

UNIVERSIDADE FEDERAL DO AMAZONAS
INSTITUTO DE CIÊNCIAS EXATAS
PÓS-GRADUAÇÃO EM MATEMÁTICA

**Séries temporais combinadas com monitoramento estatístico para uso como
ferramentas de gestão dos indicadores mensais de produtividade e resolutividade
do Ministério Público do Estado do Amazonas**

VENILTON RODRIGUES DE MELO

MANAUS-AM

2022



UFAM

SÉRIES TEMPORAIS COMBINADAS COM MONITORAMENTO ESTATÍSTICO
PARA USO COMO FERRAMENTAS DE GESTÃO DOS INDICADORES MENSAIS
DE PRODUTIVIDADE E RESOLUTIVIDADE DO MINISTÉRIO PÚBLICO DO
ESTADO DO AMAZONAS

Venilton Rodrigues de Melo

Dissertação de Mestrado apresentada ao
Programa de Pós-graduação em Matemática,
da Universidade Federal do Amazonas, como
parte dos requisitos necessários à obtenção do
título de Mestre em Matemática

Orientadores: Dr. Jeremias da Silva Leão

Dr. José Mir Justino da Costa

Manaus

Setembro de 2022

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

M528s Melo, Venilton Rodrigues de
Séries temporais combinadas com monitoramento estatístico para uso como ferramentas de gestão dos indicadores mensais de produtividade e resolutividade do Ministério Público da Estado do Amazonas : gestão da produtividade e da resolutividade no Ministério Público do Estado do Amazonas / Venilton Rodrigues de Melo . 2022
48 f.: il. color; 31 cm.

Orientador: Jeremias da Silva Leão
Coorientador: José Mir Justino da Costa
Dissertação (Mestrado em Matemática - Estatística) - Universidade Federal do Amazonas.

1. Metodologia de Box-Jenkins. 2. Método SARIMA. 3. Carta de controle CUSUM (Cumulative Sum). 4. Monitoramento estatístico. I. Leão, Jeremias da Silva. II. Universidade Federal do Amazonas III. Título

SÉRIES TEMPORAIS COMBINADAS COM
MONITORAMENTO ESTATÍSTICO PARA USO COMO
FERRAMENTAS DE GESTÃO DOS INDICADORES
MENSAIS DE PRODUTIVIDADE E RESOLUTIVIDADE
DO MINISTÉRIO PÚBLICO DO ESTADO AMAZONAS

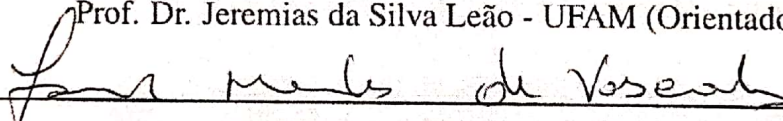
Venilton Rodrigues de Melo

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA, DA UNIVERSIDADE FEDERAL DO AMAZONAS, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM MATEMÁTICA.

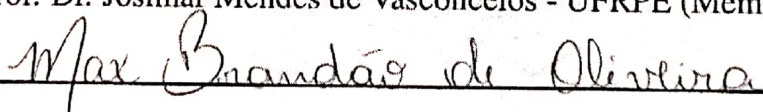
Examinado por:



Prof. Dr. Jeremias da Silva Leão - UFAM (Orientador)



Prof. Dr. Josimar Mendes de Vasconcelos - UFRPE (Membro Externo)



Prof. Dr. Max Brandão de Oliveira - UFPI (Membro Externo)

MANAUS-AM
SETEMBRO DE 2022

*À minha esposa Ednela pelo
apoio e estímulo em nossa jornada.*

Agradecimentos

Agradeço a Nossa Senhora Aparecida a quem sempre posso pedir que interceda junto a seu Filho.

Agradeço aos meus pais, esposa e filhos por serem minha principal fonte de motivação.

Agradeço aos meus colegas de Curso e aos Professores, pelo convívio saudável, embora praticamente em ambiente virtual durante todo o Curso. Em especial, aos colegas Roberto de Souza Oliveira e Antunino da Silva, agradeço pela boa amizade lembrando que todo esforço intelectual honesto vale a pena.

Agradeço ao Professor João Caldas do Lago Neto, por num passado não tão distante, ainda um jovem professor, ter apresentado no contexto da indústria de manufatura, os temas por mim tratados nesta Dissertação, agora em outro contexto.

Agradeço a todos os Professores do Departamento de Estatística da UFAM, pelo esforço coletivo para que as atividades profissionais de estatística ocupem mais espaço no mercado de trabalho, devido à sua importância incontestável.

Minha gratidão aos meus Professores Orientadores Jeremias da Silva Leão e José Mir Justino da Costa, por toda atenção, incentivo e orientações importantes e necessárias que recebi de ambos. Deus lhes proteja!

Resumo da Dissertação apresentada ao Programa de Pós-Graduação em Matemática, da Universidade Federal do Amazonas, como parte dos requisitos necessários para a obtenção do grau de Mestre em Matemática. (M.Sc.)

SÉRIES TEMPORAIS COMBINADAS COM MONITORAMENTO ESTATÍSTICO
PARA USO COMO FERRAMENTAS DE GESTÃO DOS INDICADORES MENSAIS
DE PRODUTIVIDADE E RESOLUTIVIDADE DO MINISTÉRIO PÚBLICO DO
ESTADO DO AMAZONAS

Venilton Rodrigues de Melo

Setembro/2022

Orientadores: Dr. Jeremias da Silva Leão

Dr. José Mir Justino da Costa

Área de Concentração: Estatística

Com a aprovação da Carta de Brasília em setembro de 2016, o Ministério Público Brasileiro assumiu o compromisso de ser excelente em produtividade e resolutividade, tanto na atuação perante o Judiciário, quanto atuando extrajudicialmente como intermediador da pacificação social. Neste trabalho, avaliamos somente a produtividade das Promotorias de Justiça da Capital. Aplicamos métodos estatísticos ao banco de dados Relatório de Atividades Funcionais - RAF do MPAM, no período de janeiro de 2012 a dezembro de 2019. Comparamos as performances de modelos da Metodologia de Box-Jenkins aplicados à série de dados originais, em seguida escolhemos o melhor modelo, o qual gerou uma sequência de dados apropriados que foram usados no monitoramento estatístico da Carta de Controle *CUSUM*. Os resultados obtidos comprovam que a implementação destas metodologias, desde que no ambiente online, darão suporte seguro às gestões da produtividade, da resolutividade e às tomadas de decisão no âmbito do MPAM.

Palavras-chave: Metodologia de Box-Jenkins; Método SARIMA; Carta de Controle *CUSUM*.

Abstract of Dissertation presented to Postgraduate in Mathematics, of the Federal University of Amazonas, as a partial fulfillment of the requirements for the degree of Master of Mathematics. (M.Sc.)

TIME SERIES COMBINED WITH STATISTICAL MONITORING FOR USE AS
MANAGEMENT TOOLS FOR MONTHLY MPAM PRODUCTIVITY AND
RESOLUTENESS INDICATORS

Venilton Rodrigues de Melo

September/2022

Advisors: Dr. Jeremias da Silva Leão

Dr. José Mir Justino da Costa

Research lines: Statistics

With the approval of the Brasília Charter in September 2016, the Brazilian Public Ministry assumed the commitment to be excellent in productivity and resolution, both in acting before the Judiciary, and acting extrajudicially as an intermediary of social pacification. In this work, we only evaluated the productivity of the Public Prosecutors' Offices in the Capital. We applied statistical methods to the MPAM's Functional Activities Report - RAF database, from January 2012 to December 2019. We compared the performances of Box-Jenkins Methodology models applied to the original data series, then we chose the best model, which generated a sequence of appropriate data that were used in the statistical monitoring of the Control Chart *CUSUM*. The results obtained prove that the implementation of these methodologies, as long as they are in the online environment, will provide secure support to productivity management, resolution and decision-making within the scope of the MPAM.

Keywords: Box-Jenkins Methodology; SARIMA method; CUSUM Control Chart.

Sumário

| | |
|---|-------------|
| Lista de Figuras | viii |
| Lista de Tabelas | ix |
| 1 Introdução | 1 |
| 1.1 Contextualização | 1 |
| 1.2 Formulação do problema e definição dos objetivos | 4 |
| 1.3 Estrutura da dissertação | 6 |
| 1.4 Suporte computacional | 7 |
| 2 Preliminares | 8 |
| 2.1 Metodologia de Box e Jenkins para o estudo de séries temporais | 8 |
| 2.1.1 Modelos autoregressivos de ordem p - $AR(p)$ | 9 |
| 2.1.2 Modelos de médias móveis de ordem q - $MA(q)$ | 9 |
| 2.1.3 Modelos autoregressivos e de médias móveis de ordens p e q - $ARMA(p, q)$ | 9 |
| 2.1.4 Modelos autoregressivos integrados e de médias móveis de or- dens p , d e q - $ARIMA(p, d, q)$ | 10 |
| 2.1.5 Modelos sazonais autoregressivos integrados e de médias móveis - $SARIMA(p, d, q)(P, D, Q)_s$ | 11 |
| 2.1.6 Modelos estacionários e não estacionários | 12 |
| 2.1.7 Critérios penalizadores para seleção de modelos | 13 |
| 2.2 Cartas de Controle CUSUM | 15 |
| 2.2.1 Monitoramento estatístico da média | 15 |
| 2.2.2 Carta de Controle CUSUM Tabular | 17 |
| 2.2.3 Planejamento da Carta de Controle CUSUM | 20 |

| | | |
|----------|---|-----------|
| 2.2.4 | Séries Temporais e Monitoramento Estatístico | 23 |
| 3 | Aplicação das Metodologias e Análises dos Resultados | 24 |
| 3.1 | Objeto de estudo | 24 |
| 3.2 | Procedimento de coleta de dados | 24 |
| 3.3 | Resultados obtidos | 25 |
| 3.3.1 | Metodologia de Box e Jenkins | 26 |
| 3.3.2 | Monitoramento Estatístico | 43 |
| 4 | Conclusões e proposta de novos estudos para uso de outras metodologias | 45 |
| | Referências Bibliográficas | 47 |

Lista de Figuras

| | | |
|------|---|----|
| 3.1 | Produtividade das Promotorias de Justiça da Capital | 26 |
| 3.2 | Análise da função de autocorrelação (FAC) | 27 |
| 3.3 | Análise da função de autocorrelação parcial (FACP) | 28 |
| 3.4 | Componentes sazonais | 29 |
| 3.5 | Resultado da diferenciação sazonal da série | 30 |
| 3.6 | Função de autocorrelação aplicada aos dados diferenciados | 31 |
| 3.7 | Função de autocorrelação parcial aplicada aos dados diferenciados | 32 |
| 3.8 | Indicativo de Estacionariedade | 35 |
| 3.9 | Resultado do Ajuste aos Dados | 35 |
| 3.10 | Diagnóstico dos Resíduos | 36 |
| 3.11 | Indicativo de Estacionariedade (função auto.arima) | 38 |
| 3.12 | Resultado do Ajuste aos Dados (função auto.arima) | 38 |
| 3.13 | Diagnóstico dos resíduos (função auto.arima) | 39 |
| 3.14 | Previsão da produtividade no período de jan-2020 a dez-2021 | 41 |
| 3.15 | Destaque da previsão no período de jan-2020 a dez-2021 | 42 |
| 3.16 | Carta de Controle <i>CUSUM</i> aplicada à série de resíduos | 44 |

Lista de Tabelas

| | | |
|------|---|----|
| 2.1 | Exemplo de uma tabela CUSUM tabular | 19 |
| 2.2 | Desempenho da Carta de Controle CUSUM Tabular | 20 |
| 2.3 | Valores dos Parâmetros da CUSUM Bilateral | 21 |
| 3.1 | Produtividade das Promotorias de Justiça da Capital | 25 |
| 3.2 | Evidência em favor de Sazonalidade Determinística | 29 |
| 3.3 | Evidência em favor de Aleatoriedade | 32 |
| 3.4 | Teste ADF: Evidência em favor de Estacionariedade | 33 |
| 3.5 | Teste PP: Evidência em favor de Estacionariedade | 33 |
| 3.6 | Teste KPSS: Evidência em favor de Estacionariedade | 33 |
| 3.7 | Estimação de Parâmetros | 34 |
| 3.8 | Teste z para os coeficientes | 34 |
| 3.9 | Estimação de Parâmetros (função auto.arima) | 37 |
| 3.10 | Teste z para os coeficientes (função auto.arima) | 37 |

Capítulo 1

Introdução

1.1 Contextualização

O Ministério Público do Estado do Amazonas - MPAM tem na sua estrutura física atual, vinte e quatro Procuradorias de Justiça e cento e seis Promotorias de Justiça Especializadas, todas sediadas na capital, além de uma ou mais Promotorias de Justiça sediadas em cada Comarca do interior do Estado. Todos estes órgãos ministeriais e outros, não menos importantes, têm a incumbência constitucional de defender a ordem jurídica, o regime democrático e os interesses sociais e individuais indisponíveis, quais sejam a vida, a liberdade, a saúde e a dignidade humana. Exemplificando a indisponibilidade de direitos, uma pessoa não pode vender um órgão do seu corpo, embora ele lhe pertença. Ainda, ao MPAM é atribuída a missão institucional de promover a Justiça e manter a ordem jurídica, atuando na garantia do pleno funcionamento das instituições democráticas, bem como na defesa das condições de vida e de desenvolvimento, respeitando-se as diversidades populacionais e ambientais no Estado do Amazonas.

A Carta de Brasília, acordo celebrado entre todas as Corregedorias do Ministério Público Brasileiro, assinada em setembro de 2016, no âmbito do 7º Congresso de Gestão do Conselho Nacional do Ministério Público - CNMP, objetiva modernizar o controle correicional da atividade extrajudicial, bem como fomentar a atuação resolutiva em todo o Ministério Público Brasileiro.

O fomento à atuação resolutiva é o que motiva e justifica a realização deste trabalho acadêmico, pois embora as diretrizes da Carta de Brasília para a avaliação, orientação e fiscalização das atividades extrajudiciais estejam dirigidas ao controle correicional,

na perspectiva da contribuição dos métodos estatísticos para o aprimoramento da Gestão Institucional, pode-se destacar dentre as próprias diretrizes:

- (i) Elaboração de diagnósticos e prognósticos das atividades Institucionais;
- (ii) Avaliação da duração razoável do processo, considerando, para tanto, principalmente, as necessidades do direito material, de modo a aferir se há hipótese concreta de necessidade de agilização do procedimento em situações de urgência ou se há necessidade de alargamento do procedimento nos casos em que a complexidade na matéria de fato e de direito o exigir; e
- (iii) Criação de parâmetros não meramente formais para subsidiar a avaliação e a fiscalização do tempo de tramitação dos procedimentos extrajudiciais, considerando-se, para isso, a efetividade das diligências determinadas, os intervalos entre os impulsionamentos (períodos em que o procedimento resta concluso), assim como a adoção de instrumentos resolutivos e outras medidas.

Produtividade e Resolutividade são conceitos distintos, mas não podem ser apartados. Uma Instituição Pública de existência perene, pode ser considerada produtiva segundo seus próprios critérios internos, ter sua execução orçamentária aprovada por órgãos externos de controle ao longo do tempo, sem, no entanto, ser resolutiva, quanto à finalidade para a qual foi criada, quando confrontada com os anseios legítimos da população. Por outro lado, a resolutividade, entendida como “fazer o que tem que ser feito, da melhor forma possível e no menor tempo”, tem como efeito colateral o atingimento das metas de produtividade preestabelecidas em qualquer tipo de planejamento, com destaque para o Planejamento Estratégico.

A busca por níveis aceitáveis de produtividade e resolutividade necessita de gestão apropriada e governança efetiva, suportadas em informações e conhecimento oriundos de diagnósticos precisos, prognósticos e modelos de resposta extraídos das bases de dados, via de regra, gigantescas, de propriedade das Instituições Públicas, sendo que os modelos de resposta não serão abordados neste trabalho.

Configura-se o diagnóstico em análises estatísticas descritivas precisas do presente-passado recente ou passado remoto. Porém, as análises estatísticas descritivas não são apropriadas para que, com base nelas mesmas, sejam feitas afirmações quanto ao futuro, por mais completas e abrangentes que sejam. Por sua vez, o prognóstico elaborado com métodos estatísticos computacionais de alto desempenho tem o poder de criar um ambi-

ente estratégico, dotado de previsibilidade, adequado para planejar, executar e atingir metas Institucionais pré-estabelecidas, provocando o natural giro contínuo do Ciclo PDCA, Shewhart (1980), habilitando os tomadores de decisão a concretamente trabalharem em regime de metas.

Também conhecido como Ciclo de Shewhart, o Ciclo PDCA pode ser aplicado aos mais diversos projetos profissionais e estratégias em diferentes contextos, embora tenha surgido no contexto da Gestão da Qualidade. Trata-se de um mecanismo iterativo e contínuo de administração que se baseia em quatro etapas. O nome PDCA corresponde a uma sigla emprestada do inglês, fazendo referência a estas quatro fases para a gestão: *Plan, Do, Check, Act*, que em português quer dizer Planejar, Fazer, Checar, Agir. No jargão do Ciclo PDCA: Não se gerencia o que não se mede, não se mede o que não se define, não se define o que não se entende e não há sucesso no que não se gerencia.

A seguir, explicamos melhor essas etapas.

Plan (Planejar)

Todo processo bem sucedido começa com um bom planejamento. Para construir um bom planejamento é necessário estabelecer pelo menos uma meta e um método que vai nortear a execução das atividades. Metas são as etapas que precisam ser concluídas para atingir um objetivo. O Ciclo PDCA pode ser adotado para garantir o cumprimento de metas no prazo determinado, da forma mais eficiente possível. Além de ter uma meta bem definida como alvo, o planejamento é o momento certo para escolher o método que será utilizado, ou seja, qual o exemplo ou roteiro será usado para chegar até a meta.

Do (Fazer)

Encerrada a etapa de planejamento, vem a execução do projeto. A etapa de execução envolve, antes mesmo da execução em si, essencialmente a plena capacitação dos profissionais envolvidos e a alocação dos recursos necessários. Portanto, o treinamento de pessoal e a alocação de recursos devem ser partes intrínsecas da execução de qualquer processo de trabalho no contexto do Ciclo PDCA.

Check (Checar)

Nesta terceira etapa do Ciclo PDCA é feita uma verificação quanto ao funcionamento do plano, analisando a eficácia das atividades. Os resultados positivos devem ser reconhecidos como boas práticas que podem passar a ser aplicadas por todos na Instituição. Cabe salientar que para checar é necessário estabelecer uma forma de avaliação, notadamente

através do monitoramento estatístico de Indicadores de Performance.

Act (Agir)

Chega-se assim ao momento de avaliar os resultados alcançados, tomar decisões e partir para um novo giro do Ciclo PDCA, iniciando, portanto, uma nova etapa de planejamento.

1.2 Formulação do problema e definição dos objetivos

Quando a Carta de Brasília apontou a necessidade de resolutividade na atuação ministerial, implicitamente se exigia resultados consolidados de análise estatística de dados para subsidiar os gestores e tomadores de decisão. Contudo, é necessário registrar que desde a década de 60 do século passado, principalmente devido aos trabalhos desenvolvidos, Shewhart (1980) e Deming (1990), essas ferramentas de suporte à gestão e à tomada de decisão já estão disponíveis a nível planetário, e atualmente podem se beneficiar do uso intensivo dos recursos mais avançados da Estatística Computacional.

No caso do MPAM, desde 2012 a Instituição mantém atualizados seus dados de produtividade com base em uma taxonomia nacional, o que permite produzir e publicar um relatório mensal online de diagnóstico e prognóstico da produtividade. Porém, ainda é necessária a criação de variáveis para atender a gestão da resolutividade. Portanto, este estudo trata apenas da produtividade, e tão somente no que concerne à atividade-fim nas Promotorias de Justiça da Capital, mas o propósito mais abrangente é validar e aplicar os métodos usados neste trabalho acadêmico para obter melhorias da gestão da produtividade e da resolutividade, tanto na área-fim como na área-meio de todo o MPAM.

Diante disso, podemos dizer que o objetivo geral deste estudo é aplicar aos dados de produtividade das Promotorias de Justiça da Capital, um modelo de previsão de séries temporais combinado com métodos de monitoramento estatístico para pequenas variações, aplicados aos termos aleatórios dos modelos, exatamente para gerar e monitorar estatisticamente os prognósticos dotados dos melhores padrões de previsibilidade em simultâneo com o monitoramento estatístico dos dados observados, no contexto da dinâmica do Ciclo PDCA, Shewhart (1980).

A seguir, elencamos os objetivos específicos (OE):

OE1 Identificar e aplicar modelos adequados de análise e previsão de Séries Tempo-

rais de Box e Jenkins, Box & Jenkins (1976), aos arquivos de dados de produtividade das Promotorias de Justiça da Capital no período de janeiro de 2012 a dezembro de 2019;

OE2 Comparar os modelos identificados e aplicados aos dados e escolher o que apresentar melhor aderência aos dados originais e maior capacidade de previsão dos futuros dados da série histórica.

OE3 Usar a Carta de Controle *CUSUM* para monitoramento estatístico de pequenas variações, Page (1954), nos resíduos de comportamento admitido gaussiano (o nível acadêmico deste estudo não levanta a hipótese de ausência desta admissão) gerados pelas Séries Temporais modeladas. Então, o padrão estatístico sequencial dos resíduos representa o padrão estatístico do Indicador monitorado, isso para cada novo dado registrado na Série Temporal gerada pelo próprio Indicador.

Assim, cada novo valor da Série Temporal, produz uma nova previsão de futuro, cabendo ao gestor tão somente observar se a nova previsão aponta para a meta preestabelecida ou está se afastando, por exemplo. Aí lhe resta, como gestor, fazer ajustes, mudanças, identificar e eliminar causas, até trocar as pessoas se necessário, ou seja, fazer gestão.

Quanto à Carta de Controle *CUSUM*, ela sempre vai mostrar se o padrão de desempenho melhorou ou continua o mesmo, ou piorou. Em qualquer caso, se houve mudança, a Carta de Controle *CUSUM* atualiza o modelo automaticamente. Toda vez que a performance do Indicador apresenta melhora, o que representa a capacidade do gestor e de sua equipe de mudar a realidade a seu favor, espera-se que a variabilidade diminua, tornando os ambientes tático e estratégico mais previsíveis.

A Carta de Controle *CUSUM* gera um visual de fácil entendimento, adequado para os gestores e tomadores de decisão que não lidam, via de regra, com métodos estatísticos. Na verdade, a Carta *CUSUM* funciona como uma bússola na mão dos gestores e tomadores de decisão. As Instituições que estão adotando a Cultura da Análise Estatística de Dados nos moldes atuais, conhecida no contexto da Gestão Moderna e da própria operacionalização dos Métodos Quantitativos da Ciência como *Analytics*, adquirem estas Instituições Públicas e/ou Privadas, a capacidade de mudar a realidade, atingir metas concretamente, aprendendo a trabalhar em regime de metas.

1.3 Estrutura da dissertação

A presente dissertação é constituída de quatro capítulos. No primeiro capítulo, contextualizamos o ambiente de funcionamento do MPAM, sua finalidade institucional e importância para a população do Estado do Amazonas, quanto a promoção da Justiça e manutenção da ordem jurídica, que configuram seus desafios essenciais.

Desafios estes que justificam plenamente o uso de métodos estatísticos para auxiliar na atividade de gestão institucional. No jargão da Administração Moderna, não administramos o que não conhecemos; não conhecemos o que não somos capazes de medir; no entanto, os métodos estatísticos possuem a capacidade de realizar este tipo de medição, o qual não é puramente determinístico. Sendo assim, foram definidos os objetivos a serem atingidos com a realização deste estudo. Em seguida apresentamos a própria estrutura desta dissertação. Ainda, destacamos o suporte computacional necessário para a realização deste trabalho acadêmico.

No segundo capítulo são apresentados os métodos estatísticos de análise e previsão de séries temporais (Box e Jenkins), seguidos da apresentação de um modelo de monitoramento estatístico para pequenas variações (Carta de Controle *CUSUM*), configurando as duas metodologias estatísticas usadas neste trabalho.

Em detalhes, está exposta a estrutura do modelo de Box e Jenkins, apresentando desde o modelo autoregressivo, seguido do modelo de médias móveis, ampliando para o modelo autoregressivo e médias móveis, alcançando o modelo autoregressivo integrado de médias móveis, e também o modelo includente de sazonalidade, a questão da estacionariedade das séries temporais e ainda três critérios penalizadores de modelos, quais sejam o *AIC* de Akaike, o *AICc* e o *BIC* de Schwarz.

Em seguida é apresentado o modelo para monitoramento de sequências de dados estatisticamente independentes, a Carta de Controle *CUSUM* Tabular, constando o seu necessário planejamento para aplicação e ainda o conseqüente relacionamento entre monitoramento estatístico e séries temporais.

No terceiro capítulo são apresentados os resultados das modelagens da série usando Box-Jenkins com seus respectivos usos em combinação com o monitoramento estatístico através da *CUSUM*.

Por fim, no quarto capítulo encontram-se as conclusões deste trabalho acadêmico e as sugestões consideradas pertinentes para estudos futuros.

1.4 Suporte computacional

Os resultados apresentados nesta dissertação, incluindo os gráficos, foram produzidos utilizando o software R, em sua versão 4.2.0 para o sistema operacional Windows. Esta linguagem foi criada por Ihaka & Gentleman (1996) para produzir um ambiente de programação semelhante ao S, e encontra-se disponível gratuitamente em <https://cran.r-project.org/bin/windows/base/>. Para detalhes ver (principais referências). Para a elaboração do presente texto, foi usado o sistema tipográfico \LaTeX , desenvolvido por Lamport (1985), que consiste em uma série de macros ou rotinas do sistema \TeX . O \LaTeX é distribuído de forma gratuita e maiores detalhes podem ser encontrados em Knuth (1986) e Salomon (2012), entre outros.

Capítulo 2

Preliminares

2.1 Metodologia de Box e Jenkins para o estudo de séries temporais

Uma série temporal é uma coleção de observações colhidas sequencialmente ao longo do tempo, apresentando como característica primordial a forte dependência entre os dados vizinhos. Devido a sua especificidade, ela apresenta características peculiares que dificultam sua análise. Dentre elas, de acordo com Morettin (2021) estão: presença de **tendência**, ou seja, a evidência de sentido para a série temporal ao longo dos anos; ocorrência de **ciclos** e **sazonalidade**, ou seja, movimentos ondulatórios na série de longa e curta duração respectivamente; assim como a presença de “**ruído aleatório**” ou “**erro**”, abrangendo a variabilidade intrínseca aos dados e que não pode ser modelado por modelos determinísticos.

Por esses motivos, como a maior parte dos procedimentos estatísticos foi desenvolvida para analisar observações independentes, o estudo baseado em Séries Temporais necessita da aplicação de metodologias específicas. Nesse contexto, os modelos desenvolvidos por Box & Jenkins (1976) destacam-se entre os principais métodos aplicados às atividades de **análise** e **previsão** das séries temporais.

2.1.1 Modelos autoregressivos de ordem p - AR(p)

De acordo com Werner & Ribeiro (2003), no modelo autorregressivo AR(p) a série de dados históricos Z_t é descrita por seus valores passados mais o ruído aleatório (ε_t), definido por meio da equação (2.1):

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + \varepsilon_t, \quad (2.1)$$

em que ϕ descreve (Z_t) com os valores de Z_{t-1} para $i = 1, 2, \dots, p$. Para Souza (2016), vale ressaltar que o modelo só será considerado estacionário se o parâmetro ϕ convergir para $|\phi| < 1$, também denominando-se "condição de estacionariedade".

2.1.2 Modelos de médias móveis de ordem q - MA(q)

No modelo denominado de Médias Móveis – MA(q), por sua vez, a série Z_t é resultado da combinação dos ruídos brancos ou aleatórios (ε_t) do período atual com os anteriores, sendo representado pela equação (2.2):

$$\tilde{Z}_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}, \quad (2.2)$$

em que θ descreve como (Z_t) se relaciona com o valor (ε_{t-i}) para $i = 1, 2, \dots, q$, segundo Werner & Ribeiro (2003). De acordo com Souza (2016), este modelo só será considerado estacionário se for de ordem finita, ou seja, $|\theta| < 1$, também denominando-se "condição de invertibilidade".

2.1.3 Modelos autoregressivos e de médias móveis de ordens p e q - ARMA(p, q)

De acordo com Marchezan & Souza (2010), o modelo ARMA(p, q) é a combinação dos dois modelos apresentados anteriormente, sendo utilizado em séries encontradas na prática para as quais não se deseja um grande número de parâmetros. Nesse contexto, o modelo ARMA(p, q) é dado pela equação (2.3), abaixo:

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \dots + \phi_p \tilde{Z}_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \quad (2.3)$$

em que o parâmetro ϕ representa a parte autorregressiva, assim como o parâmetro θ representa a parte de médias móveis. Ao mesmo tempo é perceptível a existência das mesmas condições de estacionariedade e invertibilidade tanto no modelo $ARMA(p, q)$, quanto no método autorregressivo $AR(p)$, quanto à estacionariedade e no modelo de médias móveis $MA(q)$, quanto à invertibilidade, ou seja, $|\phi| < 1$ e $|\theta| < 1$, respectivamente.

Segundo Morettin (2021), os três modelos apresentados anteriormente são descritos pelos chamados modelos autoregressivos integrados e de médias móveis de ordens p , d e q - $ARIMA(p, d, q)$, que podem ser generalizados pela inclusão de um operador sazonal.

2.1.4 Modelos autoregressivos integrados e de médias móveis de ordens p , d e q - $ARIMA(p, d, q)$

Uma metodologia bastante utilizada na análise de modelos paramétricos é conhecida como abordagem de Box & Jenkins (1976). Tal metodologia consiste em ajustar modelos autoregressivos integrados e de médias móveis, $ARIMA(p, d, q)$, a um conjunto de dados, ver detalhes Morettin (2021). A construção do modelo é baseada em um ciclo iterativo, no qual a escolha da estrutura do modelo se baseia nos próprios dados. As fases do ciclo iterativo são:

- (a) *Especificação*: uma classe geral de modelos é considerada para a análise;
- (b) *Identificação*: feita com base na análise de autocorrelações, autocorrelações parciais e outros critérios;
- (c) *Estimação*: os parâmetros do modelos são estimados;
- (d) *Verificação* ou *Diagnóstico*: feito através de uma análise de resíduos, para saber se o modelo é adequado para os fins em vista (previsão, por exemplo).

A utilização desse processo possibilita estimar diversos comportamentos na série, sendo fundamental para a **acurácia** nas previsões. Segundo Marchezan & Souza (2010), o método de Box & Jenkins (1976) consiste na procura por um modelo Autoregressivo Integrado de Médias Móveis - $ARIMA$, objetivando representar o processo estocástico gerador da série temporal. Nesse sentido, os modelos $ARIMA$ são métodos estatísticos que almejam captar o comportamento da correlação seriada, ou autocorrelação entre os

valores da série temporal e com base nesse comportamento, realizar previsões para observações futuras.

Os modelos ARIMA são resultados das combinações de três componentes, também, denominadas de “**filtros**”: a componente autoregressiva (AR), o filtro de integração (I) e a componente de médias móveis (MA), sendo essa interação a promotora de vários **modelos de previsão**. Adicionado a isso, segundo Souza (2016), a utilização do método ARIMA figura como uma poderosa solução para muitos problemas de previsão, proporcionando previsões extremamente acuradas para modelos de séries temporais.

Portanto, os modelos lineares não estacionários, baseados em processos lineares homogêneos que não apresentam comportamento explosivo, podem se tornar estacionários por meio de um processo finito de diferenciações d , processo descrito pelos modelos ARIMA.

Para a correta aplicação do método de Box & Jenkins (1976), há a necessidade da série ser estacionária, requerendo a transformação dos dados originais por meio de diferenças, sendo que, de acordo com eles, 1 ou 2 diferenças são suficientes para estacionarizar séries homogêneas Morettin (2021). Nesse sentido, segundo Werner & Ribeiro (2003), o modelo $ARIMA(p, d, q)$ é dado pela equação (2.4)

$$W_t = \phi_1 W_{t-1} + \dots + \phi_p W_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}, \quad (2.4)$$

em que o parâmetro ϕ representa a parte autoregressiva; θ indica a presença de médias móveis, assim como, nesse caso já se subentende a presença da série com diferença.

2.1.5 Modelos sazonais autoregressivos integrados e de médias móveis - SARIMA(p, d, q)(P, D, Q)_s

Consideradas as especificidades dos fenômenos abordados em séries temporais, a utilização de outros modelos também se faz necessária. Uma delas, de acordo com Souza (2016), é a ocorrência de periodicidade nos dados, indicando a presença de sazonalidade. Portanto, essa periodicidade pode ser incorporada ao método de Box e Jenkins por meio da utilização de um modelo *ARIMA Sazonal*, assim como é possível realizar diferenças sazonais na série, de modo a torná-la estacionária. Para Werner & Ribeiro (2003), trata-se do modelo SARIMA, o qual abrange os métodos que contemplam séries marcadas por

autocorrelação sazonal.

Nesse sentido, a notação do modelo ARIMA Sazonal é $SARIMA(p, d, q)(P, D, Q)_s$, em que (p, d, q) representam as ordens autorregressivas e de médias móveis com diferença (ARIMA), assim como (P, D, Q) representam suas ordens correspondentes do modelo sazonal em períodos determinados s . Dessa forma, o modelo SARIMA é dado pela equação (2.5):

$$(1 - \phi_p B^p)(1 - \Phi_p B^s)(1 - B^d)(1 - B^D)Z_t = (1 - \theta_q B^q)(1 - \Theta_Q B^s)\varepsilon_t, \quad (2.5)$$

em que o lado esquerdo da igualdade apresenta os componentes autorregressivos (p), seus componentes sazonais (P), as diferenças (d), assim como as diferenças sazonais (D). Assim como no lado direito da equação figuram as componentes de médias móveis (q), assim como suas equivalentes da parte sazonal (Q) em períodos determinados (s).

Ainda, (B) é o operador de defasagem, $\phi(B)$ representa os termos autoregressivos, $\theta(B)$ representa os termos de médias móveis e ε_t representa o ruído branco.

Contudo, é sabido que o método *SARIMA* também apresenta limitações, já que um único modelo SARIMA não teria a capacidade de capturar todos os padrões identificados na série. Por esse motivo, a prática recomenda a utilização do ciclo iterativo de Box & Jenkins (1976) para a identificação do modelo mais adequado.

2.1.6 Modelos estacionários e não estacionários

De acordo com Werner & Ribeiro (2003) a metodologia de Box & Jenkins (1976) ajuda a identificar diferentes modelos ARIMA que poderão ser: o modelo AR de ordem p , o modelo MA de ordem q , o modelo ARMA de ordem p e q , assim como o próprio modelo ARIMA, de ordem p, q com diferença d que, segundo os autores, figura como o principal caso dentro deste método. Nesse contexto metodológico, de acordo com Marchezan & Souza (2010), os modelos $AR(p)$, $MA(p)$ e $ARMA(p, q)$ apresentam comportamento estacionário.

Por outro lado, o modelo $ARIMA(p, d, q)$ é resultado da aplicação de diferenciação em séries com média inconstante, tornando-as por fim estacionárias. Desse modo, os modelos estacionários e os não estacionários distinguem-se pelo número de diferenças

necessárias para estacionarizar uma série, representado pela letra d . Adicionado a isso, temos os processos de memória longa, estacionários (estão fora do escopo deste estudo), os quais possuem uma função de correlação com um decaimento muito lento (hiperbólico) e cuja análise necessitará de uma diferença fracionária ($0 < d < 0,5$), fazendo-se uso do modelo ARFIMA(p, d, q).

Estes processos também são descritos de maneira adequada pelos chamados modelos autoregressivos integrados e de médias móveis de ordens p, d, q , ARIMA(p, d, q), que podem ser generalizados pela inclusão de um operador sazonal Morettin (2021), dando origem ao modelo SARIMA(p, d, q)(P, D, Q) s .

Sabe-se que para tornar uma série estacionária podemos tomar diferenças. Porém, antes, faz-se necessário, ou não, aplicar à série original alguma transformação e neste estudo destacamos uma transformação de uso geral, dada pela Equação (2.6):

$$Z_t^{(\lambda)} = \frac{Z_t^\lambda}{\lambda} \text{ se } \lambda \neq 0 \quad (2.6)$$

e pela equação (2.7).

$$Z_t^{(\lambda)} = \log Z_t \text{ se } \lambda = 0, \quad (2.7)$$

conhecidas como transformações de Box-Cox, proposta em Box & Cox (1964).

2.1.7 Critérios penalizadores para seleção de modelos

Os critérios penalizadores são utilizados dentro do ciclo iterativo de Box & Jenkins (1976), na etapa de identificação de modelos, a fim de encontrar **modelos parcimoniosos**, no sentido de que a ordem do modelo só aumente na medida do necessário, para que se obtenha um adequado ajuste aos dados. Então, para Sobral & Barreto (2011), ao ser constatado que a maioria dos dados reais não se adequam perfeitamente a nenhum modelo estatístico, uma das alternativas para a solução deste problema é o aumento de sua ordem. Para estes autores, tais problemas de imprecisão podem ser ocasionados por distúrbios nos dados (interferências) ou pela escolha de métodos limitados de modelagem.

Nesse contexto, Akaike (1973) propõe um método para viabilizar o aumento da ordem de um modelo a fim de conseguir uma melhor aderência. Para isto, sua metodologia atribui valores para determinado modelo segundo sua adequação aos dados e a ordem do

mesmo, de acordo com a Equação (2.8), a saber:

$$\text{AIC}(p, q) = \ln \hat{\sigma}_{p,q}^2 + \frac{2(p+q)}{n} \quad (2.8)$$

em que $\text{AIC}(p, q)$ é o *Akaike's Information Criterion*, $\ln \hat{\sigma}_{p,q}^2$ é o estimador de máxima verossimilhança do modelo $\text{ARMA}(p, q)$, sendo p a ordem do AR, q a ordem do modelo MA e n o número de observações.

Quando o tamanho da amostra é pequeno, há uma probabilidade substancial de que o AIC selecione modelos com muitos parâmetros, ou seja, que o AIC se ajuste demais. Para lidar com esse potencial sobreajuste, o AICc (AIC corrigido) foi desenvolvido no próprio paradigma do AIC. Portanto o AICc é o AIC com uma correção para pequenos tamanhos de amostra. A fórmula para o AICc depende do modelo estatístico. Supondo que o modelo seja univariado, linear em seus parâmetros e tenha resíduos normalmente distribuídos (condicional aos regressores), como é o caso deste trabalho acadêmico, a fórmula para o AICc é a seguinte, de acordo com a Equação (2.9), a saber:

$$\text{AICc} = \text{AIC} + \frac{2k^2 + 2k}{n - k - 1} \quad (2.9)$$

Em que n denota o tamanho da amostra e k denota o número de parâmetros. Assim, o AICc é essencialmente o AIC com um termo de penalização extra para o número de parâmetros. Observe que quando $n \rightarrow \infty$, o termo de penalidade extra converge para 0 e, portanto, o AICc converge para o AIC.

No entanto, mesmo com a larga aceitação do $\text{AIC}(p, q)$ como critério de escolha de modelos, sua utilização pode apresentar falhas em dadas circunstâncias. Por esse motivo, após a contribuição de Akaike (1973), outros métodos foram desenvolvidos a fim de eliminar estas falhas. Dentre eles, o método mais destacado foi o proposto por Schwarz (1978), o qual vem sendo utilizado até hoje em conjunto com o seu precursor e é expresso segundo a Equação (2.10) abaixo:

$$\text{BIC}(p, q) = \ln \hat{\sigma}_{p,q}^2 + (p+q) + \frac{\ln(n)}{n} \quad (2.10)$$

em que BIC é o *Bayesian Information Criterion*, $\hat{\sigma}_{p,q}^2$ é o estimador de máxima verossimilhança residual, p a ordem do modelo AR, q a ordem do modelo MA e n o número de observações.

2.2 Cartas de Controle CUSUM

2.2.1 Monitoramento estatístico da média

A carta de controle da soma cumulativa conhecida como *CUSUM* (abreviatura dos termos *cumulative sum* do inglês) incorpora diretamente toda a informação na sequência dos valores da amostra, plotando as somas acumuladas dos desvios dos valores da amostra de um valor-alvo. Por exemplo, suponha que as amostras de tamanho $n \geq 1$ sejam coletadas e que \bar{x}_j seja a média da j -ésima amostra. Então, se μ_0 é o alvo para a média do processo, a *CUSUM* é formada plotando-se a quantidade expressa através de (2.11),

$$C_i = \sum_{j=1}^i (\bar{x}_j - \mu_0) \quad (2.11)$$

versus a amostra i . C_i é chamada de soma cumulativa até e incluindo a i -ésima amostra.

Por combinarem informações de várias amostras, as Cartas de Controle da Soma Cumulativa são mais eficazes do que as Cartas de Controle de Shewhart para detectar pequenas mudanças no processo, Montgomery (2021).

Além disso, elas são particularmente eficazes com amostras de tamanho $n = 1$. Isso torna a *CUSUM* uma boa candidata para uso no MPAM e em seus diversos Órgãos de Atuação, onde subgrupos racionais são naturalmente de tamanho $n = 1$, se considerarmos as medições discretas *online* obtidas por meio do monitoramento estatístico mensal do quantitativo da produtividade e da resolutividade, denominados, aqui, indicadores de produtividade e de resolutividade.

Resolutividade esta, entendida como **fazer o que tem que ser feito, da melhor forma possível e no menor tempo** e indicada pelo tempo que leva para um Órgão Ministerial, ou o próprio MPAM, resolver extrajudicialmente ou levar à Justiça cada demanda de interesse da população no que diz respeito, constitucionalmente, à função essencial do Ministério Público Brasileiro, de defender a ordem jurídica, o regime democrático e os interesses sociais e individuais indisponíveis. Esclarece-se que Órgão Ministerial refere-se tanto a uma Procuradoria de Justiça ou uma Promotoria de Justiça, quanto a um(a) Procurador(a) de Justiça ou Promotor(a) de Justiça e também Órgãos da Administração Superior do MPAM.

As cartas de controle *CUSUM* foram propostas primeiramente por Page (1954). Se-

gundo Montgomery (2021), é possível usar essa metodologia em variáveis com distribuição normal, Poisson e binomial, para monitorar a média do processo, como também é possível construir a *CUSUM* para monitorar a variabilidade do processo. Neste trabalho tratamos, tão somente, do monitoramento estatístico da média dos indicadores de produtividade das Promotorias de Justiça da Capital. Observe que, se o processo permanecer sob controle no valor alvo μ_0 , a soma cumulativa é um passeio aleatório com média zero.

No entanto, se a média mudar para cima até algum valor, digamos, $\mu_1 > \mu_0$, então uma tendência para cima ou positiva se desenvolverá na soma cumulativa C_i . Por outro lado, se a média muda para baixo para algum $\mu_1 < \mu_0$, então uma tendência para baixo ou negativa em C_i se desenvolverá.

Portanto, se uma tendência significativa se desenvolver nos pontos plotados, tanto para cima quanto para baixo, devemos considerar isto como evidência de que a média do indicador mudou, e uma busca por alguma causa atribuível deve ser realizada, Montgomery (2021).

Se, por exemplo, tomarmos $\bar{x}_i = x_i$ (no caso o tamanho da amostra é $n = 1$) e fizermos o valor-alvo $\mu_0 = 10$, a *CUSUM* se torna

$$\begin{aligned} C_i &= \sum_{j=1}^i (x_j - 10) \\ &= (x_i - 10) + \sum_{j=1}^{i-1} (x_j - 10) \\ &= (x_i - 10) + C_{i-1} \end{aligned}$$

O valor inicial para a *CUSUM*, C_0 , é tomado como zero. Naturalmente, nestas condições, a *CUSUM* ainda não é uma Carta de Controle porque faltam limites de controle estatísticos.

Há duas maneiras de representar a *CUSUM*, quais sejam, a Tabular (ou algorítmica) e a forma máscara V da *CUSUM*. Neste estudo faz-se uso da *CUSUM* Tabular, visto que é plenamente verificável que esta é a melhor representação da *CUSUM*.

2.2.2 Carta de Controle CUSUM Tabular

Seja x_i a i -ésima observação do processo. Quando o processo está sob controle, x_i tem uma distribuição normal com média μ_0 e desvio padrão σ . Assumimos que ou σ é conhecido ou que uma estimativa confiável está disponível, Montgomery (2021).

Consideramos μ_0 como um valor-alvo para a performance do indicador monitorado estatisticamente. Portanto, se o processo flutua ou sai de seu valor-alvo, a *CUSUM* dará o sinal, o qual indica a presença de uma causa atribuível que deve ser identificada da mesma forma que no caso da Carta de Controle de Shewhart. No contexto da Gestão da Produtividade e da Resolutividade do MPAM, a ideia é **não trazer** o resultado de performance de volta ao valor-alvo, mas sim, convergir para um *benchmarking* no caso da Produtividade e para um Nível de Resolutividade Aceitável, no caso da Resolutividade.

Num ambiente de gestão bem sucedida, produtividade e resolutividade se complementam, no entanto pode ocorrer significativa produtividade sem resolutividade; já as organizações resolutivas, são naturalmente produtivas.

A *CUSUM* Tabular funciona acumulando desvios de μ_0 que estão acima do alvo com uma estatística C^+ e acumulando desvios de μ_0 que estão abaixo do alvo com outra estatística C^- . As estatísticas C^+ e C^- são chamadas de *CUSUM* Unilateral Superior e *CUSUM* Unilateral Inferior, respectivamente. Elas são calculadas das formas expressas nas equações (2.12) e (2.13), respectivamente

$$C_i^+ = \max [0, x_i - (\mu_0 + K) + C_i^+], \quad (2.12)$$

e

$$C_i^- = \max [0, (\mu_0 + K) - x_i + C_i^-], \quad (2.13)$$

onde os valores iniciais são $C_i^+ = C_i^- = 0$.

Nas Eqs. (2.12) e (2.13), K é geralmente chamado de **valor de referência** (ou **valor de tolerância** ou **valor de folga**), e é sempre escolhido a meio caminho entre o valor-alvo μ_0 e o valor da média fora de controle μ_1 que estamos interessados em detectar rapidamente. *Vide* Montgomery (2021). Assim, se a mudança for expressa em unidades de desvio padrão como $\mu_1 = \mu_0 + \delta\sigma$ (ou $\delta = \frac{|\mu_1 - \mu_0|}{\sigma}$), então K é a metade da magnitude da mudança ou equação (2.14).

$$K = \frac{\delta}{2}\sigma = \frac{|\mu_1 - \mu_0|}{2}. \quad (2.14)$$

Note que C_i^+ e C_i^- acumulam desvios a partir do valor-alvo μ_0 que são maiores do que K , com ambas as quantidades zeradas ao se tornarem negativas. Se tanto C_i^+ quanto C_i^- excederem o intervalo de decisão H , o processo é considerado fora de controle, como dito em Montgomery (2021).

A seleção adequada de K e H é muito importante, uma vez que têm impacto substancial no desempenho da *CUSUM*. Já mencionamos como escolher K e quanto ao H um valor razoável, como veremos adiante, é cinco vezes o desvio padrão do processo σ .

A Tabela 2.1 extraída de Montgomery (2021) e construída com base em uma sequência de 30 medições usadas como dados de controle da viscosidade de um processo químico industrial, apresenta o esquema da *CUSUM* Tabular, com o valor-alvo $\mu_0 = 10$, o tamanho do subgrupo sendo $n = 1$, com o desvio padrão do processo, $\sigma = 1$, e supondo que a magnitude da mudança que se esteja interessado em detectar seja $1,0\sigma = 1,0(1,0) = 1,0$. Portanto, o valor fora de controle da média do processo é $\mu_1 = 10 + 1 = 11$. Usa-se, neste caso, a *CUSUM* Tabular com $K = \frac{1}{2}$ (porque o tamanho da mudança é $1,0\sigma$ e $\sigma = 1$) e $H = 5$ (porque o valor recomendado do intervalo de decisão é $H = 5\sigma = 5(1) = 5$).

| Period i | x_i | (a) | | | (b) | | |
|------------|-------|--------------|---------|-------|-------------|---------|-------|
| | | $x_i - 10.5$ | C_i^+ | N^+ | $9.5 - x_i$ | C_i^- | N^- |
| 1 | 9.45 | -1.05 | 0 | 0 | 0.05 | 0.05 | 1 |
| 2 | 7.99 | -2.51 | 0 | 0 | 1.51 | 1.56 | 2 |
| 3 | 9.29 | -1.21 | 0 | 0 | 0.21 | 1.77 | 3 |
| 4 | 11.66 | 1.16 | 1.16 | 1 | -2.16 | 0 | 0 |
| 5 | 12.16 | 1.66 | 2.82 | 2 | -2.66 | 0 | 0 |
| 6 | 10.18 | -0.32 | 2.50 | 3 | -0.68 | 0 | 0 |
| 7 | 8.04 | -2.46 | 0.04 | 4 | 1.46 | 1.46 | 1 |
| 8 | 11.46 | 0.96 | 1.00 | 5 | -1.96 | 0 | 0 |
| 9 | 9.20 | -1.3 | 0 | 0 | 0.30 | 0.30 | 1 |
| 10 | 10.34 | -0.16 | 0 | 0 | -0.84 | 0 | 0 |
| 11 | 9.03 | -1.47 | 0 | 0 | 0.47 | 0.47 | 1 |
| 12 | 11.47 | 0.97 | 0.97 | 1 | -1.97 | 0 | 0 |
| 13 | 10.51 | 0.01 | 0.98 | 2 | -1.01 | 0 | 0 |
| 14 | 9.40 | -1.10 | 0 | 0 | 0.10 | 0.10 | 1 |
| 15 | 10.08 | -0.42 | 0 | 0 | -0.58 | 0 | 0 |
| 16 | 9.37 | -1.13 | 0 | 0 | 0.13 | 0.13 | 1 |
| 17 | 10.62 | 0.12 | 0.12 | 1 | -1.12 | 0 | 0 |
| 18 | 10.31 | -0.19 | 0 | 0 | -0.81 | 0 | 0 |
| 19 | 8.52 | -1.98 | 0 | 0 | 0.98 | 0.98 | 1 |
| 20 | 10.84 | 0.34 | 0.34 | 1 | -1.34 | 0 | 0 |
| 21 | 10.90 | 0.40 | 0.74 | 2 | -1.40 | 0 | 0 |
| 22 | 9.33 | -1.17 | 0 | 0 | 0.17 | 0.17 | 1 |
| 23 | 12.29 | 1.79 | 1.79 | 1 | -2.79 | 0 | 0 |
| 24 | 11.50 | 1.00 | 2.79 | 2 | -2.00 | 0 | 0 |
| 25 | 10.60 | 0.10 | 2.89 | 3 | -1.10 | 0 | 0 |
| 26 | 11.08 | 0.58 | 3.47 | 4 | -1.58 | 0 | 0 |
| 27 | 10.38 | -0.12 | 3.35 | 5 | -0.88 | 0 | 0 |
| 28 | 11.62 | 1.12 | 4.47 | 6 | -2.12 | 0 | 0 |
| 29 | 11.31 | 0.81 | 5.28 | 7 | -1.81 | 0 | 0 |
| 30 | 10.52 | 0.02 | 5.30 | 8 | -1.02 | 0 | 0 |

Tabela 2.1: Exemplo de uma tabela CUSUM tabular (Table 9.2 - Extraída de Montgomery D. C. Introduction to Statistical Quality Control - 6th Edition)

Os painéis (a) e (b) da Tabela 2.1 resumem os cálculos de C_i^+ e C_i^- , Equação (2.12) e Equação (2.13), respectivamente. Por sua vez, as quantidades N^+ e N^- indicam o número de períodos consecutivos em que as *CUSUMs* C_i^+ e C_i^- foram não-nulas.

A *CUSUM* é particularmente útil para determinar quando a causa atribuível ocorreu; apenas conte para trás a partir do sinal fora de controle até o período em que a *CUSUM* se tornou positiva para encontrar o primeiro período após a mudança do processo. Os contadores N^+ e N^- assumem o resultado da contagem em qualquer período considerado. A nova média da performance do indicador, quando alcançada, pode ser calculada a partir das equações (2.15) e (2.16).

$$\hat{\mu} = \mu_0 + K + \frac{C_i^+}{N^+} \text{ se } C_i^+ > H, \quad (2.15)$$

ou

$$\hat{\mu} = \mu_0 - K - \frac{C_i^-}{N^-} \text{ se } C_i^- > H. \quad (2.16)$$

Os valores sucessivos de C_i^+ e C_i^- não são independentes. Na verdade, a *CUSUM* pode ser encarada como uma média ponderada, onde os pesos são estocásticos ou aleatórios.

2.2.3 Planejamento da Carta de Controle CUSUM

A *CUSUM* Tabular é planejada escolhendo valores para o valor de referência K e o intervalo de decisão H . Normalmente, é recomendado que esses parâmetros sejam selecionados para fornecer um bom desempenho do **comprimento médio da sequência** ou *ARL* (sigla do inglês *average run lengths*). Existem muitos estudos analíticos sobre o desempenho do *ARL* da *CUSUM*. Com base nesses estudos, dispomos de algumas recomendações gerais para a seleção de H e K , Montgomery (2021).

Define-se $H = h\sigma$ e $K = k\sigma$, em que σ é o desvio padrão da variável amostral presente na formação da *CUSUM*. Usando $h = 4$ ou $h = 5$ e $k = 1/2$, geralmente resultará numa *CUSUM* que tem boas propriedades de *ARL* contra uma mudança de cerca de 1σ na média do processo.

| Mudança na média (múltiplo de σ) | $h = 4$ | $h = 5$ |
|--|---------|---------|
| 0 | 168 | 465 |
| 0.25 | 74.2 | 139 |
| 0.50 | 26.6 | 38.0 |
| 0.75 | 13.3 | 17.0 |
| 1.00 | 8.38 | 10.4 |
| 1.50 | 4.75 | 5.75 |
| 2.00 | 3.34 | 4.01 |
| 2.50 | 2.62 | 3.11 |
| 3.00 | 2.19 | 2.57 |
| 4.00 | 1.71 | 2.01 |

Tabela 2.2: Desempenho da Carta de Controle CUSUM Tabular (Table 9.3 - Extraída e adaptada de Montgomery D. C. Introduction to Statistical Quality Control - 6th Edition)

Por exemplo, para as recomendações de $h = 4$ ou $h = 5$ com $k = 1/2$, considere os **comprimentos médios das sequências** bilaterais mostrados na Tabela 2.2 extraída de

Montgomery (2021) e construída com base em uma sequência de 30 medições usadas como dados de controle da viscosidade de um processo químico industrial. Note que a mudança de 1σ seria detectada ou em 8,38 amostras (com $k = 1/2$ e $h = 4$) ou 10,4 amostras (com $k = 1/2$ e $h = 5$). Por comparação, uma Carta de Controle Shewhart para medições individuais exigiria 43,96 amostras, em média, para detectar essa mudança, ver Montgomery (2021).

Note, também, pela Tabela 2.2, que $h = 4$ resulta em um $ARL_0 = 168$ amostras sob controle, enquanto $h = 5$ resulta em um $ARL_0 = 465$ amostras sob controle. Se escolhermos $h = 4,77$, isso fornecerá uma *CUSUM* com $ARL_0 = 370$ amostras sob controle, que corresponde ao valor do ARL_0 para uma Carta de Controle Shewhart com os limites 3σ usuais.

Geralmente, nosso interesse é escolher k relativo ao tamanho do deslocamento que queremos detectar, isto é, $k = \frac{1}{2}\delta$, onde δ é o tamanho da mudança em unidades de desvio padrão. Tende-se assim a minimizar o valor do ARL_1 para detectar uma mudança de tamanho δ para um ARL_0 fixo. Como mencionado anteriormente, um valor amplamente utilizado na prática é $k = 1/2$. Então, uma vez que k é selecionado, deve-se escolher h para obter o desejado desempenho do ARL_0 sob controle. Hawkins (1993) fornece uma tabela de valores k e os valores h correspondentes que atingirão $ARL_0 = 370$. Esses valores estão reproduzidos na Tabela 2.3.

| | | | | | | |
|-----|------|------|------|------|------|------|
| k | 0.25 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 |
| h | 8.01 | 4.77 | 3.34 | 2.52 | 1.99 | 1.61 |

Tabela 2.3: Valores dos Parâmetros da *CUSUM* Bilateral (Table 9.4 - Extraída de Montgomery D. C. Introduction to Statistical Quality Control - 6th Edition)

Várias técnicas podem ser usadas para calcular o ARL_0 de uma *CUSUM*. Para uma *CUSUM* unilateral (isto é, C_i^+ ou C_i^-) com parâmetros h e k , usa-se a aproximação de Siegmund (2010), devido à sua simplicidade, a qual é expressa pela equação abaixo

$$ARL = \frac{\exp(-2\Delta b) + 2\Delta b - 1}{2\Delta^2} \quad (2.17)$$

para $\Delta \neq 0$ em que $\Delta = \delta^* - k$ para a *CUSUM* Unilateral Superior C_i^+ , $\Delta = -\delta^* - k$ para a *CUSUM* Unilateral Inferior C_i^- , $b = h + 1,66$ e $\delta^* = \frac{|\mu_1 - \mu_0|}{\sigma}$. Se $\Delta = 0$, pode-se usar $ARL = b^2$.

A quantidade δ^* representa o deslocamento da média, em unidades de σ , para o qual o ARL deve ser calculado. Portanto, se $\delta^* = 0$, calcularíamos ARL_0 da Equação (2.17), enquanto que, se $\delta^* \neq 0$ calcularíamos o valor de ARL_1 correspondente a uma mudança de tamanho δ^* . Para obter o ARL da $CUSUM$ bilateral a partir dos ARL 's das estatísticas unilaterais - digamos, ARL^+ e ARL^- usamos a equação abaixo

$$\frac{1}{ARL} = \frac{1}{ARL^+} + \frac{1}{ARL^-} \quad (2.18)$$

Para ilustrar, considere a $CUSUM$ Bilateral com $k = 1/2$ e $h = 5$. Para encontrar o ARL_0 calculamos primeiro os valores do ARL_0 para as duas bandas da $CUSUM$ - digamos, ARL_0^+ e ARL_0^- . Definir $\delta^* = 0$, então $\Delta = \delta^* - k = 0 - 1/2 = -1/2$, $b = h + 1,166 = 5 + 1,166 = 6,166$, e da equação (2.17)

$$ARL_0^+ = \frac{\exp\left[-2\left(-\frac{1}{2}\right)(6,166)\right] + 2\left(-\frac{1}{2}\right)(6,166) - 1}{2\left(-\frac{1}{2}\right)^2} = 938,2$$

Por simetria, $ARL_0^- = ARL_0^+$ e assim da Equação (2.18), o ARL sob controle para a $CUSUM$ bilateral é

$$\frac{1}{ARL_0} = \frac{1}{938,2} + \frac{1}{938,2}$$

ou

$$ARL_0 = 469,1$$

Isso está muito próximo do verdadeiro valor do ARL_0 de 465 mostrado na Tabela 3-2, de acordo com Montgomery (2021). Se a média muda em 2σ , então $\delta^* = 2$, $\Delta = 1,5$ para a $CUSUM$ Unilateral Superior, $\Delta = -2,5$ para a $CUSUM$ Unilateral Inferior e, a partir das Equações (2.17) e (2.18), podemos calcular o ARL_1 aproximado da $CUSUM$ bilateral como $ARL_1 \approx 3,89$. O valor exato mostrado na Tabela 3-2, extraída de Montgomery (2021) é 4,01.

Para atender os propósitos deste trabalho, com relação à solução da Carta de Controle $CUSUM$, foram usados os seguintes pacotes disponíveis no R: *cusum* v. 0.4.1 e *qcc* v. 2.7.

2.2.4 Séries Temporais e Monitoramento Estatístico

Uma abordagem utilizada para Monitoramento Estatístico, quando lidamos com dados autocorrelacionados, consiste em modelar diretamente a estrutura correlacional com um modelo apropriado de Séries Temporais, modelo este capaz de remover a autocorrelação dos dados e em seguida aplicar cartas de controle aos **resíduos**, ver Montgomery (2021). Confrontamos então, este entendimento, com o fato de os dados do MPAM usados neste trabalho descreverem um comportamento não-estacionário, ou seja, os resultados numéricos do indicador não comportam a ideia de haver, *a priori*, um valor fixo para a média, como é o caso das **especificações** nos processos de fabricação, compostas de um **valor nominal** e uma **tolerância**.

Então, para fazer análise de performance e gerar previsões acerca de indicadores de produtividade e resolutividade do MPAM, usamos o modelo $SARIMA(p, d, q)(P, D, Q)_s$, tanto para **estacionarizar** como para **dessazonalizar** a série do indicador considerado. Em seguida, usamos a **Série de Resíduos** para implementar o monitoramento da Carta *CUSUM*.

A ideia que norteia este trabalho é monitorar estatisticamente, simultânea e sucessivamente, a Série Temporal do Indicador, a Previsão para períodos de tempo predefinidos e a Carta *CUSUM* dos Resíduos do Indicador. Sempre ensejando o atingimento das metas estabelecidas no âmbito do Planejamento Estratégico do MPAM. Já que é imprescindível usarmos o melhor modelo que gere valores residuais mais representativos dos dados originais e também seja capaz de gerar novas previsões a cada atualização da própria Série Temporal, dando condições de antever, com bastante antecedência, se a meta traçada pelo Planejamento Estratégico vai ser alcançada, ou não, se nada for feito, quais as causas atribuíveis, ajustes e mudanças necessárias.

Embora sabendo que são importantes todas as metodologias usadas neste trabalho, o que buscamos é combinar as Cartas de Controle *CUSUM* com as Previsões Estatísticas, dando condições aos gestores de "**adentrarem**" ao futuro e tomarem decisões mais assertivas, mesmo não conhecendo métodos estatísticos. Tudo publicado nos visuais de um aplicativo de *Business Inteligency - BI*, com frequência mensal ou sob demandas, com destaque para o *Power BI* da *Microsoft*, empresa fornecedora de soluções computacionais para o MPAM na atualidade.

Capítulo 3

Aplicação das Metodologias e Análises dos Resultados

A realização deste estudo pode ser caracterizada como produção acadêmica de cunho quantitativo. Ao mesmo tempo, para sua constituição, uma ampla revisão bibliográfica foi realizada focando nos temas relacionados às metodologias Box-Jenkins de análise e previsão de Séries Temporais e monitoramento estatístico para pequenas variações.

3.1 Objeto de estudo

O arquivo de dados usado neste estudo acadêmico refere-se à Série Histórica da Produtividade Total das Promotorias de Justiça Especializadas do Ministério Público do Estado do Amazonas, sediadas na capital Manaus, no período de 1º de janeiro de 2012 a 31 de dezembro de 2019.

3.2 Procedimento de coleta de dados

Foi feita uma solicitação à Direção do MPAM para que institucionalmente autorizasse o uso de dados constantes da sua base de dados referentes ao Relatório de Atividades Funcionais da Atividade-Fim, restringindo-se tão somente aos dados das Promotorias de Justiça Especializadas da Capital, solicitação esta que foi atendida pela Direção da Instituição, para uso com os fins acadêmicos a que se pretendia.

3.3 Resultados obtidos

Apresentaremos a seguir as soluções produzidas no R com análises pertinentes e as devidas explicações, com base na Série Histórica Mensal do Indicador da Produtividade conjunta das Promotorias de Justiça Especializadas com atuação em Manaus, Capital do Estado do Amazonas, no período de janeiro de 2012 a dezembro de 2019. Na Tabela 3.1 estão os registros dos valores mensais da série. Foram carregados os seguintes pacotes: base v. 4.2.0, cusum v. 0.4.1, datasets v. 4.2.0, forecast v. 8.16, fUniRoots v. 4021.80, ggfortify v. 0.4.14, ggplot2 v. 3.3.6, gplots v. 3.1.3, graphics v. 4.2.0, grDevices v. 4.2.0, Kendall v. 2.2.1, lmtest v. 0.9-40, methods v. 4.2.0, qcc v. 2.7, randtests v. 1.0.1, stats v. 4.2.0, stats4 v. 4.2.0, tseries v. 0.10-51, urca v. 1.3-0, utils v. 4.2.0, zoo v. 1.8-10.

| PRODUTIVIDADE DAS PROMOTORIAS DE JUSTIÇA DA CAPITAL - Jan/2012 a Dez/2019 | | | | | | | |
|---|-------|--------|-------|--------|-------|--------|-------|
| data | prod | data | prod | data | prod | data | prod |
| jan/12 | 21351 | jan/14 | 43225 | jan/16 | 45356 | jan/18 | 52737 |
| fev/12 | 22670 | fev/14 | 31962 | fev/16 | 53717 | fev/18 | 58556 |
| mar/12 | 28566 | mar/14 | 51537 | mar/16 | 61805 | mar/18 | 78889 |
| abr/12 | 24757 | abr/14 | 52608 | abr/16 | 51368 | abr/18 | 71521 |
| mai/12 | 30412 | mai/14 | 52503 | mai/16 | 62022 | mai/18 | 82156 |
| jun/12 | 32016 | jun/14 | 37788 | jun/16 | 66371 | jun/18 | 77142 |
| jul/12 | 43244 | jul/14 | 60331 | jul/16 | 64219 | jul/18 | 67338 |
| ago/12 | 47410 | ago/14 | 65387 | ago/16 | 65236 | ago/18 | 99720 |
| set/12 | 34801 | set/14 | 59884 | set/16 | 54573 | set/18 | 71264 |
| out/12 | 45512 | out/14 | 57813 | out/16 | 59407 | out/18 | 84001 |
| nov/12 | 44511 | nov/14 | 49541 | nov/16 | 62698 | nov/18 | 66706 |
| dez/12 | 31387 | dez/14 | 43273 | dez/16 | 43128 | dez/18 | 51827 |
| jan/13 | 36788 | jan/15 | 41245 | jan/17 | 45860 | jan/19 | 69825 |
| fev/13 | 39875 | fev/15 | 42960 | fev/17 | 58465 | fev/19 | 88210 |
| mar/13 | 46189 | mar/15 | 60588 | mar/17 | 82829 | mar/19 | 86723 |
| abr/13 | 54897 | abr/15 | 53069 | abr/17 | 54964 | abr/19 | 88033 |
| mai/13 | 52784 | mai/15 | 55493 | mai/17 | 85528 | mai/19 | 76075 |
| jun/13 | 51568 | jun/15 | 55473 | jun/17 | 75982 | jun/19 | 67774 |
| jul/13 | 56567 | jul/15 | 60932 | jul/17 | 76648 | jul/19 | 88023 |
| ago/13 | 63469 | ago/15 | 61284 | ago/17 | 76396 | ago/19 | 91137 |
| set/13 | 43763 | set/15 | 64321 | set/17 | 62990 | set/19 | 76161 |
| out/13 | 54760 | out/15 | 62836 | out/17 | 76332 | out/19 | 72530 |
| nov/13 | 52715 | nov/15 | 64894 | nov/17 | 71231 | nov/19 | 74417 |
| dez/13 | 47395 | dez/15 | 36840 | dez/17 | 50620 | dez/19 | 48661 |

Tabela 3.1: Produtividade das Promotorias de Justiça da Capital (Fonte: RAF-MPAM - Janeiro de 2012 a Dezembro de 2019)

3.3.1 Metodologia de Box e Jenkins

A Figura 3.1 apresenta o gráfico da Série Histórica Mensal do Indicador da Produtividade conjunta das Promotorias de Justiça com atuação em Manaus, Capital do Estado do Amazonas. Em outras palavras, o Indicador apresenta a soma das produtividades mensais das referidas Promotorias de Justiça.

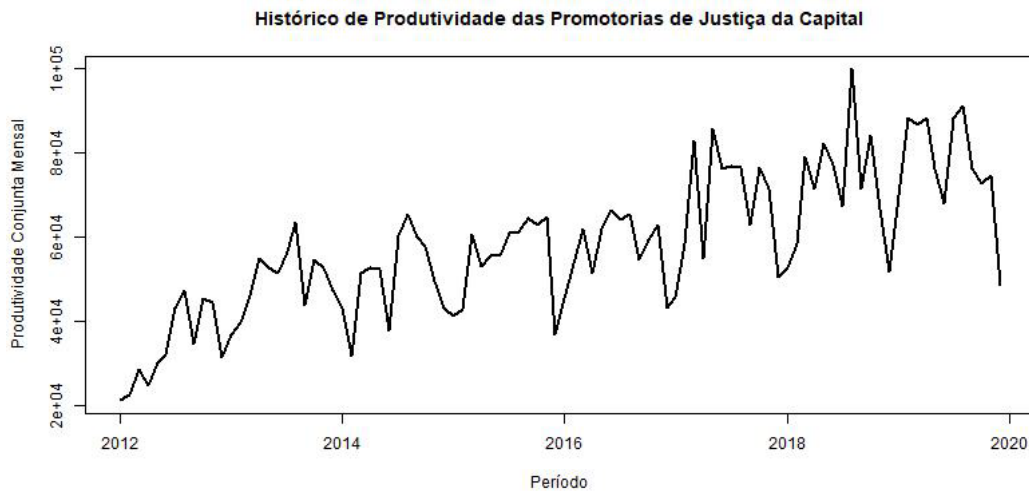


Figura 3.1: Produtividade das Promotorias de Justiça da Capital

Então, temos no eixo x o período e no eixo y a produtividade conjunta das Promotorias de Justiça de Manaus. Observamos que existe sazonalidade e a série não apresenta comportamento estacionário, ou seja, a média não é constante durante todo o período.

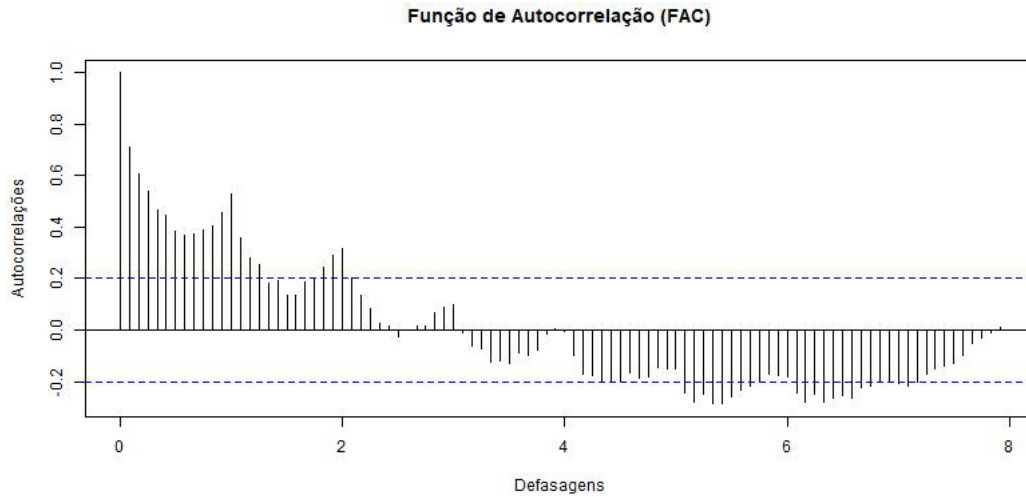


Figura 3.2: Análise da função de autocorrelação (FAC)

Na Figura 3.2 a função de autocorrelação apresenta decaimento exponencial até a segunda defasagem sazonal evidenciando a presença da componente MA(2) sazonal do modelo a ser especificado. Temos ainda um truncamento na primeira defasagem sazonal indicando a componente AR(1) sazonal.

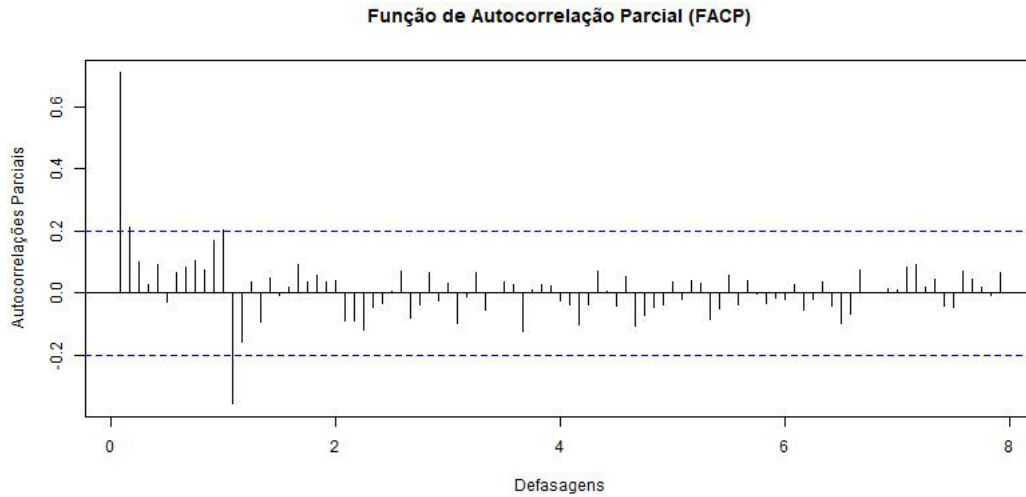


Figura 3.3: Análise da função de autocorrelação parcial (FACP)

Na Figura 3.3 a função de autocorrelação parcial apresenta decaimento exponencial na parte não sazonal evidenciando um MA(2) do modelo identificado. Completa-se assim um modelo SARIMA(0, 0, 2)(1, 1, 2)[12] que conduz à realização de uma diferenciação sazonal.

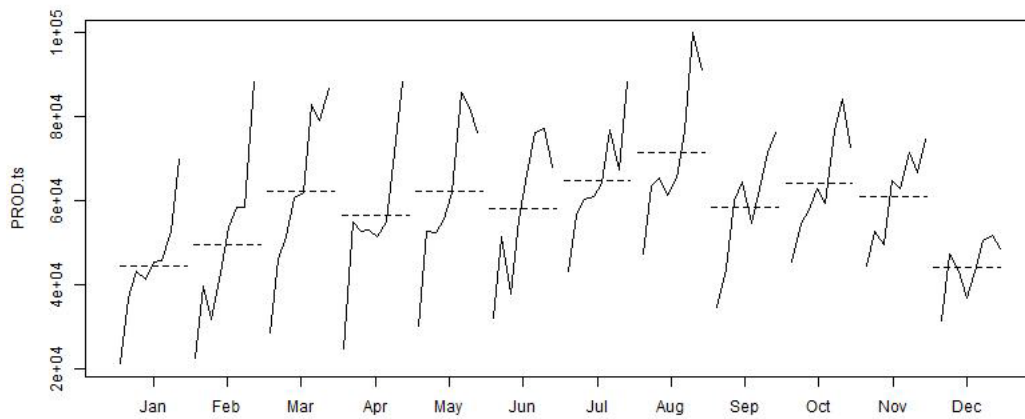


Figura 3.4: Componentes sazonais

Para a análise gráfica da Figura 3.4 o teste Qui-Quadrado de Kruskal-Wallis (pacote stats) aplicado à Série original e apresentado na Tabela 3.2 evidencia, a um nível de significância de 0,05, que nem todas as medianas do grupo são iguais, sugerindo sazonalidade determinística.

| Kruskal-Wallis rank sum test |
|--|
| data: PROD.ts and g |
| Kruskal-Wallis chi-squared = 24.321, df = 11, p-value = 0.01145 |

Tabela 3.2: Evidência em favor de Sazonalidade Determinística

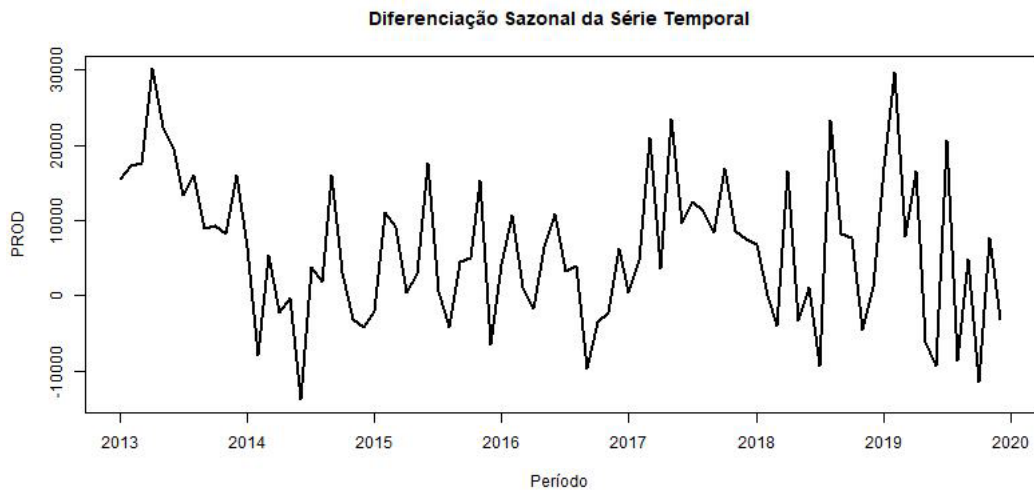


Figura 3.5: Resultado da diferenciação sazonal da série

Na Figura 3.5 verifica-se o efeito da diferenciação sazonal aplicada à Série original, com o intuito de ajustar um modelo arima adequado para previsão de resultados futuros.

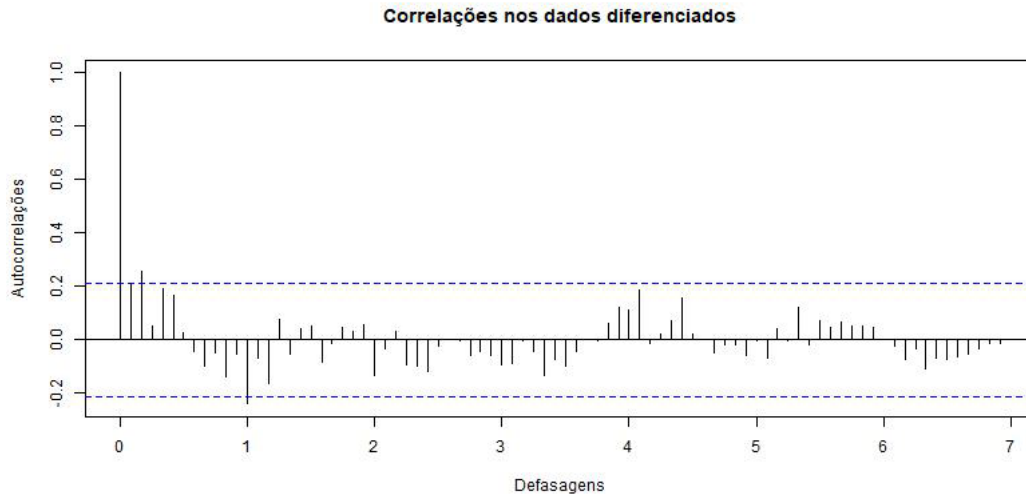


Figura 3.6: Função de autocorrelação aplicada aos dados diferenciados

A função de autocorrelação da Figura 3.6 mostra um decaimento não sazonal na segunda defasagem, indicando a componente MA(2) não sazonal. Também verifica-se uma truncagem na segunda defasagem sazonal o que representa uma componente MA(2) sazonal.

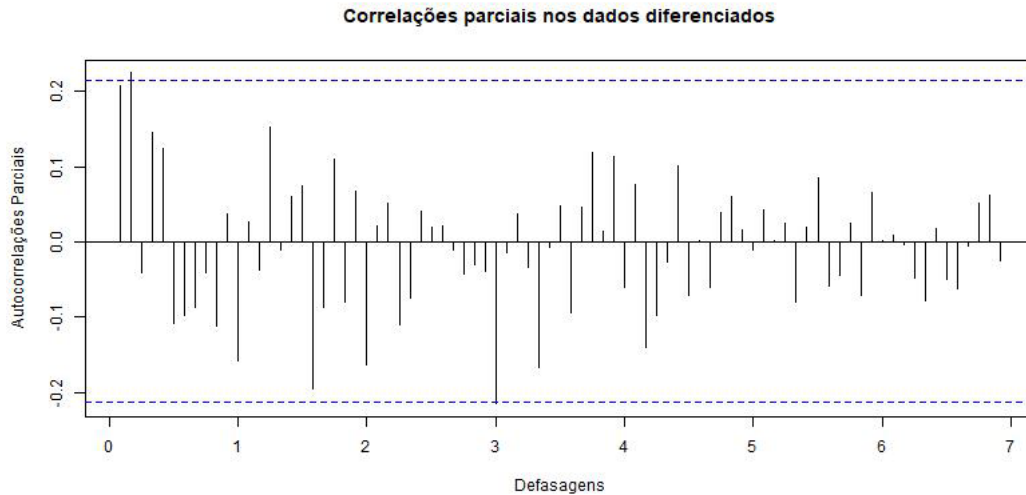


Figura 3.7: Função de autocorrelação parcial aplicada aos dados diferenciados

A função de autocorrelação parcial da Figura 3.7 está truncada na primeira defasagem sazonal, configurando a componente AR(1) sazonal. Assim continua valendo o modelo de Box-Jenkins SARIMA(0, 0, 2)(1, 1, 2)[12], para a realização de previsões e outras aplicações necessárias neste estudo acadêmico.

Na sequência, ao aplicar o Teste de Tendência de Cox-Stuart (pacote randtests) ao nível de significância de 0,05, Tabela 3.3, com implementação no R, não constatou-se evidência estatística de que a Série diferenciada apresente tendência crescente.

| Cox Stuart test |
|---|
| data: PROD.D1 |
| statistic = 20, n = 42, p-value = 0.678 |
| alternative hypothesis: increasing trend |

Tabela 3.3: Evidência em favor de Aleatoriedade da Série diferenciada

Ainda, os testes de raiz unitária (pacote tseries) Dickey-Fuller aumentado (ADF), Tabela 3.4, Phillips-Perron (PP), Tabela 3.5 e Kwiatkowski-Phillips-Schmidt-Shin(KPSS), Tabela 3.6, aplicados à Série diferenciada, não acusaram a presença de raiz unitária, o que depôs a favor da adequada identificação inicial do modelo.

| Augmented Dickey-Fuller Test |
|---|
| data: PROD.D1 |
| Dickey-Fuller = -3.1429, Lag order = 4, p-value = 0.1071 |
| alternative hypothesis: stationary |

Tabela 3.4: Teste ADF: Evidência em favor de Estacionariedade

| Phillips-Perron Unit Root Test |
|---|
| data: PROD.D1 |
| Dickey-Fuller Z(alpha) = -71.351, Truncation lag parameter = 3, p-value = 0.01 |
| alternative hypothesis: stationary |

Tabela 3.5: Teste PP: Evidência em favor de Estacionariedade

| KPSS Test for Level Stationarity |
|--|
| data: PROD.D1 |
| KPSS Level = 0.21607, Truncation lag parameter = 3, p-value = 0.1 |

Tabela 3.6: Teste KPSS: Evidência em favor de Estacionariedade

Então o ajuste arima foi usado na Série original, o qual retornou, entre outros resultados, Tabela 3.7 e Tabela 3.8, o critério de informação $AIC = 1795.61$.

| arima(x = PROD.ts, order = c(0, 0, 2), seasonal = list(order = c(1, 1, 2)), include.mean = FALSE) | | | | | | | |
|---|----------|----------|----------|----------|---------|-----------|------------|
| Coefficients: | | | | | | | |
| | ma1 | ma2 | sar1 | sma1 | sma2 | | |
| | 0,3933 | 0,3953 | -0,1208 | 0,0210 | -0,1300 | | |
| s.e. | 0,1105 | 0,1031 | 0,7062 | 0,6913 | 0,1529 | | |
| sigma^2 estimated as 96473003: log likelihood = -891,81, aic = 1795,61 | | | | | | | |
| Training set error measures: | | | | | | | |
| | ME | RMSE | MAE | MPE | MAPE | MASE | ACF1 |
| Training set | 3622,711 | 9187,711 | 6756,946 | 5,468992 | 11,233 | 0,7385082 | -0,1564186 |

Tabela 3.7: Estimação de Parâmetros

| z test of coefficients: | | | | | | | | | | | |
|-------------------------|-----------|------------|---------|-----------|----------|----|------|---|-----|---|---|
| | Estimate | Std. Error | z | value | Pr(> z) | | | | | | |
| ma1 | 0,393334 | 0,110506 | 35,594 | 0,0003717 | *** | | | | | | |
| ma2 | 0,395255 | 0,103074 | 38,347 | 0,0001257 | *** | | | | | | |
| sar1 | -0,12076 | 0,706244 | -0,1710 | 0,8642327 | | | | | | | |
| sma1 | 0,020962 | 0,691309 | 0,0303 | 0,9758101 | | | | | | | |
| sma2 | -0,130044 | 0,152924 | -0,8504 | 0,3951125 | | | | | | | |
| --- | | | | | | | | | | | |
| Signif. Codes: | 0 | **** | 0.001 | *** | 0.01 | ** | 0.05 | ' | 0.1 | " | 1 |

Tabela 3.8: Teste z para os coeficientes

Ainda podemos analisar graficamente a questão da Raiz Unitária, Figura 3.8 e a adequação do ajuste aos dados, Figura 3.9. No caso da Raiz Unitária, como os pontos encontram-se no interior dos círculos unitários, concluímos que estamos lidando com uma Série Estacionária.

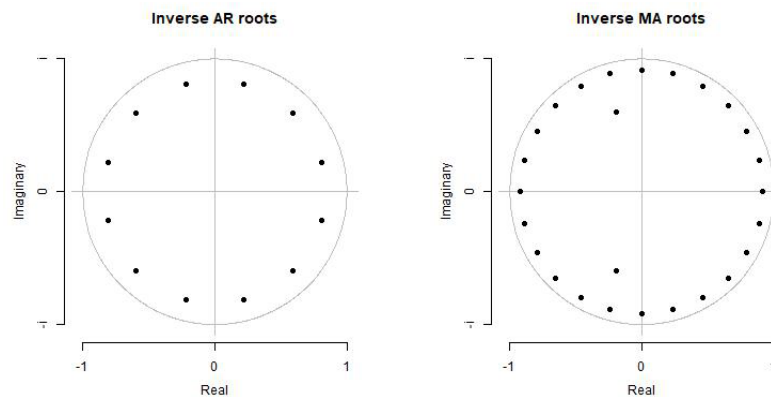


Figura 3.8: Indicativo de Estacionariedade

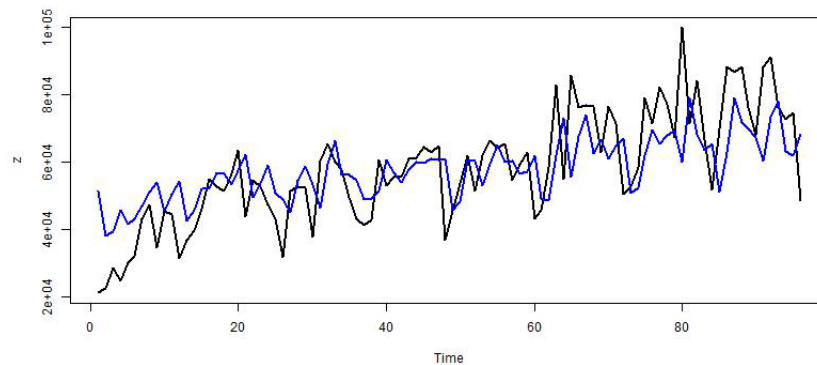


Figura 3.9: Resultado do Ajuste aos Dados

Em seguida foi feito o diagnóstico dos resíduos (pacotes stats e graphics), apresentado na Figura 3.10.

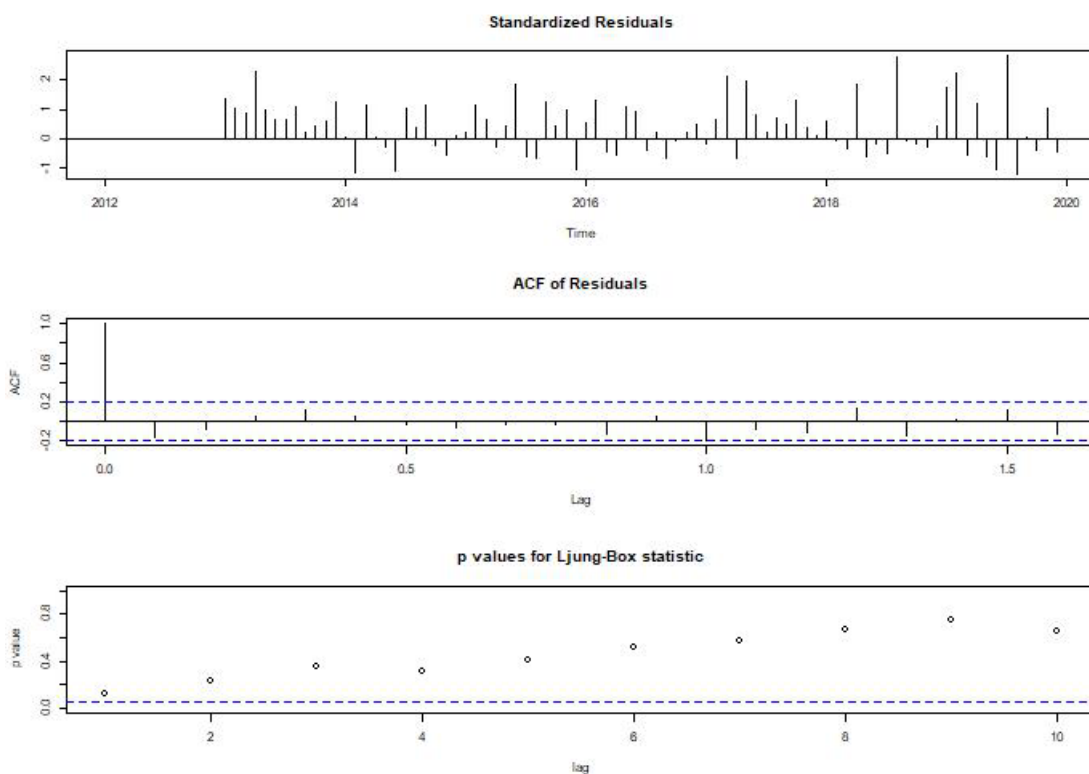


Figura 3.10: Diagnóstico dos Resíduos

Dando continuidade a este trabalho, no intuito de fazer comparação de modelos, foi aplicada à Série original a função `auto.arima` implementada no R, a qual tem capacidade de ajustar o melhor modelo ARIMA para séries temporais univariadas, sob as restrições de ordem fornecidas, de acordo com valores resultantes do AIC, do AICc ou do BIC, respectivamente.

Após testar diferentes restrições de ordem, foram fornecidas as seguintes para a função `auto.arima` aplicada na Série original: $\text{max.p} = 5$, $\text{max.q} = 5$, $\text{max.d} = 1$, $\text{max.P} = 5$, $\text{max.Q} = 5$, $\text{max.D} = 1$, $\text{max.order} = 5$, $\text{ic} = \text{c}(\text{"aicc"}, \text{"aic"}, \text{"bic"})$. Desse modo, função `auto.arima` retornou um SARIMA(1, 0, 0)(0, 1, 1)[12] com drift e com destaque para o valor do $AIC = 1765.5$, Tabela 3.9 e Tabela 3.10.

| | | | | | | | |
|---|--------------|-------------|----------|------------|----------|-----------|-------------|
| Series: PROD.ts | | | | | | | |
| ARIMA(1,0,0)(0,1,1)[12] with drift | | | | | | | |
| Coefficients: | | | | | | | |
| | ar1 | sma1 | drift | | | | |
| | 0,2261 | -0,7036 | 466,5894 | | | | |
| s.e. | 0,1092 | 0,1434 | 43,5134 | | | | |
| sigma^2 = 67217352: log likelihood = -878,75 | | | | | | | |
| AIC=1765,5 | AICc=1766,01 | BIC=1775,23 | | | | | |
| Training set error measures: | | | | | | | |
| | ME | RMSE | MAE | MPE | MAPE | MASE | ACF1 |
| Training set | 504,5493 | 7530,912 | 5667,038 | 0,01374557 | 9,635288 | 0,6240069 | -0,03310065 |

Tabela 3.9: Estimação de Parâmetros (função auto.arima)

| | | | | | |
|--------------------------------|-----------|------------|---------|-----------|--------------------------|
| z test of coefficients: | | | | | |
| | Estimate | Std. Error | z value | Pr(> z) | |
| ar1 | 0,22612 | 0,10916 | 2,0714 | 0,03832 | * |
| sma1 | -0,70356 | 0,14341 | -4,9061 | 9,29E-04 | *** |
| drift | 466,58936 | 43,51339 | 10,7229 | < 2,2e-16 | *** |
| --- | | | | | |
| Signif. codes: | 0 | **** | 0.001 | *** | 0.01 ** 0.05 . 0.1 ' ' 1 |

Tabela 3.10: Teste z para os coeficientes (função auto.arima)

Também aqui podemos analisar graficamente a questão da Raiz Unitária, Figura 3.11 e a adequação do ajuste aos dados, Figura 3.12. No caso da Raiz Unitária, como os pontos encontram-se no interior dos círculos unitários, concluímos que estamos lidando com uma Série Estacionária.

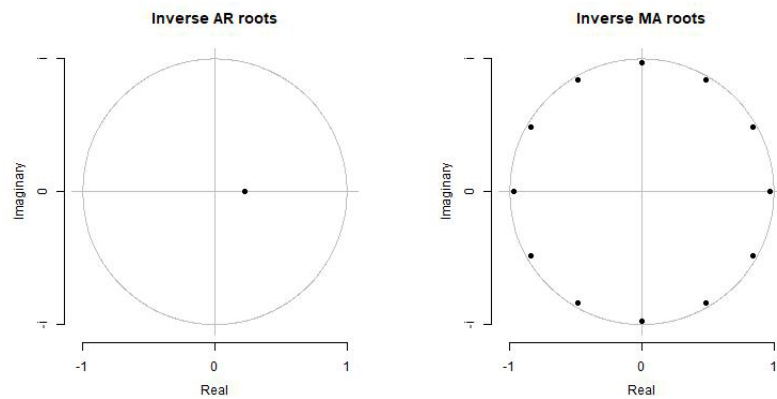


Figura 3.11: Indicativo de Estacionariedade (função auto.arima)

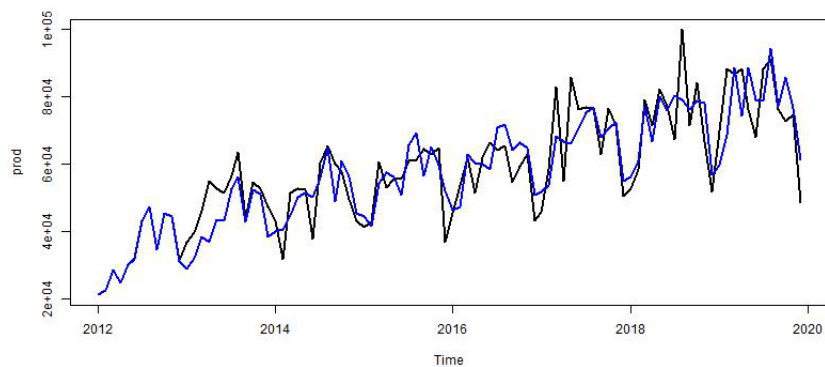


Figura 3.12: Resultado do Ajuste aos Dados (função auto.arima)

Dáí, preferiu-se continuar os estudos com o modelo otimizado `auto.arima`, o qual apresentou o menor valor de *AIC*, para dentre outras aplicações necessárias neste estudo, realizar previsões de valores futuros para a Série original. O critério de informação usado aqui, dentre outros, é o que foi retornado pelo dois modelos, podendo assim ter seus valores comparados. Em seguida foi feito o diagnóstico dos resíduos (pacotes `stats` e `graphics`).

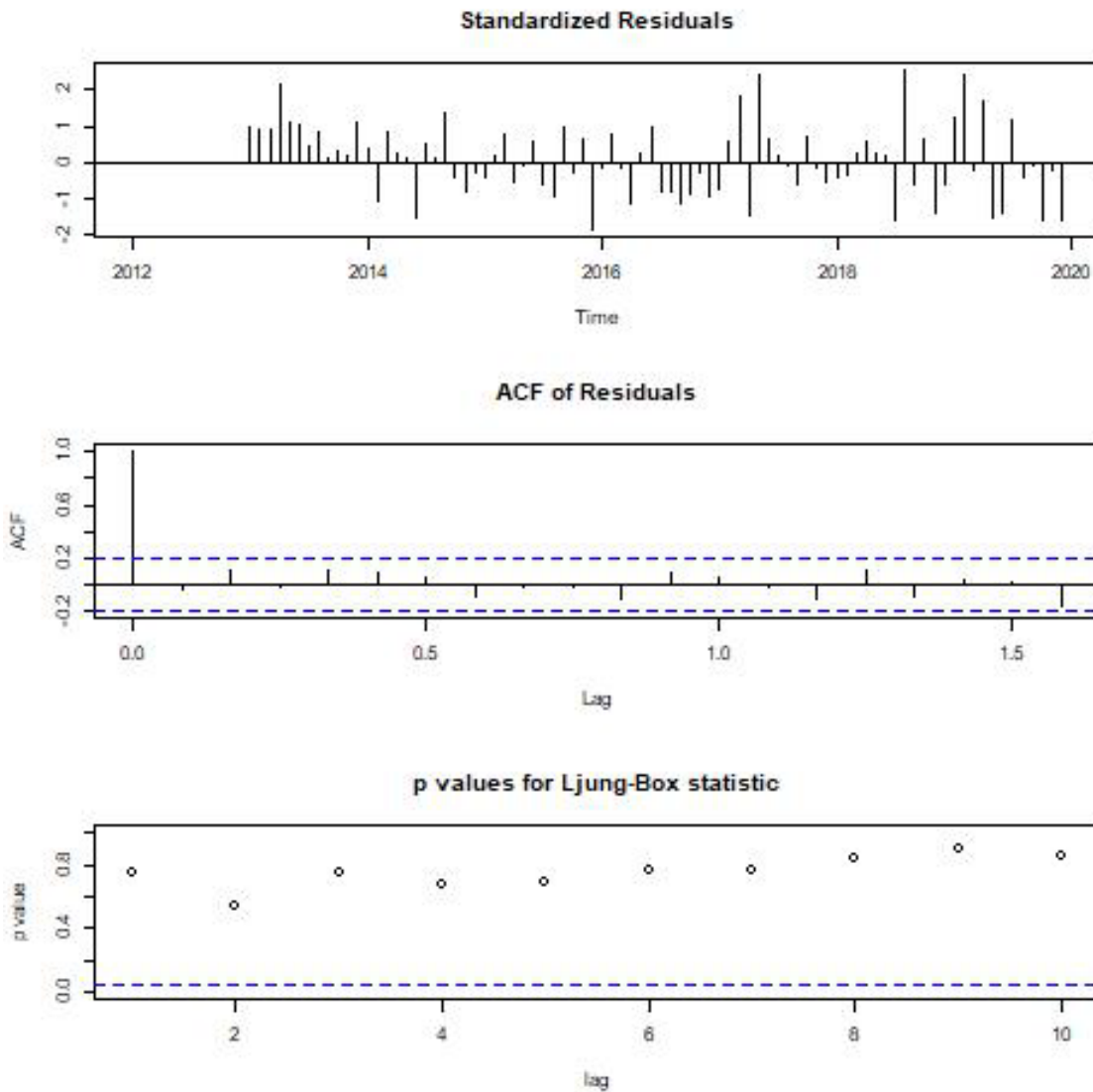


Figura 3.13: Diagnóstico dos resíduos (função `auto.arima`)

Portanto, na Figura 3.13 a análise gráfica corresponde ao diagnóstico dos resíduos do modelo otimizado, com evidências que as autocorrelações residuais podem ser consideradas nulas, condição necessária para que se realize previsões de valores em períodos futuros da Série original.

Nos dois diagnósticos de resíduos feitos anteriormente, foram usados os teste para autocorrelação residual de Box-Pierce e de Ljung-Box, ambos implementados no R.

Considerando a análise dos resíduos, os dois modelos mostraram-se adequados para uso nas etapas seguintes deste trabalho, mas o modelo otimizado apresentou menor valor de critério de informação e por essa razão foi o escolhido.

A partir desta etapa do presente estudo, o objetivo é identificar o melhor modelo que produza uma série de resíduos com comportamento de "ruído branco", o que se verifica neste caso. Conforme mostrado na Figura 3.14 a modelagem de previsão de Séries Temporais, por si só, é a parte mais importante de qualquer estudo de Série Temporais.

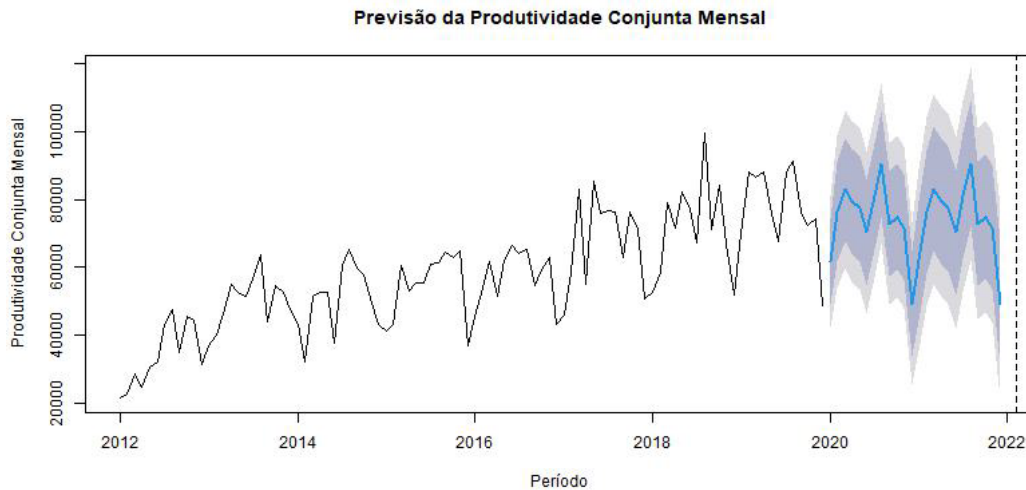


Figura 3.14: Previsão da produtividade no período de jan-2020 a dez-2021

Entretanto, para cumprirmos a principal finalidade deste trabalho acadêmico, necessitamos ir além, já que o fato mesmo de o modelo gerar previsões da Série original, atesta que este também produz em simultâneo e sucessivamente, uma medida estocasticamente válida (resíduos), a qual pode ser monitorada estatisticamente, com o poder de indiretamente também monitorar os próprios dados originais e fazer gestão apropriada do Indicador objeto da aplicação destes métodos estatísticos.

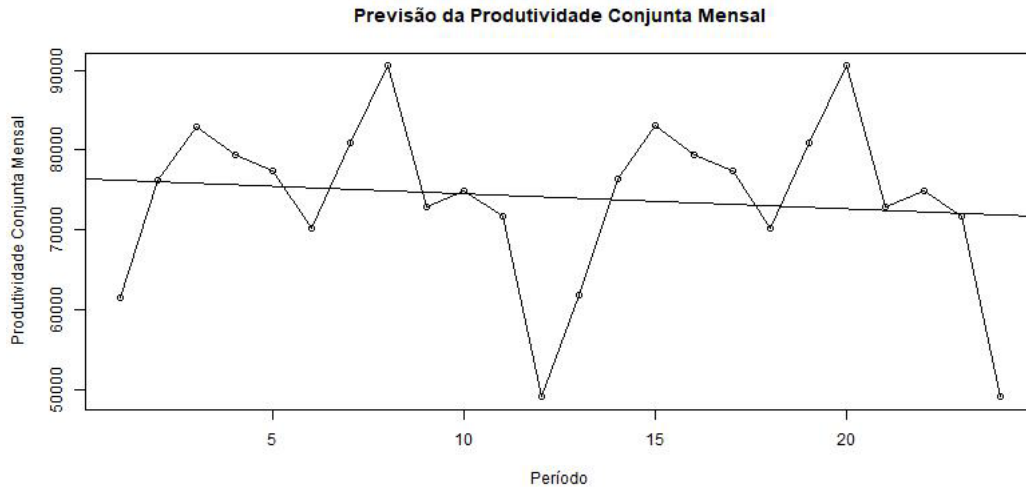


Figura 3.15: Destaque da previsão no período de jan-2020 a dez-2021

A Figura 3.15 apresenta em destaque um trecho específico do gráfico anterior, que é exatamente a produção dos valores previstos para a Série original, período de janeiro/2020 a dezembro/2021, acrescido da reta de regressão simples obtida destes valores previstos, elementos que necessariamente são analisados em conjunto com a Carta de Controle *CUSUM* que será apresentada a seguir. Veremos que cada ponto gerado na Carta *CUSUM* corresponde a um ponto que é origem de uma nova previsão de valores para a Série original.

3.3.2 Monitoramento Estatístico

A Carta de Controle *CUSUM* não monitora os dados originais da Série e sim, os resíduos gerados pela Série Temporal Ajustada. A *CUSUM* atualiza um novo valor do resíduo, o qual corresponde exatamente ao novo valor da Série Temporal original, que torna-se a origem de uma nova previsão (reta de previsão), cabendo ao gestor tão somente observar se a nova previsão aponta para a meta preestabelecida ou está se afastando. Aí lhe resta, como gestor, fazer ajustes, mudanças, identificar e atuar para eliminar causas.

Quanto à Carta de Controle *CUSUM*, ela sempre vai mostrar se o padrão de desempenho melhorou ou continua o mesmo, ou piorou. Em qualquer caso, se houve mudança, a Carta de Controle *CUSUM* atualiza o modelo automaticamente. Toda vez que a performance do Indicador apresenta melhora, o que representa a capacidade do gestor e de sua equipe de mudar a realidade a seu favor, espera-se que a variabilidade diminua, tornando os ambientes tático e estratégico mais previsíveis.

A Figura 3.16 apresenta o gráfico de Monitoramento Estatístico para pequenas variações com uso da Carta de Controle *CUSUM* (pacotes *cusum* e *qcc*).

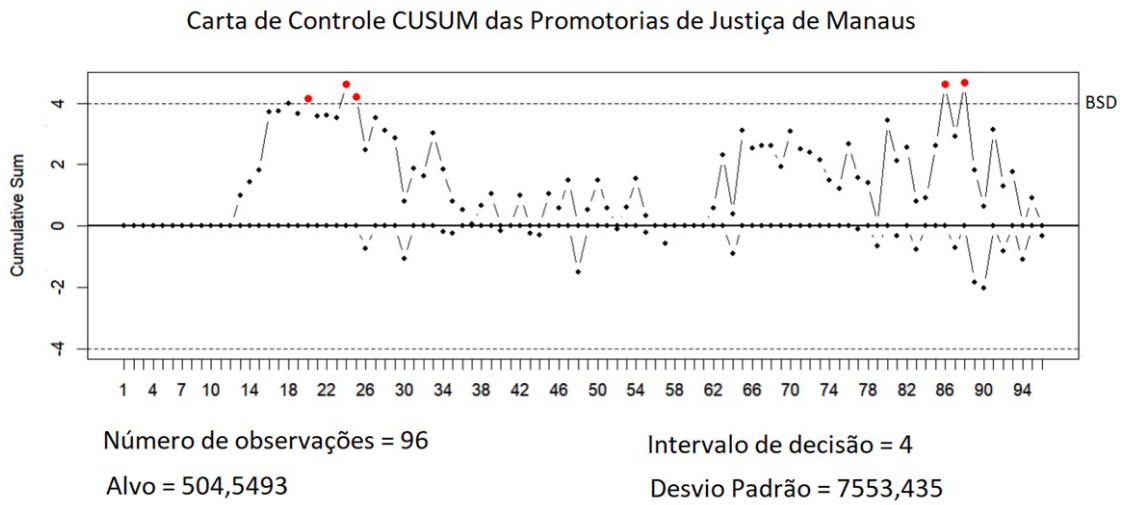


Figura 3.16: Carta de Controle *CUSUM* aplicada à serie de resíduos

Então, podemos constatar a presença de resultados indicados por pontos vermelhos plotados fora da faixa de monitoramento, significando mudança lá no padrão estatístico de desempenho do Indicador de Produtividade em estudo, permitindo que se recalcule tanto a média quanto a variabilidade da Série original no período monitorado, obtendo-se assim, um novo padrão de monitoramento da Carta *CUSUM*, isto se a mesma estivesse em uso. Simultaneamente seria atualizado o gráfico que destaca as previsões da produtividade (Figura 3.15). A Carta de Controle *CUSUM* atua no momento presente, enquanto o modelo estatístico de previsão empregado trata do futuro, obviamente, mas é necessário que os dois recursos estatísticos atuem conjuntamente.

Capítulo 4

Conclusões e proposta de novos estudos para uso de outras metodologias

Inicialmente identificamos um modelo arima sazonal, SARIMA(0,0,2)(1,1,2)[12], que se ajustou aos dados da série original e está apto para fazer previsões.

No entanto, ao aplicarmos a função `auto.arima` aos mesmos dados, identificamos outro modelo arima sazonal, um SARIMA(1,0,0)(0,1,1)[12], o qual apresentou ajuste aos dados e adequação para previsão superiores ao modelo anterior. Então usamos a série de resíduos deste modelo para fazer o monitoramento CUSUM, última etapa deste trabalho acadêmico.

Concluimos então, que os resultados obtidos comprovam que a implementação destas metodologias, desde que no ambiente *online*, darão suporte seguro às gestões da produtividade, da resolutividade e às tomadas de decisão no âmbito do MPAM.

Contudo, no contexto da Produtividade e Resolutividade do MPAM, nos deparamos com algumas centenas de indicadores, se considerarmos Órgãos e Membros Ministeriais. Além disso, em 15/09/2022 foi entregue ao Subprocurador-Geral de Justiça para Assuntos Administrativos, o Relatório Final do Grupo de Trabalho constituído pela Portaria nº 457/2022/SUBADM, relatório este que propõe entre outras soluções uma metodologia para gerar a Série Histórica do Indicador essencial para a gestão da Resolutividade no MPAM, que é exatamente o Tempo Médio de Duração do Processo.

Assim, propomos em estudos futuros, aplicarmos também, métodos de filtragem de partículas, já que são modelos dinâmicos que trabalham em tempo real em grandes massas de dados, não importando a quantidade de indicadores a serem monitorados.

Destacamos a abordagem de Liu e West (2001), referência devido à sua eficiência e generalidade nas hipóteses de construção de algoritmo, assegurando a concepção de modelos de alta capacidade descritiva, capazes de adequar a cenários bastante complexos, o que objetivamos no contexto dos indicadores do MPAM. Assim poderemos avaliar as performances das duas metodologias aplicadas no mesmo contexto institucional.

Referências Bibliográficas

- Akaike, H. (1973). *Information Theory and an Extension of the Maximum Likelihood Principle*, pages 199–213. Springer New York, New York, NY.
- Box, G. E. & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, **26**(2), 211–243.
- Box, G. E. P. & Jenkins, G. M. (1976). *Time Series Analysis*. ISBN 9780816211043.
- Deming, W. E. (1990). *Saia da crise*. ISBN 9788585238155.
- Hawkins, D. (1993). Cumulative sum control charts - an underutilized spc tool. *Quality Engineering*, **5**, 463–477.
- Ihaka, R. & Gentleman, R. (1996). R: a language for data analysis and graphics. *Journal of computational and graphical statistics*, **5**(3), 299–314.
- Knuth, D. E. (1986). *Computers & Typesetting*. Addison-Wesley Reading, MA.
- Lamport, L. (1985). *II (\LaTeX)—A Document*, volume 410. pub-AW.
- Marchezan, A. & Souza, A. M. (2010). Previsão do preço dos principais grãos produzidos no rio grande do sul. *Ciência Rural*, **40**, 2368–2374.
- Montgomery, D. C. (2021). *Introdução ao controle estatístico da qualidade*. ISBN 9788521614005.
- Morettin, P. A. (2021). *Análise de séries temporais*. ISBN 9788521203896.
- Page, E. S. (1954). Continuous inspection schemes. *Biometrika*, **41**(1/2), 100–115.
- Salomon, D. (2012). *The Advanced TEXbook*. Springer Science & Business Media.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, **6**(2), 461–464.
- Shewhart, W. A. (1980). *Economic Control of Quality of Manufactured Product*. ISBN 9780873890762.

- Siegmund, D. (2010). *Sequential Analysis*. ISBN 9781441930750.
- Sobral, T. E. & Barreto, G. (2011). Análise dos critérios de informação para a seleção de ordem em modelos auto-regressivos. In *Conferência Brasileira de Dinâmica, Controle e Aplicações*, volume 10, pages 379–382.
- Souza, F. (2016). Modelos de previsão: aplicações à energia elétrica arima-arch-ai e acp. *Curitiba: Appris*.
- Werner, L. & Ribeiro, J. L. D. (2003). Previsão de demanda: uma aplicação dos modelos box-jenkins na área de assistência técnica de computadores pessoais. *Gestão & Produção*, **10**, 47–67.