



Universidade de Brasília

IE – Departamento de Estatística

Estágio Supervisionado em Estatística 2

Utilização de modelo de previsão e de gráficos de controle combinados Shewhart-MMEP para a Arrecadação Média de Tributos da Secretaria de Fazenda do Estado do Piauí - SEFAZ/PI

Gustavo José de Carvalho

Relatório Final

Orientador: Prof. Dr. Gladston Luiz da Silva

Brasília – DF

2013

Gustavo José de Carvalho

Utilização de modelo de previsão e de gráficos de controle combinados Shewhart-MMEP para a Arrecadação Média de Tributos da Secretaria de Fazenda do Estado do Piauí - SEFAZ/PI

Relatório Final de conclusão de curso, apresentado à disciplina Estágio Supervisionada II do Curso Bacharelado em Estatística – Instituto de Ciências Exatas da Universidade de Brasília.

Orientador:

Gladston Luiz da Silva

Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Brasília – DF

2013

Dedicatória

Dedico este trabalho aos meus
pais Guilherme e Nice.

Agradecimentos

A Deus, renovador da minha força, por estar sempre acompanhando meus passos e abençoando cada dia da minha vida.

Aos meus pais por todo apoio, cuidado, dedicação e paciência. Por estarem ao meu lado em todos os momentos em que mais precisei. Ao meu irmão Vinicius Carvalho que está sempre à disposição.

Aos meus colegas de curso que, ao decorrer da minha vida acadêmica, tiveram papel primordial para o sucesso da conclusão deste ciclo da minha vida.

Ao meu orientador Professor Gladston Luiz da Silva, pelo grande esforço, dedicação e paciência para me guiar nos momentos difíceis na elaboração deste trabalho.

Ao Professor Jhames Sampaio pela disposição em me ajudar em alguns momentos específicos deste trabalho.

À Maria José por fornecer as informações utilizadas neste trabalho e pela colaboração para a realização desta monografia.

Resumo

A preocupação pela prestação de serviços de qualidade no setor público despertou à Secretaria da Fazenda do Estado do Piauí buscar o acompanhamento dos seus objetivos e a avaliação dos resultados de seus serviços oferecidos. Estas ações se fazem necessárias para se ter conhecimento se os objetivos da organização estão sendo atingidos. Acompanhar indicadores estratégicos da organização é etapa essencial para o sucesso da avaliação do andamento da Secretaria. Os indicadores são uma forma de medir uma situação atual contra um padrão previamente estabelecido e, conseqüentemente, apoiam à análise crítica dos resultados, às tomadas de decisão e ao planejamento e controle dos processos da organização. Após estudo do documento relativo ao planejamento estratégico da SEFAZ-PI foram identificados vários indicadores estratégicos, dentre os quais foi selecionado o indicador relativo à *Arrecadação Média* para ser objeto de estudo neste trabalho, cujo objetivo é ajustar modelos de séries temporais para o indicador e propor o uso combinado dos gráficos de controle de Shewhart e Média Móvel Exponencialmente Ponderada (MMEP) para os resíduos do modelo ajustado, para que a Secretaria de Fazenda do Estado do Piauí – SEFAZ possa acompanhar o comportamento do indicador e atuar de acordo com os resultados alcançados.

Palavras-Chave: Controle Estatístico de Qualidade, Gráficos de Controle de Shewhart e Médias Móveis Exponencialmente Ponderadas, Modelos de Séries Temporais, Indicadores Estratégicos.

Abstract

The concern for providing quality services in the public sector aroused Treasury of the State of Piauí to pursue the monitoring of its provided services. These actions are needed in order to make sure if the organization's objectives are being accomplished. Follow strategic indicators of the company is essential to the success of the evaluation of the services provided by the Treasury of the State of Piauí. Indicators are a way to measure a current situation against a pattern previously established and hence support the critical analysis of the results, the decision-making and the planning and control of the processes of the organization. After studying the documents related to the strategic planning of the SEFAZ-PI some strategic indicators were identified, among which the Average Storage was selected to be the object of study of this work, whose goal is to adjust the time series models to the indicator and propose the combined use of control charts from Shewhart and the Exponentially Weighted Moving Average (EWMA) for residues of the adjusted model, so that the Secretary of Treasury of the State of Piauí – SEFAZ can monitor the behavior of the indicator and chose the course of action according to the results achieved.

Keywords: Statistical Process Control, Control Chart of Shewhart, Statistical Quality Control, Moving Average, Exponentially Weighted, Auto Regressive Integrated.

Sumário

Capítulo 1	14
1.1. Introdução	14
1.2. Justificativa e Motivação.....	15
1.3. Objetivos	16
1.3.1. Gerais	16
1.3.2. Específicos	16
1.4. Metodologia	17
1.5. Estrutura	19
Capítulo 2	20
2. Revisão Bibliográfica.....	21
2.1. Análise de Séries Temporais	21
2.1.1. Estacionariedade.....	21
2.1.2. Modelos ARIMA.....	22
2.1.2.2. Modelos Médias Móveis (MA)	23
2.1.2.3. Modelos Auto-Regressivo de Médias Móveis (ARMA).....	23
2.1.2.4. Os modelos Auto-Regressivo Integrados de Médias Móveis (ARIMA).....	23
2.1.3. Identificação do modelo	24
2.1.3.3. Informações de Akaike (AIC) e Bayesiano (BIC)	25
2.1.3.4. Escolha do modelo	26
2.1.4. Modelos de suavização exponencial	26
2.1.5. Modelos e procedimentos de previsão	29
2.2. Análise dos Resíduos	31
2.3. Gráficos de Controle de Shewhart.....	32
2.3.1. Gráficos de controle para \bar{x} e R.....	32
2.3.2. Gráficos de controle para \bar{x} e S	34
2.3.3. Gráficos de Controle de Shewhart para medidas individuais.....	36
2.4. Gráfico de controle de Média Móvel Exponencialmente Ponderada.....	38
2.5. Combinado: Shewhart e MMEP.....	40
2.6. Gráficos de controle para monitorar previsões.....	41
Capítulo 3	42
3. Desenvolvimento.....	42
3.1. Seleção do Indicador da SEFAZ/PI.....	42

3.2.	Identificação do modelo de previsão.....	46
3.2.1.	Análise exploratória dos dados.....	46
3.2.2.	Modelagem de suavização exponencial de Holt-Winters	49
3.2.3.	Identificação de modelos ARIMA	51
3.3.	Preposição dos gráficos de controle	57
3.3.1.	Controle dos resíduos	57
3.3.2.	Controle da série transformada.....	59
3.4.	Implantação dos gráficos de controle.....	62
Capítulo 4.....		63
4.	Conclusão.....	63
	Referência Bibliográfica	65
	Apêndice 1: Os Quatorze Pontos do Método Deming	66
	Apêndice 2: Teste de Shapiro-Wilk	67
	Apêndice 3: Teste Ljung-Box	69
	Anexo 1: Tabela distribuição X^2	70

Lista Símbolos

A2, B3, B4, D3 e D4 - Coeficientes para construção gráficos de Shewhart

AIC - Critério de Informação de Akaike

AICc - Critério de Informação de Akaike corrigido

α_t - Ruído Branco

B - Operador de Translação para o Passado

BIC - Critério de Informação Bayesiano

CEP - Controle Estatístico de Processos

c4- Valor dependente de n para a construção de gráficos de Shewhart

d - Numero de diferenciações realizadas nos modelos ARIMA

d2 e d3-Fatores de correção para a estatística \bar{R}

H- Intervalo de decisão

e- Valor de referencia

F_t - Componente da Sazonalidade de uma Série Temporal

k - Numero de defasagens

L - Distancia em função de desvios padrões dos limites de controle para a linha central

n - Tamanho da amostra

\bar{n} - Tamanho médio da amostra

p - Ordem do modelo AR

Q - Estatística do teste de Ljung-Box

q - Ordem do modelo MA

R - Amplitude de um processo

\bar{R} - Amplitude media

S - Desvio padrão amostral

\bar{S} - Media dos desvios padrões amostral

S_t - Componente da sazonalidade da série temporal

t - Tempo de observação de uma série

T_t - Componente da tendência de uma série temporal

X_i - i-ésima observação

\bar{X} - Média das observações

$\bar{\bar{X}}$ - Média das médias as observações

Z_t - Nível de uma série temporal

Δ - Desvio do valor nominal

δ - Tamanho da mudança da média em unidades de desvio padrão

θ - Parâmetro do modelo de um processo MA(q)

λ - Fator de suavização da média móvel

μ_0 - valor médio desejado do processo, media sob controle

μ_1 - valor médio do processo para o estado fora controle

μ - Média da população

σ - Desvio padrão da população

$\bar{\sigma}$ - Desvio padrão estimado

Φ - Parâmetro do modelo de um processo AR(p)

φ_k - Função de Auto Correlação Parcial

Lista Abreviações

AR - Auto-Regressivo

ARMA -Auto-Regressivo e de Médias Móveis

ARIMA - Auto-Regressivo Integrado e de Médias Móveis ou Auto-Regressive Integrated Moving Averages

CUSUM - Somas Acumuladas ou Cumulative Sum

FAC - Função de Auto-Correlação

FACP - Função de Auto-Correlação Parcial

GQT - Gestão de Qualidade Total

LC - Linha Central

LIC - Limite Inferior de Controle

LSC - Limite Superior de Controle

MA - Médias Móveis

MMEP - Médias Móveis Exponencialmente Ponderadas ou Exponentially Weighted Moving Average

SEFAZ/PI - Secretaria da Fazenda do Governo do Piauí

Lista de Ilustrações

Lista de Gráficos

Gráfico 1 - Arrecadação por Contribuinte

Gráfico 2 - Histograma de Arrecadação por Contribuinte

Gráfico 3 - Componente Sazonal da Arrecadação Média da SEFAZ-PI

Gráfico 4 - Componente de Tendência da Arrecadação Média da SEFAZ-PI

Gráfico 5 - Ajustamento Holt-Winters “Aditivo”

Gráfico 6 - Ajustamento Holt-Winters “Multiplicativo”

Gráfico 7 - FAC da Arrecadação Por Contribuinte

Gráfico 8 - FACP da Arrecadação Por Contribuinte

Gráfico 9 - Arrecadação Média Transformada

Gráfico 10 - FAC da Arrecadação Média Transformada

Gráfico 11 - FACP da Arrecadação Média Transformada

Gráfico 12 - Modelo Ajustado ARIMA(0,1,1) para a Arrecadação Média da SEFAZ-PI

Gráfico 13 - Histograma dos Resíduos

Gráfico 14 - Gráfico de Controle Shewhart

Gráfico 15 - Gráfico de Controle “MMEP” ($\Lambda = 0,1$ E $L = 2,7$)

Gráfico 16 - Gráfico de Controle “MMEP” ($\Lambda = 0,2$ E $L = 2,7$)

Gráfico 17 - Gráfico de Controle Shewhart

Gráfico 18 - Gráfico de Controle “MMEP” ($\Lambda = 0,1$ E $L = 2,7$)

Lista de Tabelas

Tabela 1 - Comportamento das Funções FAC e FACP para Modelos Estacionários

Tabela 2 - Variáveis para Construção do Indicador

Tabela 3 - Arrecadação Média do Estado do Piauí

Tabela 4 - Coeficientes Otimizados

Tabela 5 - Comparação AIC, AICc E BIC

Tabela 6 - Teste Ljung-Box

Tabela 7 - Shapiro-Wilk

Tabela 8 - Resíduos da Série Ajustada à Arrecadação Média Da SEFAZ-PI, 2008-13

Tabela 9 - Limites de Controle de Shewart para os Resíduos do Modelo Ajustado

Tabela 10 - Série Diferenciada da Arrecadação Média

Tabela 11 - Limites de Controle de Shewart para a Arrecadação Média Diferenciada

Capítulo 1

1.1. Introdução

Os princípios de gerenciamento foram introduzidos por Frederick W. Taylor em um momento de desenvolvimento das indústrias de produção em massa, antes de 1900. Já naquela época, Taylor trouxe ideias como a divisão de trabalho em tarefas e métodos padronizados de produção e montagem que impactaram positivamente na qualidade dos bens manufaturados.

O controle estatístico de qualidade inicia-se formalmente com Walter A. Shewhart, em 1924, com a criação do gráfico de controle. Esta nova técnica permitia distinguir o momento em que o processo de produção apresentasse algum comportamento fora do padrão.

No momento em que as ferramentas estatísticas para o controle são implementadas como parte de um sistema de gerenciamento orientado para qualidade, a gestão passa a ser muito mais eficiente. O gerenciamento da qualidade total (GQT) realiza esta junção e tem como seus principais propulsores W.E. Deming, J. M. Juran e A. V. Feigenbaum.

Apesar destes três importantes nomes relacionado ao GQT, W. Edward Deming pode ser considerado o principal deles. Até sua morte em 1994, o Dr. Deming foi uma força inspiradora para a melhoria da qualidade ao redor do mundo. Ele acreditava firmemente que a responsabilidade pela qualidade está no gerenciamento e sua filosofia se resume a 14 pontos para gerenciamento (Apêndice 1).

O conceito de qualidade vem evoluindo ao longo dos anos e a gestão de qualidade tornou-se um dos focos das empresas que pretendem conquistar vantagem competitiva sobre os concorrentes.

O controle estatístico de qualidade é uma etapa de suma importância dentro da gestão de qualidade e tem como ponto chave o acompanhamento do desempenho da organização.

1.2. Justificativa e Motivação

Por um motivo de avanço cultural, tecnológico e econômico, a preocupação em prestar um serviço de qualidade na área pública está sendo cada vez mais disseminado em nosso país. Um exemplo disso são as parcerias, cada dia mais comum, do governo federal com organizações de consultorias de gestão empresarial.

A Secretaria da Fazenda do Estado do Piauí – SEFAZ/PI está realizando um trabalho de acompanhamento de seu planejamento estratégico e de avaliação dos resultados obtidos. Essas ações visam elevar seu desempenho e adequar seus objetivos, metas e ações ao Plano de Governo do Estado. Desde 2008, a SEFAZ/PI tem investido em tecnologia, principalmente em sistemas para gestão das atividades desenvolvidas pelo seu público interno e externo. Desde então a Secretaria tem armazenado uma quantidade considerável de informações relativas aos resultados alcançados, que ainda não foram exploradas.

Este trabalho teve o objetivo de identificar, a partir das informações fornecidas pela SEFAZ-PI, um indicador de desempenho da organização e propor ferramentas estatísticas para seu acompanhamento. Primeiramente, devido à não estacionaridade da série, foi ajustado um modelo de séries temporais ao indicador selecionado. Em seguida, foi proposto o uso de gráficos de controle estatístico para os resíduos do modelo ajustado à série original que permitam o acompanhamento do processo. Estas ações tem como principal finalidade de mostrar a aplicação de técnicas estatísticas a indicador de desempenho estratégico da Secretaria de Fazenda do Estado do Piauí.

1.3. Objetivos

1.3.1. Gerais

Este trabalho tem como objetivo ajustar um modelo de séries temporais para o indicador *Arrecadação Média de Tributos* da SEFAZ/PI e propor o uso combinado dos gráficos de controle de Shewhart e Média Móvel Exponencialmente Ponderada (MMEP) para os resíduos do modelo ajustado.

1.3.2. Específicos

Os objetivos específicos desta monografia são:

- Ajustar modelo de previsão para o indicador selecionado;
- Propor gráficos de controle de Shewhart e de MMEP para os resíduos do modelo ajustado;
- Parametrizar os gráficos de controle;
- Sugerir procedimentos para a implantação dos gráficos de controle.

1.4. Metodologia

Esta seção apresenta a metodologia adotada para o alcance dos objetivos propostos neste projeto.

A primeira etapa consiste na revisão bibliográfica das técnicas estatísticas que serão abordadas no presente trabalho, sendo elas: Análise de Séries Temporais e Controle Estatístico de Processo. Antes da revisão bibliográfica propriamente dita, há também um breve resumo relacionado à Gestão de Qualidade e Indicadores de Desempenho. Foram consultados livros e artigos que abordam os conteúdos citados com o intuito de revisar os conceitos essenciais e a teoria necessária para o desenvolvimento do trabalho.

Como o principal insumo dos gráficos de controle são os indicadores, fez-se um estudo minucioso do Planejamento Estratégico da Secretaria de Fazenda do Piauí, o que possibilitou selecionar, dentre os indicadores da organização, aquele que seria objeto deste estudo.

Selecionado o indicador, foi realizado um estudo descritivo dos dados disponibilizados pela SEFAZ/PI, em especial daqueles relacionados ao indicador selecionado, a *Arrecadação Média de Tributos*, que resulta da divisão da arrecadação auferida pelo número de contribuintes no período.

Em seguida, foi verificado se a série relacionada ao indicador era estacionária, isto é, se “os dados do processo variavam em torno de uma média fixa de maneira estável e previsível” (Montgomery, Douglas C., 1943, 100). Foi verificado, também, se havia correlação serial dos dados.

Em virtude do indicador selecionado representar uma série não-estacionária, este estudo propôs o uso de gráficos de controle para os resíduos do modelo de séries temporais ajustado para o indicador selecionado.

Neste estudo é proposto o uso combinado dos gráficos de controle de Shewhart e MMEP para os resíduos do modelo ajustado, visto que, quando usados conjuntamente, são eficientes para detecção de pequenas ou de grandes perturbações no processo. Enquanto o MMEP é um gráfico eficiente na detecção de pequenas perturbações, os gráficos de Shewhart são eficientes na detecção de perturbações de maior grandeza. Com a combinação desses gráficos o processo

é interrompido para avaliação sempre que um dos gráficos, ou ambos, apresentar indicativo de que o processo saiu de controle.

O passo seguinte foi a parametrização dos gráficos sugeridos, decorrente do estudo descritivo do indicador, realizado a partir dos dados disponibilizados pela SEFAZ/PI.

Para a realização das análises estatísticas do presente trabalho foi utilizado o Software R, que foi bastante útil, principalmente, para as análises de séries temporais..

1.5. Estrutura

O trabalho está estruturado em quatro capítulos, assim expostos: o Capítulo 1 apresenta a justificativa e motivação deste trabalho, os objetivos gerais e específicos almejados e a metodologia utilizada para sua realização; o Capítulo 2 apresenta o referencial bibliográfico utilizado no presente trabalho; o Capítulo 3 apresenta o desenvolvimento do trabalho; e, por fim, são apresentadas as conclusões obtidas do trabalho.

Capítulo 2

Este capítulo tem início com um breve resumo relacionado a assuntos de gestão de qualidade e de indicadores de desempenho, e segue com uma revisão bibliográfica das técnicas estatísticas que serão abordadas. A seção 2.1 apresenta o tema de Séries Temporais. Na seção 2.2 apresenta-se sobre a análises dos resíduos. Os conceitos sobre Gráficos de Controle de Shewhart e MMEP são apresentados nas seções 2.3 e 2.4, respectivamente. O uso combinado dos gráficos será apresentado na seção 2.5. Por fim, o tema de Gráficos de Controle para monitorar previsões será apresentado na seção 2.6.

A Gestão da Qualidade tem em Ishikawa um dos principais nomes. Ele propôs o conceito de gestão de qualidade total (GQT) que trazia consigo a ideia de aplicar a gestão de qualidade por toda a empresa.

Para Bonilla (1999), o GQT se baseia basicamente em duas fontes: *“uma de natureza técnica, que nasce com Taylor, se desenvolve com os métodos de controle estatístico de Shewhart, e se consolida com todo o conhecimento científico dos últimos 40 anos, através do trabalho dos grandes mestres: Feigenbaum, Deming e Juran; e outra, de natureza humana, apoiada nas pesquisas sobre comportamento desenvolvidas por Mc Gregor, Herzberg e Maslow e, mais recentemente, na abordagem holística representada, entre outros, por Capra e Ferguson.”*

A gestão de qualidade está totalmente atrelada à mensuração de indicadores e de seu acompanhamento pela a organização, pois são ferramentas integrantes do planejamento estratégico e fundamentais para o controle de processos de uma organização.

A análise dos resultados do processo é necessária para se constatar se os objetivos da organização estão sendo atingidos. Os indicadores são utilizados para verificar se o contexto está em acordo com padrões previamente estabelecidos e, conseqüentemente, para apoiar a análise crítica dos resultados, a tomada de decisão e o planejamento e controle dos processos da organização.

Este trabalho tem foco na natureza técnica da gestão de qualidade, visto que tem como objetivo propor gráficos de controle para indicador estratégico da Secretaria de Fazenda do Piauí – SEFAZ/PI.

2. Revisão Bibliográfica

2.1. Análise de Séries Temporais

Uma série temporal é um conjunto de observações ordenadas no tempo (não necessariamente igualmente espaçadas) e que apresentam dependência serial, isto é, dependência de instantes de tempo (Morettin; Toloí, 2006).

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

Um modelo de série temporal para dados observados (z_t) é uma especificação das distribuições de uma sequência de variáveis aleatórias (Z_t) da qual (z_t) é denominada uma realização (Brockwell & Davis, 1996).

2.1.1. Estacionariedade

Morettin e Toloí (2006) dizem que uma das suposições mais frequentes que se faz a respeito de uma série temporal é a de que ela é estacionária, ou seja, ela se desenvolve no tempo aleatoriamente ao redor de uma média constante, refletindo alguma forma de equilíbrio estável. Apesar da maior parte das séries encontradas na prática apresenta alguma forma de não-estacionariedade, na maioria dos procedimentos de análise estatística de séries temporais, supõe-se que estas sejam estacionárias.

Uma série temporal $\{X_t\}$ é dita ser (fracamente) estacionária se

$$(1) E(X_t) = \mu ;$$

$$(2) V(X_t) = \sigma_x^2 < \infty ;$$

$$(3) \gamma_{t,s} = \gamma_{t-s}.$$

onde $E(X_t)$ é a média da série, $V(X_t)$ é a variância e $\gamma_{t,s}$ é a covariância X_t e X_s .

Portanto, um processo estacionário apresenta média e variância constantes ao longo do tempo t e a covariância entre os dois pontos dependente da distância entre esses pontos e independente do tempo t (Granger & Newbold, 1986).

Com isso, caso a série não seja estacionária, será necessário transformar os dados originais. A transformação mais comum consiste em tomar diferenças sucessivas da série original, até se obter uma série estacionária. (Morettin; Toloi, 2006).

$$\Delta Z_t = Z_t - Z_{t-1} \quad (1.0)$$

2.1.2. Modelos ARIMA

Os modelos ARIMA (p,d,q) são resultados da combinação de três filtros: p o filtro Auto-Regressivo (AR), d o de integração (I) e q o de Médias Móveis (MA); e, para a modelagem de determinada série temporal, tanto se podem utilizar os três, quanto apenas um subconjunto destes (Fava, 2000).

A integração que é representada pela letra d do modelo ARIMA corresponde a possível diferenciação tomada na série para torná-la estacionária, citada na seção anterior. Caso o modelo não apresente esta componente, ele então poderá ser um modelo Auto-Regressivo (AR), um modelo de Médias Móveis (MA), ou então uma junção destes dois que é representado pela sigla (ARMA).

2.1.2.1. Modelo Auto-Regressivo (AR)

O modelo Auto-Regressivo (AR), de ordem p , é um processo iterativo em que há a identificação da ordem p através da função de auto-correlação; a partir daí, faz-se a estimativa de um modelo de previsão, bem como a análise dos resíduos para a avaliação da existência de vieses e/ou grandes erros de estimativas.

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

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + \alpha_t \quad (1.1)$$

onde $\phi_0, \phi_1, \dots, \phi_p$ são os parâmetros do modelo e α_t é o ruído branco.

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

2.1.2.2. Modelos Médias Móveis (MA)

Os modelos de Médias Móveis são gerados a partir de uma média ponderada das perturbações aleatórias ou erros aleatórios da defasagem q . Os modelos MA são semelhantes aos modelos AR, onde, ao invés dos valores passados da série, neste modelo, os valores considerados são os valores correntes dos erros aleatórios e seus q valores passados.

O modelo é dado pela forma

$$Z_t = \mu + \alpha_t - \theta_1 \alpha_{t-1} - \dots - \theta_q \alpha_{t-q} \quad (1.2)$$

onde μ e $\theta_1, \dots, \theta_q$ são constantes.

2.1.2.3. Modelos Auto-Regressivo de Médias Móveis (ARMA)

Um processo ARMA(p, q) é uma generalização dos modelos AR(p) e MA(q), sendo definido como

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + \alpha_t - \theta_1 \alpha_{t-1} - \dots - \theta_q \alpha_{t-q} \quad (1.3)$$

2.1.2.4. Os modelos Auto-Regressivo Integrados de Médias Móveis (ARIMA)

Se d for um número inteiro não-negativo, então uma série temporal Z_t é um processo ARIMA(p, d, q) ou um processo integrado de ordem d se

$$Y_t = (1 - B)^d Z_t = \Delta^d Z_t \quad (1.4)$$

for algum processo ARMA(p,q). Este tipo de modelo é utilizado para séries não-estacionárias pode-se dizer que um modelo ARIMA ($p,0,q$) pode ser dito como um modelo ARMA (p,q).

2.1.3. Identificação do modelo

Morettin e Tolo (2006) apresentam a estratégia para a construção do modelo que é baseada em um ciclo iterativo, no qual a escolha da estrutura do modelo ocorre nos próprios dados. Os estágios do ciclo iterativo são:

1. uma classe geral de modelos é considerada para a análise;
2. há identificação de um modelo, com base na análise de auto-correlações, auto-correlações parciais e outros critérios como o AIC, AICc e o BIC.
3. a seguir, vem a fase de estimação, na qual os parâmetros do modelo identificados são estimados.
4. finalmente, há a verificação ou diagnóstico do modelo ajustado, através de uma análise de resíduos, para se saber se este é adequado para os fins em vista.

Caso o modelo não seja adequado, o ciclo é repetido, voltando-se à fase de identificação. Um procedimento que é utilizado muitas vezes é identificar não só um único modelo, mas alguns modelos que serão então estimados e verificados. Se o propósito é previsão, escolher-se-á dentre os modelos ajustados o melhor, por exemplo, no sentido de fornecer o menor erro quadrático médio de previsão.

2.1.3.1. Função de Auto-Correlação (FAC)

A estatística de auto-correlação descreve a correlação entre dois valores de determinada variável, em diferentes períodos de tempo. Um coeficiente de auto-correlação ρ_1 , que mede a correlação entre a variável no período t em relação ao período $t - 1$ é chamado de coeficiente de auto-correlação de lag 1 , assim como um coeficiente de auto-correlação ρ_k é dito coeficiente de auto-correlação de lag k .

2.1.3.2. Função de Auto-Correlação Parcial (FACP)

A FACP é utilizada de forma conjunta com a FAC para a identificação dos modelos a serem utilizados para a modelagem e previsão da série temporal, podendo também auxiliar na identificação da condição de estacionaridade da série temporal. Para $k = 1$, os coeficientes de auto-correlação e auto-correlação parcial são idênticos. Para $k > 1$, pode ser utilizada Equação 1.5 (Matos, 2000)

$$\varphi_k = \left[\frac{t^2}{(t^2 + n - r)} \right]^{\frac{1}{2}} \quad (1.5)$$

onde t é a estatística do parâmetro \hat{a}_k na função $Y_{t-k} = \hat{a}_0 + \hat{a}_k Y_{t-k} + s_t$, r é o número de repressores incluídos na função, inclusive a constante.

A quantidade φ_k , encarada como função de k , é chamada *função de auto-correlação parcial*.

2.1.3.3. Informações de Akaike (AIC) e Bayesiano (BIC)

Burnham e Anderson (2004), enfatizam a importância de selecionar modelos baseando-se em princípios científicos. Dentre as diversas metodologias utilizadas para este fim, as mais comumente usadas são as informações de Akaike (AIC), Akaike Corrigido (AICc) e Bayesiano (BIC), quanto a sua performance na seleção de modelos.

O Critério de Informação de Akaike (AIC) é definido como

$$AIC_p = -2 \times \log(L_p) + 2 \times [(p + 1) + 1] \quad (1.6)$$

em que L_p é a função de máxima verossimilhança do modelo e p é o número de variáveis explicativas consideradas no modelo.

O Critério de Informação Bayesiano (BIC) é definido como

$$BIC_p = -2 \times \log(L_p) + 2 \times [(p + 1) + 1] \times \log(n) \quad (1.7)$$

Tanto o AIC quanto o BIC aumentam conforme a soma quadrada dos erros aumenta, ou seja, são preferíveis valores menores de AIC e BIC. Além disso, ambos penalizam modelos com muitas variáveis. Como modelos com mais variáveis tendem a produzir menor SQE mas usam mais parâmetros, a melhor escolha é balancear o ajuste com a quantidade de variáveis.

2.1.3.4. Escolha do modelo

A identificação do modelo segundo as autocorrelações, auto-correlações parciais seguem o seguinte comportamento.

Tabela 1 – Comportamento das funções FAC e FACP para modelos estacionários

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

Fonte: Adaptado de Johnson; Montgomery. (1974, p. 469).

Após a escolha das ordens do modelo a ser ajustado, pela observação de sua FAC e FACP, deve-se analisar as informações de Akaike e de Bayesiano com descrito na seção anterior.

2.1.4. Modelos de suavização exponencial

O alisamento exponencial é um procedimento geral para obtenção de algoritmos de previsão automática que produz resultados relativamente acurados, de maneira rápida e barata (Granger & Newbold, 1986).

A forma mais simples de alisamento exponencial é aquela para séries temporais que não possuem sazonalidade nem tendência crescente ou decrescente.

O objetivo é estimar o “nível” (ou a “média”) presente da série e usá-lo como previsão de valores futuros. O nível da série no tempo t é estimado como

$$\bar{Z}_t = \alpha Z_t + \alpha(1-\alpha)Z_{t-1} + \alpha(1-\alpha)^2 Z_{t-2} + \dots \quad (1.8)$$

A previsão de todos os valores futuros é dada pelo último valor exponencialmente suavizado, isto é,

$$\hat{Z}_t(h) = \bar{Z}_t, \quad \forall h > 0, \quad (1.9)$$

$$\hat{Z}_t = \alpha Z_t + \alpha(1-\alpha)Z_{t-1}(h+1), \quad (1.10)$$

que pode ser interpretada como uma equação de atualização de previsões, quando tivermos uma nova observação. Além disso, a previsão feita de acordo com a equação (1.10) reduz o problema de armazenagem de observações, pois pode ser calculada utilizando apenas uma observação mais recente, a previsão imediatamente anterior e o valor de α .

O erro da previsão é definido como

$$e_t = Z_t - \hat{Z}_{t-1},$$

Supondo $a_t \sim N(0, \sigma_a^2)$ podemos construir um intervalo de confiança assintótico para Z_{t+h} sendo ele

$$\hat{Z}_t(h) \pm z_\gamma \sigma_a \sqrt{\frac{\alpha}{2-\alpha}}$$

onde $E(\hat{Z}_t(h)) = \mu$, $Var(\hat{Z}_t(h)) = \frac{\alpha}{2-\alpha} \sigma_a^2$, z_γ é o coeficiente da $N(0,1)$ e $\hat{Z}_t(h)$ é dado por

(1.9). Normalmente σ_a^2 é desconhecida e deve ser estimada.

Caso de série temporal apresente tendência, a suavização exponencial simples (SES) não é indicada, pois este tipo de alisamento é recomendado para séries que apresentam tendência linear. A suavização exponencial de Holt leva em consideração este comportamento, pois ele utiliza uma nova constante de suavização pra modelar a tendência da série.

Os valores do nível e da tendência da série, no instante t , serão estimados por

$$\bar{Z}_t = AZ_t + (1 - A)(\bar{Z}_{t-1} + \hat{T}_{t-1}), 0 < A < 1 \text{ e } t = 2, \dots, N,$$

$$\hat{T}_t = C(\bar{Z}_t - \bar{Z}_{t-1}) + (1 - C)\hat{T}_{t-1}, 0 < C < 1 \text{ e } t = 2, \dots, N,$$

respectivamente. A e C são denominados constantes de suavização.

Para séries temporais que além de nível e tendência apresentem também sazonalidade o método mais indicado é o de Holt-Winters. Existem dois tipos de procedimentos para essa metodologia, cujo sua utilização depende das características da série considerada.

1 - Série Sazonal Multiplicativa

Considere uma série sazonal com período s . A variante mais usual do método HW considera o fator sazonal F_t como sendo multiplicativo, enquanto a tendência permanece aditiva, isto é,

$$Z_t = \mu_t F_t + T_t + a_t, \quad t = 1, \dots, N.$$

As três equações de suavização são dadas por

$$\hat{F}_t = D \left(\frac{Z_t}{\bar{Z}} \right) + (1 - D)\hat{F}_{t-s}, \quad 0 < D < 1 \text{ e } t = s + 1, \dots, N.$$

$$\bar{Z}_t = A \left(\frac{Z_t}{\hat{F}_{t-s}} \right) Z_t + (1 - A)(\bar{Z}_{t-1} + \hat{T}_{t-1}), \quad 0 < A < 1 \text{ e } t = s + 1, \dots, N.$$

$$\hat{T}_t = C(\bar{Z}_t - \bar{Z}_{t-1}) + (1 - C)\hat{T}_{t-1}, \quad 0 < C < 1 \text{ e } t = s + 1, \dots, N.$$

e representam estimativas do fator sazonal, do nível e da tendência, respectivamente; A , C e D são as constantes de suavização.

2 - Série Sazonal Aditiva

O procedimento anterior pode ser modificado para tratar com situações onde o fator sazonal é aditivo

$$Z_t = \mu_t + T_t + F_t + a_t, \quad t = 1, \dots, N.$$

As estimativas do fator sazonal, nível e tendência da série são dadas por

$$\hat{F}_t = D(Z_t - \bar{Z}_t) + (1-D)\hat{F}_{t-s}, \quad 0 < D < 1 \text{ e } t = s+1, \dots, N.$$

$$\bar{Z}_t = A(Z_t - \hat{F}_{t-s})Z_t + (1-A)(\hat{Z}_{t-1} + \hat{T}_{t-1}), \quad 0 < A < 1 \text{ e } t = s+1, \dots, N.$$

$$\hat{T}_t = C(\bar{Z}_t - \bar{Z}_{t-1}) + (1-C)\hat{T}_{t-1}, \quad 0 < C < 1 \text{ e } t = s+1, \dots, N.$$

respectivamente; A , C e D são as constantes de suavização.

2.1.4.1. Previsão

As previsões dos valores futuros da série para dois procedimentos são dadas a seguir.

1 - Série Sazonal Multiplicativa

$$\hat{Z}_t(h) = (\bar{Z}_t - h\hat{T}_{t-s})\hat{F}_{t+h-s}, \quad h = 1, 2, \dots, s,$$

$$\hat{Z}_t(h) = (\bar{Z}_t - h\hat{T}_{t-s})\hat{F}_{t+h-2s}, \quad h = s+1, 2, \dots, s,$$

e assim por diante.

2 - Série Sazonal Aditiva

$$\hat{Z}_t(h) = \bar{Z}_t + h\hat{T}_{t-s} + \hat{F}_{t+h-s}, \quad h = 1, 2, \dots, s,$$

$$\hat{Z}_t(h) = \bar{Z}_t + h\hat{T}_{t-s} + \hat{F}_{t+h-2s}, \quad h = s+1, 2, \dots, s,$$

e assim por diante.

2.1.5. Modelos e procedimentos de previsão

Os estudos teóricos sobre previsão são todos baseados no método de “mínimos quadrados”, usado para calcular valores futuros a partir de combinações lineares de valores passados.

Um modelo que descreve uma série não conduz, necessariamente, a um procedimento (ou fórmula) de previsão. Será necessário especificar uma função-perda, além do modelo, para se chegar ao procedimento. Uma função-perda, que é utilizada frequentemente, é o erro quadrático médio, embora em algumas ocasiões, outros critérios ou função-perdas sejam mais apropriadas.

Suponhamos que temos observações de uma série temporal até o instante t e queiramos prever o valor da série o instante $t+h$.

Diremos que $\hat{Z}_t(h)$ é a previsão de $Z_t(t+h)$, de origem t e horizonte h . O erro quadrático médio de previsão é:

$$E[Z(t+h) - \hat{Z}_t(h)]^2. \quad (1.11)$$

Então, dado o modelo que descreve a série temporal até o instante t e dado que queremos minimizar erro quadrático médio, obtém-se uma fórmula para $\hat{Z}_t(h)$.

É importante salientar que a previsão não constitui um fim em si, mas apenas um meio de fornecer informações para uma consequente tomada de decisão, visando a determinar objetivos.

2.2. Análise dos Resíduos

Os resíduos encontrados após o modelo para séries temporais ajustado serão controlados a partir do combinado dos gráficos de controle de Shewhart e de MMEP. Para o uso dos gráficos de Shewhart deve ser verificado se os resíduos apresentam distribuição normal, onde alguns dos testes mais conhecidos e utilizados, para atestar esta normalidade, são: Kolmogorov-Smirnov, Anderson-Darling, Shapiro-Wilk, Jarque-Bera, histograma e a curva de probabilidade normal.

Quando a distribuição dos resíduos não seguir a distribuição normal, deve-se fazer uma transformação dos dados originais. Em muitos casos, dados originais não-normais vão gerar previsões também não-normais, cujos erros também não se aproximam suficientemente da distribuição normal. Algumas transformações podem ser: diferenciação, aplicação de logaritmos, diferença de logaritmos, potenciação, transformação de Box-Cox, dentre outras.

Propriedades estatísticas e aplicação direta de outros métodos de previsão justificam estas transformações.

Uma vez que os resíduos apresentem distribuição normal, é possível traçar a curva normal que represente a distribuição da previsão do indicador de arrecadação. No entanto, para se alcançar este nível, é preciso que os processos de previsão sejam bem aplicados e os resultados apresentem-se consistentes, como discutido anteriormente, com seus resíduos possuindo além da distribuição normal, ausência de autocorrelações e de autocorrelações parciais.

2.3. Gráficos de Controle de Shewhart

Para fins de aprofundamento relacionados aos gráficos de controle MMEP os artigos de James M. Lucas e Michael S. Saccucci (1990) e Stephen V. Crowder (1987a, 1989) são indicados.

Lucas e Saccuci (1990) abordam em seu artigo a forma de uso do gráfico de controle MMEP, discutem sobre a eficiência desta metodologia, comparam o gráfico de MMEP com os gráficos de controle de Soma Cumulativa (CUSUM) e de Shewhart, comentam sobre a eficiência do uso combinado do método de MMEP e de Shewhart, e abordam também a questão da robustez do MMEP à Não-normalidade.

Em Crowder (1987a, 1989) pode se observar o detalhamento matemático por traz do gráfico de controle de médias móveis exponencialmente ponderadas.

Os gráficos de controle de Shewhart são usados para monitorar tanto o valor médio das características da qualidade, quanto a sua variabilidade. Para monitorar o valor médio é usado gráfico de controle específico para média, enquanto para monitorar a variabilidade do processo o gráfico usado é o de controle para a amplitude ou para o desvio padrão. Assumindo que numa situação inicial o processo se encontra sob controle com média μ_0 e desvio padrão σ_0 , caso haja deslocamento dessa situação inicial de qualquer uma dessas estatísticas a resultante é uma fração maior da produção fora das especificações e por esse motivo se tem a necessidade de se fazer o controle dessas duas mediadas.

2.3.1. Gráficos de controle para \bar{x} e R

Para a construção do gráfico de controle para a média e para a amplitude, que podem também ser chamados de gráfico de controle para \bar{x} e R, precisamos estimar μ e σ .

Suponha que contamos com m amostras, cada uma com n observações, onde $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m$ são as médias de cada uma das amostras. Então o melhor estimador de μ , a média do processo, é a média geral, isto é (Montgomery, 2009)

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \dots + \bar{x}_m}{m} \quad (1.12)$$

Assim $\bar{\bar{x}}$ deve ser usado como a linha central no gráfico \bar{x} .

Para construir os limites de controle, que funcionam como um intervalo de confiança, é necessário estimar o desvio padrão σ . Para este gráfico especificamente iremos estimar o desvio padrão a partir do método das amplitudes, que consiste em fazer a diferença entre a maior e a menor observação; isto é,

$$R = x_{máx} - x_{mín}$$

Sejam R_1, R_2, \dots, R_m , as amplitudes das m amostras. A amplitude média é:

$$\bar{R} = \frac{R_1 + R_2 + \dots + R_m}{m} \quad (1.13)$$

Podemos então, com as medidas encontradas, construir os limites de controle da seguinte forma:

Limites de Controle para o Gráfico de \bar{x}

$$\text{LSC} = \bar{\bar{x}} + A_2 \bar{R} \quad (1.14)$$

$$\text{Linha Central} = \bar{\bar{x}}$$

$$\text{LIC} = \bar{\bar{x}} - A_2 \bar{R}$$

onde A_2 é constante e encontra-se tabulada para vários tamanhos de amostra na Tabela VI do Apêndice do livro do Montgomery.

A variabilidade do processo pode ser monitorada plotando-se os valores das amplitudes amostrais R em um gráfico de controle. A linha central e os limites de controle para o gráfico R são os seguintes:

Limites de Controle para o Gráfico de R

$$\text{LSC} = D_4 \bar{R} \quad (1.15)$$

$$\text{Linha Central} = \bar{R}$$

$$\text{LIC} = D_3 \bar{R}$$

onde D_3 e D_4 são constantes e encontram-se tabulados para vários tamanhos de n na Tabela VI do Apêndice do livro do Montgomery.

2.3.2. Gráficos de controle para \bar{x} e S

Embora os gráficos \bar{x} e R sejam bastante utilizados, em algumas situações é preferível usar os gráficos de controle \bar{x} e S , onde S é o desvio padrão amostral. Situações essas quando o tamanho da amostra n é grande ou quando o tamanho de n é variável.

A construção dos gráficos de controle \bar{x} e S seguem praticamente a mesma sequência de etapas que dos gráficos de \bar{x} e R , exceto que, para cada amostra, temos que calcular a média amostral \bar{x} e o desvio padrão amostral S .

O desvio padrão amostral S não é um estimador não-viesado para σ , o que ele na verdade estima é $c_4\sigma$, onde c_4 é uma constante que depende do tamanho da amostra n . Além disso, o desvio padrão S é o $\sigma = \sqrt{1 - c_4^2}$. Essa informação pode ser usada para definir os gráficos de controle para \bar{x} e S .

Considere o caso onde o valor de referência é dado para σ . Como $E(S) = c_4\sigma$, a linha central para o gráfico é $c_4\sigma$. Os limites de controle três-sigma são então

$$\text{LSC} = c_4\sigma + 3\sigma\sqrt{1 - c_4^2}$$

$$\text{LIC} = c_4\sigma - 3\sigma\sqrt{1 - c_4^2}$$

Como definimos que S estima $c_4\sigma$ podemos dizer que \bar{S}/c_4 é um estimador não-viesado de σ . Então, os parâmetros para o gráfico S são \bar{S}/c_4

É costume definir as duas constantes

$$B5 = c_4 - 3\sqrt{1 - c_4^2}$$

$$B6 = c_4 + 3\sqrt{1 - c_4^2}$$

Como definimos que S estima $c_4\sigma$ podemos dizer que \bar{S}/c_4 é um estimador não-viesado de σ . Então, os parâmetros para o gráfico S são

$$LSC = \bar{S} + 3\frac{\bar{S}}{c_4}\sqrt{1 - c_4^2}$$

$$\text{Linha Central} = \bar{S}$$

$$LIC = \bar{S} - 3\frac{\bar{S}}{c_4}\sqrt{1 - c_4^2}$$

É usual definir as constantes

$$B3 = 1 - \frac{3}{c_4}\sqrt{1 - c_4^2}$$

e

$$B4 = 1 + \frac{3}{c_4}\sqrt{1 - c_4^2}$$

Consequentemente, obtém-se os parâmetros do gráfico S como

$$LSC = B_4\bar{S}$$

$$\text{Linha Central} = \bar{S}$$

$$LIC = B_3\bar{S}$$

(1.16)

Note que $B_4 = B_6/c_4$ e $B_3 = B_5/c_4$ e que os valores de B_5 e B_6 são tabulados para vários tamanhos de amostra na Tabela VI do Apêndice do livro do Montgomery.

Quando \bar{S}/c_4 é usado para estimar σ , podemos definir os limites de controle para o gráfico \bar{x} correspondente como

$$\text{LSC} = \bar{\bar{x}} + \frac{3\bar{S}}{c_4\sqrt{n}}$$

$$\text{Linha Central} = \bar{\bar{x}}$$

$$\text{LIC} = \bar{\bar{x}} - \frac{3\bar{S}}{c_4\sqrt{n}}$$

Definimos a constante $A_3 = \frac{3}{c_4\sqrt{n}}$, e por fim os parâmetros do gráfico \bar{x} se tornam

Gráfico de controle de Shewhart

$$\text{LSC} = \bar{\bar{x}} + A_3\bar{S} \tag{1.17}$$

$$\text{Linha Central} = \bar{\bar{x}}$$

$$\text{LIC} = \bar{\bar{x}} - A_3\bar{S}$$

As constantes B_3 , B_4 , e A_3 para a construção dos gráficos \bar{x} e S , obtidas através de dados passados, encontram-se listadas na Tabela VI do Apêndice no livro do Montgomery para diversos tamanhos de amostra.

2.3.3. Gráficos de Controle de Shewhart para medidas individuais

Existem situações que o tamanho da mostra consiste em uma única unidade individual, para estes casos é usado o gráfico de controle de Shewhart para medidas individuais.

Os procedimentos para a construção do gráfico de controle para a média são basicamente os mesmos, onde se diferencia simplesmente pelo fato de que, por contarmos com apenas uma observação por amostra, os cálculos dos limites se dão pela média dessas observações individuais, ou seja, em vez de se obter $\bar{\bar{x}}$ calcula-se \bar{x} de cada observação amostrada.

Para controlar a variabilidade do processo construímos um gráfico de controle para a amplitude móvel. Este procedimento consiste no uso da amplitude móvel em duas observações consecutivas e é definida como

$$MR_i = |x_i - x_{i-1}|$$

A obtenção dos limites para gráfico de controle é feita da mesma forma como descrito em 1.14 e 1.15 para o gráfico de controle de Shewhart, com a diferença de que em vez de se usar \bar{R} , para o gráfico de controle de Shewhart para medidas individuais se é utilizada a média de MR, ou seja, \overline{MR} .

2.4. Gráfico de controle de Média Móvel Exponencialmente Ponderada

Para a detecção de pequenas perturbações no processo, o gráfico de controle da média móvel exponencialmente ponderada (MMEP) é altamente recomendável e uma boa alternativa ao gráfico de controle de Shewhart. O MMEP em geral é mais utilizado para controle de amostras com observações individuais e ele será assim descrito abaixo.

O gráfico MMEP é definido como

$$z_i = \lambda x_i + (1 - \lambda)z_{i-1} \quad (1.18)$$

onde $0 < \lambda \leq 1$ é uma constante com função de dar peso para observações anteriores. O valor inicial é o alvo do processo, de modo que

$$z_0 = \mu_0$$

Em alguns casos a média de dados preliminares é usada como o valor inicial do MMEP, de modo que $z_0 = \bar{x}$.

Se as observações são variáveis aleatórias independentes com variância σ^2 , então a variância de z_i é

$$\sigma_{z_i}^2 = \sigma^2 \left(\frac{\lambda}{2 - \lambda} \right) [1 - (1 - \lambda)^{2i}]$$

Podemos então definir a linha central e os limites de controle para o gráfico de controle MMEP da seguinte forma:

O Gráfico de Controle MMEP

$$\text{LSC} = \mu_0 + L\sigma \sqrt{\frac{\lambda}{2 - \lambda} [1 - (1 - \lambda)^{2i}]} \quad (1.19)$$

$$\text{Linha Central} = \mu_0$$

$$\text{LIC} = \mu_0 - L\sigma \sqrt{\frac{\lambda}{2-\lambda} [1 - (1-\lambda)^{2i}]}$$

Note que o termo $[1 - (1-\lambda)^{2i}]$ se aproxima de 1 a medida que i se torna grande fazendo com que os limites superior e inferior de controle sejam a partir de um momento valores fixos.

Para grupos racionais de tamanho $n > 1$, então simplesmente substituímos x_i por \bar{x}_i , e σ por $\sigma_{\bar{x}} = \sigma\sqrt{n}$ nas equações anteriores.

2.5. Combinado: Shewhart e MMEP

Lucas (1973) mostra que os gráficos de controle de Shewhart detectam grandes perturbações no processo de forma mais rápida, já os gráficos MMEP, como visto anteriormente, são bastante recomendados para detectar pequenas perturbações. A combinação dessas duas metodologias oferece, na maioria dos casos, um aperfeiçoamento no controle do processo, uma vez que, tanto em casos de pequenas ou grandes perturbações elas serão detectadas.

O controle combinado dos dois gráficos se resume em fazer o controle para o gráfico de médias moveis exponencialmente ponderadas quanto para o gráfico de Shewhart e esperar o momento que qualquer um dos métodos saiam de controle para que o processo seja parado

2.6. Gráficos de controle para monitorar previsões

Como discutido na seção 1.4, os resíduos de um modelo de previsão devem ser aleatórios, centrados em zero, e tão pequenos quanto possível. Desta forma as previsões terão boa acuracidade e todo padrão possível de ser encontrado na série temporal estará modelado, deixando para os resíduos apenas uma componente aleatória.

Quando houver alguma perturbação ao processo e o indicador começar a se comportar de maneira diferente os resíduos encontrados a partir do modelo proposto também se comportará de forma distinta. Este problema poderá ser detectado a partir da proposição dos gráficos de controle aplicados aos resíduos, de forma que quando um ponto (o resíduo de um período) estiver além dos limites de controle, então houve alguma mudança no processo que deve ser verificado, pois o modelo não está sendo mais adequado aos dados.

Assim, a combinação dos gráficos de Shewhart e MMEP torna-se válida, pois se houver um resíduo grande em um único período, há boas chances de ele ser detectado pelo gráfico de Shewhart, enquanto que para um pequeno padrão constante nos resíduos o gráfico MMEP mostra-se eficaz para sua detecção. Logo, a combinação das duas teorias apresentasse como a alternativa natural para monitorar este processo.

Em Atienza et al. (1997) é realizado uma comparação entre algumas etapas dos processos de previsão e de controle estatístico com uso de gráficos de controle. Enquanto o modelo de previsão realiza a estimação de parâmetros para a equação de previsão, o gráfico de controle calcula os limites de controles. Para o controle estatístico de processos, quando um ponto é localizado além dos limites de controle considera-se que o processo está instável estatisticamente que, portanto ele deve ser parado para se observar o motivo da instabilidade do processo.

Capítulo 3

3. Desenvolvimento

3.1. Seleção do Indicador da SEFAZ/PI

A principal influência para a formulação do indicador a ser controlado foi o impacto que ele poderia causar nos principais objetivos da Secretaria de Fazenda do Piauí perante a sociedade. A partir, então, da missão da SEFAZ-PI que se compromete em “*Promover, com alto desempenho, as atividades de tributação, fiscalização e arrecadação, bem como a gestão da política financeira e contábil do Estado, primando pela excelência no atendimento e contribuindo para a elevação da consciência fiscal e do desenvolvimento do Piauí*” (Site SEFAZ-PI), foi considerado de suma importância a geração do indicador “arrecadação” para que este seja analisado e acompanhado.

A arrecadação da Secretaria de Fazenda do Piauí pode ser realizada de duas formas diferentes e seguem o seguinte processo:

a) O Contribuinte gerando o DAR – Documento de Arrecadação

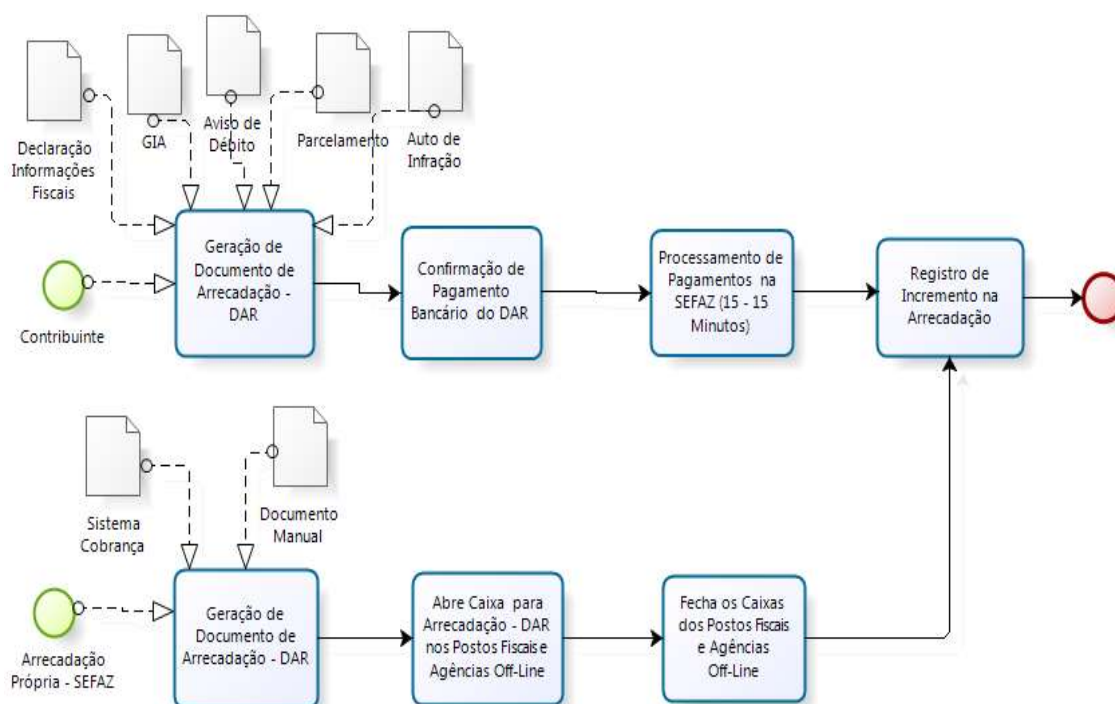
- As Aplicações como Declaração de Informações Fiscais, realizada mensalmente pelos contribuintes, GIA (Declaração para Contribuintes de Outros Estados, Aviso de Débito, Parcelamento e Auto de Infração, geram o DAR) Documento de Arrecadação para o Contribuinte.
- O contribuinte faz o pagamento desta DAR na rede bancária.
- A rede bancária envia a confirmação pagamento eletrônico à SEFAZ.
- Na SEFAZ, de forma automatizada, o Sistema Integrado faz o recebimento e processamento dos arquivos de arrecadação de 15 em 15 minutos.
- Após processados os arquivos de confirmação, ocorre o incremento da arrecadação com os pagamentos efetuados.

b) Os Postos Fiscais e Agencias Off-Line da SEFAZ gerando o DAR – Documento de Arrecadação

- Nos Postos Fiscais ou Agências Off-Line é gerado o DAR – Documento de Arrecadação para o Contribuinte via Sistema de Cobrança ou via Documento Manual.
- Uma aplicação de Caixa é utilizada para o recebimento desta arrecadação, sendo um caixa aberto diariamente.
- Todos os caixas abertos devem ser fechados diariamente.
- Os registros de arrecadação recebidos são incrementados a arrecadação

O esquema a seguir resume como está estruturado o processo de arrecadação feita pela a SEFAZ-PI.

Figura 1 – Esquema de arrecadação



Fonte: Secretaria de Fazenda do Estado do Piauí

Essas informações são armazenadas no Sistema Integrado de Administração Tributária, que é atualizado com as novas arrecadações de 15 em 15 minutos. É desse sistema que as variáveis ditas como essenciais para a composição do indicador “arrecadação” são geradas, sendo elas:

Tabela 2 – Variáveis para construção do indicador

Campo	Descrição
PERIODO	Período que houve a arrecadação Exemplo: 022012 (Fevereiro 2012)
DATA_ARREC	Data da arrecadação
TIPO_TRIBUTO	Tipo de receita arrecadada Exemplo: ICMS - Antecipação Parcial
TOT_PAGO	Total arrecadado
CONTRIBUINTE	Quantidade de contribuintes

Fonte: Secretaria de Fazenda do Estado do Piauí

Devido à variação mensal do número de contribuintes, o indicador selecionado para análise foi a arrecadação média mensal realizada pela SEFAZ-PI.

Para se chegar a quantidade de contribuintes real responsáveis pela arrecadação da SEFAZ – PI em determinado período, foram consideradas a quantidade de contribuintes “inscritos” na Secretaria e a quantidade de contribuintes “não inscritos”.

As situações de contribuintes “inscritos” são descritas no Decreto *No.13.500, de 23 de Dezembro de 2008 - Regulamento de ICMS do Estado do Piauí* no Artigo 202:

Art. 202. São obrigados á inscrição no Cadastro de Contribuintes do Estado do Piauí antes de iniciarem suas atividades:

I - o comerciante, o industrial, o produtor e o extrator de mercadorias;

II - prestador de serviços de transportes interestadual e intermunicipal e de comunicação;

III - a cooperativa;

IV - a instituição financeira e a seguradora;

V - a empresa de construção civil ou similar;

VI - a concessionária ou permissionária de serviço público de transporte de comunicação e de energia elétrica;

VII - o prestador de serviços não compreendidos na competência tributária dos municípios, que envolvam fornecimento de mercadorias;

VIII - o prestador de serviços compreendidos na competência tributária dos municípios, que envolvam fornecimento de mercadorias ressalvadas em Lei Complementar;

IX - o armazém geral e congênere;

X - as demais pessoas naturais ou jurídicos de direito público ou privado, que participarem habitualmente em nome próprio ou de terceiros de operações relativas à circulação de mercadorias ou prestação de serviços conforme definido neste regulamento;

XI - os Leiloeiros Oficiais, observado o disposto nos arts. 821 a 829;

§1o. Todo aquele que produzir em propriedade alheia e promover a saída de produtos em seu próprio nome fica também obrigado à inscrição;

Os contribuintes “não inscritos” são aqueles sem inscrição estadual na SEFAZ-PI, mas que participam da arrecadação da Secretaria de forma não regular em algumas transações como:

emissão de Nota Fiscal avulsa, pagamento de taxas DETRAN e outras taxas públicas. Portanto, esses contribuintes também devem ser considerados para o cálculo do indicador.

Para a construção deste indicador divide-se a quantidade arrecadada pelo estado no mês, pela quantidade de contribuintes que foram responsáveis pela a arrecadação de Piauí naquele mês.

$$\textit{Arrecadação Média} = \frac{\textit{Arrecadação mensal do Estado do Piauí}}{\textit{Contribuinte resposnsavei pela a arrecadação}}$$

3.2. Identificação do modelo de previsão

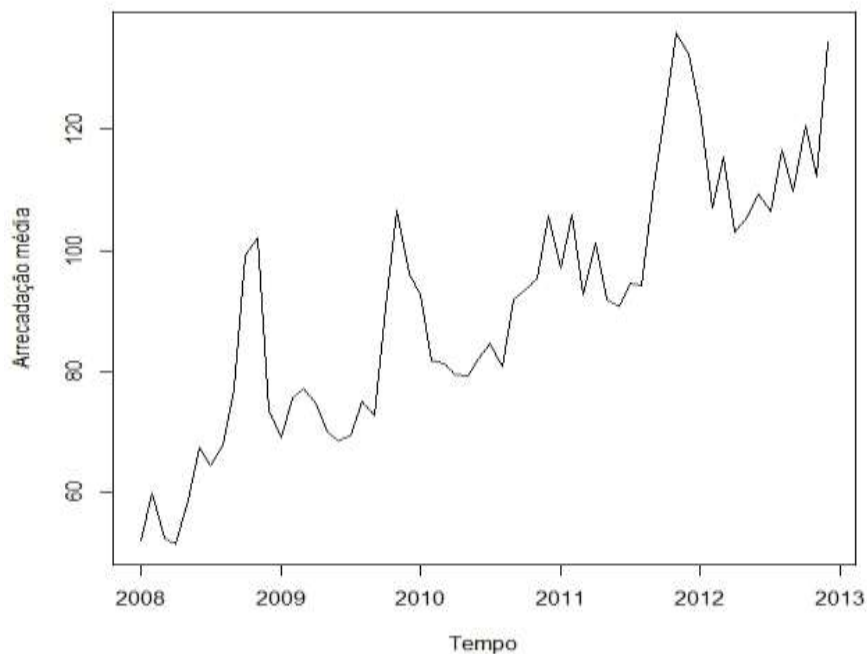
Esta seção apresenta os passos realizados para a identificação do modelo de série temporal ajustado. Esta etapa é de suma importância, pois da diferença entre a série observada e o modelo ajustado são gerados os resíduos que serão controlados para acompanhamento do processo.

3.2.1. Análise exploratória dos dados

O primeiro passo realizado foi uma análise descritiva da série para se observar suas características e, assim, se ter uma ideia inicial de quais modelos seriam possíveis candidatos para o ajuste.

O gráfico a seguir apresenta a série temporal do indicador arrecadação mensal por contribuinte, de 2008 ao final de 2012.

Gráfico 1 : Arrecadação por contribuinte



Fonte: Secretaria de Fazenda do Estado do Piauí

Percebe-se que a série apresenta uma tendência de crescimento no decorrer dos anos e um comportamento que pode representar a presença de sazonalidade.

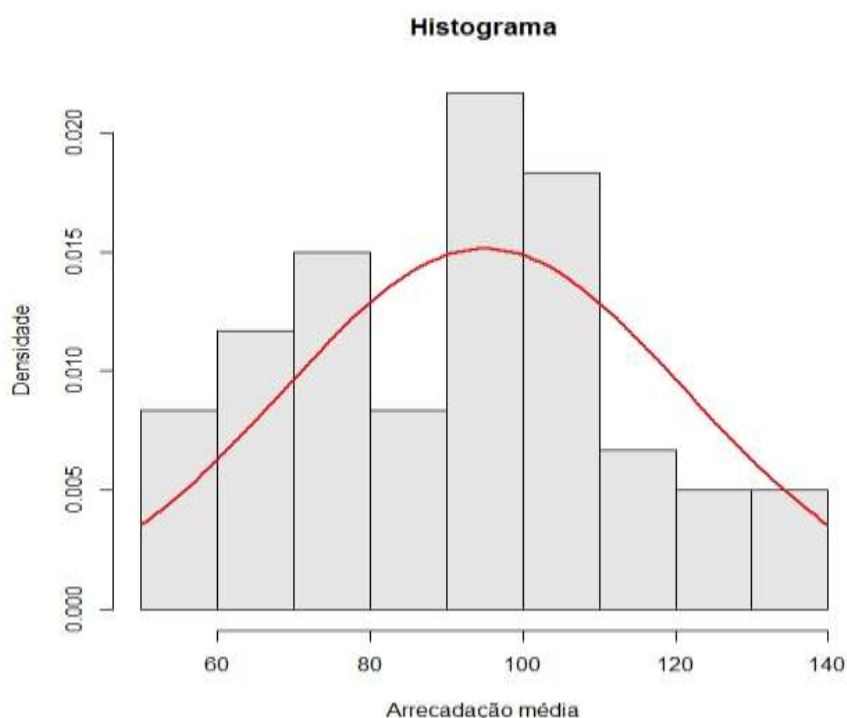
Tabela 3 : Arrecadação média do Estado do Piauí

	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
2008	51,94	59,75	52,41	51,37	58,65	67,34	64,35	67,61	77,02	99,22	102,11	73,37
2009	68,94	75,48	77,02	74,63	69,97	68,43	69,32	74,92	72,61	90,44	106,64	96,17
2010	92,56	81,65	81,23	79,34	79,16	82,12	84,47	80,75	91,91	93,49	95,23	105,71
2011	97,05	105,89	92,54	101,30	91,93	90,72	94,54	94,16	110,93	122,89	136,02	132,61
2012	123,41	106,92	115,43	103,04	105,43	109,38	106,35	116,70	109,74	120,62	112,20	134,52

Fonte: Secretaria de Fazenda do Estado do Piauí

A seguir, apresenta-se o histograma da série da arrecadação por contribuinte, que expõe a frequência das arrecadação por contribuinte em reais.

Gráfico 2 : Histograma de arrecadação por contribuinte

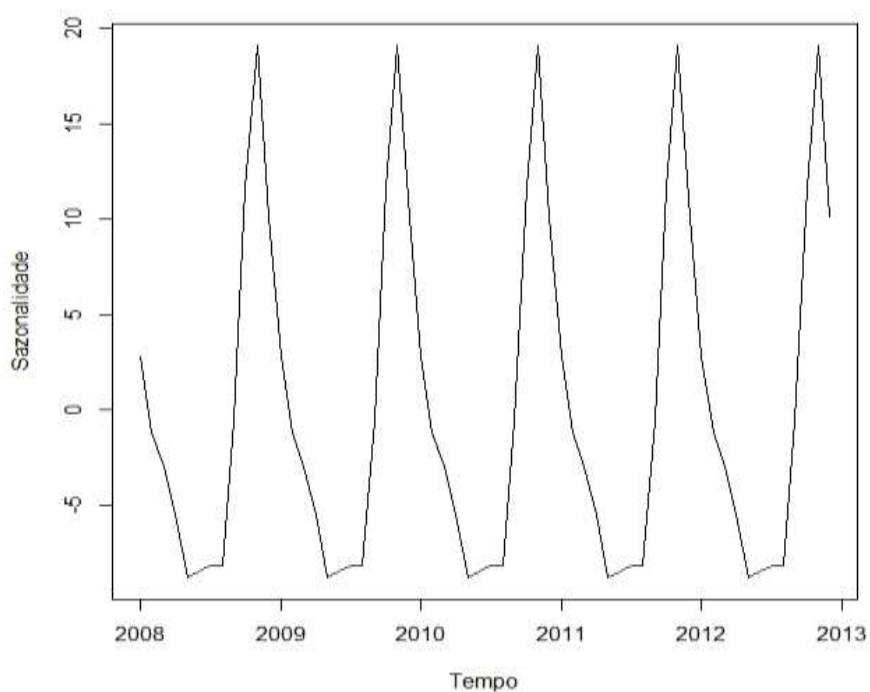


A partir da visualização do gráfico e tendo o valor do coeficiente de assimetria igual a 0,132, pode-se afirmar que a distribuição apresenta uma boa simetria com uma leve tendência

à direita. O coeficiente de curtose é de $-0,576$, sugerindo que a distribuição de “arrecadação por contribuinte” é platicúrtica. Fica clara a não normalidade nos dados.

O gráfico a seguir apresenta a componente sazonal da arrecadação média por contribuinte.

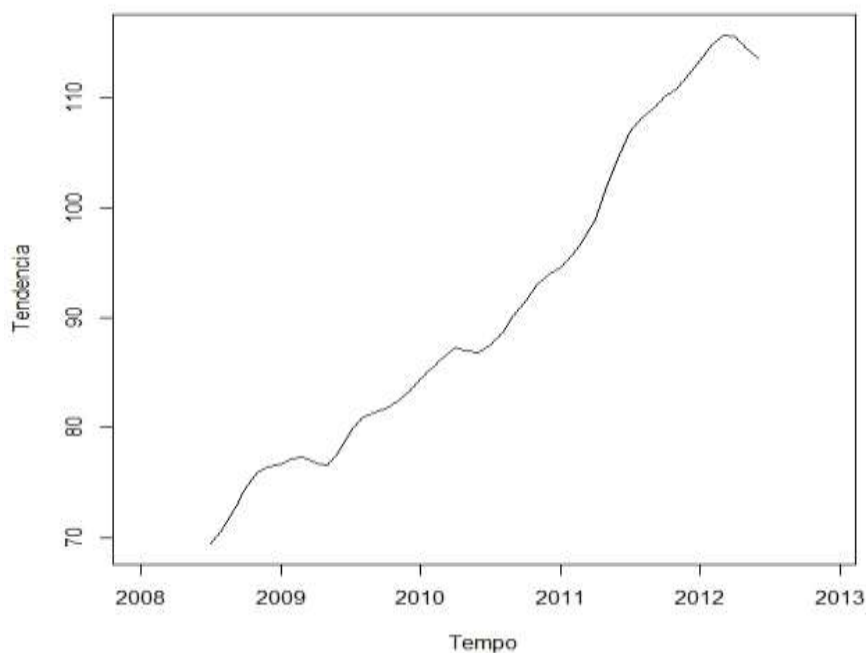
Gráfico 3 : Componente sazonal da arrecadação média da SEFAZ-PI



O Gráfico 3 mostra claramente a presença de sazonalidade na série, com um comportamento homogêneo e estável em todos os anos da série.

O Gráfico 4 apresenta a componente de tendência da arrecadação média por contribuinte do Estado do Piauí.

Gráfico 4 : Componente de tendência da arrecadação média da SEFAZ-PI



Fonte: Secretaria de Fazenda do Estado do Piauí

Pode-se perceber que a série apresenta uma clara tendência de crescimento no período observado.

A partir desses resultados, decidiu-se por realizar uma modelagem de suavização exponencial capaz de captar a sazonalidade e a tendência apresentada pelo indicador.

3.2.2. Modelagem de suavização exponencial de Holt-Winters

Um modelo que parece ser razoável para a série em estudo é o alisamento exponencial proposto por Holt-Winters, visto que considera a componente sazonal, na forma multiplicativa ou aditiva. A seguir são apresentados os coeficientes para estas duas metodologias.

Tabela 4 – Coeficientes otimizados

	alpha	beta	gamma
H.W. (aditivo)	0,20	0,00	0,62
H.W. (mult.)	0,17	0,00	0,67

Percebe-se uma pequena diferença entre os coeficientes estimados para os modelos aditivo e multiplicativo, cujos Gráficos 5 e 6 apresentam os ajustamentos realizados para a arrecadação média por contribuinte.

Gráfico 5 : Ajustamento Holt-Winters “aditivo”

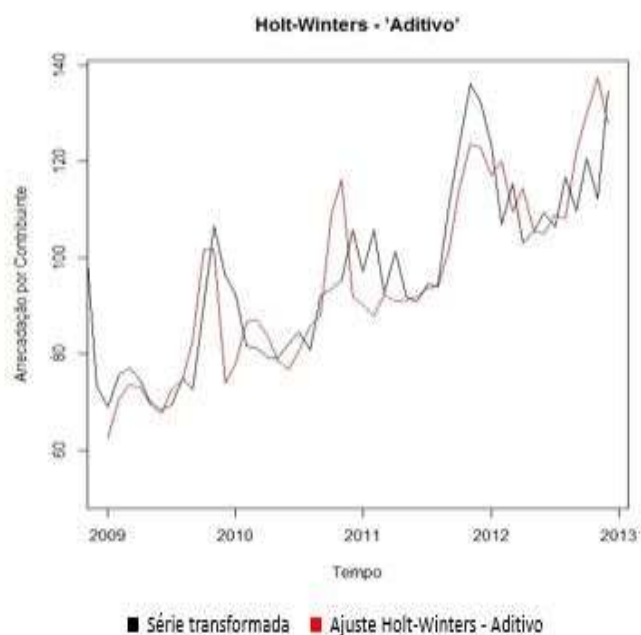
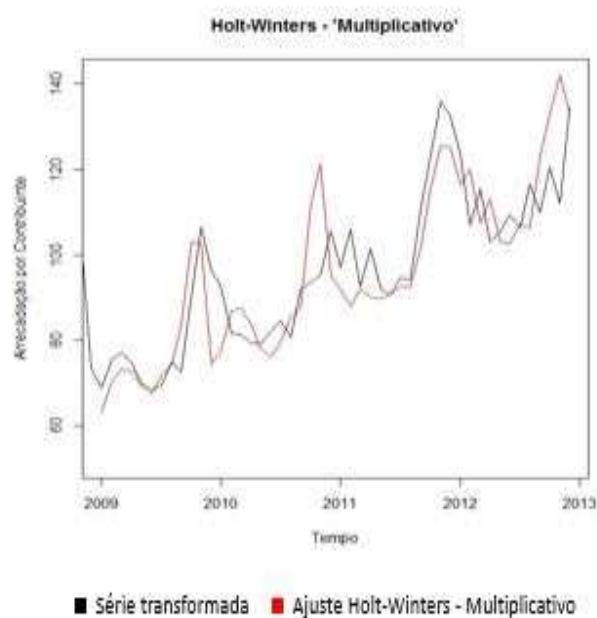


Gráfico 6 : Ajustamento Holt-Winters “multiplicativo”



A comparação visual revela uma diferença mínima entre os modelos propostos, e que os modelos não estão bem ajustados aos dados, o que motivou a consideração de um modelo ARIMA.

3.2.3. Identificação de modelos ARIMA

As etapas a serem seguidas para a identificação do melhor modelo ARIMA para série original são especificadas pela metodologia proposta por Box e Jenkins. A primeira etapa baseia-se no comportamento das funções de autocorrelação (FAC) e autocorrelação parcial (FACP) dos dados. Os Gráficos 7 e 8 apresentam a FAC e a FACP da ARRECADAÇÃO POR CONTRIBUINTE respectivamente.

Gráfico 7 : FAC da arrecadação por contribuinte

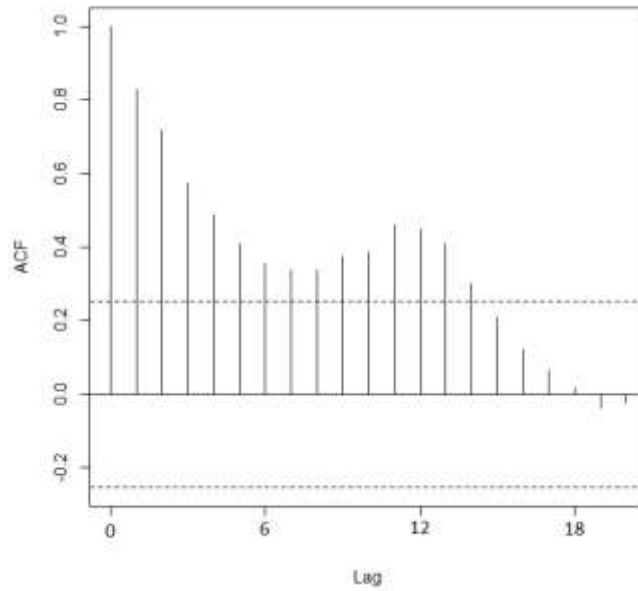
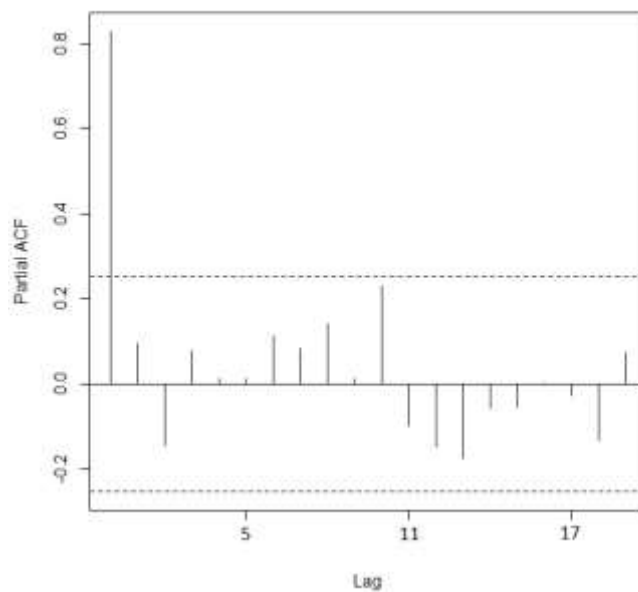
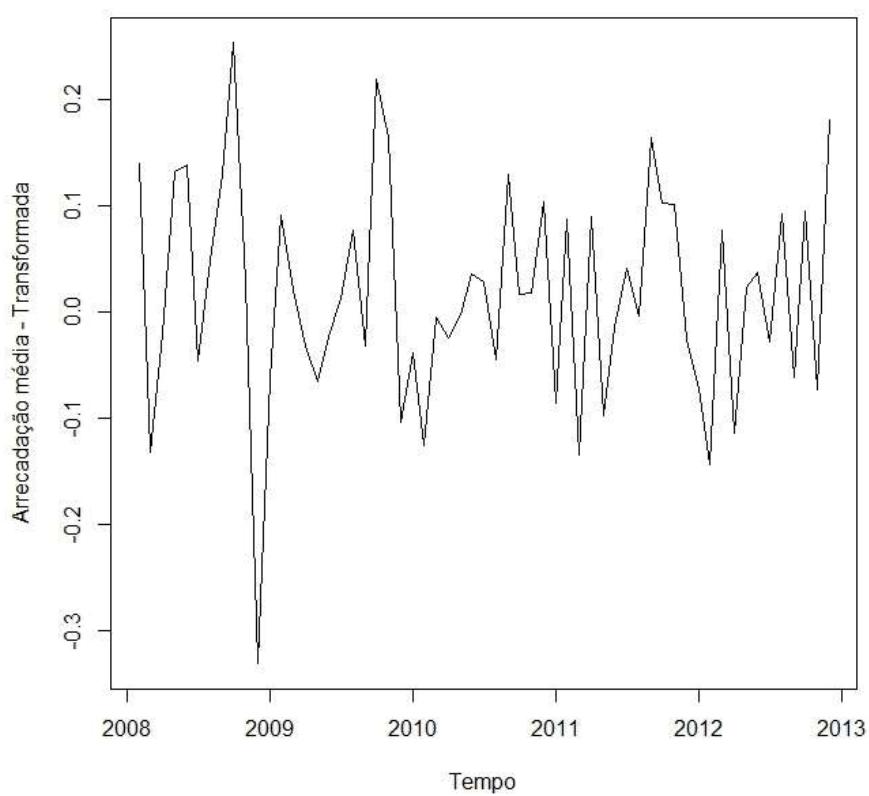


Gráfico 8 : FACP da arrecadação por contribuinte



Para alcançar a condição de estacionaridade para ajuste do modelo ARIMA, tomou-se a 1ª. diferença da série. O Gráfico 9 apresenta o modelo ajustado.

Gráfico 9 : Arrecadação média transformada



As funções de autocorrelação e de autocorrelação parcial, como dito anteriormente, ajudam na escolha do modelo ARIMA a ser selecionado para se ajustar a série. Seguem, então, os Gráficos 10 e 11 apresentam a FAC e a FACP da série transformada, respectivamente.

Gráfico 10 : FAC da arrecadação média tranformada

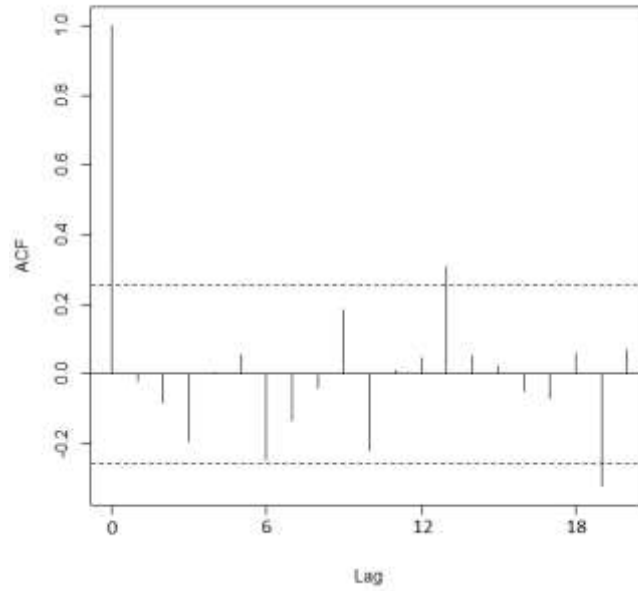
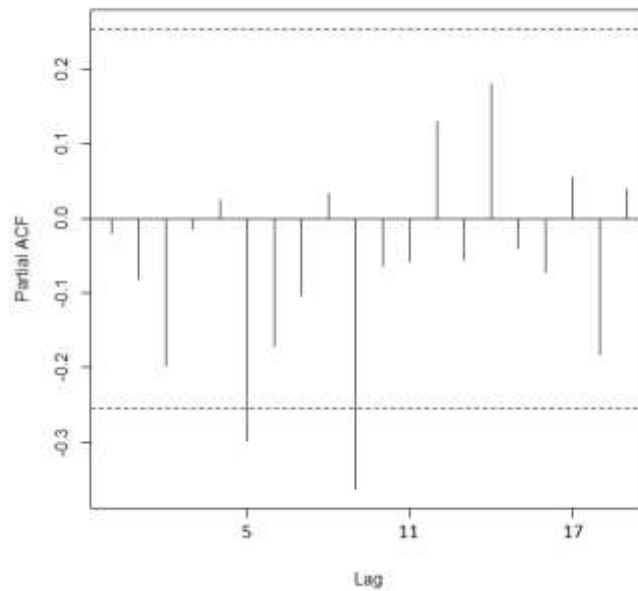


Gráfico 11 : FACP da arrecadação média tranformada



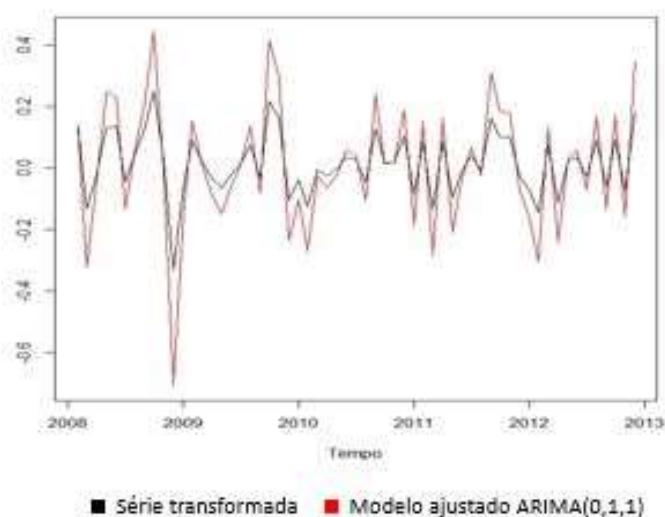
A FAC e a FACP indica que o modelo de médias móveis parece ser mais adequado para descrever o comportamento da série. Uma forma de confirmar esta afirmação é ajustando alguns modelos ARIMA e comparando suas informações de AIC, AICc e BIC, onde os menores deles recomendam o melhor modelo ajuste.

Tabela 5 : Comparação AIC, AICc e BIC

	AIC	AICc	BIC
ARIMA(1,1,0)	-67,86	-67,64	-63,74
ARIMA(2,1,0)	-69,30	-68,86	-63,12
ARIMA(0,1,1)	-90,24	-90,02	-86,12
ARIMA(0,1,2)	-88,24	-87,80	-82,06
ARIMA(1,1,1)	-88,25	-87,90	-82,07

Os menores valores de AIC, AICc e BIC retratam qual o modelo que melhor se ajustou aos dados, portanto pode-se afirmar que o modelo que deve ser escolhido é o ARIMA(0,1,1). O Gráfico 12 apresenta o modelo ajustado à série em estudo.

Gráfico 12 : Modelo Ajustado ARIMA(0,1,1) para a Arrecadação Média da SEFAZ-PI



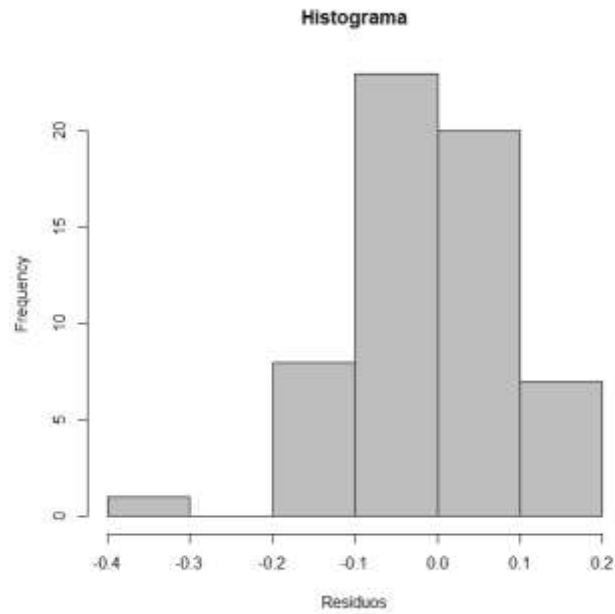
Visualmente o modelo estimado parece satisfatório. Para confirmar que o modelo foi bem ajustado, faz-se o teste de Ljung-Box e o teste de normalidade para os resíduos do modelo. A Tabela 4 apresenta os resultados do teste de Ljung-Box, que foi realizado considerando-se o lag=1, pois foi esta a ordem do modelo considerado.

Tabela 6 : Teste Ljung-Box

Qui-quadrado = 5e-04	gl = 1	p-valor = 0,9828
----------------------	--------	------------------

A normalidade dos resíduos, condição necessária para o controle do gráfico de Shewhart, foi verificada pelo teste de Shapiro-Wilk.

Gráfico 13 : Histograma dos Resíduos



O histograma dos resíduos apresenta um comportamento de normalidade destes dados, hipótese confirmada a partir do teste de Shapiro-Wilk, cuja hipótese nula remete à não normalidade dos resíduos.

Shapiro-Wilk

W = 0.9679 p-valor = 0.1214

Os resultados apresentados na Tabela 5 confirmam a normalidade dos resíduos, o que atesta, portanto, que os resíduos reúnem condições necessárias para serem controlados pelo gráfico de Shewart.

3.3. Preposição dos gráficos de controle

Foram propostas duas alternativas para a realização do controle da arrecadação média da Secretaria da Fazenda do Estado do Piauí. Primeiramente, é apresentada o controle estatístico dos resíduos gerados a partir do modelo de série temporal estimado para a série em questão. A segunda, apresentada como forma alternativa, de caráter operacional, para o controle da arrecadação média que não depende do ajuste do modelo de séries temporais para o indicador.

3.3.1. Controle dos resíduos

Os resíduos gerados a partir do modelo ARIMA(0,1,1) ajustado à série temporal da arrecadação média da SEFAZ – PI são apresentados na Tabela 5.

Tabela 8 – Resíduos da série ajustada à arrecadação média da SEFAZ-PI, 2008-13

	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
2008		0,00	-0,19	-0,02	0,12	0,10	-0,09	0,01	0,09	0,19	-0,04	-0,38
2009	-0,09	0,06	-0,01	-0,06	-0,09	-0,04	0,00	0,06	-0,05	0,20	0,14	-0,13
2010	-0,06	-0,15	-0,02	-0,04	-0,02	0,02	0,01	-0,06	0,11	0,00	0,00	0,09
2011	-0,10	0,07	-0,15	0,07	-0,11	-0,03	0,03	-0,02	0,15	0,08	0,08	-0,05
2012	-0,09	-0,16	0,06	-0,13	0,01	0,02	-0,04	0,08	-0,08	0,08	-0,09	0,17

Fonte: Secretaria de Fazenda do Estado do Piauí

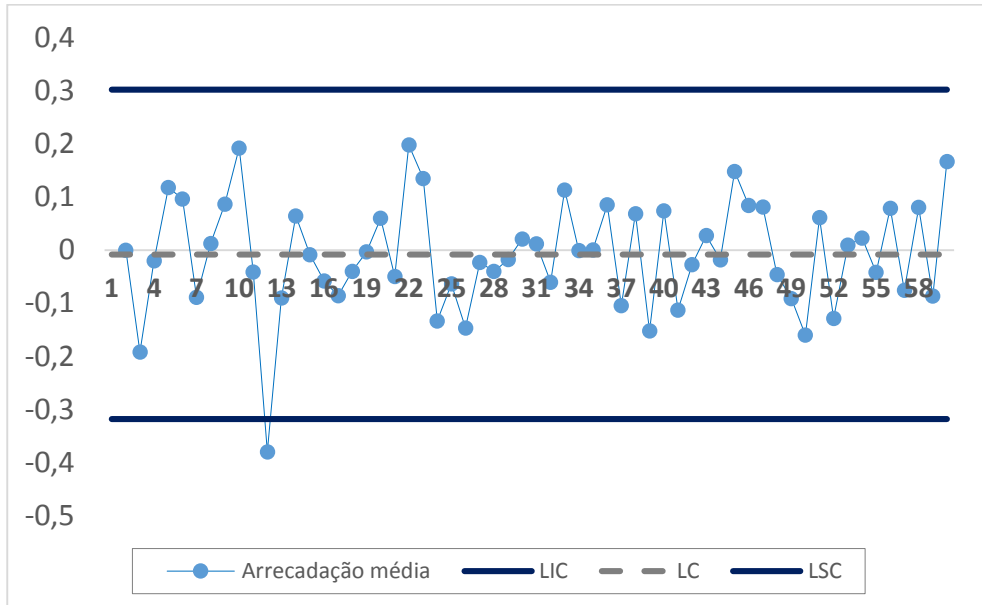
O controle será realizado a partir da combinação do gráfico de Shewhart e de MMEP. Caso seja registrado, em qualquer dos gráficos, alguma observação estiver fora dos limites de controle, a SEFAZ-PI deverá investigar a causa do fato gerador.

O primeiro gráfico a ser construído é o de Shewhart. Para a realização do controle a deve-se calcular a média dos dados e seu desvio padrão. Assim, para os resíduos tem-se como média e desvio padrão -0,0081 e 0,1033, respectivamente. A partir destas estatísticas são construídos o limite superior e o limite inferior do gráfico, obtidos a partir da soma da média a 3 desvios para o LS e da diminuição da média a 3 desvios para o LI. O limite central (LC) é a própria média do processo.

Tabela 9 : Limites de controle de Shewart para os resíduos do modelo ajustado

LIC	LC	LSC
-0,382	-0,0082	0,301

Gráfico 14 : Gráfico de controle Shewhart



Fonte: Secretaria de Fazenda do Estado do Piauí

Para o controle combinado, analisa-se paralelamente o gráfico de MMEP com $\mu_0 = 0$, $\lambda = 0,1$ ou $0,2$ e $L = 2,7$, conforme recomenda Montgomery (2009) e as fórmulas utilizadas para a construção dos limites estão em (1.19).

Gráfico 15 : Gráfico de controle “MMEP” ($\lambda = 0,1$ e $L = 2,7$)

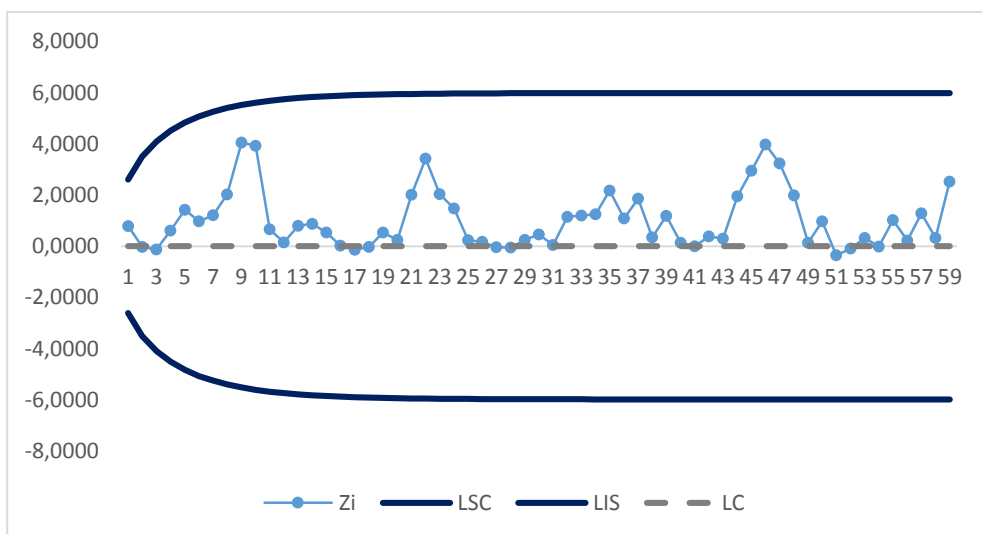
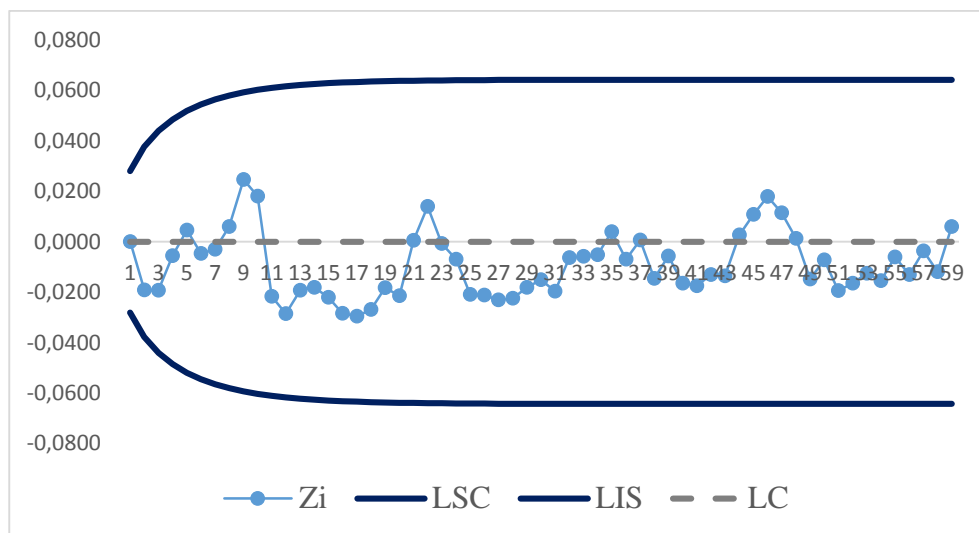


Gráfico 16 : Gráfico de controle “MMEP” ($\lambda = 0,2$ e $L = 2,7$)



No acompanhamento dos dois gráficos de forma combinada, pode-se perceber que somente no gráfico de Shewhart ocorre uma situação fora de controle. Ao analisar a causa deste ponto discrepante, percebe-se que a arrecadação média da SEFAZ – PI apresentava um crescimento de forma estável no decorrer do ano de 2008, quando no mês de dezembro verificou-se uma queda considerável, resultando num resíduo discrepante o suficiente para ser detectado como um ponto fora de controle no gráfico de Shewhart, o que remete a uma análise para este período no intuito de detectar qual foi o motivo para esta queda da arrecadação média da SEFAZ-PI.

3.3.2. Controle da série transformada

Existe a possibilidade de se fazer o controle da arrecadação média, sem a necessidade de estimação do modelo de séries temporais. Nesta situação o controle deve ser realizado a partir da simples diferenciação da série original. A vantagem desta alternativa é a possibilidade de se realizar o controle estatístico de uma forma menos trabalhosa no ponto de vista matemático o que, com certeza, facilitaria a implementação do controle estatístico de processo à Secretaria. Por outro lado, não seria estimado um modelo ao indicador estratégico, perdendo

assim todos os benefícios que um modelo estimado pode proporcionar como, por exemplo, a previsão de valores futuros. A Tabela 7 expõe os dados obtidos a partir da série diferenciada.

Tabela 10 – Série diferenciada da arrecadação média

	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
2008		7,81	-7,34	-1,04	7,27	8,69	-2,99	3,26	9,40	22,20	2,90	-28,75
2009	-4,43	6,55	1,54	-2,40	-4,66	-1,54	0,89	5,60	-2,31	17,83	16,20	-10,47
2010	-3,61	-10,90	-0,42	-1,89	-0,18	2,96	2,35	-3,72	11,16	1,58	1,74	10,48
2011	-8,65	8,84	-13,36	8,76	-9,37	-1,21	3,82	-0,38	16,77	11,96	13,13	-3,41
2012	-9,19	-16,50	8,51	-12,39	2,38	3,95	-3,03	10,34	-6,96	10,88	-8,42	22,32

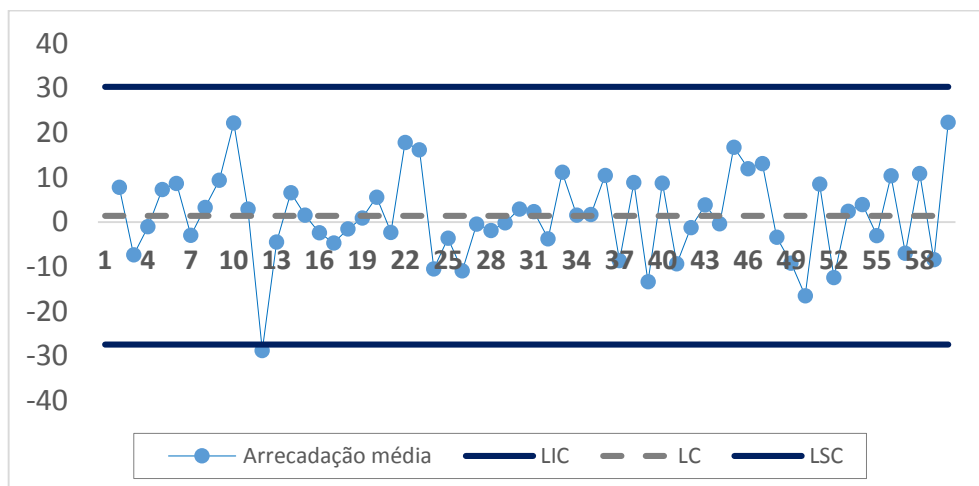
Fonte: Secretaria de Fazenda do Estado do Piauí

Para a proposição do gráfico de controle de Shewhart estimou-se a média e o desvio padrão da série, que são 1,4 e 9,62, respectivamente. O cálculo dos limites de controle foram calculados a partir da fórmula (1.17), onde a constante A_3 é igual a 3 pela convenção utilizada 3σ .

Tabela 11 :Limites de controle de Shewart da arrecadação média diferenciada

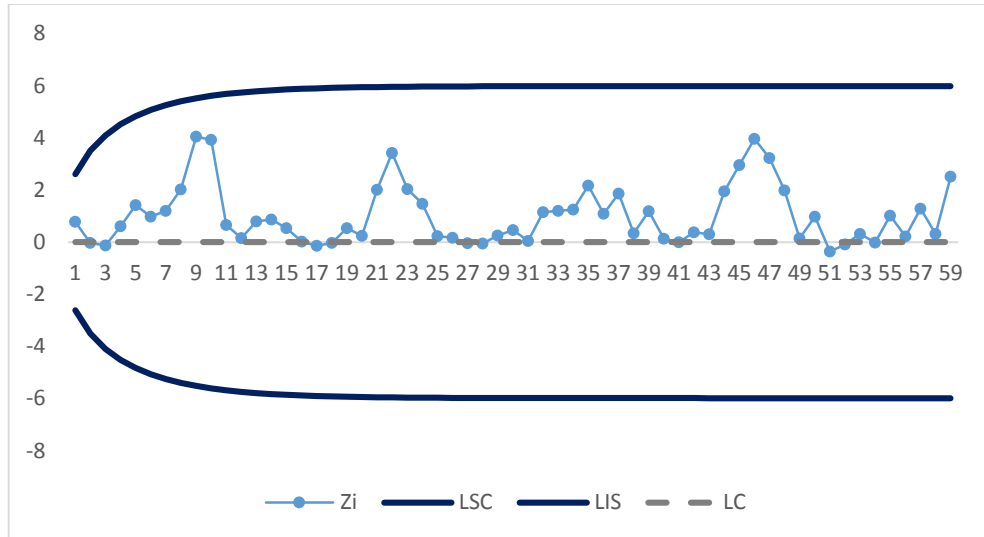
LIC	LC	LSC
-27,45	1,399	30,25

Gráfico 17 : Gráfico de Controle Shewhart



Para o controle combinado, analisa-se paralelamente o gráfico de MMEP, cuja linha central é baseada em $\mu_0 = 0$. Para a parametrização do gráfico Montgomery (2009) recomenda: $\lambda = 0,1$ ou $0,2$ e $L = 2,7$. As fórmulas utilizadas para a construção dos limites estão em (1.19).

Gráfico 18 : Gráfico de controle “MMEP” ($\lambda = 0,1$ e $L = 2,7$)



A alternativa de se controlar a diferença da série original de arrecadação média apresentou resultados similares aos obtidos pelo uso do controle dos resíduos a partir do modelo estimado. Importante observar que o ponto fora de controle se apresentou mais próximo do limite de controle do que a alternativa apresentada anteriormente. Portanto poderia acontecer deste tipo de controle não detectar, em uma situação de perturbação, onde na verdade pode sim ter havido.

3.4. Implantação dos gráficos de controle

Apresenta-se a seguir alguns passos que a SEFAZ-PI precisa seguir para o sucesso da implantação dos gráficos de controle em seus processos (Montgomery, 2009).

- Determinação de quais características dos processos controlar;
- Determinação de onde, no processo, os gráficos devem ser implementados;
- Escolha do tipo próprio de gráfico de controle;
- Seleção dos sistemas de coleta de dados e dos programas de computador;
- Ações de melhoria dos processos como resultado do CEP/análise do gráfico de controle.

Capítulo 4

4. Conclusão

A Secretaria de Estado da Fazenda do Piauí vem promovendo o acompanhamento de seus objetivos e a avaliação de seus serviços oferecidos no intuito de verificar se os objetivos da organização estão sendo atingidos.

Este trabalho ajustou modelos de séries temporais para um dos indicadores estratégicos da SEFAZ/PI e propôs o uso combinado dos gráficos de controle de Shewhart e Média Móvel Exponencialmente Ponderada (MMEP) para os resíduos do modelo ajustado.

A seleção do indicador partiu do estudo dos objetivos estratégicos da SEFAZ-PI, que tem como missão promover a arrecadação tributária no estado.

O indicador proposto para controle foi a Arrecadação Média, que consiste na arrecadação total pela quantidade de contribuintes. Os dados utilizados para análise têm início em 2008, ano que o Sistema Integrado de Administração Tributário foi implementado.

Na análise inicial do indicador, percebeu-se um comportamento de tendência de crescimento da arrecadação média, o que tornaria inviável o controle estatístico. A solução para este problema foi a estimação de modelos de séries temporais a este indicador no intuito de gerar resíduos com propriedades que tornariam o controle estatístico possível.

Por apresentar tendência e uma aparente sazonalidade, o primeiro modelo ajustado foi o alisamento exponencial de Holt-Winters. Este modelo não se adequou bem a série gerando resíduos com um comportamento que impossibilitou o controle estatístico. Foram então testados os modelos ARIMA, propostos por Box & Jenkins. A escolha do modelo ocorreu após análise das funções de auto correlação, auto correlação parcial e das informações AIC, AICc e BIC, cujo modelo selecionado foi o ARIMA(0,1,1).

Os resíduos da série original com o modelo ajustado apresentaram as propriedades necessárias para a realização do controle estatístico, ou seja, eram estacionários e seguiam a

distribuição de probabilidade da normal. Para a verificação destas propriedades foram utilizados o Teste de Ljung-Box e de Shapiro Wilk.

Foi também proposta uma metodologia mais simples para a obtenção dos dados para controle. Esta alternativa consistiu a preposição dos gráficos de controle para os dados da série de arrecadação média diferenciada, ou seja, realizou-se a diferença de uma observação com a sua anterior para obter uma série que apresentasse as condições necessárias para controle. Apesar de ser mais simples e de mais fácil implementação quando comparada à primeira proposta, deve-se entender que a estimação de um modelo para a série traz inúmeros benefícios (como previsões), embora maior complexidade envolvida em sua análise.

Para as duas alternativas propostas foram submetidos aos gráficos de controle Shewhart e de MMEP. Os resultados revelaram um ponto fora de controle e ao analisar os dados originais pode-se perceber um comportamento anormal da série naquele período (dezembro/2008), o que indica que de haver necessidade de um estudo mais minucioso da arrecadação no dado período.

Referência Bibliográfica

- ATIENZA, O. O., ANG, B. W., TANG, A. C. Statistical process control and Forecasting, **International Journal of Quality Science**, v. 2 n. 1, 1997.
- BONILLA José A. **Gestão da Qualidade Total para Produtores de Café**. Universidade Federal de Lavras. 183 p, 1999.
- BROCKWELL, P.J. e Davis, R.A. **Introduction to Time Series and Forecasting**. New York: Springer-Verlag. 1996.
- BURNHAM, K. P.; ANDERSON, D. R. **Multi model inference: understanding aic and bic in model selection**. Sociological Methods and Research. Beverly Hills, v.33, n.2, p.261-304, 2004.
- DEMING, W. E. **Qualidade: A Revolução da Administração**. Rio de Janeiro: Marques Saraiva, 1990.
- FAVA, V. L.: **Análise de Séries de Tempo**. São Paulo: Ed. Atlas, 2000
- GOMES, Paulo J. P. A evolução do conceito de qualidade: dos bens manufacturados aos serviços de informação, **Cadernos BAD**. Lisboa: APBAD, v. 2, p. 6-18, 2004.
- GRANGER, C.W.J. e Newbold, P Forecasting Economic Time Series. **New York: Academic Press**. 1986.
- ISHIKAWA, Kaoru, **Controle da qualidade total: à maneira japonesa**, tradução Iliana Torres, 2ª Ed., Rio de Janeiro: Campus, 193, 1915.
- LUCAS, J. M. e SACCUCI, M. S. Exponentially Weighted Moving Average Control Schemes. Properties and Enhancements. **Technometrics** 32, 1-12, 1990.
- LUCAS, J.M. - A Modified "V" Mascara Control Scheme, **Technometrics**, 15(4) p.833-847, 1973.
- MATOS, O. C. **Econometria básica: teoria e aplicações**. São Paulo: Atlas, 2000
- MONTGOMERY, Douglas C. **Introdução ao Controle Estatístico de Qualidade**, tradução Ana Maria Lima de Farias, Vera Regina Lima de Farias e Flores; revisão técnica Luiz da Costa Laurencel. 4ª Ed., Rio de Janeiro: LTC, 2009.
- MORETTIN, P. A. e Toloi, C. M. C. **Análise de Séries Temporais**. São Paulo: Edgar Blücher, 2006.
- S.V. CROWDER A simple method for studying run-length distributions of exponentially weighted moving average charts. **Technometrics**, 29, pp. 401–407, 1987.

Apêndice 1: Os Quatorze Pontos do Método Deming

- 1º: Estabelecer a Constância de Finalidade para Melhorar o Produto e o Serviço;
- 2º: Adotar a Nova Filosofia;
- 3º: Acabar com a Dependência da Inspeção em Massa;
- 4º: Cessar a Prática de Avaliar as Transações Apenas com Base no Preço;
- 5º: Melhorar Sempre e Constantemente o Sistema de Produção e Serviço;
- 6º: Instituir o Treinamento e o Retreinamento no Local de Trabalho;
- 7º: Instituir a Liderança;
- 8º: Afastar o Medo;
- 9º: Eliminar as Barreiras entre as Áreas e o Meio;
- 10º: Eliminar Slogans, Exortação e Metas para os Empregados;
- 11º: Eliminar as Cotas Numéricas Na Linha de Produção;
- 12º: Remover as Barreiras ao Orgulho da Execução;
- 13º: Instituir um Sólido Programa de Educação e Retreinamento;
- 14º: Agir no Sentido de Concretizar a Transformação.

Apêndice 2: Teste de Shapiro-Wilk

O teste Shapiro-Wilk tem o objetivo testar, como hipótese nula, se uma amostra x_1, \dots, x_n vem de uma população normalmente distribuída. Foi publicado 1965 por Samuel Shapiro e Martin Wilk. Suas hipóteses de teste são:

$$\begin{cases} H_0: A amostra provém de uma população Normal \\ H_1: A amostra não provém de uma população Normal \end{cases}$$

tomando assim a decisão rejeitar H_0 ao nível de significância α se $W_{\text{calculado}} < W_{\text{tabelado}}$ ou quando o p-valor é menor do que o nível de significância estabelecido, normalmente de 0,05. Sendo que, a estatística W dada por:

$$W = \frac{b^2}{\sum_{i=1}^n (x_{(i)} - \bar{x})^2} \quad (1)$$

em que x_i são os valores da amostra ordenados ($x_{(1)}$ é o menor). Menores valores de W são evidências de que os dados são normais. A constante b é determinada da seguinte forma

$$b = \begin{cases} \sum_{i=1}^{n/2} a_{n-i+1} \times (x_{(n-i+1)} - x_{(i)}) & \text{se } n \text{ é par} \\ \sum_{i=1}^{(n+1)/2} a_{n-i+1} \times (x_{(n-i+1)} - x_{(i)}) & \text{se } n \text{ é ímpar} \end{cases} \quad (2)$$

em que a_{n-i+1} são constantes geradas pelas médias, variâncias e covariâncias das estatísticas de ordem de uma amostra de tamanho n de uma distribuição Normal. Seus valores, tabelados, são dados abaixo.

$i \setminus N$	2	3	4	5	6	7	8	9	10
1	0.7071	0.7071	0.6872	0.6646	0.6431	0.6233	0.6052	0.5888	0.5739
2		0.0000	0.1877	0.2413	0.2806	0.3031	0.3164	0.3244	0.3291
3				0.0000	0.0875	0.1401	0.1743	0.1976	0.2141
4						0.0000	0.0561	0.0947	0.1224
5								0.0000	0.0399

$i \setminus N$	11	12	13	14	15	16	17	18	19	20
1	0.5601	0.5475	0.5359	0.5251	0.5150	0.5056	0.4968	0.4886	0.4808	0.4734
2	0.3315	0.3325	0.3325	0.3318	0.3306	0.3290	0.3273	0.3253	0.3232	0.3211
3	0.2260	0.2347	0.2412	0.2460	0.2495	0.2521	0.2540	0.2553	0.2561	0.2565
4	0.1429	0.1586	0.1707	0.1802	0.1878	0.1939	0.1988	0.2027	0.2059	0.2085
5	0.0695	0.0922	0.1089	0.1240	0.1353	0.1447	0.1524	0.1587	0.1641	0.1686
6	0.0000	0.0303	0.0539	0.0727	0.0880	0.1005	0.1109	0.1197	0.1271	0.1334
7			0.0000	0.0240	0.0433	0.0593	0.0725	0.0837	0.0932	0.1013
8					0.0000	0.0196	0.0359	0.0496	0.0612	0.0711
9							0.0000	0.0163	0.0303	0.0422
10									0.0000	0.0140

Apêndice 3: Teste Ljung-Box

Ljung e Box propuseram uma pequena alteração no teste de Box - Pierce, na qual torna o teste mais generalizado. Foi observado na literatura que o teste de Ljung - Box apresenta melhores resultados do que o teste de Box - Pierce.

As hipóteses do teste:

$$\begin{cases} H_0: Os\ resíduos\ são\ independentemente\ e\ identicamente\ distribuídos \\ H_1: Os\ resíduos\ não\ são\ independentemente\ e\ identicamente\ distribuídos \end{cases}$$

Calcula-se as estimativas de auto correlações por

$$\hat{r}_k = \frac{\sum_{i=k+1}^n \hat{a}_i \hat{a}_{i-k}}{\sum_{i=1}^n \hat{a}_i^2} \quad (4)$$

sendo $\hat{r}_k \sim N(0, \frac{1}{n})$.

Caso o modelo for apropriado, a estatística do teste

$$Q(k) = n(n-2) \sum_{j=1}^K \frac{\hat{r}_j^2}{(n-j)} \quad (5)$$

terá aproximadamente uma distribuição X^2 com $(K - p - q)$ graus de liberdade, onde K é o número de defasagens tomada, p e q são as ordens do modelo ajustado. Portanto, rejeitamos a hipótese nula se $Q(k) > X_{1-\alpha, K-p-q}^2$ com um nível de significância α .

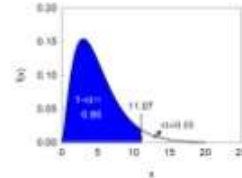
Anexo 1: Tabela distribuição χ^2

Distribuição Qui-Quadrado

Valores da Função de Distribuição

$$F(u) = \int_0^u \frac{x^{(n/2)-1} e^{-x/2}}{2^{n/2} \Gamma(n/2)} dx$$

Se $X \sim \chi^2$ então $P(X \leq \chi) = 1 - \alpha$
 Se $X \sim \chi^2_{(s)}$ então $P(X \leq 11,07) = 0,95$
 $\chi_{0,95}(s) = 11,07$



n/F	1 - α												
	0,005	0,010	0,025	0,050	0,100	0,250	0,500	0,750	0,900	0,950	0,975	0,990	0,995
1	0,000	0,000	0,001	0,004	0,016	0,102	0,455	1,323	2,706	3,841	5,024	6,635	7,879
2	0,010	0,020	0,051	0,103	0,211	0,575	1,386	2,773	4,605	5,991	7,378	9,210	10,597
3	0,072	0,115	0,216	0,352	0,584	1,213	2,366	4,108	6,251	7,815	9,348	11,345	12,838
4	0,207	0,297	0,484	0,711	1,064	1,923	3,357	5,385	7,779	9,488	11,143	13,277	14,860
5	0,412	0,554	0,831	1,145	1,610	2,675	4,351	6,626	9,236	11,070	12,833	15,086	16,750
6	0,676	0,872	1,237	1,635	2,204	3,455	5,348	7,841	10,645	12,592	14,449	16,812	18,548
7	0,989	1,239	1,690	2,167	2,833	4,255	6,346	9,037	12,017	14,067	16,013	18,475	20,278
8	1,344	1,646	2,180	2,733	3,490	5,071	7,344	10,219	13,362	15,507	17,535	20,090	21,955
9	1,735	2,088	2,700	3,325	4,168	5,899	8,343	11,389	14,684	16,919	19,023	21,666	23,589
10	2,156	2,558	3,247	3,940	4,865	6,737	9,342	12,549	15,987	18,307	20,483	23,209	25,188
11	2,603	3,053	3,816	4,575	5,578	7,584	10,341	13,701	17,275	19,675	21,920	24,725	26,757
12	3,074	3,571	4,404	5,226	6,304	8,438	11,340	14,845	18,549	21,026	23,337	26,217	28,300
13	3,565	4,107	5,009	5,892	7,042	9,299	12,340	15,984	19,812	22,362	24,736	27,688	29,819
14	4,075	4,660	5,629	6,571	7,790	10,165	13,339	17,117	21,064	23,685	26,119	29,141	31,319
15	4,601	5,229	6,262	7,261	8,547	11,037	14,339	18,245	22,307	24,996	27,488	30,578	32,801
16	5,142	5,812	6,908	7,962	9,312	11,912	15,338	19,369	23,542	26,296	28,845	32,000	34,267
17	5,697	6,408	7,564	8,672	10,085	12,792	16,338	20,489	24,769	27,587	30,191	33,409	35,718
18	6,265	7,015	8,231	9,390	10,865	13,675	17,338	21,605	25,989	28,869	31,526	34,805	37,156
19	6,844	7,633	8,907	10,117	11,651	14,562	18,338	22,718	27,204	30,144	32,852	36,191	38,582
20	7,434	8,260	9,591	10,851	12,443	15,452	19,337	23,828	28,412	31,410	34,170	37,566	39,997
21	8,034	8,897	10,283	11,591	13,240	16,344	20,337	24,935	29,615	32,671	35,479	38,932	41,401
22	8,643	9,542	10,982	12,338	14,041	17,240	21,337	26,039	30,813	33,924	36,781	40,289	42,796
23	9,260	10,196	11,689	13,091	14,848	18,137	22,337	27,141	32,007	35,172	38,076	41,638	44,181
24	9,886	10,856	12,401	13,848	15,659	19,037	23,337	28,241	33,196	36,415	39,364	42,980	45,559
25	10,520	11,524	13,120	14,611	16,473	19,939	24,337	29,339	34,382	37,652	40,646	44,314	46,928
26	11,160	12,198	13,844	15,379	17,292	20,843	25,336	30,435	35,563	38,885	41,923	45,642	48,290
27	11,808	12,879	14,573	16,151	18,114	21,749	26,336	31,528	36,741	40,113	43,195	46,963	49,645
28	12,461	13,565	15,308	16,928	18,939	22,657	27,336	32,620	37,916	41,337	44,461	48,278	50,993
29	13,121	14,256	16,047	17,708	19,768	23,567	28,336	33,711	39,087	42,557	45,722	49,588	52,336
30	13,787	14,953	16,791	18,493	20,599	24,478	29,336	34,800	40,256	43,773	46,979	50,892	53,672
31	14,458	15,655	17,539	19,281	21,434	25,390	30,336	35,887	41,422	44,985	48,232	52,191	55,003
32	15,134	16,362	18,291	20,072	22,271	26,304	31,336	36,973	42,585	46,194	49,480	53,486	56,328
33	15,815	17,074	19,047	20,867	23,110	27,219	32,336	38,058	43,745	47,400	50,725	54,776	57,648
34	16,501	17,789	19,806	21,664	23,952	28,136	33,336	39,141	44,903	48,602	51,966	56,061	58,964
35	17,192	18,509	20,569	22,465	24,797	29,054	34,336	40,223	46,059	49,802	53,203	57,342	60,275
36	17,887	19,233	21,336	23,269	25,643	29,973	35,336	41,304	47,212	50,998	54,437	58,619	61,581
37	18,586	19,960	22,106	24,075	26,492	30,893	36,336	42,383	48,363	52,192	55,668	59,893	62,883
38	19,289	20,691	22,878	24,884	27,343	31,815	37,335	43,462	49,513	53,384	56,896	61,162	64,181
39	19,996	21,426	23,654	25,695	28,196	32,737	38,335	44,539	50,660	54,572	58,120	62,428	65,476
40	20,707	22,164	24,433	26,509	29,051	33,660	39,335	45,616	51,805	55,758	59,342	63,691	66,766
41	21,421	22,906	25,215	27,326	29,907	34,585	40,335	46,692	52,949	56,942	60,561	64,950	68,053
42	22,138	23,650	25,999	28,144	30,765	35,510	41,335	47,766	54,090	58,124	61,777	66,206	69,336
43	22,859	24,398	26,785	28,965	31,625	36,436	42,335	48,840	55,230	59,304	62,990	67,459	70,616
44	23,584	25,148	27,575	29,787	32,487	37,363	43,335	49,913	56,369	60,481	64,201	68,710	71,893
45	24,311	25,901	28,366	30,612	33,350	38,291	44,335	50,985	57,505	61,656	65,410	69,957	73,166
46	25,041	26,657	29,160	31,439	34,215	39,220	45,335	52,056	58,641	62,830	66,617	71,201	74,437
47	25,775	27,416	29,956	32,268	35,081	40,149	46,335	53,127	59,774	64,001	67,821	72,443	75,704
48	26,511	28,177	30,755	33,098	35,949	41,079	47,335	54,196	60,907	65,171	69,023	73,683	76,969
49	27,249	28,941	31,555	33,930	36,818	42,010	48,335	55,265	62,038	66,339	70,222	74,919	78,231
50	27,991	29,707	32,357	34,764	37,689	42,942	49,335	56,334	63,167	67,505	71,420	76,154	79,490
51	28,735	30,475	33,162	35,600	38,560	43,874	50,335	57,401	64,295	68,669	72,616	77,386	80,747
52	29,481	31,246	33,968	36,437	39,433	44,808	51,335	58,468	65,422	69,832	73,810	78,616	82,001
53	30,230	32,018	34,776	37,276	40,308	45,741	52,335	59,534	66,548	70,993	75,002	79,843	83,253
54	30,981	32,793	35,586	38,116	41,183	46,676	53,335	60,600	67,673	72,153	76,192	81,069	84,502
55	31,735	33,570	36,398	38,958	42,060	47,610	54,335	61,665	68,796	73,311	77,380	82,292	85,749
56	32,490	34,350	37,212	39,801	42,937	48,546	55,335	62,729	69,919	74,468	78,567	83,513	86,994

(c) 2012, Pedro Casquilho