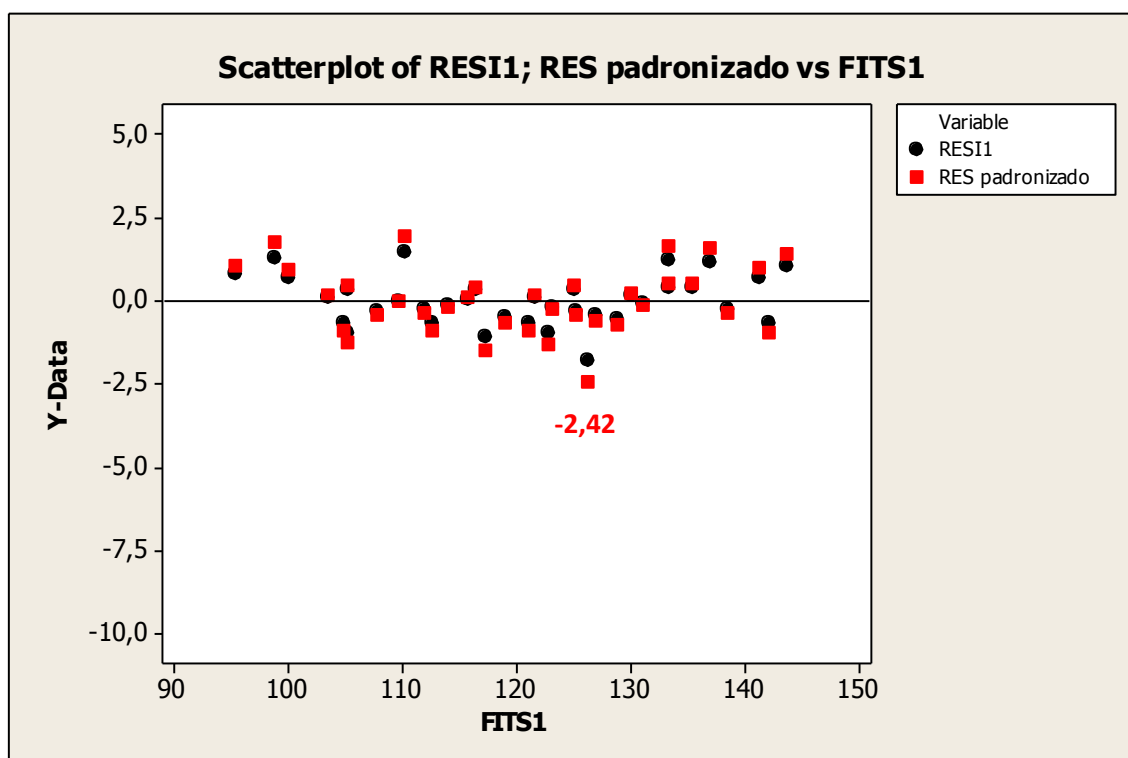


Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Aparentemente, o vetor de resíduos também se distribuiu aleatoriamente, pois não parece existir uma configuração especial na relação entre os resíduos versus os valores ajustados. Outro aspecto favorável ao modelo de regressão dinâmica é que o mesmo conseguiu mensurar melhor a ocorrência da crise econômica mundial no quarto trimestre de 2008. Tanto é verdade que, no modelo construído, o maior valor observado para o resíduo padronizado foi de -2,42 e faz referência ao 2º trimestre de

2007 (e não ao último trimestre de 2008). Se considerar como critério o intervalo resíduo padronizado $\leq -3,1$ ou resíduo padronizado $\geq 3,1$ para detecção de outliers, o modelo de regressão dinâmica não apresenta elementos tão discrepantes quanto os observados nos modelos anteriores (SARIMA e de alisamento exponencial). O Gráfico 25 traduz esses resultados:

Gráfico 25 – Resíduos regulares e padronizados versus valores ajustados pelo modelo de Regressão Dinâmica (dentro da amostra)



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Por outro lado, o modelo apresentou indícios relevantes de multicolinearidade ao combinar o passado da série de índice de volume do PIB com as variáveis *proxys* utilizadas, tendo em vista que as variáveis podem ser entendidas como partes importantes da composição do índice de volume da variável de interesse. Mesmo assim, ao visualizar as correlações entre a variável de interesse (Y_t) e as *proxys* (A_t , I_t e S_t) constata que as mesmas foram significativas na relação entre a variável de interesse com as *proxys* I_t e S_t . No caso da variável *proxy* A_t a correlação com o índice de volume do PIB não foi significativa, mas ficou bem próxima de ser. Portanto, as variáveis *proxys* construídas foram importantes para o entendimento da variável a ser explicada e, por isso, mostraram-se significativas no modelo.

Apesar dos indícios relevantes de multicolinearidade, decidiu-se manter o modelo de regressão dinâmica construído como um bom modelo para representar o que acontece com o índice de volume do PIB. Em primeiro lugar, o problema da multicolinearidade pode ser relativizado pelo fato de que a amostra contém poucas observações e o tamanho da amostra pode afetar substancialmente o modelo. A contraposição do problema de multicolinearidade em razão do tamanho reduzido da amostra é conhecida como micronumerosidade. A citação abaixo traduz este efeito:

“[...] A multicolinearidade não desrespeita nenhuma das premissas da regressão. Serão obtidas estimativas não tendenciosas, consistentes, e seus erros-padrão serão estimados corretamente. O único efeito da multicolinearidade será dificultar a obtenção de estimativas dos coeficientes com erros-padrão pequenos. Mas um pequeno número de observações também provoca esse efeito, como também o fazem as variáveis independentes com pequenas variâncias. Na verdade, no nível teórico, multicolinearidade, poucas observações e pequenas variâncias da variável independente são essencialmente o mesmo problema. Assim, “O que devo fazer a respeito da multicolinearidade?” é uma pergunta como “O que devo fazer quando só conto com poucas observações?” Não há respostas estatísticas para isso” (GUJARATI, 2006, p. 280).

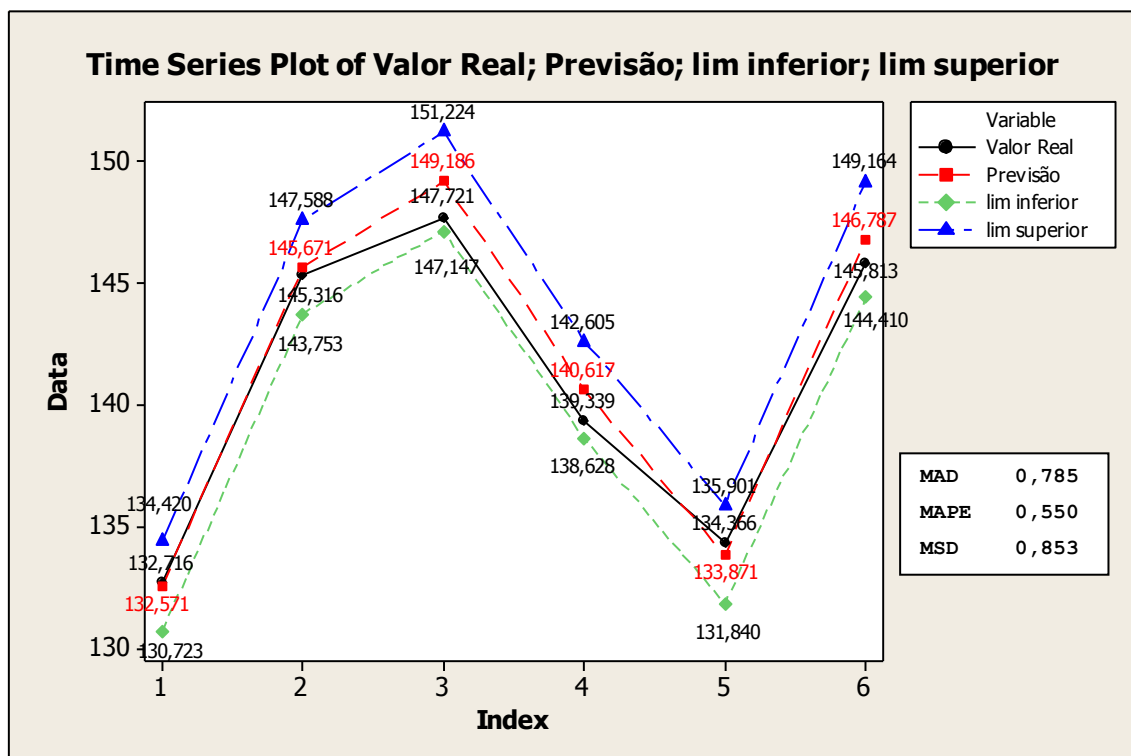
Em segundo lugar, existem situações em que a multicolinearidade não cria problemas graves. Isto ocorre quando o R^2 é elevado e os coeficientes da regressão são individualmente significativos, com altos valores de t (GUJARATI, 2006). Esse é exatamente o caso do modelo de regressão dinâmica construído. Além disso, não parece existir um problema generalizado de multicolinearidade em todas as variáveis. De fato, o problema dimensionado pelo Fator de Inflação da Variância (VIF), parece ser mais grave no caso do índice de volume do produto agregado defasado quatro trimestres no tempo (Y_{t-4}) e no caso da *proxy* construída para o setor de serviços (S_t). Nas demais variáveis não se registrou um VIF tão elevado que não possa ser entendido como aceitável.

A exclusão de variáveis do modelo, sobretudo essas que se mostraram altamente colineares, não se mostrou uma boa solução e contribuiu para a ocorrência de um viés de especificação. Isto ocorre quando se retira variáveis importantes do modelo quando a teoria econômica diz exatamente o contrário: essas variáveis são importantes para o entendimento do índice de volume do PIB e antecipação de situações inesperadas e de crise econômica. Ademais, a retirada de variáveis ou a estimação do modelo de regressão dinâmica apenas com as variáveis *proxys* explicativas, apesar de resolver o problema dos VIF's elevados, piorou substancialmente a performance do modelo no que se refere a sua capacidade de previsão. Gujarati (2006) argumenta que se o propósito da análise de regressão for a

realização de previsão ou prognóstico, então a multicolinearidade pode não ser um problema tão grave.

Em terceiro lugar, o autor argumenta que existem situações em que um VIF elevado não é necessário nem suficiente para a ocorrência de altas variâncias e altos erros-padrão. Isto significa dizer que, a alta multicolinearidade, mensurada por um alto VIF, pode não gerar altos erros-padrão e que o VIF é um medidor “imperfeito” de multicolinearidade. De toda forma, os indícios de multicolinearidade ajudam a entender porque o modelo de regressão dinâmica apresentou um R^2 tão elevado, embora não conte toda a história. Lembre-se de que nos modelos univariados (SARIMA), construídos anteriormente, também foram obtidos R^2 tão elevados quanto o do modelo de regressão dinâmica.

Gráfico 26 – Previsões do modelo de Regressão Dinâmica seis passos a frente e medidas de desempenho



Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Com relação à análise do modelo de regressão dinâmica fora da amostra, percebe-se que os valores previstos acompanharam a trajetória dos valores efetivamente observados. Além disso, as medidas de desempenho do modelo (MAD = 0,785; MAPE = 0,550 e MSD = 0,853) foram tão boas quanto as observadas no modelo SARIMA (0, 1, 1)(1, 1, 0)₄ (MAD = 0,808; MAPE = 0,575 e MSD = 0,782).

Porém, no caso do modelo de regressão dinâmica, o intervalo de confiança preditivo foi muito mais estreito do que o observado no modelo SARIMA $(0, 1, 1)(1, 1, 0)_4$. O Gráfico 26 acima traz os valores previstos pelo modelo de regressão dinâmica, os valores efetivamente observados, o limite inferior e superior de predição, além das medidas de desempenho do modelo no período de validação.

CAPÍTULO 6: CONCLUSÃO

Esta monografia procurou preencher uma lacuna no que se refere ao acompanhamento e previsão da série de índice de volume do PIB de Minas Gerais utilizando tanto modelos univariados (através da própria variável de interesse) quanto modelos multivariados (a variável dependente de interesse e outras variáveis explicativas intervenientes no processo).

Dentro das limitações relativas à construção dos modelos, merece destaque o tamanho reduzido do número de observações utilizadas no período amostral (apenas 40 observações). Apesar dessa dificuldade, foram construídos quatro modelos univariados com alguma qualidade estatística para acompanhar e descrever a série de índice de volume do PIB de Minas Gerais: dois modelos SARIMA e dois modelos de alisamento exponencial.

Em relação à análise dos modelos univariados, pode-se dizer que o modelo SARIMA $(2, 0, 0)(0, 1, 1)_4$ aplicado sobre a série original (Y_t) teve o melhor desempenho no período amostral, pois apresentou os menores valores para as medidas de performance dentro da amostra (MAD, MAPE e MSD). Por outro lado, esse mesmo modelo apresentou o pior resultado no período de validação, indicando que, apesar de ter representado razoavelmente bem a série dentro da amostra, não foi um modelo satisfatório para a realização de previsões. O Quadro 4 traduz estes resultados.

Quadro 4 – Medidas de desempenho (MAD, MAPE e MSD) dos cinco modelos construídos tanto no período amostral quanto no período de validação

Medidas de Desempenho	SARIMA (0, 1, 1)(1, 1, 0) ₄	SARIMA (2, 0, 0)(0, 1, 1) ₄	Alisamento Exponencial Multiplicativo ($\alpha = 0,9$; $\gamma = 0,2$; $\delta = 0,2$)	Alisamento Exponencial Aditivo ($\alpha = 0,9$; $\gamma = 0,2$; $\delta = 0,2$)	Regressão Dinâmica
Período amostral					
MAD	1,956	1,451	2,161	1,992	0,586
MAPE	1,603	1,197	1,878	1,710	0,492
MSD	7,321	4,279	8,149	8,209	0,531
Período de validação					
MAD	0,808	1,871	1,236	1,229	0,785
MAPE	0,575	1,333	0,889	0,876	0,550
MSD	0,782	5,806	1,958	2,568	0,853

Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

De forma geral, os modelos univariados construídos, tanto via modelos SARIMA quanto através da técnica de alisamento exponencial, apresentaram um problema comum: não foram capazes de descrever bem a série no período que coincide com o início da crise econômica mundial – o quarto trimestre de 2008. Esta afirmação é corroborada pela presença de um ponto outlier em todos os modelos univariados.¹⁹ Como o período de crise é marcado pela inflexão não esperada no produto agregado e os resíduos dos modelos foram calculados subtraindo o verdadeiro valor divulgado pela Fundação João Pinheiro do valor estimado, em todos os modelos univariados encontrou-se um ponto outlier menor ou igual a - 3,1 e que faz referência ao quarto trimestre de 2008.

De qualquer maneira, já era esperado que os modelos univariados não fossem capazes de capturar a ocorrência da crise econômica. O argumento é que o passado da série de interesse não possui nenhuma informação com capacidade de prever a inflexão no produto agregado (ou no índice de volume do PIB de Minas Gerais), como a que aconteceu no quarto trimestre de 2008.

Esse problema foi contornado com a construção do modelo de regressão dinâmica, que ao inserir variáveis *proxys* do comportamento dos valores adicionados setoriais capturou, o efeito da crise econômica no estado de Minas Gerais. De fato, a inserção das “novas causalidades”, indicadores de produção física de setores chaves e de nível de emprego, assimilou a redução no índice de volume do PIB de Minas Gerais ocorrida no último trimestre de 2008. Tanto é verdade que o modelo de regressão dinâmica não registrou a ocorrência de pontos outliers como os observados nos modelos univariados: nenhum dos resíduos padronizados foi menor/igual a - 3,1 ou maior/igual a 3,1. Em virtude disso, o modelo de regressão dinâmica foi o que apresentou o melhor resultado no período amostral, com valores extremamente baixos para as medidas de desempenho (MAD, MAPE e MSD).

Com relação à capacidade dos modelos em realizar previsões (análise no período de validação), o desempenho do modelo de regressão dinâmica foi seguido de perto pelo modelo SARIMA (0, 1, 1)(1, 1, 0)₄ aplicado sobre a série original. O Quadro 5 traz a taxa de crescimento real da economia de Minas Gerais divulgada pela Fundação João Pinheiro e a calculada conforme as previsões dos dois melhores modelos na análise fora da amostra. A taxa de crescimento real em cada período é

¹⁹ A identificação de um ponto considerado outlier é controversa, no entanto, resolveu-se adotar nesta monografia o limite de - 3,1 e 3,1 do resíduo padronizado apenas como critério de referência.

calculada dividendo-se o índice de volume de um período em relação ao mesmo período do ano anterior.

Quadro 5 – Taxas de crescimento real (%) da economia mineira segundo o desempenho dos dois melhores modelos no período de validação em diferentes períodos analisados comparativamente a divulgada pela Fundação João Pinheiro

Período	Divulgada pela Fundação João Pinheiro	SARIMA (0, 1, 1)(1, 1, 0) ₄	Regressão Dinâmica
1º trimestre de 2012/1º trimestre de 2011	1,85%	1,10%	1,74%
2º trimestre de 2012/ 2º trimestre de 2011	2,70%	1,77%	2,96%
1º semestre de 2012/1º semestre de 2011	2,30%	1,45%	2,37%
3º trimestre de 2012/3º trimestre de 2011	2,10%	1,90%	3,11%
4º trimestre de 2012/4º trimestre de 2011	2,55%	2,23%	3,49%
2º semestre de 2012/ 2º semestre de 2011	2,32%	2,06%	3,29%
Ano de 2012/Ano de 2011	2,31%	1,76%	2,84%
1º trimestre de 2013/1º trimestre de 2012	1,24%	2,58%	0,98%
2º trimestre de 2013/2º trimestre de 2012	0,34%	2,00%	0,77%
1º semestre de 2013/1º semestre de 2012	0,77%	2,28%	0,87%

Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Como se pode notar pelo Quadro 5, o modelo de regressão dinâmica apresentou melhores resultados para previsão do índice de volume do PIB de Minas Gerais no 1º semestre de 2012 e 1º semestre de 2013 e o modelo SARIMA (0, 1, 1)(1, 1, 0)₄ aplicado sobre a série original foi melhor no 2º semestre de 2012. Para o ano de 2012, o modelo de regressão dinâmica superestimou o crescimento real (2,84%), ao passo que o modelo SARIMA (0, 1, 1)(1, 1, 0)₄ aplicado na série de interesse subestimou o crescimento real (1,76%), tendo em vista que o “verdadeiro” crescimento real da economia mineira foi de 2,31% naquele ano (considerando o último dado divulgado pela Fundação João Pinheiro). Todavia, o modelo de regressão dinâmica foi muito melhor para prever o resultado do 1º semestre de 2013 (0,87%) do que o modelo SARIMA (0, 1, 1)(1, 1, 0)₄ (2,28%), levando-se em conta que a economia estadual cresceu “verdadeiramente” 0,77% no período, conforme a última divulgação pela Fundação João Pinheiro.

Tudo indica que o modelo de regressão dinâmica construído é mais robusto para realizar previsões devido à inserção das “novas causalidades” que antecipam a ocorrência de algum evento inesperado (como o de uma crise econômica), mesmo com os indícios de multicolinearidade. Em todo caso, seria interessante que todos os

modelos construídos neste trabalho fossem reestimados no futuro, já que um número maior de observações contribuiria para a construção de modelos mais próximos da realidade da economia mineira. De fato, um número maior de observações possibilitaria a obtenção de modelos ainda mais robustos com um período amostral ampliado, além de uma análise mais refinada para validação do modelo ou de sua capacidade preditiva (análise fora da amostra). Mesmo assim, apesar de todas as limitações, a realização do trabalho contribuiu para o entendimento do comportamento e a dinâmica da economia estadual e apontou direcionamentos para os diversos atores interessados no resultado do índice de volume do PIB de Minas Gerais: governo estadual, políticos, empresários, consultores, instituições financeiras, estudantes e a sociedade civil de forma geral.

REFERÊNCIAS

ARAÚJO NETO, Djalma Leite; COSTA, Ecio de Farias. **Dimensionamento do PIB do agronegócio em Pernambuco**. Brasília: Ver. Econ. Sociol. Rural, 2005. v.43, n.4, p. 1-24.

BECKER, Marcelo Miranda. **Ano eleitoral apressa inaugurações; veja obras inacabadas**. Terra/Eleições 15 mai. 2012.

CALAZANS, Roberto Balau. **A lógica de um discurso: O empresário Schumpeteriano**. Porto Alegre: Ensaios/Fundação de Economia e Estatística (FEE), 1992, p. 1-28.

CARVALHO, Fernando Cardim de. Políticas econômicas para Economias Monetárias. In: LIMA, Gilberto Tadeu; SICSÚ, João; PAULA, Luiz Fernando de (Org). **Macroeconomia Moderna: Keynes e a Economia Contemporânea**. Rio de Janeiro: Ed. Campus LTDA, 1999. p.258-283.

CLARK, Jeffrey; DOWNING, Douglas. Indicadores Econômicos. In:_____. **Estatística Aplicada**. Ed. Saraiva, 1º edição, 1998, cap. 18, p.383-419.

CORDEIRO JÚNIOR, Olívio Armando. **Utilizando Séries Temporais na Previsão da Arrecadação do Imposto de Renda**. Monografia (Especialização em Orçamento Público). Brasília/Distrito Federal: ISC/ Centro de Formação, Treinamento e Aperfeiçoamento – CEFOR, 2007.

DAVIDSON, Paul. Colocando as evidências em Ordem: Macroeconomia de Keynes versus Velho e Novo Keynesianismo. In: LIMA, Gilberto Tadeu; SICSÚ, João; PAULA, Luiz Fernando de (Org). **Macroeconomia Moderna: Keynes e a Economia Contemporânea**. Rio de Janeiro: Ed. Campus LTDA, 1999. p.35-64.

DIAS, Eduardo Dessupoio Moreira. **Previsão de médio prazo do Consumo de Energia Elétrica no Brasil: estimacão via metodologia Box & Jenkins e Regressão Dinâmica**. Dissertação (Mestrado em Economia Aplicada). Juiz de Fora: Faculdade de Economia e Administração da Universidade Federal de Juiz de Fora, 2008.

FÁVERO, Luiz Paulo [et al.]. **Análise de dados: modelagem multivariada para tomada de decisões**. Rio de Janeiro: Ed. Campus-Elsevier, 2009.

FUNDAÇÃO JOÃO PINHEIRO. **Informativo CEI – PIB MG 2010**. Belo Horizonte: Fundação João Pinheiro – Centro de Estatística e Informações, Resultados Anuais, 2010. Disponível em: <<http://www.fjp.gov.br/index.php/analise-economica/pib-produto-interno-bruto-de-minas-gerais>>. Acesso em: jul. 2013.

FUNDAÇÃO JOÃO PINHEIRO. **Boletim de Conjuntura Econômica de Minas Gerais – 2º trimestre de 2012**. Belo Horizonte: Fundação João Pinheiro – Centro de Estatística e Informações, 2012. ano 5, n. 3, p. 1-54. Disponível em: <<http://www.fjp.mg.gov.br/index.php/analise-economica/boletim-de-conjuntura-de-minas-gerais>>. Acesso em: jul. 2013.

FUNDAÇÃO JOÃO PINHEIRO. **Informativo CEI – PIB 2º trimestre de 2013**. Belo Horizonte: Fundação João Pinheiro – Centro de Estatística e Informações, Planilhas referentes ao PIB-MG (Anexo Estatístico), 2013. Disponível em: <<http://www.fjp.gov.br/index.php/analise-economica/pib-produto-interno-bruto-de-minas-gerais>>. Acesso em: set. 2013.

GIAMBIAGI, Fabio. **Demografia (VII): população e PIB**. Valor Econômico 10 jul. 2013.

GUJARATI, Damodar N. **Econometria Básica**. São Paulo: Ed. Makron Books, 3º edição, 2000.

GUJARATI, Damodar N. **Econometria Básica**. Rio de Janeiro: Ed. Campus-Elsevier, 4º edição, 2006.

IBGE. **Contas Regionais do Brasil**. Ministério do Planejamento, Orçamento e Gestão. Rio de Janeiro: Diretoria de Pesquisas (DPE) e Coordenação de Contas Nacionais (CONAC) – Série Relatórios Metodológicos nº 37, 2008.

IBGE. **Sistemas de Contas Nacionais Brasil 2004-2008**. Ministério do Planejamento, Orçamento e Gestão. Rio de Janeiro: Diretoria de Pesquisas (DPE) e Coordenação de Contas Nacionais (CONAC) – Relatório Contas Nacionais nº 31, 2010.

KRUGMAN, Paul; WELLS, Robin. Monitorando a macroeconomia. In:_____. **Introdução à Economia**. Rio de Janeiro: Ed. Campus-Elsevier, 2007. cap. 24, p.485-508.

LAPORTA, Taís. **PIB deve elevar salário mínimo para cerca de R\$ 725 em 2014, diz Dieese**. São Paulo: IG Economia 01 mar. 2013.

LIMA, Flavia. **Com PEA menor, PIB exige mais produtividade**. São Paulo: Valor Econômico 11 jul. 2013.

MARTÍNEZ, Raydonal Ospina; ZAMPROGNO, Bartolomeu. **Comparação de algumas técnicas de previsão em análise de séries temporais**. Revista Colombiana de Estadística, 2003. vol. 26, p.129-157.

MIGLIOLI, Jorge. Kalecki: A dinâmica das Economias Capitalistas. In:_____. **Acumulação de Capital e Demanda Efetiva**. São Paulo: T. A. Queiroz, Editor, LTDA, 1986. p. 275-296.

MORETTIN, Pedro Alberto; TOLOI, Clélia Maria de Castro. **Previsão de Séries Temporais**. São Paulo: Ed. Atual, 2º edição, 1987.

NASCIMENTO, Monike Gomes do. **Investimento em capital fixo no Brasil: Uma análise de 1991 a 2003**. Monografia (Graduação em Ciências Econômicas). João Pessoa: Universidade Federal da Paraíba, 2006.

PEREZ, Antonio; ROSA, Sílvia; CASTRO, José de. **Dólar fecha a R\$ 2,225, após ata do Copom e nova atuação do BC**. São Paulo: Valor Econômico 19 jul. 2013.

RESENDE, Thiago; SIMÃO, Edna. **Medidas somam estímulos ao consumo e ao investimento**. Valor Econômico 30 ago. 2012.

SEI. **PIB Estadual Anual – Notas Metodológicas**. Salvador: Superintendência de Estudos Econômicos e Sociais da Bahia, 2013, p. 1-12. Disponível em: <http://www.sei.ba.gov.br/index.php?option=com_content&view=article&id=131&Itemid=213>. Acesso em: jul. 2013.

SILVA, Luciano Ferreira da. **O PIB como indicador de qualidade das economias**. Ideias & Números/ Valor Econômico 17 nov. 2011.

SIQUEIRA, Marcelo Lettieri. **Melhorando a Previsão de Arrecadação Tributária Federal através da utilização de Modelos de Séries Temporais**. Monografia (Tributação, Orçamentos e Sistemas de Informação sobre a Administração Financeira Pública). Brasília: ESAF, 2002.

TOSCANO, Ela Mercedes Medrano de. **Alisamento Exponencial**. Belo Horizonte: Apostila do Curso de Especialização em Estatística da UFMG, 2012.

TOSCANO, Ela Mercedes Medrano de. **Análise de Séries Temporais – Modelos ARMA**. Belo Horizonte: Apostila do Curso de Especialização em Estatística da UFMG, 2012.

TOSCANO, Ela Mercedes Medrano de. **Análise de Séries Temporais – Processo Estocástico**. Belo Horizonte: Apostila do Curso de Especialização em Estatística da UFMG, 2012.

TOSCANO, Ela Mercedes Medrano de. **Conceitos Básicos de Processos Estocásticos – Como tratar a não estacionariedade das séries?** Belo Horizonte: Apostila do Curso de Especialização em Estatística da UFMG, 2012.

ZANINI, Alexandre. **Redes Neurais e regressão dinâmica: um modelo híbrido para previsão de curto prazo da demanda de gasolina automotiva no Brasil**. Dissertação (Mestrado em Engenharia Elétrica: Teoria de Controle e Estatística). Rio de Janeiro: Pontifícia Universidade Católica do Rio de Janeiro, 2000.

APÊNDICE

Apêndice 1 – Série encadeada do índice de volume do PIB e do valor adicionado dos setores da economia de Minas Gerais (Agropecuária, Indústria e Serviços) e a arrecadação de impostos (média 2002 = 100) – 1º Trimestre de 2002-2º Trimestre de 2013

Trimestre	Valor Adicionado			Impostos	PIB
	Agropecuária	Indústria	Serviços		
2002:1	76,17	93,34	96,43	92,65	93,40
2002:2	143,75	98,06	99,64	103,03	103,55
2002:3	119,44	105,20	101,99	103,60	104,52
2002:4	60,64	103,40	101,94	100,72	98,53
2003:1	79,43	98,16	97,95	97,30	96,27
2003:2	135,48	100,41	100,69	104,14	104,21
2003:3	106,38	106,81	102,88	103,98	104,33
2003:4	61,14	106,55	104,37	102,09	100,75
2004:1	81,27	101,49	102,92	99,70	100,09
2004:2	155,17	106,84	106,29	110,12	111,68
2004:3	118,71	114,69	109,20	112,28	111,97
2004:4	62,41	111,98	109,82	107,91	105,61
2005:1	70,39	106,71	107,59	105,14	103,66
2005:2	159,03	112,96	111,39	117,52	116,82
2005:3	127,47	117,78	113,06	117,14	116,20
2005:4	64,67	116,31	113,48	111,55	109,67
2006:1	67,10	111,23	111,24	111,50	107,44
2006:2	150,76	114,53	114,74	121,58	118,62
2006:3	143,41	121,05	117,80	124,99	121,80
2006:4	70,05	120,20	120,00	120,84	115,83
2007:1	69,13	117,71	117,27	120,27	113,88
2007:2	141,40	124,10	121,07	130,23	124,59
2007:3	140,93	130,31	123,89	134,62	128,27
2007:4	67,44	129,64	126,04	131,31	122,96
2008:1	78,60	124,08	123,21	129,26	120,55
2008:2	173,54	131,69	128,03	142,29	134,48
2008:3	163,01	139,89	131,35	147,21	138,19
2008:4	70,07	119,63	128,76	131,55	121,85
2009:1	83,80	97,78	122,69	116,88	111,66
2009:2	161,33	108,60	127,40	130,40	125,48
2009:3	149,99	120,05	132,20	137,49	131,02
2009:4	84,64	121,38	134,02	135,95	126,48
2010:1	86,83	119,09	131,32	136,10	124,90
2010:2	172,65	128,31	135,28	150,54	138,23
2010:3	162,23	133,99	139,32	157,74	141,94
2010:4	84,41	129,64	140,27	147,49	133,68
2011:1	87,45	125,00	136,76	143,13	130,30
2011:2	166,77	130,56	140,67	154,06	141,49
2011:3	167,96	134,95	143,27	157,96	144,69
2011:4	84,86	130,34	143,27	150,77	135,87
2012:1	89,77	125,60	140,51	146,27	132,72
2012:2	184,29	130,41	143,86	158,40	145,32
2012:3	172,81	137,12	146,21	160,83	147,72
2012:4	86,27	135,32	146,55	154,48	139,34
2013:1	106,24	121,90	142,75	148,64	134,37
2013:2	176,79	129,36	146,74	160,11	145,81

Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013).

Apêndice 2 – Produção Anual em toneladas 2002-2013 (A) e proporção de safra colhida por trimestre no último ano completo (B) conforme as quatro culturas de Minas Gerais: Batata, Café, Feijão e Milho

(A)

Produção (PAM)/Previsão de safra Agrícola (LSPA)	Batata - inglesa (1ª Safra) (Toneladas)	Batata - inglesa (2ª Safra) (Toneladas)	Batata - inglesa (3ª Safra) (Toneladas)	Café total (arábica e canephora) (Toneladas)	Feijão (1ª Safra) (Toneladas)	Feijão (2ª Safra) (Toneladas)	Feijão (3ª Safra) (Toneladas)	Milho (1ª Safra) (Toneladas)	Milho (2ª Safra) (Toneladas)
2002	456.742	261.910	225.143	1.301.029	215.568	161.011	119.862	4.567.362	240.808
2003	501.417	276.691	248.242	886.925	231.301	185.592	127.254	5.214.660	111.458
2004	471.701	282.533	211.774	1.228.124	169.619	169.027	125.644	5.841.538	110.634
2005	467.734	262.833	273.054	1.002.672	240.895	186.803	131.872	6.124.980	118.893
2006	455.438	273.389	265.304	1.325.238	176.510	187.009	113.121	5.066.548	85.652
2007	602.372	285.208	238.726	987.292	217.002	156.776	107.085	5.964.244	101.833
2008	578.484	412.084	215.368	1.416.106	215.490	214.190	154.612	6.407.426	203.674
2009	532.029	333.963	268.207	1.195.488	243.742	207.379	151.153	6.366.522	170.023
2010	531.794	366.543	245.296	1.504.188	213.999	211.937	197.838	5.925.619	164.322
2011	618.503	374.119	282.466	1.335.738	223.374	178.482	181.110	6.208.835	327.352
2012	507.420	392.914	281.283	1.597.341	218.781	199.740	215.306	7.091.516	533.626
2013	540.572	431.753	300.924	1.535.382	152.844	194.037	218.007	6.828.644	619.555

(B)

Trimestre	Batata - inglesa (1ª Safra) (Toneladas)	Batata - inglesa (2ª Safra) (Toneladas)	Batata - inglesa (3ª Safra) (Toneladas)	Café total (arábica e canephora) (Toneladas)	Feijão (1ª Safra) (Toneladas)	Feijão (2ª Safra) (Toneladas)	Feijão (3ª Safra) (Toneladas)	Milho (1ª Safra) (Toneladas)	Milho (2ª Safra) (Toneladas)
1º Trimestre	100%	0%	0%	0%	100%	0%	0%	13%	0%
2º Trimestre	0%	66%	0%	34%	0%	65%	0%	68%	0%
3º Trimestre	0%	34%	15%	65%	0%	35%	60%	19%	95%
4º Trimestre	0%	0%	85%	2%	0%	0%	40%	0%	5%

Fonte: Elaboração Própria com base na Pesquisa Agrícola Municipal (PAM)/Levantamento Sistemático de Produção Agrícola (LSPA) do IBGE e no Grupo de Coordenação de Estatísticas Agropecuárias de Minas Gerais (GCEA-MG).

Apêndice 3 – Índice Mensal de Produção Física Industrial em Minas Gerais (média de 2002 = 100) (A) e Pessoas de 10 anos ou mais de idade, ocupadas na semana de referência, na Região Metropolitana de Belo Horizonte – RMBH (Mil pessoas) (B) – Janeiro de 2002-Junho de 2013

(A)

Ano/Mês	Janeiro	Fevereiro	Março	Abril	Mai	Junho	Julho	Agosto	Setembro	Outubro	Novembro	Dezembro
2002	96,50	88,43	99,44	96,18	99,54	97,91	103,59	104,89	105,52	109,78	103,70	94,45
2003	95,91	94,52	99,66	95,75	101,59	98,07	102,85	105,14	106,87	109,97	105,30	101,55
2004	97,98	95,05	104,94	101,19	106,79	106,46	112,79	117,16	113,62	116,99	113,22	103,76
2005	106,38	100,62	111,69	110,95	112,78	117,61	119,48	122,84	119,02	123,12	117,43	109,68
2006	111,75	107,59	119,46	112,59	122,36	116,72	122,31	127,42	124,20	127,45	124,99	116,95
2007	118,60	111,12	129,01	123,49	132,77	130,07	136,42	140,84	132,44	140,32	136,77	125,86
2008	130,38	122,79	131,93	132,26	138,67	138,00	148,36	145,97	142,96	142,17	117,72	91,29
2009	92,00	91,10	108,95	103,62	111,73	117,04	124,42	125,99	124,99	131,62	125,42	118,19
2010	117,46	115,06	132,95	129,72	136,84	132,34	138,24	139,73	139,79	140,68	132,88	125,91
2011	121,27	125,16	135,86	127,92	138,49	134,36	137,99	139,06	131,81	135,74	136,33	122,21
2012	118,06	123,72	134,83	127,86	135,38	131,97	138,25	145,16	137,54	149,11	139,92	127,26
2013	130,69	112,15	128,70	131,15	138,03	130,85						

(B)

Ano/Mês	Janeiro	Fevereiro	Março	Abril	Mai	Junho	Julho	Agosto	Setembro	Outubro	Novembro	Dezembro
2002*			1.730	1.739	1.748	1.783	1.819	1.836	1.865	1.901	1.884	1.856
2003	1.871	1.862	1.860	1.904	1.897	1.899	1.875	1.921	1.964	1.941	1.968	1.957
2004	1.903	1.929	1.943	1.961	1.964	1.990	2.001	2.041	2.044	2.043	2.052	2.033
2005	1.973	1.980	1.981	2.006	2.063	2.069	2.035	2.051	2.103	2.094	2.105	2.121
2006	2.069	2.067	2.089	2.090	2.154	2.175	2.187	2.215	2.237	2.220	2.215	2.221
2007	2.188	2.183	2.204	2.225	2.209	2.231	2.276	2.288	2.288	2.290	2.323	2.318
2008	2.282	2.294	2.291	2.345	2.320	2.338	2.353	2.381	2.373	2.408	2.406	2.365
2009	2.312	2.294	2.307	2.318	2.354	2.369	2.404	2.380	2.401	2.428	2.425	2.467
2010	2.412	2.387	2.451	2.449	2.483	2.480	2.494	2.515	2.521	2.524	2.517	2.525
2011	2.465	2.492	2.510	2.519	2.565	2.568	2.574	2.573	2.571	2.565	2.579	2.564
2012	2.561	2.576	2.595	2.602	2.634	2.584	2.586	2.602	2.660	2.652	2.642	2.646
2013	2.592	2.565	2.551	2.573	2.574	2.538						

Fonte: Elaboração Própria com base na Pesquisa Industrial Mensal de Produção Física (PIM-PF) e na Pesquisa Mensal de Emprego (PME).

Nota: * A PME inicia em Março de 2002 e, em razão disso, utilizou-se a informação de Março como referência para o 1º Trimestre de 2002.

Apêndice 4 – Série encadeada do índice de volume do PIB de Minas Gerais e variáveis explicativas adaptadas e utilizadas como proxy para o comportamento dos valores adicionados setoriais (média de 2002 = 100) –
1º Trimestre de 2002-2º Trimestre de 2013

Trimestre	Série de índice de volume do PIB de Minas Gerais (Y)	Variáveis Explicativas		
		Proxy para o comportamento do VA da Agropecuária	Proxy para o comportamento do VA da Indústria	Proxy para o comportamento do VA de Serviços
2002:1	93,40	67,08	94,79	95,26
2002:2	103,55	202,36	97,88	96,73
2002:3	104,52	116,21	104,67	101,32
2002:4	98,53	14,35	102,64	103,54
2003:1	96,27	74,74	96,70	102,66
2003:2	104,21	219,69	98,47	104,62
2003:3	104,33	103,10	104,95	105,72
2003:4	100,75	14,88	105,61	107,67
2004:1	100,09	74,22	99,32	106,00
2004:2	111,68	247,97	104,81	108,57
2004:3	111,97	120,57	114,52	111,70
2004:4	105,61	13,47	111,32	112,48
2005:1	103,66	79,73	106,23	108,91
2005:2	116,82	254,10	113,78	112,66
2005:3	116,20	116,74	120,45	113,60
2005:4	109,67	16,20	116,74	116,00
2006:1	107,44	68,38	112,93	114,26
2006:2	118,62	222,07	117,22	117,82
2006:3	121,80	115,05	124,64	121,85
2006:4	115,83	15,63	123,13	122,17
2007:1	113,88	84,50	119,58	120,68
2007:2	124,59	247,78	128,78	122,33
2007:3	128,27	112,52	136,57	125,76
2007:4	122,96	14,08	134,32	127,21
2008:1	120,55	86,20	128,37	126,04
2008:2	134,48	277,78	136,31	128,54
2008:3	138,19	141,55	145,76	130,44
2008:4	121,85	14,64	117,06	131,77
2009:1	111,66	84,96	97,35	126,88
2009:2	125,48	269,42	110,80	129,23
2009:3	131,02	130,62	125,13	131,88
2009:4	126,48	16,68	125,08	134,35
2010:1	124,90	80,33	121,82	133,07
2010:2	138,23	260,31	132,97	136,04
2010:3	141,94	138,50	139,25	138,21
2010:4	133,68	16,87	133,16	138,87
2011:1	130,30	87,37	127,43	137,05
2011:2	141,49	266,64	133,59	140,45
2011:3	144,69	143,04	136,29	141,66
2011:4	135,87	18,49	131,43	141,48
2012:1	132,72	87,32	125,54	141,92
2012:2	145,32	304,47	131,74	143,53
2012:3	147,72	173,12	140,32	144,04
2012:4	139,34	19,91	138,76	145,73
2013:1	134,37	83,78	123,85	141,48
2013:2	145,81	295,06	133,34	141,05

Fonte: Elaboração Própria com base em FUNDAÇÃO JOÃO PINHEIRO (2013) e com base nas adaptações efetuadas nos dados.