**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

**Modelos de séries temporais para temperatura em painéis de
cimento-madeira**

Valiana Alves Teodoro

Dissertação apresentada para obtenção do título
de Mestra em Ciências. Área de concentração:
Estatística e Experimentação Agronômica

**Piracicaba
2015**

Valiana Alves Teodoro
Licenciatura em Matemática

Modelos de séries temporais para temperatura em painéis de cimento-madeira
versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientadora:
Prof. Dr. **TACIANA VILLELA SAVIAN**

Dissertação apresentado para obtenção do título de Mestra em
Ciências. Área de concentração: Estatística e Experimentação
Agronômica

Piracicaba
2015

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP

Teodoro, Valiana Alves

Modelos de séries temporais para temperatura em painéis de cimento-madeira /
Valiana Alves Teodoro. - - versão revisada de acordo com a resolução CoPGr 6018 de
2011. - - Piracicaba, 2015.
96 p. : il.

Dissertação (Mestrado) - - Escola Superior de Agricultura "Luiz de Queiroz".

1. Modelo ARIMA 2. Comparação de séries 3. Painéis particulados I. Título

CDD 674.835
T314m

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"

DEDICATÓRIA

*Aos meus pais e minha família pelo apoio e compreensão. Por todo amor e carinho que
sempre encontro em vocês.*

AGRADECIMENTOS

A Deus pelo Dom da vida.

Aos meus pais Valdete e Eliana, agradeço por muitas vezes deixaram seus sonhos de lado para realizar os meus. Agradeço ainda aos meus pais e ao meu irmão Marcelo pelo carinho, apoio, amor e dedicação que sempre tiveram comigo, me dando coragem e ensinando a viver com o intuito de me transformarem em uma pessoa melhor. Aos demais familiares que estiveram carinhosamente presentes.

A minha orientadora, Prof. Dra. Taciana Villela Savian, pela orientação, dedicação, incentivo e paciência ao longo do desenvolvimento desta pesquisa. Que não mediu esforços em ajudar e também em contribuir para este trabalho sempre com o melhor, com novas idéias e experiências.

A Professora Rosimeire Cavalcante dos Santos, pelo apoio, concessão do banco de dados utilizado neste trabalho e sugestões.

Aos amigos do Mestrado pela ajuda em momentos de dificuldade e, principalmente, por vossas amizades.

As amigas “ Virginetes” que me proporcionaram a amizade, momentos de alegria e diversão na casa das n mulheres.

Aos professores e funcionários do Departamento de Ciências Exatas - LCE.

A todos aqueles que indiretamente colaboraram para a conclusão desta dissertação;

Estendo meus agradecimentos a Djair, por me deixar compartilhar parte da sua vida.

A CNPQ pela concessão da bolsa de mestrado para a realização da pesquisa e ao projeto temático.

SUMÁRIO

RESUMO	9
ABSTRACT	11
LISTA DE FIGURAS	13
LISTA DE TABELAS	17
LISTA DE ABREVIATURAS	19
1 INTRODUÇÃO	21
2 REVISÃO BIBLIOGRÁFICA	23
2.1 Candeia	23
2.2 Aproveitamento dos Resíduos de Madeira	23
2.3 Séries Temporais	24
2.4 Estacionariedade	25
2.5 Ruído Branco	26
2.6 Tendência	26
2.7 Processo estocástico	27
2.8 Modelos de Séries Temporais	28
2.8.1 Modelo auto regressivo (AR)	29
2.8.2 Modelo de média móvel (MA)	29
2.8.3 Modelo auto regressivo e de média móvel (ARMA)	29
2.8.4 Modelo auto regressivo integrado de média móvel (ARIMA)	30
2.9 Construção dos Modelos ARIMA	30
2.10 Função de autocorrelação (fac) e função de autocorrelação parcial (fACP)	31
2.11 Testes de Raiz Unitária	33
2.11.1 Dickey-Fuller	34
2.12 Critério de Seleção de Modelos	35
2.13 Teste de Box-Pierce	38
2.14 Testes para comparação de séries temporais	38
2.14.1 Teste de Igualdade das Funções de Autocorrelação	39
2.14.2 Método de comparação de Séries Temporais proposto por Silva, Ferreira e Sáfiadi (2000)	40
3 METODOLOGIA	41
3.1 Material	41

3.2 Método	42
4 RESULTADOS E DISCUSSÃO	45
4.1 Tratamento utilizando 25% Candeia e 75% Eucalipto com água	45
4.2 Tratamento utilizando 50% Candeia e 50% Eucalipto com água	52
4.3 Tratamento utilizando 75% Candeia e 25% Eucalipto com água	59
4.4 Tratamento utilizando 25% Candeia e 75% Eucalipto sem água	66
4.5 Teste de comparação de séries temporais	72
5 CONCLUSÕES	77
REFERÊNCIAS	79
APÊNDICE	83

RESUMO

Modelos de séries temporais para temperatura em painéis de cimento-madeira

Por meio do monitoramento da evolução da temperatura da mistura cimento-madeira, pode-se utilizar esta informação como uma série temporal. O objetivo deste estudo foi utilizar modelos de séries temporais para descrever as séries de temperatura do experimento constituído por diferentes espécies associadas a resíduos de Candeia na produção de painéis particulado e compara-las duas a duas para averiguar se foram geradas pelo mesmo processo estocástico. Inicialmente foi realizado um estudo para avaliar a estacionariedade das séries utilizando o correlograma e o teste da raiz unitária de Dickey-Fuller, na qual todas as séries apresentaram não estacionariedade, para o tratamento de 25% Candeia e Eucalipto com tratamento prévio de água foi dita uma série $I(2)$ e pelos critérios AIC, BIC e MAPE o melhor modelo foi $ARIMA(2, 2, 2)$, para o tratamento de 50% Candeia e Eucalipto também com tratamento prévio de água foi dita uma série $I(1)$ e pelos critérios o melhor modelo foi $ARIMA(4, 2, 2)$, para o tratamento de 75% Candeia e Eucalipto com tratamento prévio de água foi dita uma série $I(1)$ com o modelo $ARIMA(5, 1, 0)$, e para o tratamento de 25% Candeia e Eucalipto sem tratamento prévio de água foi dita uma série $I(1)$ com o modelo $ARIMA(2, 1, 2)$. Em relação à comparação das séries temporais contempladas neste trabalho é possível concluir que as mesmas são diferentes entre si, ou seja, não foram geradas pelo mesmo processo estocástico.

Palavras-chave: Modelo ARIMA; Comparação de séries; Painéis particulados

ABSTRACT

Time series models for temperature cement-wood panels

By monitoring the temperature evolution of the cement-wood mixture, one can utilize this information as a time series. The objective of this study was to utilize time series models to describe the temperature series from an experiment, consisting of different species associated to Candeia residuals in the production of particleboard panels, and do a pairwise comparison to verify if they were generated from the same stochastic process. Initially it was realized the Dickey-Fuller unit root test to verify series stationarity, which indicated that all series were not stationary. For the 25% Candeia and Eucalyptus treatment with previous water treatment the series was best modelled by an ARIMA(2, 2, 2) as evidenced by the AIC, BIC and MAPE criteria. For the 50% Candeia and Eucalyptus treatment also with previous water treatment the series was best modelled by an ARIMA(4, 2, 2) as indicated by the same criteria. Finally for the 75% Candeia and Eucalyptus treatment with previous water treatment and the 25% Candeia and Eucalyptus treatment without previous water treatment the best models were the ARIMA(5, 1, 0) and the ARIMA(2, 1, 2) respectively. In relation to the comparison of the time series contemplated in this study it is possible to conclude that they are different, that is, they were not generated by the same stochastic process.

Keywords: ARIMA model; Series of comparison; Particulate panels

LISTA DE FIGURAS

Figura 1 - Caixa térmica contendo as amostras	42
Figura 2 - Gráficos da temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água	45
Figura 3 - Gráficos da série diferenciada da temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água. (a) série na primeira diferença e (b) série na segunda diferença	47
Figura 4 - Gráficos das funções fac e facp da série, com uma e duas diferenças, de temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água. (a) fac da série; (b) facp da série; (c) fac da série com uma diferença; (d) facp da série com uma diferença; (e) fac da série com duas diferença e (f) facp da série com duas diferença	49
Figura 5 - Gráficos dos dados observados e preditos com o modelo ARIMA(2, 2, 2) da série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água	51
Figura 6 - Gráficos das funções fac (a) e facp (b) dos resíduos do modelo ARIMA(2, 2, 2) para a série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água	52
Figura 7 - Gráfico da temperatura média de 50% Candeia e 50% Eucalipto com água	53
Figura 8 - Gráficos da série diferenciada da temperatura média do tratamento de 50% Candeia e 50% Eucalipto com água. (a) série na primeira diferença e (b) série na segunda diferença	55
Figura 9 - Gráficos das funções fac e facp da série, com uma e duas diferenças, de temperatura média do tratamento de 50% Candeia e 50% Eucalipto com água. (a) fac com uma diferença; (b) facp com uma diferença; (c) fac com duas diferença e (d) facp com duas diferença	56
Figura 10 - Gráficos dos dados observados e dados preditos com o modelo ARIMA(4, 2, 2) da série de temperatura médias do tratamento de 50% Candeia e 50% Eucalipto com água	58
Figura 11 - Gráficos das funções fac (a) e facp (b) dos resíduos do modelo ARIMA(4, 2, 2) para a série de temperatura média do tratamento de 50% Candeia e 50% Eucalipto com água	59

Figura 12 - Temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água	60
Figura 13 - Gráficos da série com uma diferença da temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água	62
Figura 14 - Gráficos das funções fac e facp da série, com uma e duas diferenças, de temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água. (a) fac com uma diferença; (b) facp com uma diferença; (c) fac com duas diferença e (d) facp com duas diferença	63
Figura 15 - Gráficos dos dados observados e dados preditos com o modelo ARIMA(5, 1, 0) da série de temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água	65
Figura 16 - Gráficos das funções fac (a) e facp (b) dos resíduos do modelo ARIMA(5, 1, 0) para a série de temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água	66
Figura 17 - Temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água	67
Figura 18 - Gráficos da série diferenciada da temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água. (a) série na primeira diferença e (b) série na segunda diferença	69
Figura 19 - Gráficos das funções fac e facp da série, com uma e duas diferenças, de temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água. (a) fac com uma diferença; (b) facp com uma diferença; (c) fac com duas diferença e (d) facp com duas diferença	70
Figura 20 - Gráficos dos dados observados e dados preditos com o modelo ARIMA(2, 1, 2) da série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água	71
Figura 21 - Gráficos das funções fac e facp dos resíduos do modelo ARIMA(2, 1, 2) para a série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água. (a) fac dos resíduos do modelo ARIMA(2, 1, 2); (b) facp dos resíduos do modelo ARIMA(2, 1, 2)	72

Figura 22 - Gráficos das séries diferenciadas duas a duas (a) série 1 - série 2, (b) série 1 - série 3, (c) série 1 - série 4, (d) série 2 - série 3 (e) série 2 - série 4 e (f) série 3 - série 4	74
Figura 23 - Gráficos das funções da fac e facp (a) s1-s2, (b) s1-s3, (c) s1-s4, (d) s2-s3 (e) s2-s4 e (f) s3-s4	75

LISTA DE TABELAS

Tabela 1 - Propriedades teóricas da fac e facp	32
Tabela 2 - Correlograma da série temporal da temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água	46
Tabela 3 - Modelos ajustados para a série de temperatura médias do tratamento de 25% Candeia e 75% Eucalipto com água	50
Tabela 4 - Ajuste do modelo ARIMA(2,2,2) para a série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água	51
Tabela 5 - Medidas de erros do modelo ARIMA(2,2,2) da série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água	52
Tabela 6 - Correlograma da série temporal da temperatura média dos tratamentos de 50% Candeia e 50% Eucalipto com água	54
Tabela 7 - Modelos ajustados para a série de temperatura média do tratamento de 50% Candeia e 50% de Eucalipto com água	57
Tabela 8 - Ajuste do modelo ARIMA(4,2,2) para a série de temperatura média do tratamento de 50% Candeia e 50% Eucalipto com água	57
Tabela 9 - Medidas de erros do modelo ARIMA da série de temperatura média do tratamento de 50% Candeia e 50% Eucalipto com água	59
Tabela 10 -Correlograma da série temporal da temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água	61
Tabela 11 -Modelos ajustados para a série das temperatura médias do tratamento de 75% Candeia e 25% Eucalipto com água	64
Tabela 12 -Ajuste do modelo ARIMA(5,1,0) para a série de temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água	64
Tabela 13 -Medidas de erros do modelo ARIMA da série de temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água	66
Tabela 14 -Correlograma da série temporal da temperatura de 25% Candeia e 75% Eucalipto sem água	68
Tabela 15 -Modelos ajustados para a série das temperatura médias do tratamento de 25% Candeia e 75% Eucalipto sem água	71
Tabela 16 -Ajuste do modelo ARIMA(2,1,2) para a série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água	71

Tabela 17 -Medidas de erros do modelo ARIMA da série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água	72
Tabela 18 -Teste de Quenouille nas séries diferenciadas duas a duas	73
Tabela 19 -Teste da raiz unitária para as séries diferenciadas	75

LISTA DE ABREVIATURAS

ESALQ – Escola Superior de Agricultura “Luiz de Queiroz”

AR – auto-regressivo

MA – média móvel

ARMA – auto-regressivo e de média móvel

ARIMA – auto-regressivo integrado de média móvel

fac – função de autocorrelação

facp – função de autocorrelação parcial

DF – Dickey-Fuller

AIC – critério de informação de Akaike

BIC – critério de informação Bayesiano

SBC – critério Bayesiano Schwartz

MAPE – média percentual de erro absoluto

MPE – erro percentual médio

MSE – erro quadrático média

MAE – erro médio absoluto

UFPA – Universidade Federal de Lavras

1 INTRODUÇÃO

A questão dos resíduos vem ganhando importância no cenário mundial. Assuntos como tratamento, destinação e reciclagem de resíduos começam a fazer parte do vocabulário de pessoas comuns (MACEDO, 2008). Nos últimos anos, tem-se verificado um aumento do descarte de resíduos sólidos, que, juntamente com os problemas advindos da exaustão de matérias-primas naturais, vêm impulsionando os estudos sobre o aproveitamento de resíduos industriais na forma de novos materiais, com menor impacto ambiental (LOPES, 2009).

A indústria madeireira vem ao longo dos anos utilizando os resíduos madeireiros na produção de painéis de madeira, tanto na quantidade produzida como na diversidade de matérias-primas empregadas, oriundas de florestas nativas e plantadas. Tem ocorrido também uma ampliação do leque de outros materiais (orgânicos e inorgânicos) que são misturados à madeira para a fabricação de novos produtos tais como, cimento-madeira, gesso-madeira, madeira-plástico etc., aumentando a sua participação no mercado de compostos para utilização em fins específicos, onde a madeira sozinha não responde às necessidades de uma determinada aplicação. O aproveitamento do resíduo da madeira vem sendo estudada por alguns autores como Abreu (2005), Fontes (1994) e Paula (2006).

A candeia é uma espécie florestal de múltiplos usos, por exemplo, na produção de óleos essenciais. Outras razões para se investir na candeia é que é uma espécie nativa e os produtos dela obtidos alcançam preços relativamente altos no mercado. No Brasil existem cinco indústrias de óleo de candeia natural bruto, com uma produção anual estimada em cerca de 170 mil quilos, gerando assim, uma quantidade enorme de resíduos madeireiros. Apesar da exploração da candeia e da comercialização de seus produtos serem atividades geradoras de renda, o aproveitamento dos resíduos madeireiro não são largamente utilizados na produção de novos materiais. Possíveis soluções para a utilização do grande volume de resíduo madeireiro são fornecer energia às caldeiras durante o processo de extração e a produção de painéis particulados.

Santos (2008), estudou o aproveitamento do resíduo da madeira de candeia após a extração do óleo para a produção de painéis particulados, no qual um dos estudos consistiu em mensurar, de forma sistemática (a cada minuto), a evolução da temperatura da mistura cimento-madeira durante 24h, caracterizando uma série temporal. Nesse tipo

de experimento, o conhecimento da evolução da temperatura é de suma importância pois, são nas elevadas temperaturas que ocorrem a “pega” ou o endurecimento da mistura. Por meio da análise de série temporal, é possível verificar se duas séries diferentes são resultados de um mesmo processo natural. Esse conhecimento é de grande importância nesse tipo de experimento pois, um dos objetivos é analisar a influência do tratamento prévio com água e as diferentes proporções de mistura sobre a inibição da pega do cimento, na confecção dos painéis cimento-madeira.

Nesta dissertação, objetiva-se utilizar modelos de séries temporais descrever cada série de temperatura do experimento constituído por diferentes espécies associadas a resíduos de Candeia na produção de painéis particulados e compara-las duas a duas para averiguar se foram geradas pelo mesmo processo estocástico.

2 REVISÃO BIBLIOGRÁFICA

2.1 Candeia

A Candeia é uma espécie florestal de múltiplos usos. Sua madeira é utilizada predominantemente como moirão de cerca, em decorrência de sua alta durabilidade e, também, na extração de óleo essencial, cujo princípio ativo é o α -bisabolol, empregado na fabricação de medicamentos e cosméticos (cremes, bronzeadores, protetores solares, veículo para medicamentos), além do uso na profilaxia e cuidados da pele de bebês e adultos, entre outros (SCOLFORO, 2012).

2.2 Aproveitamento dos Resíduos de Madeira

O desenvolvimento de novos produtos e a reciclagem de materiais são dois campos importantes da pesquisa, que podem ser combinados para atender às necessidades sociais, ambientais e do mercado, tendo-se uma Tecnologia Ambientalmente Saudável - TAS (MACEDO, 2008).

A questão dos resíduos vem ganhando importância no cenário mundial. Assuntos como tratamento, destinação e reciclagem de resíduos começam a fazer parte do vocabulário de pessoas comuns. O tema é de interesse premente, pois trata da melhoria da qualidade de vida da população e da menor agressão ao meio ambiente (MACEDO, 2008). Dentre os benefícios da utilização de resíduos de madeira estão a melhoria das condições sociais, através da criação de postos de trabalho e do emprego de mão de obra não qualificada; compatibilidade com pequenos investimentos; conservação do meio ambiente em virtude da valorização do resíduo; e, incentivo a economia, pois, esta técnica agrega valor aos resíduos, proporcionando novas fontes de renda para a cultura local (LOPES, 2009).

Os laminados e os particulados constituídos pelas lâminas, compensados, aglomerados, painéis minerais e chapas de fibra são produtos com maior valor agregado quando comparados com a madeira “in natura” ou com os produtos sem industrialização ou semi-industrializados. Estas duas modalidades de uso da madeira vêm substituindo os produtos tradicionalmente usados e vários tipos de painéis vêm ganhando espaço comercial, em virtude da melhor relação preço/desempenho e da crescente conscientização dentro da sociedade moderna de que não é mais viável a convivência com processos que utilizam reservas florestais, com níveis elevados de perdas (SANTOS, 2008).

Ampliar pesquisas com a utilização de diferentes espécies associadas a resíduos madeireiros na produção de painéis particulados pode resultar numa produção satisfatória e numa inserção considerável desses painéis no mercado consumidor, gerando novas alternativas tecnológicas. O aproveitamento desses resíduos para a produção de painéis particulados torna-se uma alternativa viável para a produção de produtos de maior valor agregado, evitando diversos problemas de ordem ambiental, além de possibilitar a inserção desses produtos no mercado consumidor (SANTOS, 2008).

2.3 Séries Temporais

Seja uma série de tempo uma sequência de observações ordenadas, em intervalos de tempo igualmente espaçados, a ordem também pode ser tomada através de outras dimensões, como espaço. Séries temporais ocorrem em uma variedade de campos, como na agricultura, em que observa-se a produção e os preços de safra anual; na economia, pode-se observar das cotações diárias de fechamento das ações, taxas de juros, índices de preços semanais mensais, as vendas trimestrais e ganhos anuais; em meteorologia, observam-se as velocidades dos ventos, temperatura diária e precipitação anual; no controle de qualidade (WEI, 2006).

Existem vários objetivos para o estudo de séries temporais, dentre eles destacam-se os seguintes (EHLERS, 2009):

- Descrição: Descrever propriedades da série, os padrões de tendência, existência de variação sazonal ou cíclica, observações discrepantes, alterações estruturais;
- Explicação: Usar a variação em uma série para explicar a variação em outra série;
- Predição: prever valores futuros com base em valores passados. Aqui assume-se que o futuro envolve incerteza, ou seja as previsões não são perfeitas. Porém devemos tentar reduzir os erros de previsão;
- Controle: os valores da série temporal medem a qualidade de um processo de manufatura e o objetivo é o controle do processo. Um exemplo é o controle estatístico de qualidade aonde as observações são representadas em cartas de controle;

A natureza inerente de uma série temporal consiste que as suas observações são dependentes ou correlacionadas, e a ordem das observações é relevante, enquanto em

modelos de regressão por exemplo a ordem das observações é irrelevante para a análise. Assim, os procedimentos estatísticos e técnicas que dependem de suposição de independência não são mais aplicáveis, e são necessários métodos diferentes, exigindo o uso de técnicas especiais (EHLERS, 2009).

As séries temporais podem ser estacionárias ou não estacionárias. Além disso, podem ser estocásticas ou determinísticas. Segundo Bueno (2011), a série será estacionária e estocástica se a ela for acrescentado um componente aleatório independente e identicamente extraído de uma distribuição normal, $\varepsilon_t \sim N(0, \sigma^2)$.

Em resumo, uma série estacionária é aquela que flutua em torno de uma mesma média e a série não estacionária tem uma tendência, que pode ter uma natureza determinística ou estocástica. A série não estacionária determinística, acrescida de um componente aleatório extraído de uma distribuição normal, flutua em torno de uma tendência temporal. A série não estacionária com tendência estocástica move-se em torno de médias flutuantes (BUENO, 2011).

2.4 Estacionariedade

Uma das suposições mais frequentes que se faz a respeito da série temporal é a de que ela seja estacionária, ou seja, ela se desenvolve no tempo aleatoriamente em torno de uma média constante, mostrando alguma forma de equilíbrio estável. Entretanto, na prática, a maioria das séries temporais encontradas não são estacionárias e os componentes tendência e sazonalidade são algumas formas de não estacionariedade (MORETTIN; TOLOI, 2006).

Segundo Morettin e Toloi (2006), um processo estocástico diz-se fracamente estacionário ou estacionário de segunda ordem se:

- i. $E\{z(t)\} = \mu(t) = \mu$, constante, $\forall t$;
- ii. $E\{z^2(t)\} < \infty$, $\forall t$;
- iii. $\gamma(t_1, t_2) = \text{Cov}\{z(t_1)z(t_2)\}$ é uma função de $|t_1 - t_2|$

A primeira condição assegura que a média é igual para todos os períodos, mesmo que a distribuição da variável aleatória se altere ao longo do tempo. A segunda condição afirma apenas que o segundo momento não centrado deve ser finito, ainda que

desigual em diferentes períodos. A terceira condição é que a autocovariância não depende do tempo, mas da distancia temporal entre as observações (BUENO, 2011).

Visualmente, observa-se estacionariedade se uma série flutua em torno de uma média fixa e se a variância da série é constante ao longo do tempo. Não obstante, serão necessários testes estatísticos para verificar ou não a estacionariedade da série (BUENO, 2011).

O procedimento mais simples utilizado para tornar uma série temporal estacionária consiste em tomar diferenças consecutivas na série original até obter uma série estacionária. Normalmente, uma ou duas diferenças são suficientes para retirar a tendência.

2.5 Ruído Branco

Um processo fundamental das séries temporais estocástica discretas é o ruído branco. Uma sequencia ε_t é um ruído branco se cada valor nela tiver média zero, variância constante e não correlacionado a qualquer realização da própria série (GUJARATI, 2000).

Seja uma sequência $\{\varepsilon_t\}_{t=-\infty}^{\infty}$ de variáveis aleatórias. Se:

- i. $E(\varepsilon_t) = 0, \forall t$
- ii. $E(\varepsilon_t^2) = \sigma^2, \forall t$
- iii. $E(\varepsilon_{t_1}\varepsilon_{t_2}) = 0$

diz-se que o processo é um ruído branco cuja representação é RB $(0, \sigma^2)$ (BUENO, 2011).

É preciso ter em mente a ideia de que sempre se buscam resíduos estacionários, pois só nesse caso é possível confiar nos testes estatísticos de coeficientes e da regressão. Uma vez que as inferências estatísticas só terão validades se os resíduos da série temporal estimada forem estacionários (BUENO, 2011).

2.6 Tendência

Segundo Morettin e Toloi (2004), uma série temporal é considerada estacionária quando ela se desenvolve no tempo aleatoriamente em torno de uma média constante, mostrando alguma forma de equilíbrio estável. Frequentemente, supõe-se que a série temporal é estacionária. Entretanto, na prática, a maioria das séries temporais encontradas

não são estacionárias. Os componentes tendência e sazonalidade são formas de não estacionariedade.

O procedimento mais simples utilizado para tornar uma série temporal estacionária consiste em tomar diferenças consecutivas na série original até obter uma série estacionária. Normalmente, uma ou duas diferenças são suficientes para retirar a tendência.

A primeira diferença de uma série temporal $z(t)$ é dada por:

$$\Delta z_t = z_t - z_{t-1}$$

E a segunda diferença é:

$$\Delta^2 z_t = z_t - 2z_{t-1} + z_{t-2}$$

Quando se diferencia uma série temporal uma única vez e a série resultante é considerada estacionária, diz-se que a série original (com caminho aleatório, ou seja, não estacionária) é integrada de ordem 1, apresentada na literatura como $I(1)$. Similarmente, se a série original tiver de ser diferenciada duas vezes antes de se tornar estacionária, a série é integrada de ordem 2, ou $I(2)$. Em geral, se a série temporal tiver de ser diferenciada d vezes, ela é integrada de ordem d , ou $I(d)$. Assim, sempre que tiver uma série temporal integrada de ordem 1 ou maior, tem-se uma série temporal não estacionária, e se $d = 0$, o processo $I(0)$ resultante representa uma série temporal estacionária (GUJARATI, 2000).

De acordo com Margarido (1994) a utilização da aplicação dos operadores de diferenças reside no fato de que eles são capazes de deixar a série estacionária, o que significa que esses operadores não somente estabilizam a variância, como também removem a tendência que está por trás das séries originais, tomando-as estáveis.

A análise gráfica é uma ferramenta importante em séries temporais. Por meio dessa, é possível identificar características inerentes aos dados, como variabilidade, observações atípicas, sazonalidade e tendência, dentre outros.

2.7 Processo estocástico

Observa-se uma série temporal $\{z_1, z_2, \dots, z_n\}$, decorrente de uma variável aleatória Z . A série temporal observada é uma possível realização do processo estocástico gerador de dados. Suponha que se queira gerar K sequências de séries, cada uma com

infinitas observações. Então, haveria um conjunto com as seguintes sequências (BUENO, 2011):

$$\left(\left\{ z_t^{(1)} \right\}_{t=-\infty}^{\infty}, \left\{ z_t^{(2)} \right\}_{t=-\infty}^{\infty}, \dots, \left\{ z_t^{(K)} \right\}_{t=-\infty}^{\infty} \right).$$

Isso significa que, em cada instante de tempo t , há K estados da natureza. Se fosse possível selecionar as k observações associadas à data t , chegaríamos ao conjunto de dados $\{z_t^{(1)}, z_t^{(2)}, \dots, z_t^{(K)}\}$ cuja distribuição, por construção, é normal. Nessas situações, vários momentos da série poderiam ser estimados. Com essa noção intuitiva, pode-se mencionar a seguinte definição, sendo Ω representa o conjunto amostral e $\mathfrak{S}$, os subconjuntos de Ω , cuja probabilidade associada a cada um deles é P (BUENO, 2011):

Definição 2.1 *Sejam $(\Omega, \mathfrak{S}, P)$ um espaço de probabilidade e $\mathbb{Z}$ um conjunto de índices de números reais. Defina a função $z(\cdot, \cdot)$ por $z(\cdot, \cdot) : K \times \mathbb{Z} \rightarrow \mathbb{R}$. A sequência ordenada de variáveis aleatórias $\{z(\cdot, \cdot), t \in \mathbb{Z}\}$ é chamado processo estocástico (aleatório).*

A definição anterior indica que, para $t \in \mathbb{Z}$, $z(\cdot, t)$ representa uma variável aleatória sobre K . Além disso, para cada $k \in K$, a sequência $z(k, \cdot)$ representa uma realização de um processo estocástico. Desse modo, $z(k, t)$ é simplesmente um número real (BUENO, 2011).

2.8 Modelos de Séries Temporais

Os modelos utilizados para descrever séries temporais são processos estocásticos, isto é, processos controlados por leis probabilísticas. Qualquer que seja a classificação que façamos para os modelos de séries temporais, podemos considerar um número muito grande de modelos diferentes para descrever o comportamento de uma série particular.

A construção destes modelos depende de vários fatores, tais como o comportamento do fenômeno ou o conhecimento a priori que temos de sua natureza e do objetivo da análise. Na prática, depende, também, da existência de métodos apropriados de estimação e da disponibilidade de software (MORETTIN; TOLOI, 2006).

Os modelos propostos por Box e Jenkins: Auto regressivo (AR), Médias Móveis (MA), Auto Regressivo com Médias Móveis (ARMA), Auto Regressivo Integrado com Médias Móveis (ARIMA), (BOX et al., 1994) serão apresentados a seguir.

2.8.1 Modelo auto regressivo (AR)

Segundo Wei (2006), uma série z_t que segue um processo auto regressivo de ordem p , $AR(p)$, é dado por:

$$z_t = \theta_1 z_{t-1} + \theta_2 z_{t-2} + \cdots + \theta_p z_{t-p} + \varepsilon_t.$$

em que z_t é a observação no tempo t ; θ_i é o parâmetro auto regressivo do modelo com $i=1, \dots, p$; z_{t-i} é a observação defasada no tempo $t-i$, com $i=1, \dots, p$; ε_t é o termo de erro aleatório não correlacionado com média zero e variância constante, supondo σ^2 , isto é ruído branco que podemos representar por $\varepsilon_t \sim N(0, \sigma^2)$.

2.8.2 Modelo de média móvel (MA)

A série z_t é a combinação dos erros passados. Assim, dizemos que z_t segue um processo de médias móveis de ordem q , ou $MA(q)$, é dado por:

$$z_t = \varepsilon_t - \varphi_1 \varepsilon_{t-1} - \varphi_2 \varepsilon_{t-2} - \cdots - \varphi_q \varepsilon_{t-q}.$$

em que z_t é a observação no tempo t ; φ_j é o parâmetro de médias móveis do modelo com $j=1, \dots, q$; ε_{t-j} é o erro no tempo $t-j$, com $j=1, \dots, q$; ε_t é o termo de erro aleatório não correlacionado com média zero e variância constante, supondo σ^2 , isto é ruído branco que podemos representar por $\varepsilon_t \sim N(0, \sigma^2)$ (WEI, 2006).

2.8.3 Modelo auto regressivo e de média móvel (ARMA)

Facilmente, é bem possível que a série z_t tenha características tanto de AR como de MA, e seja, portanto, ARMA. Assim z_t segue um processo auto regressivo de médias móveis, $ARMA(p, q)$, sendo escrito:

$$z_t = \theta_1 z_{t-1} + \theta_2 z_{t-2} + \cdots + \theta_{p+d} z_{t-p-d} + \varepsilon_t - \varphi_1 \varepsilon_{t-1} - \varphi_2 \varepsilon_{t-2} - \cdots - \varphi_q \varepsilon_{t-q}.$$

em que z_t é a observação no tempo t ; z_{t-i} é a observação no tempo $t-i$, com $i=1, \dots, p$; θ_i é o parâmetro do modelo da parte auto regressiva com $i=1, \dots, p$; ε_{t-j} é o erro no tempo $t-j$, com $j=1, \dots, q$; φ_j é o parâmetro do modelo da parte de médias móveis, com $j=1, \dots, q$; ε_t é o termo de erro aleatório não correlacionado com média zero e variância

constante, supondo σ^2 , isto é ruído branco que podemos representar por $\varepsilon_t \sim N(0, \sigma^2)$ (WEI, 2006).

2.8.4 Modelo auto regressivo integrado de média móvel (ARIMA)

Os modelos de séries temporais que já foram discutidos se baseiam na hipótese de que as séries temporais sejam estacionárias, ou seja, média e a variância para uma série temporal são constantes e sua covariância é invariável no tempo. Mas sabemos que muitas séries temporais são não estacionárias. Para isso é necessário remover as fontes de variação não estacionária da série temporal observada. Se a série observada for não estacionária pode-se tentar remover a tendência tomando diferenças sucessivas da série original, até obter uma série estacionária. Este modelo é conhecido por auto regressivo integrado de média móvel, ARIMA (p,d,q) , dado por:

$$w_t = \theta_1 w_{t-1} + \theta_2 w_{t-2} + \cdots + \theta_{p+d} w_{t-p-d} + \varepsilon_t - \varphi_1 \varepsilon_{t-1} - \varphi_2 \varepsilon_{t-2} - \cdots - \varphi_q \varepsilon_{t-q}.$$

em que $w_t = \Delta^n z(t)$; z_t é a observação no tempo t ; d é a diferença dos processos de combinação linear de valores passados; θ_i é o parâmetro do modelo da parte auto regressiva com $i=1, \dots, p$; z_{t-i} é a observação no tempo $t-i$, com $i=1, \dots, p$; ε_{t-j} é o erro no tempo $t-j$; φ_j é o parâmetro do modelo da parte de médias móveis; ε_t representa o erro, supondo $\varepsilon_t \sim N(0, \sigma^2)$ (WEI, 2006).

Os modelos de séries temporais, em especial os modelos ARIMA é bastante utilizado na literatura, em diferentes áreas aplicadas, vazão em microbacias pariadas (AMARAL, 2014), previsão de vazão fluvial (BATISTA, 2009), temperatura do solo (CARNEIRO, 2014), focos de calor no Brasil (SILVA, 2009).

2.9 Construção dos Modelos ARIMA

A seguir serão apresentados os estágios do ciclo iterativo do método de Box e Jenkins para construção dos modelos ARIMA, que são: 1. identificação; 2. estimação; 3. verificação. Dentre estes estágios o que demanda maior empenho do pesquisador é a fase de identificação do modelo ARIMA, ou seja, a determinação dos valores p , d e q do modelo. Esta escolha é feita a partir das autocorrelações e autocorrelações parciais estimadas, dentre

as quais espera-se que representem de forma adequada as verdadeiras quantidades teóricas, que são desconhecidas.

O objetivo da identificação é determinar os valores de p , d e q do modelo ARIMA (p,d,q) , além de estimativas preliminares dos parâmetros a serem usadas no estágio de estimação (MORETTIN; TOLOI, 2006).

O procedimento de identificação consiste em três partes:

- a - verificar se existe necessidade de uma transformação na série original, com o objetivo de estabilizar sua variância. Tal identificação pode ser realizada utilizando o auxílio de gráficos;
- b - tomar diferenças da série, obtida no item (a), tantas vezes quantas necessárias para se obter uma série estacionária, de modo que o processo $\Delta^d z_t$ seja reduzido a uma processo ARMA(p,q). O número de diferenças, d , necessárias para que o processo se torne estacionário, é alcançado quando a fac amostral de $W_t = \Delta^d z_t$ decresce rapidamente para zero. Neste estágio a utilização de um teste para verificar a existência de raiz unitária no polinômio auto regressivo, pode ser de grande utilidade;
- c - identificar o processo ARMA(p,q) resultante, através da análise das autocorrelações e autocorrelações parciais estimadas, cujos comportamento devem imitar os comportamentos das respectivas quantidades teóricas.

2.10 Função de autocorrelação (fac) e função de autocorrelação parcial (facp)

Considere uma série estacionária z_t , na qual sua função de autocorrelação ρ_k para $k = 1, \dots, n-1$, é calculado as correlações amostrais entre os pares (z_1, z_{k+1}) , $(z_2, z_{k+2}), \dots, (z_{t-k}, z_t)$. A função de autocorrelação (fac) amostral r_k é definido como:

$$r_k = \frac{c_k}{c_0}$$

em que $c_k = \frac{1}{n} \sum_{t=1}^{n-k} (z_t - \bar{z})(z_{t+k} - \bar{z})$, $k = 0, 1, \dots, n-1$ e $\bar{z}$ é a média amostral, tem-se ainda que $c_{-k} = c_k$ e $r_{-k} = r_k$. O gráfico de $r_k \times k$ é denominado de correlograma.

A correlação entre z_t e z_{t-k} pode ser definido removendo o efeito das variáveis $z_{t-1}, z_{t-2}, \dots, z_{t-k+1}$. Esta medida, para séries estacionárias, é chamada de autocorrelação

parcial até a posição k e será denotada por ϕ_{kk} ou seja,

$$\phi_{kk} = \text{corr}(z_t, z_{t-k} | z_{t-1}, z_{t-2}, \dots, z_{t-k+1})$$

em que ϕ_{kk} é o coeficiente de correlação da distribuição conjunta de z_t, z_{t-k} condicional a $z_{t-1}, z_{t-2}, \dots, z_{t-k+1}$.

Durbin (1960) demonstraram que o coeficiente de autocorrelação pode ser obtido, recursivamente, por:

$$\hat{\phi}_{kk} = \frac{\rho_k - \sum_{i=1}^{k-1} \varphi_{k-1,i} \rho_{k-i}}{1 - \sum_{i=1}^{k-1} \varphi_{k-1,i} \rho_i}$$

em que $\phi_{kj} = \phi_{kk} \phi_{k-1,k-j}$ para $j = 1, \dots, k-1$, e nas estimativas de $\hat{\rho}_k$ por r_k .

Uma importante ferramenta para verificar estacionariedade da série é a análise gráfica do autocorrelograma, que é a representação gráfica do coeficiente da autocorrelação.

Para determinar a ordem adequada de um modelo é recomendado utilizar as funções de autocorrelação (fac) e função autocorrelação parcial (fapc). As propriedades teóricas das funções de autocorrelação e autocorrelação parcial para alguns processos estacionários auxiliam na determinação da ordem (p, q) do modelo ARIMA, em que p é a ordem do modelo AR e q é a ordem do modelo MA (EHLERS, 2009). A tabela 1 apresenta as propriedades teóricas da fac e fapc.

Tabela 1 – Propriedades teóricas da fac e fapc

Processo	fac	fapc
AR(p)	Declina exponencialmente ou padrão de onda senoidal amortecidas, ou ambas	Picos (lag) significativos através das defasagens p
MA(q)	Picos(lag) significativos através das defasagens q	Declina exponencialmente
ARMA(p, q)	Declínio exponencial	Declínio exponencial

Fonte: GUJARATI (2000)

Algumas observações serão uteis no procedimento de identificação dos modelos (MORETTIN; TOLOI, 2004):

- i. Um processo $AR(p)$ tem fac que decai de acordo com exponenciais e/ou senóide amortecidas, infinitas em extensão. Segundo Box et al. (1994), a facp ajuda na identificação da ordem do processo auto regressivo, que se adequa à série temporal em estudo.
- ii. Um processo $MA(q)$ tem fac finita, no sentido que representa um corte após o lag q e tem facp que se comporta de maneira similar à fac de um processo $AR(p)$.
- iii. Um processo $ARMA(p,q)$ tem fac infinita em extensão, a qual decai de acordo com exponenciais e/ou senóides amortecidos após o lag $q - p$ e tem uma facp de um processo MA puro.

Quando há uma forma complicada da fac e facp de um modelo ARMA, estas funções não são muitos uteis para identificação de tais modelos. O que recomenda, neste caso é ajustar alguns modelos de baixa ordem e utilizar critérios que permitam escolher o modelo mais adequado (MORETTIN; TOLOI, 2006). O que importa, na prática, é extrair o máximo de informações da série, a ponto de gerar um resíduo que seja estatisticamente um ruído branco. Essa seria a melhor modelagem possível de ser obtida a partir dos dados observados. Supostamente, essa modelagem levaria às melhores previsões estatísticas (BUENO, 2011).

Após a identificação de um modelo para uma determinada série, o próximo passo é o de estimar os parâmetros que pode ser feita pelo método de máxima verossimilhança ou dos mínimos quadrados.

2.11 Testes de Raiz Unitária

A não estacionariedade da série temporal poderá muitas vezes reduzir a uma série com diferenças. Um teste alternativo de estacionariedade que recentemente se tornou popular é o teste de raiz unitária. Para decidir se precisa ou não de diferenças, não tendo somente que depender de da análise da gráfico da série e da função de autocorrelação (WEI, 2006; GUJARATI, 2000).

De acordo com Box-Jenkins (1994), o número de diferenças necessárias para que uma série se torne estacionária é conhecido como ordem de integração da série. Os testes de raízes unitárias são capazes, em geral, de detectar se a série foi suficientemente diferenciada para se tornar estacionária. Para tanto, testa-se a hipótese nula de que a série

não é estacionária, ou seja, possui raiz unitária, contra a alternativa de que a série é estacionária. Os testes clássicos mais utilizados para se testar essa hipótese foram, inicialmente, estudados por Fuller (1976) e Dickey e Fuller (1979). Outros trabalhos foram publicados posteriormente, como os testes apresentados por Said e Dickey (1984) e Phillips e Perron (1988).

O problema de raiz unitária em modelos auto regressivos e de médias móveis (ARMA) aparece quando o polinômio auto regressivo apresenta pelo menos uma raiz sobre o círculo unitário. Isso implica que deve-se tomar um número de diferenças apropriadas da série original antes de ajustar o modelo (MORETTIN; TOLOI 2004).

2.11.1 Dickey-Fuller

Para entender o teste de Dickey-Fuller (DF), considere o seguinte modelo AR(1):

$$z_t = \theta z_{t-1} + \varepsilon_t \quad (1)$$

em que ε_t é o ruído branco do processo $\varepsilon_t \sim N(0, \sigma^2)$, $t = 1, \dots, n$. Se o coeficiente θ do z_{t-1} for de fato igual a 1, defrontamo-nos com o que é conhecido como o problema da raiz unitária, isto é, uma situação de não estacionariedade. Portanto se $\theta = 1$, diz-se então que a variável estocástica z tem uma raiz unitária. Em séries temporais, uma série que tenha uma raiz unitária é conhecida como uma série de caminho aleatório (GUJARATI, 2000).

A tendência inicial é estimar os parâmetros deste modelo 1 e usar a estatística t sobre θ calculada de modo convencional é conhecida como estatística τ (tau), cujos valores críticos foram tabulados por Dickey e Fuller (1979) com base na simulação de Monte Carlo. Na literatura, o teste tau é conhecido como o teste de Dickey-Fuller (DF). Tendo como hipótese $H_0 : \theta = 1$ e como hipótese alternativa $H_1 : |\theta| < 1$. Então, pode-se alterar o testes subtraindo z_{t-1} de ambos os lados (BUENO, 2011; GUJARATI, 2000; WEI, 2006):

$$\Delta z_t = (\theta - 1) z_{t-1} + \varepsilon_t = \alpha z_{t-1} + \varepsilon_t$$

em que se define $\alpha = (\theta - 1)$. Assim, $H_0 : \theta = 1$ é equivalente a $H_0 : \alpha = 0$.

Por estas razões teóricas e práticas citadas acima, o teste de Dickey-Fuller é aplicado nos modelos:

$$\Delta z_t = \alpha z_{t-1} + \varepsilon_t \quad (2)$$

$$\Delta z_t = \beta_0 + \alpha z_{t-1} + \varepsilon_t \quad (3)$$

$$\Delta z_t = \beta_0 + \beta_1 t + \alpha z_{t-1} + \varepsilon_t \quad (4)$$

em que t é a variável tempo ou tendência. Em cada caso, a hipótese nula é $H_0 : \alpha = 0$, ou seja, há raiz unitária. O modelo 2 desconsidera a existência de constante e tendência determinística; o modelo 3, a existência de tendência determinística, o modelo 4, há um intercepto e tendência determinística.

Segundo Wei (2006) os modelos acima são nomeados como 2 o modelo AR (1) sem constante, 3 o modelo AR (1) com constante e 4 como o modelo AR (1) com tendência.

Em sua forma mais simples, estimamos os parâmetros de uma regressão do tipo 2, 3 ou 4, dividimos o coeficiente $\hat{\alpha}$ por seu erro-padrão para calcular a estatística τ de Dickey-Fuller e consultamos as tabelas de Dickey-Fuller, se o valor absoluto calculados da estatística τ excede os valores críticos absolutos de τ então não aceitamos a hipótese de que a série tenha raiz unitária. Se, por outro lado, ele for menor que o valor crítico, aceita-se a hipótese que a série tem raiz unitária, ou seja, a série é não-estacionária (GUJARATI, 2000).

2.12 Critério de Seleção de Modelos

Não existem modelos verdadeiros, há apenas modelos aproximados da realidade que, causam perda de informações. Deste modo, é necessário fazer a seleção do melhor modelo, dentre aqueles que foram ajustados, para explicar o fenômeno sob estudo.

Admitindo um modelo com k parâmetro para uma série temporal. Para avaliar a qualidade do modelo ajustado, Akaike (1973, 1974) introduziu este critério de informação. O critério é chamado de AIC (critério de informação de Akaike), que é expresso por:

$$AIC = 2\ln\hat{\sigma}_a^2 + 2k.$$

em que $\hat{\sigma}_a^2$ é a estimativa de máxima verosimilhança de σ_a^2 ; k é o número de parâmetros. A ordem do modelo é o valor de k , com as funções de p e q , no qual o AIC é mínimo (WEI, 2006).

Akaike (1974), Rissanen (1978) e Schwarz (1978) desenvolveram uma extensão Bayesiana, que minimiza o critério de informação de Akaike, chamado de BIC (critério

de informação Bayesiano), este penaliza mais fortemente modelos com um maior número de parâmetros do que o AIC tendendo, dessa forma, a selecionar modelos com um menor número de parâmetros. O critério BIC é dado pela seguinte expressão:

$$\text{BIC}(k) = n \ln \hat{\sigma}_a^2 - (n - k) \ln \left(1 - \frac{k}{n} \right) + k \ln \left[\frac{\left(\frac{\hat{\sigma}_z^2}{\hat{\sigma}_a^2} - 1 \right)}{k} \right].$$

em que $\hat{\sigma}_a^2$ é a estimativa de máxima verosimilhança de σ_a^2 ; $\hat{\sigma}_z^2$ é a variância da série; k é o número de parâmetros; n é o tamanho da série (MORETTIN; TOLOI, 2006; WEI, 2006).

A acurácia é um critério muito importante na seleção e avaliação das técnicas de previsão e predição. A fim de avaliar o desempenho (acurácia) absoluto e relativo dos diversos modelos de previsão, diversas ferramentas matemáticas foram desenvolvidas ao longo do tempo, existindo diversas medidas de erros de previsões. Assim, a seguir apresenta-se algumas medidas de erros que são utilizados na literatura.

O ME (erro médio) é a medida mais simples de erro de previsão, definido como o somatório dos erros dividido pelo número de observações, que é dado por (ACZEL, 1993):

$$ME = \frac{1}{n} \sum_{i=1}^n e_i.$$

em que $e_i = z_i - \hat{z}_i$ é o erro de predição; z_i é a observação no tempo i , sendo $i=1, \dots, n$; n é o tamanho da série.

O MAE (erro médio absoluto), é o erro médio tomado em termos absolutos, para que o erro positivo não seja anulado por outro negativo, que é dado por:

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i|.$$

em que $e_i = z_i - \hat{z}_i$ é o erro de predição; z_i é a observação no tempo i , sendo $i=1, \dots, n$; n é o tamanho da série.

O MSE (erro quadrático médio), foi definido com a mesma finalidade do MAE de não anular os erros durante sua somatório, que é dado por:

$$MSE = \frac{1}{n} \sum_{i=1}^n e_i^2.$$

em que $e_i = z_i - \hat{z}_i$ é o erro de predição; z_i é a observação no tempo i , sendo $i=1, \dots, n$; n é o tamanho da série.

O MPE (erro percentual médio), que também é conhecido como viés, pois mede o viés de previsão, é dado por:

$$MPE = \left(\frac{1}{n} \sum_{i=1}^n \frac{e_i}{z_i} \right) 100\%.$$

em que z_i é a observação no tempo i , sendo $i=1, \dots, n$; $e_i = z_i - \hat{z}_i$ é o erro de predição; n é o tamanho da série.

A raiz do erro quadrático médio (RMSE), é definido por:

$$RMSE = \left(\frac{1}{n} \sum_{i=1}^n e_i^2 \right)^{1/2}.$$

em que $e_i = z_i - \hat{z}_i$ é o erro de predição; n é o tamanho da série.

A média absoluta dos erros escalonados (MASE), é definido por:

$$MASE = \frac{1}{n} \sum_{i=1}^n \left(\frac{|e_t|}{\frac{1}{n-1} \sum_{i=2}^n |z_i - z_{i-1}|} \right)$$

em que z_i é a observação no tempo i , sendo $i=1, \dots, n$; $e_i = z_i - \hat{z}_i$ é o erro de predição; n é o tamanho da série.

Makridakis e Hibon (1979) e Lawrence et al. (1985) sugerem o MAPE, média percentual de erro absoluto ou erro percentual absoluto médio, trata-se do percentual da diferença entre o previsto e o realizado, em módulo, devido a algumas características, tais como ser independente de escala e menos afetado por “outliers” da previsão. Além disso, esta medida pode ser utilizada para verificar o bom desempenho do ajuste. O MAPE é expresso pela fórmula:

$$MAPE = \left(\frac{1}{n} \sum_{t=1}^n \left| \frac{e_t}{z_t} \right| \right) 100\%.$$

em que z_i é a observação no tempo i , sendo $i=1, \dots, n$; $e_i = z_i - \hat{z}_i$ é o erro de predição; n é o tamanho da série.

Na concepção de Eaves (2002), para comparar previsões de séries diferentes, deve ser utilizado o MAPE, que por ser uma unidade métrica livre, consegue relacionar

o tamanho do erro para uma observação real proporcionalmente. E para Edmundson e O'Connor (1995) também sugere a utilização deste, pois é uma medida mais apropriada nos estudos empíricos, sendo uma medida que menos é afetada por valores extremos, utilizando percentuais do erro e não depende da unidade dos dados.

Segundo Kutner et. al. (2005), sugere-se tomar o modelo mais parcimonioso, ou seja, com o menor número de parâmetros e de acordo com Legarra et al. (2004) o ajuste perfeito aumenta a demanda computacional e esta suscetível a erros numéricos.

2.13 Teste de Box-Pierce

Após a estimação dos modelos, é feita uma análise dos resíduos para verificação adequada dos modelos ajustados. Box e Jenkins (2008) sugere o uso do teste de Box e Pierce (1970), para verificar se os resíduos apresentam um processo de ruído branco. O teste proposto consiste em examinar a $\hat{r}_j$ e a $\hat{r}_j^2$ dos resíduos. Um modelo adequado não deve apresentar valores estatisticamente diferentes de zero para as autocorrelação. O teste de Box e Pierce (1970) é definido por:

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

A hipótese nula para o teste Box-Pierce é de $H_0 : \rho_1 = \dots = \rho_n = 0$ contra a hipótese alternativa de $H_0 : \rho_i \neq 0$ para pelo menos um $i \in \{1, \dots, n\}$. A estatística Q do testes é aproximadamente distribuída como uma χ^2 com $k - p - q$ graus de liberdade, onde k é o número de defasagens tomada. A escolha de k tem grande influência sobre o desempenho do teste de Box-Pierce (TSAY, 2005).

O teste de Box-Pierce é uma alternativa para a verificação se os resíduos são ruído branco de fácil interpretação. É um teste bastante utilizado na literatura, em diferentes áreas aplicadas, no setor financeiro como Guimarães (2014) e Durán (20014) , vazão de microbacia (AMARAL, 2014), alturas de árvores (SILVA, 2008) e com vazão fluvial (BATISTA, 2009).

2.14 Testes para comparação de séries temporais

Em áreas práticas do conhecimento a utilização da teoria de séries temporais nem sempre é suficiente para determinados estudos. A comparação dessas séries é muitas

vezes o maior interesse de estudos de monitoramento de séries coletadas em locais próximos. A verificação da igualdade entre duas séries temporais pode acarretar numa redução de custos operacionais na coleta de dados, e para tanto, existem maneiras de comparar se duas séries diferentes, foram geradas pelo mesmo processo estocástico.

Para tanto Quenouille (1958) e Coates e Diggle (1986), propuseram testes univariados para a comparação da função de autocorrelação e função de densidade espectral, respectivamente. Echeverry e Tolo (2000), aplicaram este teste em séries de temperatura e salinidade provenientes de estações de monitoramento instaladas na Colômbia.

Um segundo teste para comparação de séries temporais, proposto por Silva, Ferreira e Sáfi (2000), utilizado por Costa e Sáfi (2010) para a comparação de séries de material particulado, também será descrito.

2.14.1 Teste de Igualdade das Funções de Autocorrelação

O teste proposto por Quenouille (1958), tem a finalidade de verificar se duas séries temporais distintas tem a mesma estrutura de correlação. Para tanto, considerando o ajuste de um modelo auto regressivo para duas séries $z_1(t)$ e $z_2(t)$, e suas respectivas funções de autocorrelação, $\rho_1(j)$ e $\rho_2(j)$ as hipóteses testadas são:

$$H_0 : \rho_1(j) = \rho_2(j), \text{ para todo } j = 1, 2, \dots, J;$$

$$H_1 : \rho_1(j) \neq \rho_2(j), \text{ para algum } j = 1, 2, \dots, J.$$

O procedimento para a aplicação do teste de Quenouille, como descrito segundo Echeverry e Tolo (2000), consiste em inicialmente calcular as funções de autocorrelação $\hat{\rho}_1(j)$ e $\hat{\rho}_2(j)$ das séries $z_1(t)$ e $z_2(t)$, com $j = 1, 2, \dots, J$.

Em seguida, obtém-se a função de autocorrelação comum de $\hat{\rho}_1(j)$ e $\hat{\rho}_2(j)$ por meio de uma média ponderada de forma $\rho(j) = \frac{n_1 \hat{\rho}_1(j) + n_2 \hat{\rho}_2(j)}{n_1 + n_2}$, onde n_1 e n_2 são os tamanhos das séries $z_1(t)$ e $z_2(t)$.

A partir de $\rho(j)$, calcula-se a função de autocorrelação parcial estimada comum entre as duas séries, $\hat{\Phi}_{kk}$, com a qual identifica-se a ordem p do modelo auto regressivo ajustado.

Resolvendo as equações de *Yule – Walker*, estima-se os p coeficientes do modelo AR (p), ajustando para $z_1(t)$ e $z_2(t)$, o modelo AR (p) correspondente aos coeficientes

comuns encontrados, obtendo-se assim, as chamadas séries residuais $\hat{a}_1$ e $\hat{a}_2$.

Por fim, calcula-se as facp v_j e v'_j para as séries residuais $\hat{a}_1$ e $\hat{a}_2$, respectivamente, para a utilização da estatística teste dada por:

$$SQ = \sum_{j=1}^J \frac{(v_j - v'_j)^2}{\frac{1}{n_1-j} + \frac{1}{n_2-j}} \sim \chi_J^2 \quad (5)$$

A hipótese H_0 é rejeitada, a um determinado nível de significância α , caso $SQ > C_\alpha$, em que C_α é tal que $P(\chi_J^2 > C_\alpha) = \alpha$, com J igual ao número de graus de liberdade da distribuição qui-quadrado.

2.14.2 Método de comparação de Séries Temporais proposto por Silva, Ferreira e Sáfadi (2000)

O procedimento para comparação de duas séries temporais $z_1(t)$ e $z_2(t)$, como proposto por Silva, Ferreira e Sáfadi (2000), utilizado por Costa e Sáfadi (2010).

Inicialmente calcula-se a diferença entre as duas séries dada por $z_d = z_1(t) - z_2(t)$, para $t = 1, 2, \dots$. Aplica-se em seguida a z_d o teste de tendência para verificar a existência de não estacionariedade, o teste de Fisher para verificar se existe sazonalidade e para verificar se os resíduos são independentes e identicamente distribuídos, com média zero e variância constante, aplica-se o teste de Box e Pierce.

Caso z_d seja estacionária, ou seja, não apresente tendência nem sazonalidade, e os resíduos sejam ruído branco, conclui-se que as duas séries são resultado do mesmo processo estocástico, ou seja, são iguais no período analisado.

3 METODOLOGIA

3.1 Material

Os dados experimentais, utilizados neste trabalho, foram obtidos por Santos (2008) na qual avaliou-se a evolução da temperatura da mistura cimento-madeira durante 24 horas em intervalos de 1 minuto. Foram realizados quatro tratamentos para a análise de aptidão das diferentes porcentagens de candeia e eucalipto com o cimento. Do tratamento T1 ao tratamento T3, as partículas receberam tratamento prévio em água fria, durante 24 horas e a massa para cada espécie a ser utilizada no teste foi calculada de acordo com as proporções que seriam utilizadas na confecção dos painéis. Os tratamentos T1, T2 e T3 foram de eucalipto e candeia nas proporções (25% : 75%), (50% : 50%) e (75% : 25%), respectivamente, com tratamento prévio de água e o tratamento T4 foi eucalipto e candeia (25 : 75) sem tratamento prévio de água.

Para análise de aptidão foram colocados dentro de um saco plástico (4×20cm), 200g de cimento Portland CP V - ARI Plus, 15g de partículas secas ao ar em proporções equivalentes às que foram utilizadas na confecção dos painéis, durante 24 horas e para cada mistura foram adicionados 90ml de água. Sem exceder 5 minutos, o material foi misturado e, em cada saco, foi colocado um cabo termopar tipo "J", acoplado a um sistema de aquisição de dados de fabricação da YOCOGAWA, modelo DC100 - 12 - 11 - 1D - Data Collector, no qual os dados foram gerados e armazenados em intervalos de um minuto, pelo período de 24 horas.

Os sacos plásticos contendo a mistura cimento-madeira-água foi envolvido em papel alumínio e colocados em garrafas térmicas que foram revestidas com lã de vidro e acomodadas em uma caixa térmica hermeticamente fechada, a fim de evitar a perda de calor para o ambiente, uma vez que as reações que ocorrem com a mistura cimento-madeira e água são exotérmicas.

A caixa térmica continha quatro garrafas como mostra na Figura 1, cada uma delas com um cabo termopar. Para cada tratamento foram feitas quatro repetições do ensaio de termometria.

Figura 1 – Caixa térmica contendo as amostras

Fonte: SANTOS (2008)

3.2 Método

Dentre as metodologias disponíveis para análise de dados ao longo do tempo, utilizou-se a metodologia séries temporais para o estudo do monitoramento das temperaturas deste experimento, considerando cada tratamento como uma série temporal, ou seja:

- Série 1: média aritmética das temperaturas ($^{\circ}\text{C}$) das quatro repetições do tratamento utilizando 25% de Candeia e de Eucalipto com tratamento prévio de água, com 1448 observações.
- Série 2: média aritmética das temperaturas ($^{\circ}\text{C}$) das quatro repetições tratamento utilizando 50% de Candeia e de Eucalipto com tratamento prévio de água, com 1265 observações.
- Série 3: média aritmética das temperaturas ($^{\circ}\text{C}$) das quatro repetições tratamento

utilizando 75% de Candeia e de Eucalipto com tratamento prévio de água, com 1317 observações.

- Série 4: média aritmética das temperaturas ($^{\circ}\text{C}$) das quatro repetições tratamento utilizando 25% de Candeia e de Eucalipto sem tratamento prévio de água, com 1439 observações.

Para avaliar as séries anteriormente descritas, foram propostos os modelos de Box e Jenkins: Auto-Regressivo (AR), Médias Móveis (MA), Auto-Regressivo com Médias Móveis (ARMA), Auto-Regressivo Integrado com Médias Móveis (ARIMA).

Inicialmente foi realizado um estudo para avaliar a estacionariedade das séries utilizando o correlograma, o teste da raiz unitária de Dickey-Fuller, caso o pressuposto de estacionariedade não fosse satisfeito fez-se o teste da raiz unitária nas séries com diferenças até obter uma série estacionária com d diferenças. Uma vez obtida as séries estacionárias foram avaliadas as funções de autocorrelação (fac) e autocorrelação parcial (fACP), com base nas propriedades teóricas apresentadas na Secção 2.10, foram ajustados modelos propostos para as séries.

Os critérios utilizados na seleção do melhor modelo foram: AIC, BIC e MAPE. Um outro propósito da dissertação foi comparar as séries duas a duas para averiguar se foram geradas pelo mesmo processo estocástico. Para essas análises foram utilizados o teste de Quenoulli e o teste proposto por Silva, Ferreira e Safádi (2000) que consiste em fazer a diferença de séries duas a duas e verificar se estas novas séries são estacionárias.

Os cálculos foram realizados no programa R versão 3.0.2 (R Development Core Team, 2008). Neste software existe um pacote **forecast** específico para séries temporais e contém a função **accuracy** que permite calcular as medidas de erros do modelo ajustado. O pacote **urca** específico para séries temporais e econometria que contém a função **ur.df** que são obtidos o teste de regressão e os valores críticos para o teste de Dickey e Fuller.

4 RESULTADOS E DISCUSSÃO

Nesta seção serão apresentados e discutidos os resultados das análises das séries de temperatura obtidas do experimento de diferentes espécies associadas a resíduos madeireiros na produção de painéis particulados.

4.1 Tratamento utilizando 25% Candeia e 75% Eucalipto com água

Na Figura 2 é apresentada a série de temperatura ($^{\circ}\text{C}$) média do tratamento utilizando 25% de Candeia e 75% de Eucalipto com tratamento prévio das partículas em água. Pode-se observar por meio de uma inspeção visual que a série não apresenta um comportamento estacionário e não se repete em intervalos relativamente constantes de tempo, indicando também ausência de sazonalidade.

Figura 2 – Gráficos da temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água

A aparência de não estacionariedade da série temporal do tratamento em questão, foi inicialmente avaliado por meio de seu correlograma amostral considerando 31 defasagens (Tabela 2). Um aspecto saliente deste correlograma da amostra é que ele inicia com um valor bem alto, cerca de 0,99 na defasagem 1, e diminui muito pouco ao longo das demais defasagens. Mesmo na defasagem 31, isto é, correlação entre os valores de temperatura à distância de 31 minutos, o coeficiente de autocorrelação é ainda alto

sendo 0,96. Este tipo de padrão é geralmente um indicador de que a série temporal é não estacionária.

Tabela 2 – Correlograma da série temporal da temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água

Defasagem	FAC amostral	Defasagem	FAC amostral
0	1	16	0,98
1	0,99	17	0,98
2	0,99	18	0,98
3	0,99	19	0,98
4	0,99	20	0,97
5	0,99	21	0,97
6	0,99	22	0,97
7	0,99	23	0,97
8	0,99	24	0,97
9	0,99	25	0,97
10	0,99	26	0,97
11	0,99	27	0,96
12	0,99	28	0,96
13	0,98	29	0,96
14	0,98	30	0,96
15	0,98	31	0,96

Como foi observado o lento decaimento da autocorrelação sugerindo que a série é não estacionária, testou-se esta conjectura por meio do teste de Dickey-Fuller (DF) em suas três especificações. Para o modelo AR (1) com o termo da constante, os valores críticos para a estatística τ considerando os níveis de significância de 1%, 5% e 10%, respectivamente, foram $-3,43$, $-2,86$, $-2,57$. Assim, o valor calculado de τ foi de $-2,50$ para a série não sendo estatisticamente significativa para nenhum dos níveis de significância considerados, sugerindo que há uma raiz unitária nos dados de temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água.

Para o modelo AR(1) com o termo da tendência, os valores críticos para a estatística τ considerando os níveis de significância de 1%, 5% e 10% , respectivamente, foram $-3,96$, $-3,41$, $-3,12$. Assim, o valor calculado de τ foi de $-2,48$ para a série não é estatisticamente significativa, tendo uma raiz unitária nos dados de temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água.

Para a série na primeira diferença apresentada seu gráfico na Figura 3(a), o modelo do AR(1) sem o termo da constante, os valores críticos para a estatística τ considerando os níveis de significância de 1%, 5% e 10%, respectivamente, foram $-2,58$,

$-1,95$, $-1,62$. Em valores absolutos, o valor de τ foi de $-1,38$, é menor que os valores críticos, de modo que aceita a hipótese nula de que a série na primeira diferença exibe uma raiz unitária.

Para o modelo AR(1) com o termo da constante para a série na primeira diferença, os valores críticos para a estatística τ considerando os níveis de significância de 1%, 5% e 10%, respectivamente, foram $-3,43$, $-2,86$, $-2,57$. Assim, o valor calculado de τ foi de $-1,35$ para a série diferenciada não é estatisticamente significativa. Confirmando que os dados de temperatura do tratamento de 25% Candeia e 75% Eucalipto com água na primeira diferença exibe uma raiz unitária, o que significa que os dados diferenciados são não-estacionários.

Para a série com duas diferenças apresenta-se o gráfico na Figura 3(b), o teste do AR(1) sem o termo da constante, os valores críticos para a estatística τ considerando os níveis de significância de 1%, 5% e 10%, respectivamente, foram $-2,58$, $-1,95$, $-1,62$. Como o valor calculado de τ foi de $-21,95$, que em valores absolutos é maior que os valores críticos, assim rejeita-se a hipótese nula de que a série exibe uma raiz unitária, que é um outro modo de dizer que a série com duas diferenças é estacionária. Confirmando que os dados de temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água na segunda diferença é estacionária.

Figura 3 – Gráficos da série diferenciada da temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água. (a) série na primeira diferença e (b) série na segunda diferença

Como foi visto que a série na segunda diferença, pelo teste da raiz unitária, de temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água é uma

processo estocástico, o que significa que os dados desta série propriamente dita é uma série temporal $I(2)$, essencialmente, é um caminho aleatório.

Com base nos testes de tendência, analisaremos a função de autocorrelação (fac) e autocorrelação parcial ($facp$) na série original e nas diferenças como mostra na Figura 4. Nota-se que a fac para a série original o decaimento é lento para zero indicando que a série não é estacionária. Para fac para a primeira diferença (Figura 4(c)) há um decaimento oscilatório e para segunda diferença (Figura 4(e)) há um decaimento no terceiro lag. E a $facp$ para a primeira (Figura 4(d)) e segunda diferença (Figura 4(e)) há um decaimento exponencial.

Figura 4 – Gráficos das funções fac e facp da série, com uma e duas diferenças, de temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água. (a) fac da série; (b) facp da série; (c) fac da série com uma diferença; (d) facp da série com uma diferença; (e) fac da série com duas diferença e (f) facp da série com duas diferença

Observando a fac e a facp para estudar modelos ARIMA. Verificou-se que a fac (Figura 4(c)) na primeira diferença tem um decaimento lento para zero indicado que a série não é estacionária na primeira diferença como foi confirmado pelo teste da raiz unitária. Analisando estas funções na segunda diferença nota-se que a fac (Figura 4(e)) tem quatro picos significativos e a facp (Figura 4(f)) tem um decaimento exponencial. Para ajustar modelos foi necessário fazer duas diferenças entre observações subsequentes e então ajustou-se modelos do tipo ARIMA considerando termos autoregressivo variando de 0 a 10, e termos de médias móveis variando de 0 a 3.

Foram ajustados diversos modelos do tipo ARIMA para a série das temperatura médias do tratamento de 25% Candeia e 75% Eucalipto com água, nos quais os que apresentaram resíduos que foram estatisticamente um ruído branco estão na Tabela 3. Nesta tabela apresenta-se os modelos com seus respectivos AIC, BIC, MAPE e o p-valor do teste Box-Pierce. Na concepção de Eaves (2002), Edmundson e O'Connor (1995), para comparar previsões de séries diferentes, deve ser utilizado o MAPE, sendo que todos os modelos tiveram um MAPE com valores bem baixos, indicando um bom desempenho do ajuste e apresentaram valores bem próximos para os critérios AIC e BIC. O modelo ARIMA(5, 2, 2) apresentou, em valores absolutos, os menores valores dos critérios de seleção AIC e BIC. Contudo, existe pouca diferença nos valores do AIC e BIC do modelo ARIMA(2, 2, 2), sendo ele mais parcimonioso, segundo Box e Jenkins (1976) sugerem o emprego parcimonioso dos modelos ARIMA, com isso propõe-se como modelo de predição da temperatura de 25% Candeia e 75% Eucalipto com água, o modelo ARIMA(2, 2, 2).

Tabela 3 – Modelos ajustados para a série de temperatura médias do tratamento de 25% Candeia e 75% Eucalipto com água

Modelo	AIC	BIC	MAPE (%)	Teste Box-Pierce (p-valor)
ARIMA(2, 2, 2)	-4945,76	-4917,38	0,08	0,36
ARIMA(3, 2, 2)	-4947,85	-4914,19	0,08	0,77
ARIMA(4, 2, 2)	-4946,33	-4907,40	0,08	0,80
ARIMA(5, 2, 2)	-4944,40	-4900,19	0,08	0,79
ARIMA(2, 2, 3)	-4948,16	-4914,51	0,08	0,78
ARIMA(3, 2, 3)	-4946,18	-4907,25	0,08	0,78
ARIMA(4, 2, 3)	-4944,71	-4900,51	0,08	0,91
ARIMA(5, 2, 3)	-4976,35	-4926,87	0,08	0,41
ARIMA(6, 2, 3)	-4963,44	-4910,68	0,08	0,06

Identificou-se o modelo ARIMA(2, 2, 2) para a predição da série da tempe-

raturas médias do tratamento utilizando 25% de Candeia e 75% de Eucalipto com água. Na Tabela 4 pode-se observar o modelo ARIMA(2,2,2) com seus respectivos parâmetros, segundo Gomes (1987), pode-se afirmar que os erros padrões de cada parâmetros é uma estimativa extremamente confiável já que seus erros padrões (e.p) são bastante pequeno.

Tabela 4 – Ajuste do modelo ARIMA(2,2,2) para a série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água

Parâmetros	ARIMA(2,2,2)	e.p
θ_1	0,45	0,03
θ_2	0,14	0,03
φ_1	-1,84	0,01
φ_2	0,88	0,01

Para ilustrar uma boa aderência do modelo apresenta-se na Figura 5 a respectiva predição com o modelo ARIMA(2,2,2) para a série de temperatura, com a predição de 1 minuto a 1448 minutos, comparadas com temperatura média do tratamento utilizando 25% de Candeia e 75% de Eucalipto com água.

Figura 5 – Gráficos dos dados observados e preditos com o modelo ARIMA(2,2,2) da série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água

Para a predição com o modelo ARIMA(2,2,2) da série analisada, teve como temperatura máxima de 61,78 °C com 552 minutos, e na série observada foi 61,77 °C com

556 minutos.

Foi confirmado que os resíduos são “ruído aleatório” pela estatística Box-Pierce. Essa afirmação é reconhecida com a análise dos gráficos da fac (Figura 6(a)) e da facp (Figura 6(b)) dos resíduos do modelo ARIMA(2, 2, 2), apresentados na Figura 6, observando que fac não teve nenhum lag significativo indicando que os erros são independentes.

Figura 6 – Gráficos das funções fac (a) e facp (b) dos resíduos do modelo ARIMA(2, 2, 2) para a série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água

Na Tabela 5 são apresentados os valores de ME(erro médio), RMSE(erro quadrático médio), MAE(erro absoluto médio) e MPE(erro percentual médio), sendo considerados valores baixos de medidas de erros, indicando uma boa qualidade de ajuste.

Tabela 5 – Medidas de erros do modelo ARIMA(2,2,2) da série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água

ME	RMSE	MAE	MPE	MASE
> -0,01	0,04	0,03	> -0,01	> 0,01

4.2 Tratamento utilizando 50% Candeia e 50% Eucalipto com água

Na Figura 7 é apresentada a série de temperatura ($^{\circ}\text{C}$) média do tratamento utilizando 50% Candeia e 50% Eucalipto com tratamento prévio das partículas em água. Pode-se observar por meio de uma inspeção visual que a série não apresenta um comportamento estacionário e não se repete em intervalos relativamente constantes de tempo, indicando também ausência de sazonalidade.

Figura 7 – Gráfico da temperatura média de 50% Candeia e 50% Eucalipto com água

A série em questão, tem uma aspecto de não estacionariedade, na Tabela 6 fez-se um estudo inicial avaliando por meio de seu correlograma amostral com 31 defasagens. Um aspecto relevante deste correlograma da amostra é que ele inicia com um valor bem alto, cerca de 0,99 na defasagem 1, e diminui muito pouco ao longo das demais defasagens. Mesmo na defasagem 31, isto é, correlação entre os valores de temperatura à distância de 31 minutos, o coeficiente de autocorrelação é ainda alto sendo 0,97. Este tipo de padrão é geralmente um indicador de não estacionária da série temporal.

Este lento declínio da autocorrelação é característica de uma série não estacionária, com isso testou-se esta pressuposição por meio do teste DF em suas especificações. Para o modelo AR(1) com o termo da constante, os valores críticos para a estatística τ considerando os níveis de significância de 1%, 5% e 10%, respectivamente, foram $-3,43$, $-2,86$, $-2,57$. Assim, o valor calculado de τ foi de $-0,50$ para a série não sendo estatisticamente significativa para nenhum dos níveis de significância considerados, sugerindo que há uma raiz unitária nos dados de temperatura do tratamento de 50% Candeia e 50% Eucalipto com água.

Para o modelo AR(1) com o termo da tendência os valores críticos para a estatística τ considerando os níveis de significância de 1%, 5% e 10%, respectivamente, foram $-3,96$, $-3,41$, $-3,12$. Assim, o valor calculado de τ foi de $-0,89$ para a série não

Tabela 6 – Correlograma da série temporal da temperatura média dos tratamentos de 50% Candeia e 50% Eucalipto com água

Defasagem	FAC amostral	Defasagem	FAC amostral
0	1	16	0,99
1	0,99	17	0,98
2	0,99	18	0,98
3	0,99	19	0,98
4	0,99	20	0,98
5	0,99	21	0,98
6	0,99	22	0,98
7	0,99	23	0,98
8	0,99	24	0,98
9	0,99	25	0,98
10	0,99	26	0,97
11	0,99	27	0,97
12	0,99	28	0,97
13	0,99	29	0,97
14	0,99	30	0,97
15	0,99	31	0,97

sendo estatisticamente significativa para nenhum dos níveis de significância considerados, tendo uma raiz unitária nos dados de temperatura do tratamento de 50% Candeia e 50% Eucalipto com água. Indicando, que os valores críticos de DF dependem se há um termo constante e/ou tendência.

Para a série com uma diferença apresentada na Figura 8(a) o teste do AR(1) sem constante os valores críticos para a estatística τ considerando os níveis de significância de 1%, 5% e 10%, respectivamente, foram $-2,58$, $-1,95$, $-1,62$, respectivamente. Como o valor calculado de τ foi de $-4,30$, que em valores absolutos é maior que os valores críticos, rejeitou-se a hipótese nula, assim a série diferenciada não exibe uma raiz unitária, que é um outro modo de dizer que a série diferenciada é estacionária.

Para o modelo AR(1) com constante os valores críticos para a estatística τ considerando os níveis de significância de 1%, 5% e 10%, respectivamente, foram $-3,43$, $-2,86$, $-2,57$. Assim, o valor calculado de τ foi de $-4,63$ para a série diferenciada é estatisticamente significativa, sugerindo que não há uma raiz unitária nos dados diferenciados de temperatura do tratamento de 50% Candeia e 50% Eucalipto com água. Confirmando que os dados de temperatura média do tratamento de 50% Candeia e 50% Eucalipto com água na primeira diferença é estacionária.

Como foi visto que a série na primeira diferença, pelo teste da raiz unitária,

Figura 8 – Gráficos da série diferenciada da temperatura média do tratamento de 50% Candeia e 50% Eucalipto com água. (a) série na primeira diferença e (b) série na segunda diferença

de temperatura do tratamento de 50% Candeia e 50% Eucalipto com água é uma processo estocástico, o que significa que os dados desta série de modo preciso é dita uma série temporal $I(1)$, essencialmente, é um caminho aleatório.

De acordo com os testes de estacionariedade, a seguir é mostrado na Figura 9 a função de autocorrelação (fac) e autocorrelação parcial ($facp$) da série em questão nas diferenças, sendo que a fac na primeira diferença há um decaimento oscilatório (Figura 9(a)) e para segunda diferença há um decaimento no segundo lag (Figura 9(c)). E a $facp$ para a a série original não teve nenhum pico significativo, e para primeira diferença da série a um decaimento oscilatório (Figura 9(b)) e para a segunda diferença há um decaimento exponencial (Figura 9(d)).

Observou a fac e $facp$ para estudar modelos ARIMA, mostrando que estas funções na primeira diferença, Figura 9(a) e Figura 9(b), tem uma forma muito complicada de se interpretar, não sendo útil para identificação de modelos, sendo assim ajustou alguns modelos ARIMA de baixa ordem na primeira diferença. Analisando estas funções na segunda diferença nota-se que a fac (Figura 9(c)) tem dois lag significativos e a $facp$ (Figura 9(d)) tem um decaimento exponencial. Para ajustar modelos foi necessário fazer duas diferenças entre observações subsequentes e então ajustou-se modelos do tipo ARIMA considerando vários termos autoregressivo, e termos de médias móveis variando de 0 a 2.

Foram ajustados diversos modelos do tipo ARIMA para a série de temperatura média do tratamento de 50% Candeia e 50% Eucalipto com água, nos quais os que

Figura 9 – Gráficos das funções fac e facp da série, com uma e duas diferenças, de temperatura média do tratamento de 50% Candeia e 50% Eucalipto com água. (a) fac com uma diferença; (b) facp com uma diferença; (c) fac com duas diferença e (d) facp com duas diferença

apresentaram resíduos que foram estatisticamente um ruído branco estão na Tabela 7. Nesta tabela apresenta-se os modelos com seus respectivos AIC, BIC, MAPE e o p-valor do teste Box-Pierce. Na concepção de Eaves (2002), Edmundson e O'Connor (1995), para comparar previsões de séries diferentes, deve ser utilizado o MAPE, sendo que todos os modelos tiveram um MAPE com valores bem baixos, indicando um bom desempenho do ajuste e apresentaram valores bem próximos para os critérios AIC e BIC. O modelo ARIMA(4, 2, 2) apresentou, em valores absolutos, os menores valores dos critérios de seleção AIC e BIC, sugerindo-o como modelo de predição para a série de temperatura média do tratamento de 50% Candeia e 50% Eucalipto com água.

Tabela 7 – Modelos ajustados para a série de temperatura média do tratamento de 50% Candeia e 50% de Eucalipto com água

Modelo	AIC	BIC	MAPE (%)	Teste Box-Pierce (p-valor)
ARIMA(4, 2, 2)	-4057,73	-4019,75	0,10	0,35
ARIMA(6, 2, 2)	-4060,89	-4012,62	0,10	0,96

Constatou-se o modelo ARIMA(4, 2, 2) para serem utilizada na predição da série média aritmética das temperaturas das quatro repetições do tratamento utilizando 50% de Candeia e 50% de Eucalipto com tratamento prévio das partículas em água. Na Tabela 8 pode-se observar o modelo com seus respectivos parâmetros.

Tabela 8 – Ajuste do modelo ARIMA(4, 2, 2) para a série de temperatura média do tratamento de 50% Candeia e 50% Eucalipto com água

Parâmetros	ARIMA(4, 2, 2)	e.p
θ_1	0,70	0,03
θ_2	0,26	0,03
θ_3	-0,02	0,03
θ_4	-0,11	0,02
φ_1	-1,93	0,01
φ_2	0,94	0,01

Para expor uma boa aderência do modelo apresenta-se na Figura 10 a respectiva predição com o modelo ARIMA(4, 2, 2) para a série de temperatura, com a predição de 1 minuto a 1265 minutos, comparadas com a média aritmética das temperaturas das quatro repetições do tratamento utilizando 50% de Candeia e 50% de Eucalipto com tratamento prévio das partículas em água.

Figura 10 – Gráficos dos dados observados e dados preditos com o modelo ARIMA(4,2,2) da série de temperatura médias do tratamento de 50% Candeia e 50% Eucalipto com água

Para a predição com o modelo ARIMA(4, 2, 2) da série de temperatura média do tratamento de 50% Candeia e 50% Eucalipto com água teve como temperatura máxima de 50,57 °C com 1049 minutos, e na série observada foi de 50,45 °C com 1049 minutos.

Apresenta-se na Figura 11 abaixo, os gráficos da fac (Figura 11(a)) e da facp (Figura 11(b)) dos resíduos do modelo ARIMA(4, 2, 2), sendo que a fac teve poucos lag's significativos, indicando que os erros são independentes e confirmado pela estatística Box-Pierce.

Figura 11 – Gráficos das funções fac (a) e facp (b) dos resíduos do modelo ARIMA(4, 2, 2) para a série de temperatura média do tratamento de 50% Candeia e 50% Eucalipto com água

Na Tabela 9 são apresentados os valores de ME(erro médio), RMSE(erro quadrático médio), MAE(erro absoluto médio) e MPE(erro percentual médio) , sendo considerados valores baixos de medidas de erros, sendo um indicativo da boa qualidade das estimativas.

Tabela 9 – Medidas de erros do modelo ARIMA da série de temperatura média do tratamento de 50% Candeia e 50% Eucalipto com água

ME	RMSE	MAE	MPE	MASE
$> -0,01$	0,04	0,03	$> -0,01$	$> -0,01$

4.3 Tratamento utilizando 75% Candeia e 25% Eucalipto com água

Na Figura 12 é apresentada a série de temperatura ($^{\circ}\text{C}$) média do tratamento utilizando 75% Candeia e 25% Eucalipto com tratamento prévio das partículas em água. Pode-se observar por meio de uma inspeção visual que a série não apresenta um comportamento estacionário e não se repete em intervalos relativamente constantes de tempo, indicando também ausência de sazonalidade.

Figura 12 – Temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água

Na Tabela 10 apresenta o correlograma amostral com 31 defasagens, apresentando uma não estacionariedade da série temporal em estudo. Um aspecto considerável deste correlograma da amostra é que ele inicia com um valor bem alto, cerca de 0,99 na defasagem 1, e decresce muito pouco ao longo das demais defasagens. Mesmo na defasagem 31, isto é, correlação entre os valores de temperatura à distância de 31 minutos, o coeficiente de autocorrelação é ainda alto sendo 0,96. Este tipo de padrão é geralmente um indicador de que a série temporal é não estacionária.

Como foi constatado o lento decaimento da autocorrelação averiguando que a série é não estacionaria, testou-se esta conjectura por meio do teste DF em suas especificações. Para o modelo AR(1) com o termo da constante para a série original, os valores críticos para estatística τ considerando os níveis de significância de 1%, 5% e 10%, respectivamente, foram $-3,43$, $-2,86$, $-2,57$. Assim, o valor calculado de τ foi de $-1,22$ para a série não sendo estatisticamente significativa para nenhum dos níveis de significância considerados, sugerindo que há uma raiz unitária nos dados de temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água.

Para o modelo AR(1) com o termo da tendência, os valores críticos para estatística τ considerando os níveis de significância de 1%, 5% e 10%, respectivamente, foram $-3,96$, $-3,41$, $-3,12$. Assim, o valor calculado de τ foi de $0,79$ para a série não é estatis-

Tabela 10 – Correlograma da série temporal da temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água

Defasagem	FAC amostral	Defasagem	FAC amostral
0	1	16	0,99
1	0,99	17	0,99
2	0,99	18	0,99
3	0,99	19	0,98
4	0,99	20	0,98
5	0,99	21	0,98
6	0,99	22	0,98
7	0,99	23	0,98
8	0,99	24	0,98
9	0,99	25	0,98
10	0,99	26	0,98
11	0,99	27	0,98
12	0,99	28	0,97
13	0,99	29	0,97
14	0,99	30	0,97
15	0,99	31	0,97

ticamente significativa para nenhum dos níveis de significância considerados, confirmando também que a série tem uma raiz unitária nos dados de temperatura do tratamento de 75% Candeia e 25% Eucalipto com água. Indicando, que os valores críticos de DF dependem se há um termo constante e/ou tendência.

Para a série com uma diferença o teste do AR(1) sem o termo da constante, os valores críticos para estatística τ considerando os níveis de significância de 1%, 5% e 10%, respectivamente, foram $-2,58$, $-1,95$, $-1,62$, respectivamente. Como o valor calculado de τ foi de $-7,20$, que em valores absolutos é maior que os valores críticos, rejeita-se a hipótese nula, assim a série diferenciada não exibe uma raiz unitária, que é um outro modo de dizer que a série diferenciada é estacionária.

Figura 13 – Gráficos da série com uma diferença da temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água

Como foi examinado que a série na primeira diferença de temperatura do tratamento de 75% Candeia e 25% Eucalipto com água é uma processo estocástico, o que significa que os dados desta série são propriamente dita que é uma série temporal $I(1)$, essencialmente, é um caminho aleatório.

A seguir apresenta Figura 14 com a função de autocorrelação (fac) e autocorrelação parcial (facp) para a primeira e segunda diferença. Nota-se que a fac para a primeira diferença há um decaimento oscilatório (Figura 14(a)) e para segunda diferença há um decaimento no segundo lag (Figura 14(c)). E a facp para a primeira diferença a um decaimento oscilatório (Figura 14(b)).

Utilizou a fac e facp para estudar modelos ARIMA, mostrando na Figura 14(a) que a fac para da série na primeira diferença tem padrão de ondas senoidais amortecidas e na Figura 14(b) que a facp tem vários picos significativos, segundo Morettin e Tolo (2006) quando há uma forma complicada da fac e facp, o recomendado é ajustar alguns modelos de baixa ordem, no qual ajustou-se modelos do tipo ARIMA com uma diferença, considerando vários termos autoregressivo. Analisando estas funções na segunda diferença nota-se que a fac (Figura 14(c)) tem dois lag significativos e a facp (Figura 14(d)) tem um decaimento exponencial. Também ajustou modelos ARIMA que foi necessário fazer duas diferenças entre observações subsequentes, considerando vários termos autoregressivo, e termos de médias móveis variando de 0 a 2.

Foram ajustados diversos modelos do tipo ARIMA para a série das temperatura médias do tratamento de 75% Candeia e 25% Eucalipto com água, nos quais os que

Figura 14 – Gráficos das funções fac e facp da série, com uma e duas diferenças, de temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água. (a) fac com uma diferença; (b) facp com uma diferença; (c) fac com duas diferença e (d) facp com duas diferença

apresentaram resíduos que foram estatisticamente um ruído branco estão na Tabela 11. Nesta tabela apresenta-se os modelos com seus respectivos AIC, BIC, MAPE e o p-valor do teste Box-Pierce. Na concepção de Eaves (2002), Edmundson e O'Connor (1995), para comparar previsões de séries diferentes, deve ser utilizado o MAPE, sendo que todos os modelos tiveram um MAPE com valores bem baixos, indicando um bom desempenho do ajuste e apresentaram valores bem próximos para os critérios AIC e BIC. O modelo ARIMA(5, 1, 0) apresentou, em valores absolutos, os menores valores dos critérios de seleção AIC e BIC, sugerindo-o como modelo de predição para a série das temperatura médias do tratamento de 75% Candeia e 25% Eucalipto com água.

Tabela 11 – Modelos ajustados para a série das temperatura médias do tratamento de 75% Candeia e 25% Eucalipto com água

Modelo	AIC	BIC	MAPE (%)	Teste Box-Pierce (p-valor)
ARIMA(5,1,0)	-2830,58	-2797,50	0,15	0,14
ARIMA(2,1,1)	-2885,01	-2862,29	0,14	0,37
ARIMA(2,1,2)	-2884,38	-2856,47	0,14	0,47
ARIMA(0,2,2)	-2877,71	-2860,17	0,14	0,20

Identificou-se o modelo ARIMA(5, 1, 0) para serem utilizada na predição da série média aritmética das temperaturas das quatro repetições do tratamento utilizando 75% de Candeia e 25% de Eucalipto com tratamento prévio das partículas em água. Na Tabela 12 pode-se observar o modelo com seus respectivos parâmetros.

Tabela 12 – Ajuste do modelo ARIMA(5, 1, 0) para a série de temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água

Parâmetros	ARIMA(5, 1, 0)	e.p
θ_1	-0,06	0,02
θ_2	0,12	0,02
θ_3	0,10	0,02
θ_4	0,09	0,02
θ_5	0,09	0,02

Para ilustrar uma boa aderência do modelo apresenta-se na Figura 15 a respectiva predição com o modelo ARIMA(5, 1, 0) para a série de temperatura, com a predição de 1 minuto a 1317 minutos, comparadas com a média aritmética das temperaturas das quatro repetições do tratamento utilizando 75% de Candeia e 25% de Eucalipto com água.

Figura 15 – Gráficos dos dados observados e dados preditos com o modelo ARIMA(5,1,0) da série de temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água

Para a predição com o modelo ARIMA(4, 2, 2) da série de temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água teve como temperatura máxima predita de 45,80 °C com 1010 minutos, e na série observada de 45,80 °C com 1015 minutos

A seguir apresenta-se na Figura 16 os gráficos da fac (Figura 16(a)) e da facp (Figura 16(b)) dos resíduos do modelo ARIMA(5, 1, 0), sendo que a fac teve poucos lag's significativos, indicando que os erros são independentes e confirmado pela estatística Box-Pierce.

Figura 16 – Gráficos das funções fac (a) e facp (b) dos resíduos do modelo ARIMA(5, 1, 0) para a série de temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água

Na Tabela 13 são apresentados os valores de ME(erro médio), RMSE(erro quadrático médio), MAE(erro absoluto médio) e MPE(erro percentual médio), sendo considerados valores baixos de medidas de erros, um indicativo da boa qualidade das estimativas.

Tabela 13 – Medidas de erros do modelo ARIMA da série de temperatura média do tratamento de 75% Candeia e 25% Eucalipto com água

ME	RMSE	MAE	MPE	MASE
$> -0,01$	0,08	0,05	0,01	$> -0,01$

4.4 Tratamento utilizando 25% Candeia e 75% Eucalipto sem água

Na Figura 17 é apresentada a série de temperatura ($^{\circ}\text{C}$) média do tratamento utilizando 25% Candeia e 75% Eucalipto sem tratamento prévio das partículas em água. Pode-se observar por meio de uma inspeção visual que a série não apresenta um comportamento estacionário e não se repete em intervalos relativamente constantes de tempo, indicando também ausência de sazonalidade.

Figura 17 – Temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água

Com o aspecto de não estacionariedade da série temporal em estudo, apresenta-se na Tabela 14 a avaliação por meio de seu correlograma amostral considerando 31 defasagens, na qual o aspecto deste correlograma é que ele inicia com um valor bem alto, cerca de 0,99 na defasagem 1, e decresce muito pouco ao longo das demais defasagens. Mesmo na defasagem 31, isto é, correlação entre os valores de temperatura à distância de 31 minutos, o coeficiente de autocorrelação é ainda alto sendo 0,96. Este tipo de padrão é geralmente um indicador de que a série temporal é não estacionária.

Atentando ao lento decaimento da autocorrelação, no qual indicou que a série é não estacionária, testou-se esta conjectura por meio do teste DF em suas três especificações. Para o modelo AR(1) com constante, os valores críticos para a estatística τ considerando os níveis de significância de 1%, 5% e 10%, respectivamente, foram $-3,43$, $-2,86$, $-2,57$. Assim, o valor calculado de τ foi de $-1,29$ para a série não é estatisticamente significativa para nenhum dos níveis de significância considerados, sugerindo que há uma raiz unitária nos dados de temperatura do experimento de 25% Candeia e 75% Eucalipto sem água.

Para o modelo AR(1) com tendência, os valores críticos para a estatística τ considerando os níveis de significância de 1%, 5% e 10%, respectivamente, foram $-3,96$, $-3,41$, $-3,12$. Assim, o valor calculado de τ foi de $-0,68$ para a série não é estatisticamente

Tabela 14 – Correlograma da série temporal da temperatura de 25% Candeia e 75% Eucalipto sem água

Defasagem	FAC amostral	Defasagem	FAC amostral
0	1	16	0,99
1	0,99	17	0,98
2	0,99	18	0,98
3	0,99	19	0,98
4	0,99	20	0,98
5	0,99	21	0,98
6	0,99	22	0,98
7	0,99	23	0,98
8	0,99	24	0,98
9	0,99	25	0,97
10	0,99	26	0,97
11	0,99	27	0,97
12	0,99	28	0,97
13	0,99	29	0,97
14	0,99	30	0,97
15	0,99	31	0,97

significativa para nenhum dos níveis de significância considerados, tendo uma raiz unitária nos dados de temperatura do experimento de 25% Candeia e 75% Eucalipto sem água. Indicando, que os valores críticos de DF dependem se há um termo constante e/ou tendência.

Para a série na primeira diferença o modelo do AR(1) sem constante os valores críticos para a estatística τ considerando os níveis de significância de 1%, 5% e 10%, respectivamente, foram $-2,58$, $-1,98$, $-1,62$, respectivamente. Em valores absolutos, o valor de τ foi de $-3,18$, excede qualquer um desses valores críticos, de modo que rejeita-se a hipótese que a série tem raiz unitária, ou seja, os dados de temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água na primeira diferença não exibem uma raiz unitária, o que significa que os dados diferenciados são estacionários.

Como foi averiguado que a série na primeira diferença, pelo teste da raiz unitária, de temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água é um processo estocástico, o que significa que os dados de temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água propriamente dita é uma série temporal I(1), essencialmente, é um caminho aleatório.

Com base nos testes de tendência, analisaremos a função de autocorrelação (fac) e autocorrelação parcial (facp) na primeira e segunda diferença como mostra na Figura 19. Nota-se que a fac para a primeira diferença há um decaimento oscilatório e para segunda diferença há um decaimento no segundo lag. E a facp para a primeira diferença a

Figura 18 – Gráficos da série diferenciada da temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água. (a) série na primeira diferença e (b) série na segunda diferença

um decaimento oscilatório e para segunda há um decaimento exponencial.

Observou a facp e facp para estudar modelos ARIMA, mostrando que estas funções na primeira diferença, Figura 19(a) e Figura 19(b), tem uma forma muito complicada de se interpretar, não sendo útil para identificação de modelos, sendo assim ajustou alguns modelos ARIMA de baixa ordem na primeira diferença. Analisando estas funções na segunda diferença nota-se que a fac (Figura 19(c)) tem dois picos significativos e a facp (Figura 19(d)) tem um decaimento exponencial. Para ajustar modelos foi necessário fazer duas diferenças entre observações subsequentes e então ajustou-se modelos do tipo ARIMA considerando vários termos autoregressivo, e termos de médias móveis variando de 1 a 2.

Foram ajustados diversos modelos do tipo ARIMA para a série das temperatura médias do tratamento de 25% Candeia e 75% Eucalipto sem água, nos quais os que apresentaram resíduos que foram estatisticamente um ruído branco estão na Tabela 15. Nesta tabela apresenta-se os modelos com seus respectivos AIC, BIC, MAPE e o p-valor do teste Box-Pierce. Na concepção de Eaves (2002), Edmundson e O'Connor (1995), para comparar previsões de séries diferentes, deve ser utilizado o MAPE, sendo que todos os modelos tiveram um MAPE com valores bem baixos, indicando um bom desempenho do ajuste e apresentaram valores bem próximos para os critérios AIC e BIC.

O modelo ARIMA(2, 2, 2) apresentou, em valores absolutos, os menores valores dos critérios de seleção AIC e BIC. Contudo, existe pouca diferença nos valores do AIC e BIC do modelo ARIMA(2, 1, 2), sendo ele mais parcimonioso, segundo Box e Jenkins

Figura 19 – Gráficos das funções fac e facp da série, com uma e duas diferenças, de temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água. (a) fac com uma diferença; (b) facp com uma diferença; (c) fac com duas diferença e (d) facp com duas diferença

(1976), Kutner et. al. (2005) e Legarra et al. (2004) sugerem o emprego parcimonioso dos modelos ARIMA, com isso propõe-se como modelo de predição da temperatura de 25% Candeia e 75% Eucalipto sem água, o modelo ARIMA(2, 2, 2).

Identificou-se o modelo ARIMA(2, 1, 2) para serem utilizada na predição da série média aritmética das temperaturas das quatro repetições do tratamento utilizando 25% de Candeia e 75% de Eucalipto sem tratamento prévio das partículas em água. Na Tabela 16 pode-se observar o modelo com seus respectivos parâmetros.

Tabela 15 – Modelos ajustados para a série das temperatura médias do tratamento de 25% Candeia e 75% Eucalipto sem água

Modelo	AIC	BIC	MAPE (%)	Teste Box-Pierce (p-valor)
ARIMA(2, 1, 2)	-3309,96	-3281,61	0,15	0,06
ARIMA(2, 2, 2)	-3304,48	-3276,13	0,15	0,06

Tabela 16 – Ajuste do modelo ARIMA(2, 1, 2) para a série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água

Parâmetros	ARIMA(2, 1, 2)	s.e
θ_1	1,07	0,07
θ_2	-0,07	0,07
φ_1	-1,33	0,06
φ_2	0,39	0,06

Para ilustrar uma boa aderência do modelo apresenta-se na Figura 20 a respectiva predição com o modelo ARIMA(2, 1, 2) para a série de temperatura, com a predição de 1 minuto a 1439 minutos, comparadas com a média aritmética das temperaturas das quatro repetições do tratamento utilizando 25% de Candeia e 75% de Eucalipto sem tratamento prévio das partículas em água.

Figura 20 – Gráficos dos dados observados e dados preditos com o modelo ARIMA(2, 1, 2) da série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água

Para a predição com o modelo ARIMA(2, 1, 2) da série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto com água teve como temperatura máxima predita de 57,63 °C com 718 minutos, e na série observada de 57,62 °C com 727 minutos

A seguir apresenta-se na Figura 21 os gráficos da fac (Figura 21(a)) e da facp (Figura 21(b)) dos resíduos do modelo ARIMA(2, 1, 2), sendo que a fac teve poucos lag's significativos indicando que os erros são independentes, conferida pela estatística Box-Pierce.

Figura 21 – Gráficos das funções fac e facp dos resíduos do modelo ARIMA(2, 1, 2) para a série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água. (a) fac dos resíduos do modelo ARIMA(2, 1, 2); (b) facp dos resíduos do modelo ARIMA(2, 1, 2)

Na Tabela 17 são apresentados os valores de ME(erro médio), RMSE(erro quadrático médio), MAE(erro absoluto médio) e MPE(erro percentual médio), sendo considerados valores baixos de medidas de erros sendo um indicativo da qualidade das estimativas.

Tabela 17 – Medidas de erros do modelo ARIMA da série de temperatura média do tratamento de 25% Candeia e 75% Eucalipto sem água

ME	RMSE	MAE	MPE	MASE
> -0,01	0,07	0,05	> 0,01	> 0,01

4.5 Teste de comparação de séries temporais

As séries utilizadas para esse teste foram constituídos por:

- Série 1 (s1): temperaturas média do tratamento de 25% de Candeia e 75% de Eucalipto com água, com 1448 observações. ;
- Série 2 (s2): temperaturas média do tratamento de 50% de Candeia e 50% de Eucalipto com água, com 1265 observações ;
- Série 3 (s3): temperaturas média do tratamento de 75% de Candeia e 25% de Eucalipto com água, com 1317 observações;
- Série 4 (s4): temperaturas média do tratamento de 25% de Candeia e 75% de Eucalipto sem água, com 1439 observações.

Com o intuito de verificar se duas séries são formadas pelo mesmo processo estocástico, foram utilizadas duas abordagens. Na primeira, foi utilizado o teste de Quenouille, tem a finalidade de verificar se duas séries temporais distintas tem a mesma estrutura de correlação. Os resultados obtidos pelos testes estão presentes na Tabela 18, na qual pela estatística do teste de Quenouille são significativas, ou seja, as séries tem diferentes processos estocástico.

Tabela 18 – Teste de Quenouille nas séries diferenciadas duas a duas

Séries	Estatística	p-valor
Série 1 e Série 2	238,74	0
Série 1 e Série 3	522,53	0
Série 1 e Série 4	78,83	>0,005
Série 2 e Série 3	184,19	0
Série 2 e Série 4	41,68	>0,005
Série 3 e Série 4	50,99	>0,005

Na segunda abordagem, foi utilizado o método proposto por Silva, Ferreira e Sáfadi (2000), na qual propuseram uma alternativa de comparação que foi utilizada para avaliar séries de preços ao consumir na cidade de Lavras. O procedimento de realização do teste proposto pelos autores consiste, inicialmente, em tomar a diferença entre as série 1 pela série 2, série 1 pela série 3, e assim até a diferença da série 3 pela série 4. Formando seis séries resultantes destas diferenças, sendo apresentadas os gráficos destas séries resultantes na Figura 22. Pode se observar por meio de uma inspeção visual que as séries formadas não apresentam um comportamento estacionário.

Figura 22 – Gráficos das séries diferenciadas duas a duas (a) série 1 - série 2, (b) série 1 - série 3, (c) série 1 - série 4, (d) série 2 - série 3 (e) série 2 - série 4 e (f) série 3 - série 4

A aparência de não estacionariedade das séries formadas pela diferenças duas a duas das séries originais, com isso utilizou-se o teste de Dickey e Fuller nas suas três especificações para verificar se tem raiz unitária, em que os resultados obtidos pelos testes de DF estão presentes nas Tabelas 19 no qual os valores críticos para a estatística τ considerando o nível de 5%, respectivamente, foram de $-1,95$, $-2,86$ e $-3,41$. Para cada teste apresenta também os valores de τ calculados para as séries diferenciadas duas a duas, na qual estes valores de τ calculados foram menores que os valores críticos da estatística τ , com isso aceita a hipótese que tem raiz unitária, ou seja, todas as séries diferenciadas duas a duas são não estacionárias.

Como as séries diferenciadas duas a duas são não estacionárias, conclui-se que as séries não são resultantes do mesmo processo estocástico, ou seja, os tratamentos produzem resultados diferentes nesta análise.

Tabela 19 – Teste da raiz unitária para as séries diferenciadas

Teste	Estatística						τ Crítico (0,05)
	s1 e s2	s1 e s3	s1 e s4	s2 e s3	s2 e s4	s3 e s4	
DF: AR(1) sem constante	-1,50	-1,00	-0,16	-0,01	-1,09	-0,83	-1,95
DF: AR(1) com constante	1,49	-1,04	-0,02	-0,10	-1,05	-0,87	-2,86
DF: AR(1) com tendência	1,80	-1,44	-1,30	-0,83	-1,07	-0,77	-3,41

São apresentadas as funções de autocorrelação (fac) e autocorrelação parcial (facp) na Figura 23 para as séries obtidas pela diferença das séries duas a duas, na qual observa-se que as fac's tem decaimento lento para zero indicando a não estacionariedade.

Figura 23 – Gráficos das funções da fac e facp (a) s1-s2, (b) s1-s3, (c) s1-s4, (d) s2-s3 (e) s2-s4 e (f) s3-s4

5 CONCLUSÕES

O presente trabalho avaliou quatro conjuntos de dados experimentais referentes a séries temporais de dados da evolução da temperatura da mistura cimento-madeira durante 24 horas em intervalos de 1 minuto. As análises destes dados permitiram concluir que séries são não estacionárias e que as mesmas são diferentes entre si, ou seja, não foram geradas pelo mesmo processo estocástico.

Os modelos ARIMA são capazes de prever com boa acurácia as temperaturas da mistura cimento-madeira, mostrando-se boa ferramenta para profissionais que trabalham com painéis de resíduos madeireiros e necessitam de informações sobre as temperaturas de aderência do cimento com a madeira.

O presente estudo não contempla a caracterização dos resíduos atípicos, possibilitando futuros estudos que compreendam a avaliação de pontos extremos, assim como o estudo da longa dependência temporal.

REFERÊNCIAS

- ABREU, P. H. F. de. **Formas de aproveitamento dos resíduos da madeira**. 2005. 45 p. Monografia (Trabalho de Conclusão do Curso de Engenharia de Produção) - Universidade Estadual de Maringá, Maringá, 2005.
- ACZEL, A. D. **Complete Business Statistics**, New York: Irwin Homewood, 1993. 877 p.
- AKAIKE, H. Maximum likelihood identification of Gaussian autoregressive moving average models. **Biometrika**, Oxford, v. 60, n. 2, p. 255-265, 1973.
- _____. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, Piscataway, v. 19, n. 6, p. 716-723, 1974.
- AMARAL, M. V. S. G. do. **Ajustes de modelos e comparação de séries temporais para dados de vazão específica em microbacias pareadas**. 2014. 76 p. Dissertação (Mestrado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2014.
- BATISTA, A. L. F. **Modelos de séries temporais e redes neurais artificiais na previsão de vazão**. 2009. 79 p. Dissertação (Mestrado em Engenharia de Sistemas) - Universidade Federal de Lavras, Lavras, 2009.
- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time series analysis: forecasting and control**. New Jersey: Prentice Hall, 1994. 598 p.
- BOX, G. E. P.; JENKINS, G. **Time Series Analysis: Forecasting and Control**. 4. ed. Hoboken: John Wiley & Sons, 2008. 746 p.
- BOX, G. E. P.; PIERCE, D. A. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. **Journal of the American Statistical Association**, Alexandria, v. 65, n. 332, p. 1509 - 1526, 1970.
- BUENO, R. L. S. **Econometria de Séries Temporais**. São Paulo: Cengage, 2011. 360 p.
- CARNEIRO, R. G. **Perfil da temperatura do solo nos biomas florestais da amazônia e mata atlântica com aplicação da transformada em ondeletas**. 2014. 79 p. Dissertação (Mestrado em Meteorologia) - Universidade Federal de Campina Grande, Campina Grande, 2014.
- COATES, D. S.; DIGGLE, P. J. Tests for comparing two estimated spectral densities. **Journal of Time Series Analysis**, Hoboken, v. 7, n. 1, p. 7 - 20, 1986. Disponível em: < <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9892.1986.tb00482.x/abstract> >. Acesso em: 01 dez. 2014.
- COSTA, F. M.; SÁFADI, T. Comparação estatística de duas séries de material particulado (mp10) na cidade de São Paulo. **Revista Brasileira de Biometria**, São Paulo, v. 28, n. 3, p. 23 - 38, 2010.
- DICKEY, D.; FULLER, W. Distribution of the estimators for autorregressive time series with a unit root. **Journal of the American Statistical Association**, Alexandria, v. 74, n. 366, p. 427-431, 1979.

DURÁN, W. G. R. **Modelo GARCH com mudanças de regime markoviano para séries financeiras**. 2014. 56 p. Dissertação (Mestrado em Matemática e Estatística) - Universidade de São Paulo, São Paulo, 2014.

DURBIN, J. The fitting of Time-Series models. **Review of the International Statistical Institute**, Lyon, v. 28. n. 3, p. 233-244, 1960.

DUTRA, R. C.; FERRAZ, S. O.; PIMENTA, D. S.; SOUSA, O. V. Caracterização morfoanatômica das folhas de *Eremanthus erythropappus* (DC.) MacLeish, Asteraceae. **Revista Brasileira de Farmacognosia**, Curitiba, v. 20, n. 6, p. 818-824, 2010.

EAVES, A. H. C. **Forecasting for the ordering and stock-holding of consumable spare parts**. 2002. 355 p. Tese (Doutorado em Ciência da Administração) - Universidade de Lancaster, Lancaster, Inglaterra, 2002.

ECHEVERRY, G. H. S. **Métodos de comparação de séries temporais**. 1999. 131 p. Dissertação (Mestrado em Estatística) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 1999.

ECHEVERRY, G. H. S.; TOLOI, C. M. C. Testes para comparação de séries temporais: uma aplicação a séries de temperatura e salinidade da água, medidas em profundidades diferentes. **Revista Brasileira de Estatística**, São Paulo, v. 61, n. 215, p. 51-80, 2000.

EDMUNDSON, B.; O'CONNOR, M. Forecasting support systems for the incorporation of event information: an Empirical investigation. **International Journal of forecasting**, Sydney, v. 21, n. 3, p. 411-423, 2004.

EHLERS, R. S. **Análise de Séries Temporais**. 2009. 114 p. Apostila. Disponível em: <<http://www.icmc.usp.br/ehlers/stemp/stemp.pdf>>. Acesso em: 01 dez. 2014.

FONTES, P. J. P. de. **Auto-suficiência energética em serraria de *Pinus* e aproveitamento de resíduos**. 1994. 140 p. Dissertação (Mestrado em Ciências Florestais) - Universidade Federal do Paraná, Curitiba, 1994.

FULLER, W. A. **Introduction to the statistical time series**. Hoboken: John Wiley & Sons, 1976. 728 p.

GOMES, F. C. Os modelos Arima e a abordagem de Box-Jenkins: uma aplicação na previsão do IBOVESPA a curtíssimo prazo. **Revista de Administração de Empresas**, São Paulo, v. 29, n. 2, p. 63-70, 1989.

GUIMARÃES, P. H. S. **Modelos de volatilidade com inovações skew-T**. 2014. 130 p. Tese (Doutorado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2014.

GUJARATI, D. N. **Econometria básica**. São Paulo: Markron, 2000. 846 p.

KUTNER, M. H.; NACHTSHEIM, C. J.; NETER, J. W. **Applied Linear Statistical Models**. Michigan: McGrawHill, 2005. 1408 p.

LAWRENCE, M.; EDMUNDSON, B.; O'CONNOR, M. An Examination of the accuracy of judgemental extrapolation of time series. **International Journal of Forecasting**, Sydney, v. 1, n. 1, p. 25-35, 1985.

- LEGARRA, A.; MISZTAL, I.; BERTRAND, J. K. Constructing covariance functions for random regression models for growth in Gelbvieh beef cattle. **Journal of Animal Science**, Champaign, v. 82, n. 6, p. 1564-1571, 2004.
- LOPES, C. S. D. Desenho de pequenos objetos de madeira com resíduo da indústria de processamento mecânico da madeira. **Revista de Gestão Integrada em Saúde do Trabalho e Meio Ambiente - INTERFACEHS**. São Paulo, v. 4, n. 3, p. 105-132, 2009.
- MACEDO, D. G. **Compósitos de serragem de madeira e resíduos de recauchutagem de pneu**. 2008. 144 p. Tese (Doutorado em Engenharia Florestal) - Faculdade de Tecnologia, Universidade de Brasília, Brasília, 2008.
- MAKRIDAKIS, S.; M. HIBON. Accuracy of forecasting: an empirical investigation. **Journal of the Royal Statistical Society**, Hoboken, v. 142, n. 2, p. 97-145, 1979.
- MARGARIDO, M. A.; ANEFALOS, L. C. Teste da raiz unitária e o software SAS. **Instituto de Economia Agrícola**, São Paulo, v. 46, n. 2, p. 19-45, 1999.
- MORETTIN, P. A.; TOLOI, C. M. C. **Análises de séries temporais**. São Paulo: Edgard Blucher, 2006. 556 p.
- PAULA, J. C. M. de. **Aproveitamento de resíduos de madeira para confecção de briquetes**. 2006. 37 p. Monografia (Trabalho de Conclusão do Curso de Engenharia Florestal) - Universidade Federal Rural do Rio de Janeiro, Rio de Janeiro, 2006.
- PHILLIPS, P. C. B.; PERRON, P. Testing for a unit root in time series regression. **Biometrika**, Oxford, v. 75, n. 2, p. 335-346, 1988.
- QUENOUILLE, M. H. The comparison of correlations in time-series. **Journal of the Royal Statistical Society**, London, v. 20, n. 1, p. 158-164, 1958.
- R Development Core Team. **R: A language and environment for statistical computing**. Vienna, Austria, 2008. Disponível em: <<http://www.R-project.org>>. Acesso em: 18 ago. 2014.
- RISSANEN, J. Modelling by shortest data description. **Automatica**, Oxford, v. 14, n. 5, p. 465-471, 1978.
- RIZZINI, C. T. **Árvores e madeiras úteis do Brasil: manual de dendrologia brasileira**. São Paulo: Premier, 1979. 312p.
- SAID. S.; DICKEY, D. Testing for unit root in autoregressive-moving average models of unknown order. **Biometrika**, Oxford, v. 73, n. 3, p. 599-607, 1984.
- SANTOS, R. C. **Aproveitamento de resíduos da madeira de candeia (*Eremanthus erythropappus*) para produção de chapas de partículas**. 2008. 159 p. Dissertação (Mestrado em Ciência e Tecnologia da Madeira) - Universidade Federal de Lavras, Lavras, 2008.
- SCHWARZ, G. Estimating the demension of a model. **Annals of Statistics**, Cleveland, v. 6, n. 2, p. 461-464, 1978.
- SCOLFORO, J. R. S; OLIVEIRA, A. D.; DAVIDE, A. C. **O Manejo sustentável da Candeia: o caminhar de uma nova experiência florestal em Minas Gerais**. Lavras: UFLA, 2012. 329 p.

SILVA, L. R. da. **Correlações de longo alcance em séries temporais de focos de calor no Brasil**. 2009. 55 p. Dissertação (Mestrado em Biometria e Estatística Aplicada) - Universidade Federal Rural de Pernambuco, Recife, 2009.

SILVA, R. B. V.; FERREIRA, D. F.; SÁFADI, T. Modelos de séries temporais aplicados à série dos índices de preços ao consumidor na região de lavras, mg, no período de 1992 a 1999. **Organizações Rurais & Agroindustriais**, Lavras, v. 2, n. 2, p. 44-55, 2000. Disponível em: <<http://revista.dae.ufla.br/index.php/ora/article/viewArticle/283> >. Acesso em: 20 set. 2014.

TSAY, R. S. **Analysis of financial time series**. 2. ed. New York: John Wiley & Sons, 2005. 712 p.

WEI, W. W. S. **Time Series Analysis**: univariate and multivariate methods. New York: Pearson, 2006. 624 p.

APÊNDICE

APÊNDICE A - Script do *software* R

```

rm(list=ls(all=TRUE))
library (forecast)
require(urca)
library(randtests)

#####
#Candeia + Eucalipto 25%:75% #Com Agua
#####
dados=read.table("ComAgua_Candeia+eucalipto_25-75.txt",h=T)
dados
s1=dados$X ;s1
n1=length(s1) ;n1
#####
#Candeia + Eucalipto 50%:50% #Com Agua
#####

dados=read.table("ComAgua_Candeia+eucalipto_50-50.txt",h=T)
dados
s2=dados$X ;s2
n2=length(s2) ;n2
#####
#Candeia + Eucalipto 75%:25% #Com Agua
#####
dados=read.table("ComAgua_Candeia+eucalipto_75-25.txt",h=T)
dados
s3=dados$X ;s3
n3=length(s3) ;n3
#####
#Candeia + Eucalipto 25%:75% #Sem Agua
#####
dados=read.table("SemAgua.Candeia+Eucalipto.txt",h=T)
dados
s4=dados$X ;s4
n4=length(s4) ;n4
#####
####                                Gráficos                                #####
#####

plot(s1,xlab = "minutos", ylab= "Temperatura řC",pch=16,
     type="o",lwd=3,las=1)
plot(diff(s1),xlab = "minutos", ylab= "Temperatura řC",
     pch=16,type="o",lwd=3,las=1)
plot(diff(diff(s1)),xlab = "minutos", ylab= "Temperatura řC",
     pch=16,type="o",lwd=3,las=1)

plot(s2,xlab = "minutos", ylab= "Temperatura řC",
     pch=16,
     type="o",lwd=3,las=1)
plot(diff(s2),xlab = "minutos", ylab= "Temperatura řC",
     pch=16,type="o",lwd=3,las=1)
plot(diff(diff(s2)),xlab = "minutos", ylab= "Temperatura řC",
     pch=16,type="o",lwd=3,las=1)

plot(s3,xlab = "minutos", ylab= "Temperatura řC",

```

```

    pch=16,type="o",lwd=3,las=1)
plot(diff(s3),xlab = "minutos", ylab= "Temperatura řC",
    pch=16,type="o",lwd=3,las=1)
plot(diff(diff(s3)),xlab = "minutos", ylab= "Temperatura řC",
    pch=16,type="o",lwd=3,las=1)

plot(s4,xlab = "minutos", ylab= "Temperatura řC",
    pch=16,type="o",lwd=3,las=1)
plot(diff(s4),xlab = "minutos", ylab= "Temperatura řC",
    pch=16,type="o",lwd=3,las=1)
plot(diff(diff(s4)),xlab = "minutos", ylab= "Temperatura řC",
    pch=16,type="o",lwd=3,las=1)

#####
####                               Teste da Raiz Unitária                               ####
#####

(adfn=summary(ur.df(s1,type="none",lags=10,selectlags=c("AIC"))))
(adfd=summary(ur.df(s1,type="drift",lags=10,selectlags=c("AIC"))))
(adft=summary(ur.df(s1,type="trend",lags=10,selectlags=c("AIC"))))
(adfn_d=summary(ur.df(diff(s1),type="none",lags=10,selectlags=c("AIC"))))
(adfd_d=summary(ur.df(diff(s1),type="drift",lags=10,selectlags=c("AIC"))))
(adfn_2d=summary(ur.df(diff(diff(s1)),type="none",lags=10,selectlags=c("AIC"))))
(adfd_2d=summary(ur.df(diff(diff(s1)),type="drift",lags=10,selectlags=c("AIC"))))

(adfn=summary(ur.df(s2,type="none",lags=10,selectlags=c("AIC"))))
(adfd=summary(ur.df(s2,type="drift",lags=10,selectlags=c("AIC"))))
(adft=summary(ur.df(s2,type="trend",lags=10,selectlags=c("AIC"))))
(adfn_d=summary(ur.df(diff(s2),type="none",lags=10,selectlags=c("AIC"))))
(adfd_d=summary(ur.df(diff(s2),type="drift",lags=10,selectlags=c("AIC"))))
(adft_d=summary(ur.df(diff(s2),type="trend",lags=10,selectlags=c("AIC"))))

(adfn=summary(ur.df(s3,type="none",lags=10,selectlags=c("AIC"))))
(adfd=summary(ur.df(s3,type="drift",lags=10,selectlags=c("AIC"))))
(adft=summary(ur.df(s3,type="trend",lags=10,selectlags=c("AIC"))))
(adfn_d=summary(ur.df(diff(s3),type="none",lags=10,selectlags=c("AIC"))))
(adfd_d=summary(ur.df(diff(s3),type="drift",lags=10,selectlags=c("AIC"))))
(adft_d=summary(ur.df(diff(s3),type="trend",lags=10,selectlags=c("AIC"))))

(adfn=summary(ur.df(s4,type="none",lags=10,selectlags=c("AIC"))))
(adfd=summary(ur.df(s4,type="drift",lags=10,selectlags=c("AIC"))))
(adft=summary(ur.df(s4,type="trend",lags=10,selectlags=c("AIC"))))
(adfn_d=summary(ur.df(diff(s4),type="none",lags=10,selectlags=c("AIC"))))
(adfd_d=summary(ur.df(diff(s4),type="drift",lags=10,selectlags=c("AIC"))))
(adft_d=summary(ur.df(diff(s4),type="trend",lags=10,selectlags=c("AIC"))))

#####
####                               ACP e PACF                               ####
#####

par(mfrow=c(2,1))
acf(s1)
pacf(s1)
par(mfrow=c(2,2))
acf(diff(s1),main="fac")
pacf(diff(s1),main="facp")#,cex.main=0.1)
acf(diff(diff(s1)),main="fac")

```

```

pacf(diff(diff(s1)),main="facp")

par(mfrow=c(3,2))
acf(s2,main="fac")
pacf(s2,main="facp")
acf(diff(s2),main="fac")
pacf(diff(s2),main="facp")
acf(diff(diff(s2)),main="fac")
pacf(diff(diff(s2)),main="facp")

par(mfrow=c(2,1))
acf(s3)
pacf(s3)
par(mfrow=c(2,2))
acf(diff(s3),main="fac")
pacf(diff(s3),main="facp")
acf(diff(diff(s3)),main="fac")
pacf(diff(diff(s3)),main="facp")

par(mfrow=c(2,1))
acf(s4)
pacf(s4)
par(mfrow=c(2,2))
acf(diff(s4),main="fac")
pacf(diff(s4),main="facp")
acf(diff(diff(s4)),main="fac")
pacf(diff(diff(s4)),main="facp")

#=====#
####                                Modelos ARIMA                                ####
#=====#
#S1:
#ARIMA(2,2,2)
arima_222 <- arima(s1, order=c(2,2,2)); arima_222
s1_est_arima_222=s1-arima_222$res
plot(s1,xlab = "minutos", ylab= "Temperatura ºC",
main="Predição com o Modelo ARIMA(2,2,2)",pch=16,type="o",lwd=3,las=1)#, col="blue")
lines(s1_est_arima_222, col="red",lwd=2)
legend(600,32,legend=c("Predição ", "Série observada"),col=c("red", "black"),
      lty=1:1,bty="n",lwd=c(2,2))
Box.test(arima_222$residuals,lag=10, type="Box-Pierce")
bic_s1_arima222=AIC(arima_222,k = log(length(s1)-3));bic_s1_arima222
aic_s1_arima222=arima_222$aic ;aic_s1_arima222
mape_s1_arima222=100*(sum(abs(arima_222$res)/s1))/length(s1) ;mape_s1_arima222
acuracida_s1_arima222=100-mape_s1_arima222 ;acuracida_s1_arima222
acf(arima_222$residuals,main="fac dos resíduos")
pacf(arima_222$residuals, main="facp dos resíduos")
tsdiag(arima_222)
accuracy(arima_222)

#ARIMA(3,2,2)
arima_322 <- arima(s1, order=c(3,2,2)); arima_322
s1_est_arima_322=s1-arima_322$res
plot(s1,main="Predição com o Modelo ARIMA(3,2,2)", col="blue",las=1)
lines(s1_est_arima_322, col="red",main="")
legend(370,32, legend=c("Predição ", "Série Candeia + Eucalipto 25%:75%"),

```



```

col=c("red", "blue"), lty=1:1 )
Box.test(arima_322$residuals,lag=10, type="Box-Pierce")
bic_s1_arima322=AIC(arima_322,k = log(length(s1)-3));bic_s1_arima322
aic_s1_arima322=arima_322$aic;aic_s1_arima322
mape_s1_arima322=100*(sum(abs(arima_322$res)/s1))/length(s1) ;mape_s1_arima322
acuracida_s1_arima322=100-mape_s1_arima322 ;acuracida_s1_arima322

#ARIMA(4,2,2)
arima_422 <- arima(s1, order=c(4,2,2)); arima_422
s1_est_arima_422=s1-arima_422$res
plot(s1,main="Predição com o Modelo ARIMA(4,2,2)", col="blue",las=1)
lines(s1_est_arima_422, col="red",main="")
legend(370,32, legend=c("Predição ", "Série Candeia + Eucalipto 25%:75%"),
      col=c("red", "blue"), lty=1:1 )
Box.test(arima_422$residuals,lag=10, type="Box-Pierce")
bic_s1_arima422=AIC(arima_422,k = log(length(s1)-3));bic_s1_arima422
aic_s1_arima422=arima_422$aic;aic_s1_arima422
mape_s1_arima422=100*(sum(abs(arima_422$res)/s1))/length(s1) ;mape_s1_arima422
acuracida_s1_arima422=100-mape_s1_arima422 ;acuracida_s1_arima422

#ARIMA(5,2,2)
arima_522 <- arima(s1, order=c(5,2,2)); arima_522
s1_est_arima_522=s1-arima_522$res
plot(s1,main="Predição com o Modelo ARIMA(5,2,2)", col="blue")
lines(s1_est_arima_522, col="red",main="")
legend(370,32, legend=c("Predição ", "Série Candeia + Eucalipto 25%:75%"),
      col=c("red", "blue"), lty=1:1 )
Box.test(arima_522$residuals,lag=10, type="Box-Pierce")
bic_s1_arima522=AIC(arima_522,k = log(length(s1)-3));bic_s1_arima522
aic_s1_arima522=arima_522$aic;aic_s1_arima522
mape_s1_arima522=100*(sum(abs(arima_522$res)/s1))/length(s1) ;mape_s1_arima522
acuracida_s1_arima522=100-mape_s1_arima522 ;acuracida_s1_arima522

#ARIMA(2,2,3)
arima_223 <- arima(s1, order=c(2,2,3)); arima_223
s1_est_arima_223=s1-arima_223$res
plot(s1,main="Predição com o Modelo ARIMA(2,2,3)", col="blue",las=1)
lines(s1_est_arima_223, col="red",main="")
legend(370,32, legend=c("Predição ", "Série Candeia + Eucalipto 25%:75%"),
      col=c("red", "blue"), lty=1:1 )
Box.test(arima_223$residuals,lag=10, type="Box-Pierce")
bic_s1_arima223=AIC(arima_223,k = log(length(s1)-3));bic_s1_arima223
aic_s1_arima223=arima_223$aic;aic_s1_arima223
mape_s1_arima223=100*(sum(abs(arima_223$res)/s1))/length(s1) ;mape_s1_arima223
acuracida_s1_arima223=100-mape_s1_arima223 ;acuracida_s1_arima223

#ARIMA(3,2,3)
arima_323 <- arima(s1, order=c(3,2,3)); arima_323
s1_est_arima_323=s1-arima_323$res
plot(s1,main="Predição com o Modelo ARIMA(3,2,3)", col="blue",las=1)
lines(s1_est_arima_223, col="red",main="")
legend(370,32, legend=c("Predição ", "Série Candeia + Eucalipto 25%:75%"),
      col=c("red", "blue"), lty=1:1 )
Box.test(arima_323$residuals,lag=10, type="Box-Pierce")
bic_s1_arima323=AIC(arima_323,k = log(length(s1)-3));bic_s1_arima323
aic_s1_arima323=arima_323$aic;aic_s1_arima323
mape_s1_arima323=100*(sum(abs(arima_323$res)/s1))/length(s1) ;mape_s1_arima323
acuracida_s1_arima323=100-mape_s1_arima323 ;acuracida_s1_arima323

```

```

#ARIMA(4,2,3)
arima_423 <- arima(s1, order=c(4,2,3)); arima_423
s1_est_arima_423=s1-arima_423$res
plot(s1,main="Predição com o Modelo ARIMA(4,2,3)", col="blue",las=1)
lines(s1_est_arima_423, col="red",main="")
legend(370,32, legend=c("Predição ", "Série Candeia + Eucalipto 25%:75%"),
      col=c("red", "blue"), lty=1:1 )
Box.test(arima_423$residuals,lag=10, type="Box-Pierce")
bic_s1_arima423=AIC(arima_423,k = log(length(s1)-3));bic_s1_arima423
aic_s1_arima423=arima_423$aic;aic_s1_arima423
mape_s1_arima423=100*(sum(abs(arima_423$res)/s1))/length(s1) ;mape_s1_arima423
acuracida_s1_arima423=100-mape_s1_arima423 ;acuracida_s1_arima423

#ARIMA(5,2,3)
arima_523 <- arima(s1, order=c(5,2,3)); arima_523
s1_est_arima_523=s1-arima_523$res
plot(s1,main="Predição com o Modelo ARIMA(5,2,3)", col="blue",las=1)
lines(s1_est_arima_523, col="red",main="")
legend(370,32, legend=c("Predição ", "Série Candeia + Eucalipto 25%:75%"),
      col=c("red", "blue"), lty=1:1 )
Box.test(arima_523$residuals,lag=10, type="Box-Pierce")
bic_s1_arima523=AIC(arima_523,k = log(length(s1)-3));bic_s1_arima523
aic_s1_arima523=arima_523$aic;aic_s1_arima523
mape_s1_arima523=100*(sum(abs(arima_523$res)/s1))/length(s1) ;mape_s1_arima523
acuracida_s1_arima523=100-mape_s1_arima523 ;acuracida_s1_arima523

#ARIMA(6,2,3)
arima_623 <- arima(s1, order=c(6,2,3)); arima_623
s1_est_arima_623=s1-arima_623$res #predição
plot(s1,main="Predição com o Modelo ARIMA(6,2,3)", col="blue",las=1)
lines(s1_est_arima_623, col="red",main="")
legend(370,32, legend=c("Predição ", "Série Candeia + Eucalipto 25%:75%"),
      col=c("red", "blue"), lty=1:1 )
Box.test(arima_623$residuals,lag=10, type="Box-Pierce")
bic_s1_arima623=AIC(arima_623,k = log(length(s1)-3));bic_s1_arima623
aic_s1_arima623=arima_623$aic;aic_s1_arima623
mape_s1_arima623=100*(sum(abs(arima_623$res)/s1))/length(s1) ;mape_s1_arima623
acuracida_s1_arima623=100-mape_s1_arima623 ;acuracida_s1_arima623

tabela_s1 <- data.frame("AIC"=c( aic_s1_arima322, aic_s1_arima422, aic_s1_arima522,
                                aic_s1_arima223, aic_s1_arima323, aic_s1_arima423,
                                aic_s1_arima523, aic_s1_arima623),
                        "BIC"=c( bic_s1_arima322, bic_s1_arima422, bic_s1_arima522,
                                bic_s1_arima223, bic_s1_arima323, bic_s1_arima423,
                                bic_s1_arima523, bic_s1_arima623),
                        "Acuracidade"=c( acuracida_s1_arima322, acuracida_s1_arima422,
                                           acuracida_s1_arima522, acuracida_s1_arima223,
                                           acuracida_s1_arima323, acuracida_s1_arima423,
                                           acuracida_s1_arima523,acuracida_s1_arima623),
                        row.names=c("ARIMA(3,2,2)","ARIMA(4,2,2)","ARIMA(5,2,2)","ARIMA(2,2,3)",
                                    "ARIMA(3,2,3)","ARIMA(4,2,3)","ARIMA(5,2,3)","ARIMA(6,2,3)"))

tabela_s1
#-----#

#S2:
#ARIMA(4,2,2)
arima_422 <- arima(s2, order=c(4,2,2)); arima_422
s2_est_arima_422=s2-arima_422$res

```

```

plot(s2,xlab = "minutos", ylab= "Temperatura řC",
     main="Prediřão com o Modelo ARIMA(4,2,2)",pch=16,type="o",lwd=3,las=1)
lines(s2_est_arma_422,col="red",lwd=2)
legend(120,45,legend=c("Prediřão ", "Sřrie observada"),col=c("red", "black"),
      lty=1:1,bty="n",lwd=c(2,2))
Box.test(arma_422$residuals,lag=10, type="Box-Pierce")
bic_s2_arma422=AIC(arma_422,k = log(length(s2)-3));bic_s2_arma422
aic_s2_arma422=arma_422$aic;aic_s2_arma422
mape_s2_arma422=100*(sum(abs(arma_422$res)/s2))/length(s2) ;mape_s2_arma422
acuracida_s2_arma422=100-mape_s2_arma422 ;acuracida_s2_arma422
acf(arma_422$residuals,main="fac dos resřduos")
pacf(arma_422$residuals, main="facp dos resřduos")
accuracy(arma_422)

#ARIMA(6,2,2)
arma_622 <- arima(s2, order=c(6,2,2)); arma_622
s2_est_arma_622=s2-arma_622$res
plot(s2,main="Prediřão com o Modelo ARIMA(6,2,2)", col="black")
lines(s2_est_arma_622, col="red",main="")
legend(800,32, legend=c("Prediřão ", "Sřrie observada"),col=c("red", "black"),
      lty=1:1,bty="n",lwd=c(2,2))
Box.test(arma_622$residuals,lag=10, type="Box-Pierce")
bic_s2_arma622=AIC(arma_622,k = log(length(s2)-3));bic_s2_arma622
aic_s2_arma622=arma_622$aic;aic_s2_arma622
mape_s2_arma622=100*(sum(abs(arma_622$res)/s2))/length(s2) ;mape_s2_arma622
acuracida_s2_arma622=100-mape_s2_arma622 ;acuracida_s2_arma622

tabela_s2 <- data.frame("AIC"=c(aic_s2_arma422, aic_s2_arma622),
                       "BIC"=c(aic_s2_arma422, aic_s2_arma622),
                       "Acuracidade"=c(acuracida_s2_arma422, acuracida_s2_arma622),
                       row.names=c("ARIMA(4,2,2)","ARIMA(6,2,2)"))

tabela_s2

#-----#

#S3:
#ARI(5,1)
arma_510 <- arima(s3, order=c(5,1,0)); arma_510
s3_est_arma_510=s3-arma_510$res
plot(s3,main="Prediřão com o Modelo ARIMA(5,1,0)",xlab = "minutos",
     ylab= "Temperatura řC",pch=16,type="o",lwd=3,las=1)
lines(s3_est_arma_510, col="red",lwd=2)
legend(800,32, legend=c("Prediřão ", "Sřrie observada"),col=c("red", "black"),
      lty=1:1,bty="n",lwd=c(2,2))
Box.test(arma_510$residuals,lag=10, type="Box-Pierce")
bic_s3_arma510=AIC(arma_510,k = log(length(s3)-3));bic_s3_arma510
aic_s3_arma510=arma_510$aic;aic_s3_arma510
mape_s3_arma510=100*(sum(abs(arma_510$res)/s3))/length(s3) ;mape_s3_arma510
acuracida_s3_arma510=100-mape_s3_arma510 ;acuracida_s3_arma510
acf(arma_510$residuals,main="fac dos resřduos")
pacf(arma_510$residuals,main="facp dos resřduos")
tsdiag(arma_510)
accuracy(arma_510)

#ARIMA(2,1,1)
arma_211 <- arima(s3, order=c(2,1,1)); arma_211
s3_est_arma_211=s3-arma_211$res

```

```

plot(s3,main="Predição com o Modelo ARIMA(2,1,1)", col="black")
lines(s3_est_arima_211, col="red",main="")
legend(800,32, legend=c("Predição ", "Série observada"),col=c("red", "black"),
      lty=1:1,bty="n",lwd=c(2,2))
Box.test(arima_211$residuals,lag=10, type="Box-Pierce")
bic_s3_arima211=AIC(arima_211,k = log(length(s3)-3));bic_s3_arima211
aic_s3_arima211=arima_211$aic;aic_s3_arima211
mape_s3_arima211=100*(sum(abs(arima_211$res)/s3))/length(s3) ;mape_s3_arima211
acuracida_s3_arima211=100-mape_s3_arima211 ;acuracida_s3_arima211
acf(arima_211$residuals,main="fac dos resíduos")
pacf(arima_211$residuals,main="facp dos resíduos")

#ARIMA(2,1,2)
arima_212 <- arima(s3, order=c(2,1,2)); arima_212
s3_est_arima_212=s3-arima_212$res #predição
plot(s3,main="Predição com o Modelo ARIMA(2,1,2)", col="black")
lines(s3_est_arima_212, col="red",main="")
legend(800,32, legend=c("Predição ", "Série observada"),col=c("red", "black"),
      lty=1:1,bty="n",lwd=c(2,2))
Box.test(arima_212$residuals,lag=10, type="Box-Pierce")
bic_s3_arima212=AIC(arima_212,k = log(length(s3)-3));bic_s3_arima212
aic_s3_arima212=arima_212$aic;aic_s3_arima212
mape_s3_arima212=100*(sum(abs(arima_212$res)/s3))/length(s3) ;mape_s3_arima212
acuracida_s3_arima212=100-mape_s3_arima212 ;acuracida_s3_arima212
acf(arima_212$residuals,main="fac dos resíduos")
pacf(arima_212$residuals,main="facp dos resíduos")

#IMA(2,2)
arima_022 <- arima(s3, order=c(0,2,2)); arima_022
arima_022$res #residuo
s3_est_arima_022=s3-arima_022$res #predição
plot(s3,main="Predição com o Modelo IMA(2,2)", col="blue")
lines(s3_est_arima_022, col="red",main="")
legend(0,40, legend=c("Predição ", "Série observada"),col=c("red", "blue"), lty=1:1 )
Box.test(arima_022$residuals,lag=10, type="Box-Pierce")
bic_s3_arima022=AIC(arima_022,k = log(length(s3)-3));bic_s3_arima022
aic_s3_arima022=arima_022$aic;aic_s3_arima022
mape_s3_arima022=100*(sum(abs(arima_022$res)/s3))/length(s3) ;mape_s3_arima022
acuracida_s3_arima022=100-mape_s3_arima022 ;acuracida_s3_arima022

tabela_s3 <- data.frame("AIC"=c(aic_s3_arima510, aic_s3_arima211, aic_s3_arima212, aic_s3_arima022)
                        "BIC"=c(bic_s3_arima510, bic_s3_arima211, bic_s3_arima212, bic_s3_arima022)
                        "Acuracidade"=c(acuracida_s3_arima510, acuracida_s3_arima211,
                                         acuracida_s3_arima212, acuracida_s3_arima022),
                        row.names=c("ARIMA(5,1,0)", "ARIMA(2,1,1)", "ARIMA(2,1,2)", "ARIMA(0,2,2)"))

tabela_s3
#-----#

#S4:
#ARIMA(2,1,2)
arima_212 <- arima(s4, order=c(2,1,2)); arima_212
s4_est_arima_212=s4-arima_212$res

plot(s4,main="Predição com o Modelo ARIMA(2,1,2)",xlab = "minutos",
     ylab= "Temperatura řC",pch=16,type="o",lwd=3,las=1)
lines(s4_est_arima_212, col="red",lwd=2)
legend(800,32, legend=c("Predição ", "Série observada"),col=c("red", "black"),
      lty=1:1,bty="n",lwd=c(2,2) )

```

```

Box.test(arima_212$residuals,lag=10, type="Box-Pierce")
bic_s4_arima212=AIC(arima_212,k = log(length(s4)-3));bic_s4_arima212
aic_s4_arima212=arima_212$aic;aic_s4_arima212
mape_s4_arima212=100*(sum(abs(arima_212$res)/s4))/length(s4) ;mape_s4_arima212
acuracida_s4_arima212=100-mape_s4_arima212 ;acuracida_s4_arima212
acf(arima_212$residuals,main="fac dos resíduos")
pacf(arima_212$residuals,main="facp dos resíduos")
par(mfrow=c(2,1))
tsdiag(arima_212$residuals)
accuracy(arima_212)

#ARIMA(2,2,2)
arima_222 <- arima(s4, order=c(2,2,2)); arima_222
s4_est_arima_222=s4-arima_222$res
plot(s4,main="Predição com o Modelo ARIMA(2,2,2)", col="blue")
lines(s4_est_arima_222, col="red",main="")
legend(370,32, legend=c("Predição ", "Série Candeia + Eucalipto 25%:75%"),
      col=c("red", "blue"), lty=1:1 )
AIC_arima_222= n*log((t(arima_222$res )%*%arima_222$res) /n)+2*(length(arima_222)+1)
AIC_arima_222
Box.test(arima_222$residuals,lag=10, type="Box-Pierce")
bic_s4_arima222=AIC(arima_222,k = log(length(s4)-3));bic_s4_arima222
aic_s4_arima222=arima_222$aic;aic_s4_arima222
mape_s4_arima222=100*(sum(abs(arima_222$res)/s4))/length(s4) ;mape_s4_arima222
acuracida_s4_arima222=100-mape_s4_arima222 ;acuracida_s4_arima222
tsdiag(arima_222)
accuracy(arima_222)

tabela_s4 <- data.frame("AIC"=c(aic_s4_arima212, aic_s4_arima222),
                        "BIC"=c(bic_s4_arima212, bic_s4_arima222),
                        "Acuracidade"=c(acuracida_s4_arima212,
                                          acuracida_s4_arima222),
                        row.names=c("ARIMA(2,1,2)","ARIMA(2,2,2)"))

tabela_s4
#=====#
###                               Teste de Quenouille                               ###
#=====#
#A função do teste:
quenouille.test <- function(sx, sy, j=20, P=j, p, plot=F) {
  n1 <- length(sx)
  n2 <- length(sy)
  acfx <- acf(sx, j, type="correlation", plot=F)$acf
  acfy <- acf(sy, j, type="correlation", plot=F)$acf
  acfcom <- ((n1*acfx)+(n2*acfy))/(n1+n2)
  acf <- acfcom
  zeros <- rep(0,j)
  acf <- c(zeros, acf)
  phi <- numeric(j)
  phi[1] <- acf[j+2]
  vetmat <- numeric(j)
  vetmat[1] <- list(1)
  for (m1 in 2:j) {
    mat1 <- numeric(m1*m1)
    mat1 <- matrix(mat1, m1, m1)
    for (m2 in 1:m1) {
      lim1 <- j+2-m2
      lim2 <- j+m1+1-m2
      fila <- acf[lim1:lim2]

```

```

    mat1[m2,] <- fila
  }
  diagsup <- mat1
  matden <- diagsup+t(diagsup)
  matden <- matden-diag(m1)
  vetmat[m1] <- list(matden)
  denom <- determinant(matden, logarithm=F)
  denom <- (denom$modulus)*(denom$sign)
  lim3 <- j+2
  lim4 <- j+m1+1
  col <- acf[lim3:lim4]
  col <- matrix(col,m1,1)
  mat2 <- matden[,-c(m1)]
  mat3 <- cbind(mat2,col)
  num <- determinant(mat3, logarithm=F)
  num <- (num$modulus)*(num$sign)
  phi[m1] <- num/denom
}
facp <- c(phi[1:j])
facp <- abs(facp)
maximo <- max(n1,n2)
inter <- 2*(1/sqrt(maximo))
vetorlag <- numeric(j)
for (m in 1:j){
  if (facp[m]>inter){vetorlag[m]<-m} else {vetorlag[m]<-1}
}
vetorlag <- c(vetorlag[1:j])
if(missing(p)) p <- max(vetorlag)
if(plot) {
  plot(1:j, facp, type="h", xlab="lag", ylab="PACF Conjunta")
  abline(h=0)
  abline(h=c(-inter,inter), lty=2, col=4)
  cat("Digite o valor de p")
  p <- scan(n=1)
}
lim1 <- j+2
lim2 <- j+1+p
r <- acf[lim1:lim2]
r <- matrix(r,p,1)
R <- vetmat[[p]]
R <- matrix(R,p,p)
phipar <- solve(R)%*%r
phipar <- matrix(phipar,p,1)
p1 <- p+1
np1 <- n1-p
np2 <- n2-p
x <- sx[c(p1:n1)]
x <- matrix(x,np1,1)
y <- sy[c(p1:n2)]
y <- matrix(y,np2,1)
dim1 <- np1*p
dim2 <- np2*p
xmat <- numeric(dim1)
xmat <- matrix(xmat,np1,p)
ymat <- numeric(dim2)
ymat <- matrix(ymat,np2,p)
lim1 <- p1
lim2 <- n1

```

```

lim3 <- n2
for (l in 1:p){
  lim1 <- lim1-1
  lim2 <- lim2-1
  lim3 <- lim3-1
  colx <- sx[lim1:lim2]
  coly <- sy[lim1:lim3]
  xmat[,1] <- colx
  ymat[,1] <- coly
}
xmat <- xmat
ymat <- ymat
res1 <- x-xmat%%phipar
res2 <- y-ymat%%phipar
fac1 <- acf(res1, P, type="partial", plot=F, main = "Série Residual 1")
fac2 <- acf(res2, P, type="partial", plot=T, main = "Série Residual 2")
S <- fac1$lag
fac1 <- fac1$acf
fac2 <- fac2$acf
DEN <- (1/(n1-S))+(1/(n2-S))
NUM <- (fac1-fac2)^2
sq <- sum(NUM/DEN)
pv1 <- 1-pchisq(sq,P)
data.frame("estatistica"=sq,"gl"=P,"p-valor"=pv1)
}

quenouille.test(s1, s2, j=30, P=30, plot=F)
quenouille.test(s1, s3, j=30, P=30, plot=F)
quenouille.test(s1, s4, j=30, P=30, plot=F)
quenouille.test(s2, s3, j=30, P=30, plot=F)
quenouille.test(s2, s4, j=30, P=30, plot=F)
quenouille.test(s3, s4, j=30, P=30, plot=F)

#=====#
####                                Comparações                                ####
#=====#
#s1:s2
s12=s1[1:n2]-s2
plot(s12,xlab = "minutos", ylab= "Diferença entre as séries s1 e s2",
     pch=16,type="o",lwd=3)
# Teste da Raiz Unitária
(adfn=summary(ur.df(s12,type="none",lags=10,selectlags=c("AIC"))))
(adfd=summary(ur.df(s12,type="drift",lags=10,selectlags=c("AIC"))))
(adft=summary(ur.df(s12,type="trend",lags=10,selectlags=c("AIC"))))
#Gráfico da ACP e PACF
par(mfrow=c(2,1))
acf(s12,main="ACP série s1-s2");pacf(s12,main="PACF série s1-s2")

#-----#

#s1:s3
s13=s1[1:n3]-s3
plot(s13,xlab = "minutos", ylab= "Diferença entre as séries s1 e s3",
     pch=16,type="o",lwd=3)
#Teste da Raiz Unitaria
(adfn=summary(ur.df(s13,type="none",lags=10,selectlags=c("AIC"))))
(adfd=summary(ur.df(s13,type="drift",lags=10,selectlags=c("AIC"))))
(adft=summary(ur.df(s13,type="trend",lags=10,selectlags=c("AIC"))))

```

```

#Gráfico ACP e PACP
par(mfrow=c(2,1))
acf(s13,main="ACP série s1-s3");pacf(s13,main="PACF série s1-s3")

#-----#

#s1:s4
s14=s1[1:n4]-s4
plot(s14,xlab = "minutos", ylab= "Diferença entre as séries s1 e s4",
     pch=16,type="o",lwd=3)

#Teste da Raiz Unitária
(adfn=summary(ur.df(s14,type="none",selectlags=c("AIC"))))
(adfd=summary(ur.df(s14,type="drift",selectlags=c("AIC"))))
(adft=summary(ur.df(s14,type="trend",selectlags=c("AIC"))))
#Gráfico da ACP e PACF
par(mfrow=c(2,1))
acf(s14,main="ACP série s1-s4");pacf(s14,main="PACF série s1-s4")

#-----#

#s2:s3
s23=s2-s3[1:n2]
plot(s23,xlab = "minutos", ylab= "Diferença entre as séries s2 e s3",
     pch=16,type="o",lwd=3)
#Teste da Raiz Unitaria
(adfn=summary(ur.df(s23,type="none",lags=10,selectlags=c("AIC"))))
(adfd=summary(ur.df(s23,type="drift",lags=10,selectlags=c("AIC"))))
(adft=summary(ur.df(s23,type="trend",lags=10,selectlags=c("AIC"))))
#Gráfico da ACP e PACF
par(mfrow=c(2,1))
acf(s23);pacf(s23)

#s2:s4
s24=s2-s4[1:n2]
plot(s24,xlab = "minutos", ylab= "Diferença entre as séries s2 e s4",
     pch=16,type="o",lwd=3)

#teste da Raiz Unitaria
(adfn=summary(ur.df(s24,type="none",lags=10,selectlags=c("AIC"))))
(adfd=summary(ur.df(s24,type="drift",lags=10,selectlags=c("AIC"))))
(adft=summary(ur.df(s24,type="trend",lags=10,selectlags=c("AIC"))))
#Gráfico da ACP e PACF
par(mfrow=c(2,1))
acf(s24);pacf(s24)

#-----#

#s3:s4
s34=s3-s4[1:n3]
plot(s34,xlab = "minutos", ylab= "Diferença entre as séries s3 e s4",
     pch=16,type="o",lwd=3)
#teste da Raiz Unitaria
(adfn=summary(ur.df(s34,type="none",lags=10,selectlags=c("AIC"))))
(adfd=summary(ur.df(s34,type="drift",lags=10,selectlags=c("AIC"))))
(adft=summary(ur.df(s34,type="trend",lags=10,selectlags=c("AIC"))))
#Gráfico da ACP e PACF
par(mfrow=c(2,1))

```



```
acf(s34);pacf(s34)
```

```
par(mfrow=c(3,2))  
plot(s12,xlab = "minutos", ylab= "Diferença entre as séries s1 e s3",  
     pch=16,type="o",lwd=3)  
plot(s13,xlab = "minutos", ylab= "Diferença entre as séries s1 e s3",  
     pch=16,type="o",lwd=3)  
plot(s14,xlab = "minutos", ylab= "Diferença entre as séries s1 e s3",  
     pch=16,type="o",lwd=3)  
plot(s23,xlab = "minutos", ylab= "Diferença entre as séries s1 e s3",  
     pch=16,type="o",lwd=3)  
plot(s24,xlab = "minutos", ylab= "Diferença entre as séries s1 e s3",  
     pch=16,type="o",lwd=3)  
plot(s34,xlab = "minutos", ylab= "Diferença entre as séries s1 e s3",  
     pch=16,type="o",lwd=3)
```