



TÁBATA SATURNINA TRINDADE DE MORAIS

**ESTUDO TEMPORAL DO NÍVEL MÉDIO DO MAR
EM DIFERENTES OCEANOS**

LAVRAS - MG

2012

TÁBATA SATURNINA TRINDADE DE MORAIS

**ESTUDO TEMPORAL DO NÍVEL MÉDIO DO MAR EM DIFERENTES
OCEANOS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

Orientadora
Dra. Thelma Sáfydi

**LAVRAS - MG
2012**

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca da UFLA**

Morais, Tábata Saturnina Trindade de.

Estudo temporal do nível médio do mar em diferentes oceanos /
Tábata Saturnina Trindade de Moraes. – Lavras : UFLA, 2012.

109 p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2012.

Orientador: Thelma Sáfadi.

Bibliografia.

1.Séries temporais. 2. Testes de comparação. 3. Modelo de regressão. 4. Raiz unitária sazonal. I. Universidade Federal de Lavras.
II. Título.

CDD - 519.55

TÁBATA SATURNINA TRINDADE DE MORAIS

**ESTUDO TEMPORAL DO NÍVEL MÉDIO DO MAR EM DIFERENTES
OCEANOS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

APROVADA em 27 de fevereiro de 2012.

Dra. Thelma Sáfadi

UFLA

Dr. João Domingos Scalon

UFLA

Dra. Airlane Pereira Alencar

IME-USP

Dra. Thelma Sáfadi
Orientadora

**LAVRAS - MG
2012**

A meus pais, Pedro e Leuza;
minhas irmãs, Talita e Keila;
e aos meus sobrinhos, Lara e Gabriel.

DEDICO

AGRADECIMENTOS

A Deus, pelo aprendizado constante e a força nos momentos difíceis.

À minha família, a qual amo muito, pela dedicação e amor a mim destinado.

À minha orientadora, professora Thelma Sáfadi, pela confiança, empenho, sabedoria e compreensão.

Aos colegas do curso de Pós-Graduação em Estatística e Experimentação Agropecuária, pela convivência durante este tempo de estudo. Em especial, à melhor turma de mestrado e grandes amigos: Adriele, Adriana, André, Carolina, Danielle, Elayne, Iábita, Jair Prado, Juliano, Juracy, Juscélia, Larissa, Leandro, Lourenço, Marcelo, Mariele e Siomara.

A todos os professores e funcionários do Departamento de Ciências Exatas (DEX).

À Universidade Federal de Lavras, e ao curso de Pós-Graduação em Estatística e Experimentação Agropecuária.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, pela concessão de bolsa de estudos.

Muitas foram as pessoas que diretamente e indiretamente me auxiliaram para realizar este trabalho, citar todas aqui não seria possível, porém algumas tiveram um diferencial para que eu pudesse chegar ao fim desta etapa, a elas, a minha mais profunda gratidão.

RESUMO

O comportamento do nível do mar é um fator muito importante para toda a população, ele influencia as principais atividades costeiras e um grande aumento ou diminuição poderá causar muitos prejuízos, dessa forma é de grande interesse um estudo temporal desses níveis. Neste trabalho, foi feito tal estudo, e as séries consideradas foram de mares de Aburatsu - Japão (oceano Pacífico), Key West - Estados Unidos (oceano Atlântico) e Ko Taphao Noi - Tailândia (oceano Índico), no período de janeiro de 1985 a novembro de 2011. Os objetivos foram ajustar modelos às séries, verificando a presença de tendência e sazonalidade, por meio da metodologia de Box e Jenkins, comparar as séries e, para a série de Cananéia - Brasil (oceano Atlântico) foi utilizado o modelo de regressão múltipla tendo como variáveis independentes variáveis climáticas desse local (no período de janeiro de 2005 a dezembro de 2006). Observou-se que as três séries apresentaram tendência crescente e, através do teste de raiz unitária sazonal, temos que as séries apresentam sazonalidade de natureza determinística. O modelo que melhor se ajustou à série de Aburatsu foi o ARIMA (3,1,1), para a série da Key West tivemos o modelo ARIMA (1,1,1) e em Ko Taphao Noi o modelo ARIMA (1,1,1) com intervenção em outubro de 1997. O teste das somas acumuladas e o teste de igualdade das funções de autocorrelação foram utilizados na comparação, levando a conclusão de que apenas as séries de Aburatsu e Key West podem ser consideradas geradas pelo mesmo processo estocástico. O modelo de regressão mostrou que apenas temperatura da água e temperatura do psicômetro possuem influência no nível mar de Cananéia.

Palavras-chave: Séries temporais. Raiz unitária sazonal. Testes de comparação. Séries de níveis médios do mar. Modelo de regressão.

ABSTRACT

The behavior of the sea level is a very important factor for the entire population, it influences the main coastal activities and a large increase or decrease can cause much damage, it is very important a temporal study of levels. This is the goal of this study, we analysed series of the sea levels of Aburatsu - Japan (Pacific Ocean), Key West - United States (Atlantic Ocean) and Ko Taphao Noi - Thailand (Indian Ocean), from January 1985 to November 2011. The analysis consists of identifying and fitting models evaluating the goodness of fit and comparing the series, verifying the presence of trend and seasonality, using the methodology of Box and Jenkins, to compare the series, and, for the series of Cananéia - Brazil (Atlantic Ocean), was used the multiple regression model with climatic variables (from January 2005, to December 2006) as independent variables. The three series showed an increasing trend and, in the seasonal unit root test, the series exhibit seasonality deterministic nature. The better fitting for Aburatsu series was ARIMA (1,1,1), for the series of Key West was ARIMA (3,1,1) and for Ko Taphao Noi, the ARIMA model (1,1,1) with intervention in October 1997. The accumulated sums test and the test of equality of the autocorrelation functions were used to compare, only the series of Aburatsu and Key West can be considered generated by the same stochastic process. The regression model showed that only the water temperature and psychrometer temperature have influence on the sea level of Cananéia.

Keywords: Times series. Seasonal unit root. Comparison tests. Series of average levels of the sea. Regression model.

LISTA DE FIGURAS

Figura 1	<i>Cape Hatteras, Carolina do Norte, EUA, em 1999 e 2004. Foto de Gary Braasch</i>	19
Figura 2	Mapa mundi com destaque nos locais pesquisados (Key West, Cananéia, Ko Taphao Noi e Aburatsu)	55
Figura 3	Série mensal de níveis médios do mar de Aburatsu, no período de janeiro de 1985 a dezembro de 2010	58
Figura 4	Decomposição sazonal usando STL aplicado à série mensal de níveis médios do mar de Aburatsu, no período de janeiro de 1985 a dezembro de 2010. Os círculos cinzas mostram a série original, a linha azul é a componente tendência, a vermelha mostra a componente sazonal, e as barras pretas são o componente aleatório	59
Figura 5	Série mensal de níveis médios do mar de Aburatsu: amplitude vs média	60
Figura 6	Periodograma da série de níveis médios do mar de Aburatsu	61
Figura 7	Subséries mensais dos níveis médios do mar de Aburatsu	62
Figura 8	Série diferenciada de níveis médios do mar de Aburatsu	64
Figura 9	Funções de autocorrelação (a) e autocorrelação parcial (b) da série diferenciada de níveis médios do mar de Aburatsu	65
Figura 10	Função de autocorrelação (a) e autocorrelação parcial (b) do resíduo do modelo ajustado para a série de níveis do mar de Aburatsu	66
Figura 11	Séries de valores observados e estimados de acordo com o modelo ajustado ARIMA (3,1,1), de nível médio do mar de Aburatsu, no período de janeiro de 1985 a dezembro de 2010	68
Figura 12	Série mensal de níveis médios do mar de Key West, no período de janeiro de 1985 a dezembro de 2010	69
Figura 13	Série diferenciada de níveis médios do mar da Key West	69
Figura 14	Funções de autocorrelação (a) e autocorrelação parcial (b) da série diferenciada de níveis médios do mar da Key West	70
Figura 15	Função de autocorrelação e autocorrelação parcial do resíduo do modelo ajustado para a série de níveis do mar de Key West	71
Figura 16	Séries de valores observados e estimados de acordo com o modelo ajustado ARIMA (1,1,2), de nível médio do mar de Key West, no período de janeiro de 1985 a dezembro de 2010	72
Figura 17	Série mensal de níveis médios do mar de Ko Taphao Noi, no período de janeiro de 1985 a dezembro de 2010	73
Figura 18	Série diferenciada de níveis médios do mar de Ko Taphao Noi	74

Figura 19	Funções de autocorrelação (a) e autocorrelação parcial (b) da série diferenciada de níveis médios do mar de Ko Taphao Noi . . .	74
Figura 20	Função de autocorrelação (a) e autocorrelação parcial (b) do resíduo do modelo ajustado para a série de níveis do mar de Ko Taphao Noi	76
Figura 21	Séries de valores observados e estimados de acordo com o modelo ajustado ARIMA (1,1,1)-CI, de nível médio do mar de Ko Taphao Noi, no período de janeiro de 1985 a dezembro de 2010 .	77
Figura 22	Séries de níveis do mar de Aburatsu, Key West e Ko Taphao Noi, no período de janeiro de 1985 a dezembro de 2010	78
Figura 23	Periodograma da diferença das séries livres de sazonalidade de Aburatsu (a) e Key West (b)	79
Figura 24	Periodograma das séries diferenciadas de Aburatsu (a) e Ko Taphao Noi (b)	80
Figura 25	Periodograma das séries diferenciadas de Ko Taphao Noi (a) e Key West (b)	81
Figura 26	Função de autocorrelação das séries diferenciadas de Aburatsu (a) e Key West (b)	82
Figura 27	Função de autocorrelação parcial estimada comum ($\widehat{\Phi}(k)$) . . .	82
Figura 28	Função de autocorrelação das séries residuais de Aburatsu (a) e Key West (b)	83
Figura 29	Função de autocorrelação parcial das séries residuais de Aburatsu (a) e Key West (b)	84
Figura 30	Função de autocorrelação das séries diferenciadas de Aburatsu (a) e Ko Taphao Noi (b)	85
Figura 31	Função de autocorrelação parcial estimada comum ($\widehat{\Phi}(k)$) . . .	85
Figura 32	Função de autocorrelação das séries residuais de Aburatsu (a) e Ko Taphao Noi (b)	86
Figura 33	Função de autocorrelação parcial das séries residuais de Aburatsu (a) e Ko Taphao Noi (b)	86
Figura 34	Função de autocorrelação das séries diferenciadas de Key West (a) e Ko Taphao Noi (b)	87
Figura 35	Função de autocorrelação parcial estimada comum ($\widehat{\Phi}(k)$) . . .	88
Figura 36	Função de autocorrelação das séries residuais de Key West (a) e Ko Taphao Noi (b)	89
Figura 37	Função de autocorrelação parcial das séries residuais de Key West (a) e Ko Taphao Noi (b)	89
Figura 38	Série mensal de níveis do mar de Cananéia, no período de janeiro de 2005 a dezembro de 2006	90

Figura 39	Funções de autocorrelação e autocorrelação parcial (Cananéia) . . .	91
Figura 40	Dispersão dos resíduos em função dos valores estimados pelo modelo	93
Figura 41	Correlograma dos resíduos do modelo de regressão ajustado . . .	93
Figura 42	Gráfico da normalidade dos resíduos do modelo de regressão ajustado	94
Figura 43	Decomposição sazonal usando STL aplicado à série mensal de níveis médios do mar de Key West, no período de janeiro de 1985 a dezembro de 2010. Os círculos cinzas mostram a série original, a linha azul é a componente tendência, a vermelha mostra a componente sazonal, e as barras pretas são o componente aleatório	100
Figura 44	Série mensal de níveis médios do mar de Key West: amplitude vs média	101
Figura 45	Periodograma da série de níveis médios do mar da Key West . . .	102
Figura 46	Subséries mensais dos níveis médios do mar de Key West	103
Figura 47	Decomposição sazonal usando STL aplicado à série mensal de níveis médios do mar de Ko Taphao Noi, no período de janeiro de 1985 a dezembro de 2010. Os círculos cinzas mostram a série original, a linha azul é a componente tendência, a vermelha mostra a componente sazonal, e as barras pretas são o componente aleatório	105
Figura 48	Série mensal de níveis médios do mar da Ko Taphao Noi: amplitude vs média	106
Figura 49	Periodograma da série de níveis médios do mar da Ko Taphao Noi	107
Figura 50	Subséries mensais dos níveis médios do mar de Key West	108

LISTA DE TABELAS

Tabela 1	Resultado do teste de raiz unitária sazonal (teste HEGY) para a série de nível médio do mar de Aburatsu.	63
Tabela 2	Significância dos parâmetros do modelo de regressão com <i>dummies</i> sazonais ajustado para eliminar sazonalidade determinística. . .	64
Tabela 3	Estimativas dos parâmetros do modelo ARIMA proposto para a série mensal de níveis médios do mar de Aburatsu, no período de janeiro de 1985 a dezembro de 2010.	65
Tabela 4	Valores observados (Z_{t+h}), valores preditos com o modelo ARIMA (3, 1, 1) ($\hat{Z}_t(h)$) e erro de previsão ($e_t(h)$) para a série mensal de níveis do mar de Aburatsu, no período de janeiro a novembro de 2011.	67
Tabela 5	Estimativas dos parâmetros do modelo ARIMA proposto para a série mensal de níveis médios do mar de Key West, no período de janeiro de 1985 a dezembro de 2010	70
Tabela 6	Valores observados (Z_{t+h}), valores preditos com o modelo ARIMA (1, 1, 1) ($\hat{Z}_t(h)$) e erro de previsão ($e_t(h)$) para a série mensal de níveis do mar de Key West, no período de janeiro a novembro de 2011.	72
Tabela 7	Estimativas dos parâmetros do modelo ARIMA proposto para a série mensal de níveis médios do mar de Ko Taphao Noi, no período de janeiro de 1985 a dezembro de 2010.	75
Tabela 8	Valores observados (Z_{t+h}), valores preditos com o modelo ARIMA (1, 1, 2)-CI ($\hat{Z}_t(h)$) e erro de previsão ($e_t(h)$) para a série mensal de níveis do mar de Ko Taphao Noi, no período de janeiro a novembro de 2011.	77
Tabela 9	Análise de variância do modelo de regressão ajustado para Cananéia. 92	92
Tabela 10	Estimativas dos parâmetros do modelo de regressão, desvio padrão e estatística <i>t</i> de Student.	92
Tabela 11	Resultado do teste de raiz unitária sazonal (teste HEGY) para a série de nível médio do mar de Key West.	103
Tabela 12	Significância dos parâmetros do modelo de regressão com <i>dummies</i> sazonais ajustado para eliminar sazonalidade determinística. . .	104
Tabela 13	Valores críticos e da estatística do teste HEGY	107
Tabela 14	Significância dos parâmetros do modelo de regressão com <i>dummies</i> sazonais ajustado para eliminar sazonalidade determinística . . .	109

SUMÁRIO

1	INTRODUÇÃO	14
1.1	Objetivos	15
1.1.1	Objetivo geral	15
1.1.2	Objetivos específicos	16
2	REFERENCIAL TEÓRICO	16
2.1	O nível do mar	16
2.1.1	Medições	17
2.1.2	Estudos sobre o comportamento do nível do mar	18
2.2	Fenômeno climático El Niño	20
2.3	Séries temporais	20
2.3.1	Estacionariedade	21
2.3.2	Definições importantes	22
2.3.3	Decomposição clássica - Tendência e Sazonalidade	23
2.3.4	Análise espectral	24
2.3.5	Teste de Fisher	25
2.3.6	Teste de sequência (Wald-Wolfowitz)	26
2.3.7	Teste da raiz unitária sazonal	27
2.3.8	Estimação da sazonalidade - método de regressão	31
2.3.9	Metodologia Box-Jenkins	33
2.3.10	Modelos lineares estacionários	33
2.3.10.1	Modelos médias móveis de ordem q, MA(q)	34
2.3.10.2	Modelos auto-regressivos de ordem p, AR(p)	34
2.3.10.3	Função de autocorrelação parcial	35
2.3.11	Modelos ARIMA	36
2.3.12	Identificação, estimação e previsão com modelos ARIMA	37
2.3.13	Critérios para seleção do modelo	38
2.3.14	Análise de resíduos	39
2.3.15	Análise de intervenção	40
2.4	Procedimentos para comparação de séries temporais	41
2.4.1	Teste das Somas Acumuladas (Coates e Diggle, 1986)	42
2.4.2	Teste de Igualdade das Funções de Autocorrelação (Quenouille, 1958)	44
2.5	Modelo de Regressão Linear Múltipla	46
2.5.1	Introdução	46
2.5.2	Estimação dos parâmetros do modelo	47
2.5.3	Métodos de seleção	49
2.5.4	O coeficiente de determinação múltipla	50

2.5.5	Teste de hipótese em regressão linear múltipla	50
2.5.5.1	Teste de significância da regressão	50
2.5.5.2	Teste sobre os coeficiente individuais de regressão	51
2.5.6	Estudo da adequação do modelo	51
2.5.6.1	Verificação dos pressupostos e análise dos resíduos	51
2.5.6.2	Teste de Kolmogorv-Smirnov (normalidade dos resíduos)	52
2.5.6.3	Teste de Durbin-Watson	52
3	MATERIAL E MÉTODOS	54
4	RESULTADOS E DISCUSSÃO	58
4.1	Análise inicial e ajustes	58
4.1.1	Aburatsu	58
4.1.2	Key West	68
4.1.3	Ko Taphao Noi	73
4.2	Comparação das séries	78
4.2.1	Teste de Coates e Diggle (1986) aplicado às séries diferenciadas de níveis médios dos mares de Aburatsu e Key West.	78
4.2.2	Teste de Coates e Diggle (1986) aplicado às séries diferenciadas de níveis médios dos mares de Aburatsu e Ko Taphao Noi.	80
4.2.3	Teste de Coates e Diggle (1986) aplicado às séries estacionárias de níveis médios dos mares de Ko Taphao Noi e Key West.	81
4.2.4	Teste de Quenouille (1958) aplicado às séries diferenciadas de níveis médios do mar de Aburatsu e Key West.	81
4.2.5	Teste de Quenouille (1958) aplicado às séries diferenciadas de níveis médios do mar de Aburatsu e Ko Taphao Noi.	84
4.2.6	Teste de Quenouille (1958) aplicado às séries diferenciadas de níveis médios do mar de Ko Taphao Noi e Key West.	87
4.3	Modelo de regressão para a série de níveis do mar de Cananéia	90
5	CONCLUSÃO	95
	REFERÊNCIAS	96
	APÊNDICE	100

1 INTRODUÇÃO

Uma melhor compreensão do nível do mar e sua variabilidade ajudará a reduzir as incertezas associadas ao aumento do nível do mar e assim, contribuir para o planejamento mais eficaz, ações acertivas e gestão do litoral. Considerando a área de abrangência do presente estudo, a análise e interpretação de dados apresenta extrema relevância, visto que a navegação, do ponto de vista socioeconômico, possibilitou desde os tempos mais remotos, o desenvolvimento de diversos povos, atualmente favorecendo entre outros, o turismo. Por outro prisma, pode-se considerar a influência sobre os fatores ambientais, ciclos biogeoquímicos, fauna e flora, dada a relação entre marés e alterações em outros meios. A população, por sua vez, concentrada nas áreas litorâneas, pode sofrer com inundações e danos em consequência desses fenômenos.

Além disso, a variação do nível do mar assume particular relevância no contexto das alterações climáticas: mudanças de longo prazo desses níveis podem estar associadas com a variabilidade climática regional, e podem enfrentar um risco maior de inundação.

O planejamento e desenvolvimento de projetos de obras costeiras, atividades petrolíferas desenvolvidas em áreas oceânicas, operações portuárias por dependerem de condições adequadas para a sua realização, observando que grandes oscilações negativas são indesejáveis impossibilitando as operações, atividades esportivas ligadas ao mar, entre outros, são representações de processos que podem se beneficiar diretamente desse tipo de estudo.

Com isso, a análise das séries temporais de níveis do mar tem sido cada vez mais importante. A previsão dos valores dos níveis médios do mar, bem como o estudo do seu comportamento é fundamental para evitar maiores problemas com relação às oscilações do nível do mar.

Este trabalho consiste na modelagem de séries temporais dos níveis médios do mar de Aburatsu (Japão), Key West (Flórida - Estados Unidos) e Ko Taphao Noi (Tailândia), com a finalidade de analisar o seu comportamento e fazer previsões. Utilizar métodos estatísticos para comparar estas séries duas a duas, com o objetivo de verificar se as observações em uma das séries são suficientes para explicar o comportamento da outra, e desenvolver um modelo de regressão na tentativa de identificar as variáveis que contribuem para a variação do nível do mar.

Este trabalho está organizado em cinco seções, incluindo esta introdução. Na segunda seção, é apresentado o referencial teórico, no qual são abordados conceitos sobre o nível do mar, séries temporais, testes de comparação e análise de regressão.

A terceira seção consiste em material e métodos, onde são descritas as séries utilizadas no estudo e a metodologia usada para o estudo. Na quarta seção, os resultados e discussões são apresentados, e na quinta seção, temos as conclusões referente ao estudo realizado.

1.1 Objetivos

1.1.1 Objetivo geral

Realizar um estudo temporal de dados dos níveis médios mensais de mares de diferentes oceanos verificando a tendência de crescimento, fazer comparação entre eles, e verificar a relação de dependência entre o nível do mar e algumas variáveis climáticas.

1.1.2 Objetivos específicos

- a) Identificar características das séries em estudo, como tendência e sazonalidade;
- b) Analisar a presença e os efeitos de possíveis intervenções, levando-se em conta fenômenos como El Niño;
- c) Buscar modelos que melhor se ajustem às séries de dados dos níveis do mar;
- d) Comparação entre níveis de mares de diferentes oceanos;
- e) Buscar associação entre o nível do mar e séries climáticas.
- f) Fazer previsões.

2 REFERENCIAL TEÓRICO

2.1 O nível do mar

O nível médio do mar é definido pelo IPCC (Intergovernmental Panel on Climate Change) como "a altura do mar em relação a uma referência de terra da região, em média, durante um período de tempo como um mês ou um ano, tempo suficiente para que as flutuações causadas por ondas e marés sejam removidas". Alterações no nível médio do mar medidas pelo marégrafo são chamadas de "mudanças de nível relativo do mar", porque podem acontecer tanto pelo movimento da terra em que o marégrafo está situado, ou por alterações na altura da superfície do mar (ambos considerados em relação ao centro da Terra como uma referência fixa). Tais termos têm unidades similares (mm/mês).

2.1.1 Medições

A medição do nível do mar tem uma longa história. Os povos antigos foram capazes de relacionar, em muitos locais, os movimentos regulares do mar com os movimentos do Sol e da Lua. De acordo com UNESCO (2006), estudos do século XIX referiam-se aos movimentos verticais dos continentes, na suposição de que o nível médio do mar se mantinha constante durante longos períodos de tempo. Atualmente sabemos que o nível do mar não se mantém constante. As variações do nível médio do mar estão relacionadas com mudanças no volume da água nos oceanos e também com as variações das correntes oceânicas.

O nível do mar necessita ser monitorado com "marégrafos" em muitos locais ao longo das faixas costeiras, especialmente em áreas densamente povoadas. As medições também podem ser realizadas nas profundezas oceânicas com o uso de sensores de pressão instalados no fundo do mar e, com os dados obtidos dos radares altimétricos satelitais. Pugh et al. (2002) alertam para o fato de que a base da análise científica do nível do mar tem que ser constituída por longas séries de medições. Segundo Harari; França; Camargo (2004), interesse especial se encontra nas determinações dos níveis médios mensais e anuais, os quais, em geral, fornecem importantes indicações sobre variações hidrodinâmicas sazonais e tendências de longo período.

Dada a fraca cobertura global, a qualidade razoável de registros de marégrafos e as incertezas nas correções para os movimentos de terra, as taxas encontradas para o crescimento do nível, durante este século, também devem ser questionadas (CHURCH et al., 2004), sendo necessário certo cuidado na escolha dos dados e em sua análise.

Deve-se também ter cuidado com relação às estimativas globais, segundo

Douglas (1997, *apud* CHURCH et al., 2001), a tendência de um único local é uma estimativa imprecisa da média global, tendo em vista que ele calculou a média de nove regiões e constatou um desvio padrão de cerca de 0,3 mm/ano.

2.1.2 Estudos sobre o comportamento do nível do mar

No relatório do IPCC, Solomon et al. (2007) afirmam que durante o período de 1961 a 2003, a taxa média de aumento do nível global médio do mar, estimado a partir dos dados dos marégrafos, é de $1,8 \pm 0,5$ mm/ano. Já Church et al. (2004) obtiveram 1,75, 1,89, e 1,91 mm/ano como média global estimada de aumento do nível do mar ao longo do período 1950 a 2000, calculadas com a utilização de três modelos distintos.

Wöppelmann et al. (2009) começam seu artigo afirmando que "as estimativas de crescimento em escala global do nível do mar, ao longo do século passado, são baseadas principalmente em longos registros. A gama de estimativas publicadas na literatura é bastante ampla (1 a 3 mm/ano), com valores que variam em torno de 1,8 mm/ano ". Eles ressaltam que "a origem das diferenças residem nos métodos utilizados para corrigir os registros do marégrafos e os deslocamentos verticais do terreno em que estão localizados".

Titus e Narayanan (1996) propõem um método simples para calcular projeções locais do nível médio do mar. Para considerar os movimentos de terra da região, e devido à incapacidade de mudar o modelo atual, eles propõem extrapolar linearmente o registro histórico e acrescentar a este uma projeção média global. No entanto eles apontam que, para evitar dupla contagem, é necessário fazer a correção da projeção global para a tendência correspondente de elevação do nível do mar. Church et al. (2004) alertam para o fato de que é necessário ter cuidado

na aplicação deste método, primeiro porque as projeções atuais indicam um modelo com variabilidade espacial substancial. Em segundo lugar, porque existem incertezas na precisão da tendência a partir do registro histórico. E em terceiro lugar, além de alterações no nível médio do mar, podem haver mudanças meteorológicas locais. Além disso, esses mesmos autores afirmam que "a taxa de aumento do nível do mar no século 21 é projetada para ser maior do que no século 20 na grande maioria das zonas costeiras". Na Figura 1, temos uma foto de uma praia na Carolina Norte (Estados Unidos) que reforça o que esses autores concluíram em seus estudos.



Figura 1 *Cape Hatteras, Carolina do Norte, EUA, em 1999 e 2004. Foto de Gary Braasch*

Diversos são os trabalhos realizados envolvendo o nível do mar. No entanto, em várias pesquisas realizadas até o momento, o trabalho mais próximo encontrado foi o estudo de Das e Radhakrishna (1991), que realizou a análise es-

pectral e ajustou um modelo integrado médias móveis (ARIMA) para dados de níveis do mar da Índia.

2.2 Fenômeno climático El Niño

De acordo com o INPE (Instituto Nacional de Pesquisas Espaciais), a ocorrência do fenômeno climático El Niño pode ter como efeito alterações consideráveis no nível do mar. O Sistema de Observação Global do Nível do Mar (GLOSS) também afirma que algumas das maiores mudanças no nível do mar ocorrem no Pacífico, a cada 3 ou 4 anos, como consequência de tal fenômeno. Dessa forma, é interessante analisar possíveis intervenções desse fenômeno nas séries estudadas.

O El Niño é definido pelo INPE como sendo uma alteração significativa de curta duração (12 a 18 meses), e representa o aquecimento anormal das águas superficiais e sub-superficiais do Oceano Pacífico Equatorial. O termo El Niño se referia originalmente a uma corrente oceânica que ocorria ao longo da costa do Peru e do Equador na época do Natal (de onde surgiu Niño em referência ao "menino Jesus"). Só posteriormente passou a ser associado aos incomuns aquecimentos que ocorrem a cada poucos anos e causam grandes alterações climáticas.

2.3 Séries temporais

Conforme Morettin e Toloi (2006), uma série temporal é um conjunto de observações coletadas de forma sequencial ao longo do tempo. A análise de uma série temporal pode ser feita no domínio do tempo ou no domínio da frequência, sendo os modelos propostos, respectivamente, paramétricos e não paramétricos.

Os modelos para descrever séries temporais são processos estocásticos, controlados por leis probabilísticas. Um processo estocástico é definido como sendo uma coleção de variáveis aleatórias sequenciadas no tempo e definidas em

um conjunto de pontos T , que pode ser contínuo ou discreto. A variável aleatória no tempo t é denotada por Z_t em que $t = 0, \pm 1, \pm 2, \dots \in T$. Um processo estocástico fica bem descrito por meio das funções média, variância e autocovariância (Morettin e Toloi, 2006).

Os objetivos da análise de série temporal são os seguintes, de acordo com Morettin e Toloi (2006):

1. investigar o mecanismo gerador da série temporal; por exemplo, analisando uma série de alturas de ondas, podemos querer saber como estas ondas foram geradas;
2. fazer previsões de valores futuros da série; estas podem ser a curto prazo, como para séries de vendas, produção ou estoque, ou a longo prazo, como para séries populacionais, de produtividade etc. No caso de séries de níveis do mar, a previsão a longo prazo nem sempre é confiável;
3. descrever apenas o comportamento da série; neste caso, a construção do gráfico, a verificação da existência de tendências, ciclos e variações sazonais, a construção de histogramas e diagramas de dispersão etc., podem ser ferramentas úteis;
4. procurar periodicidades relevantes nos dados; aqui, a análise espectral, pode ser de grande utilidade.

2.3.1 Estacionariedade

Uma das suposições, que se faz a respeito de uma série temporal, é a de que ela seja estacionária, ou seja, ela se desenvolve no tempo ao longo de uma média constante, refletindo alguma forma de equilíbrio estável. Existem duas formas

de estacionariedade, a fraca (ou de segunda ordem) e a forte. Como a estacionariedade forte é uma propriedade exigente, faz sentido definir um conceito de estacionariedade baseado nos momentos de primeira e segunda ordem de uma série temporal, que é dado pela estacionariedade fraca.

Morettin e Toloi (2006) definem um processo estocástico como sendo fracamente estacionário se:

- i $E[Z(t)] = \mu(t) = \mu, \forall t \in T$;
- ii $E[Z^2(t)] < \infty, \forall t \in T$;
- iii $\gamma(t_1, t_2) = cov[Z_1(t), Z_2(t)]$ é uma função de $|t_1 - t_2|$.

Um processo com estacionariedade forte (também conhecido como "ruído branco"), consiste numa seqüência de variáveis aleatórias (a_t), mutuamente independentes e identicamente distribuídas, possuindo média e variância constantes.

2.3.2 Definições importantes

As seguintes definições abaixo foram obtidas de acordo com Morettin e Toloi (2006).

Definição 2.3.2.1 A *função de autocovariância* (facv) é dada por

$$\gamma_\tau = E \{Z_t Z_{t+\tau}\} \quad (2.1)$$

em que Z_t é um processo estacionário real discreto de média zero e τ representa a defasagem no tempo.

O estimador da função de autocovariância γ_τ é dado por

$$c_\tau = \frac{1}{N} \sum_{t=1}^{N-\tau} (Z_t - \bar{Z})(Z_{t+\tau} - \bar{Z}), \quad (2.2)$$

em que $\tau = 1, 2, \dots, N - 1$, N é o número de observações da série e \bar{Z} é a média amostral.

Definição 2.3.2.2 A *função de autocorrelação (fac)* é definida por

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0} \quad (2.3)$$

em que $\tau \in \mathbb{Z}$.

O estimador da função de autocorrelação ρ_τ é dada por

$$r_\tau = \frac{c_\tau}{c_0} \quad (2.4)$$

sendo c_τ a função de autocovariância na defasagem τ e c_0 a variância.

2.3.3 Decomposição clássica - Tendência e Sazonalidade

De acordo com Morettin e Tolo (2006), um conjunto de observações de uma série temporal, representada por Z_t , pode ser decomposta como uma soma de três componentes não observáveis:

$$Z_t = T_t + S_t + a_t \quad (2.5)$$

em que T_t representa a tendência, S_t a sazonalidade, e a_t é uma componente aleatória, de média zero e variância constante σ_a^2 . Tomando a_t como processo

aleatório segue que Z_t , será em geral um processo não estacionário.

A tendência pode ser entendida como um aumento ou diminuição gradual das observações ao longo do tempo, a sazonalidade indica possíveis flutuações ocorridas sempre em períodos menores ou iguais a doze meses e a componente aleatória mostra oscilações aleatórias irregulares. Dessa forma, o interesse em estudar um modelo desse tipo será o de estimar S_t e T_t , e construir a série livre dessas componentes, uma série estacionária.

As componentes T_t e S_t são, em geral, bastante relacionadas, assim a especificação de uma depende da especificação da outra. Para a especificação da tendência, pode-se observar se existe tendência polinomial, bem como optar por métodos de suavização. Pode-se também aplicar uma diferença para eliminar a tendência. Para a sazonalidade determinística, tem-se o método de regressão como meio de estimá-la, e para a sazonalidade estocástica tem-se a suavização.

Existem vários testes para tendência e sazonalidade, neste trabalho, será considerado o teste do sequência (Wald-Wolfowitz) para verificar a presença de tendência, o teste de Fisher para verificar a presença de sazonalidade e o teste HEGY para verificar a natureza da sazonalidade. Esses testes serão aplicados nas séries estudadas de modo a obter posteriormente uma série estacionária.

2.3.4 Análise espectral

De acordo com Morettin e Toloí (2006), temos a seguinte definição:

Definição 2.3.4.1 *Considere a série temporal estacionária $Z(t)$, $t = 0, \pm 1, \dots$ com função de autocovariância γ_τ , satisfazendo*

$$\sum_{k=-\infty}^{\infty} |\gamma_k| < \infty, \quad (2.6)$$

*então o **espectro**, $f(\lambda)$, de $Z(t)$ na frequência λ existe e é dado pela transformada*

de Fourier de γ_τ .

Seja $Z_t, t = 0, 1, \dots, N$ um processo estacionário com média zero. A transformada de Fourier discreta finita dos valores (Z_1, \dots, Z_N) é dada pela expressão:

$$d_j^{(N)} = \frac{1}{\sqrt{2\pi N}} \sum_{t=1}^N Z_t \cos\left(\frac{2\pi jt}{N}\right) + i \frac{1}{\sqrt{2\pi N}} \sum_{t=1}^N Z_t \sin\left(\frac{2\pi jt}{N}\right), j = 0, 1, \dots, \left[\frac{N}{2}\right]$$

Suponha uma realização de um processo estacionário $Z_t, t = 0, 1, \dots, N$.

Então,

$$I_j^{(N)} = |d_j^{(N)}|^2 = \frac{1}{2\pi N} \left| \sum_{t=1}^N Z_t e^{-i\lambda_j t} \right|^2$$

é denominado periodograma, cuja distribuição assintótica é dado pelo teorema a seguir, cuja demonstração pode ser encontrada em Priestley (1989).

Teorema 1: As ordenadas do periodograma $I_j^{(N)}$ são variáveis aleatórias assintoticamente independentes e têm distribuição assintótica múltipla de uma variável aleatória qui-quadrado, isto é,

$$I_j^{(N)} \xrightarrow{\mathcal{D}} \begin{cases} \frac{1}{2} f(\lambda_j) \chi_2^2 & j \neq 0, N/2, \\ f(\lambda_j) \chi_1^2 & j = 0, N/2. \end{cases}$$

2.3.5 Teste de Fisher

Para testar a presença de sazonalidade utiliza-se o teste de Fisher, que foi proposto inicialmente para testar o maior período. As hipóteses a serem testadas são:

H_0 : não existe sazonalidade

H_1 : existe sazonalidade

A estatística do teste é dada por:

$$g = \frac{\max I_p}{\sum_{p=1}^{N/2} I_p}$$

em que I_p é o valor do periodograma no período p e N é o número de observações da série. A estatística do teste de Fisher é dada por:

$$z_\alpha = 1 - \left(\frac{\alpha}{n}\right)^{\frac{1}{n-1}}$$

em que α é o nível de significância do teste e $n = \frac{N}{2}$.

Se $g > z_\alpha$, rejeita-se H_0 , e se pode concluir que existe sazonalidade correspondente ao período p .

2.3.6 Teste de sequência (Wald-Wolfowitz)

O teste consiste em considerar as N observações $Z_t, t = 1, \dots, N$ de uma série temporal, e seja m a mediana destes valores. Atribuímos a cada valor Z_t o símbolo A , se ele for maior ou igual a m , e B se ele for menor que m , teremos então, $N = (n_1 \text{ pontos } A) + (n_2 \text{ pontos } B)$. A estatística do teste é dada por:

$T_1 =$ número total de sequências (isto é, grupos de símbolos iguais)

Rejeitamos a hipótese nula H_0 (de existência da tendência) se há poucas sequências. Para um dado α , rejeitamos H_0 se $T_1 < w_\alpha$, onde w_α é o quantil da distribuição de T_1 , que é tabelado. A tabela pode ser encontrada em Conover (1990).

2.3.7 Teste da raiz unitária sazonal

Após detectar a presença da componente sazonal na série, é importante detectar a natureza dessa componente. Ao se trabalhar com dados com periodicidade mensal ou trimestral, geralmente a sazonalidade é tratada através das variáveis *dummies* ou diferenciação sazonal. Cada um desses tratamentos será correto se for aplicado a um determinado processo gerador da série. Se usarmos as variáveis *dummies* onde o correto seria diferenciar, teremos um modelo mal especificado que poderá levar a conclusões erradas (ALONSO; SEMAÁN, 2010).

Dessa forma torna-se necessário o uso de algum método para determinar a natureza da componente sazonal. Este método é representado pelo teste de verificação da raiz unitária sazonal: o teste HEGY. As primeiras aplicações deste teste foram em séries temporais com sazonalidade trimestral, e foram realizadas por Hasza-Fuller, em seguida, por Dickey-Fuller-Hasza e desenvolvido logo depois por Hylleberg et al. (1990).

A componente sazonal pode ser tanto determinística como estocástica, sendo que será estocástica se as raízes unitárias não estiverem presentes e determinística caso contrário. O teste HEGY permite identificar a natureza das variações sazonais. A decisão correta com relação aos resultados do teste faz com que seja possível escolher os métodos adequados de eliminação das variações sazonais.

Inicialmente, temos que o operador diferença $(1 - B)^{12}$ pressupõe a presença de 12 raízes no círculo unitário:

$$\begin{aligned}
(1 - B^{12}) &= (1 - B)(1 + B)(1 - iB)(1 + iB) \\
&\quad \times \left[1 + (\sqrt{3} + i)B/2\right] \left[1 + (\sqrt{3} - i)B/2\right] \\
&\quad \times \left[1 - (\sqrt{3} + i)B/2\right] \left[1 - (\sqrt{3} - i)B/2\right] \\
&\quad \times \left[1 + (i\sqrt{3} + i)B/2\right] \left[1 - (i\sqrt{3} - i)B/2\right] \\
&\quad \times \left[1 - (i\sqrt{3} + i)B/2\right] \left[1 + (i\sqrt{3} - i)B/2\right]
\end{aligned} \tag{2.7}$$

onde todos os termos diferentes de $(1 - B)$ correspondem a raízes unitárias sazonais.

Assim, o fato de $\Delta_{12} y_t = y_t - y_{t-12} = (1 - B^{12}) y_t$ ser estacionário envolve as seguintes condições:

1. $(1 - B) = 0$, neste caso $y_t - y_{t-1}$ será estacionário e $y_t \approx y_{t-1}$, implicando numa *tendência estocástica*.
2. $(1 + B) = 0$, neste caso $y_t + y_{t-1}$ será estacionário e $y_t \approx y_{t-2}$, implicando numa *raiz unitária bimestral*.
3. $(1 - iB)(1 + iB) = (1 + B^2) = 0$, neste caso $y_t + y_{t-2}$ será estacionário e $y_t \approx y_{t-4}$, implicando numa *raiz unitária quadrimestral*.
4. $[1 + (1 + \sqrt{3}i)L/2][1 + (1 - \sqrt{3}i)L/2] = 0$, que equivale a $1 + L + L^2 = 0$, neste caso $y_t + y_{t-1} + y_{t-2}$ será estacionário e $y_t \approx y_{t-3}$, implicando numa *raiz unitária trimestral*.
5. $[1 - (1 + \sqrt{3}i)L/2][1 - (1 - \sqrt{3}i)L/2] = 0$, que equivale a $1 - L + L^2 = 0$, neste caso $y_t - y_{t-1} + y_{t-2}$ será estacionário e $y_t \approx y_{t-6}$, implicando numa

raiz unitária semestral.

6. $[1 + (\sqrt{3} + i)L/2][1 + (\sqrt{3} - i)L/2] = 0$, que equivale a $1 + \sqrt{3}L + L^2 = 0$, neste caso $y_t + \sqrt{3}y_{t-1} + y_{t-2}$ será estacionário, implicando numa *raiz unitária com 5 ciclos por ano.*
7. $[1 - (\sqrt{3} + i)L/2][1 - (\sqrt{3} - i)L/2] = 0$, que equivale a $1 - \sqrt{3}L + L^2 = 0$, neste caso $y_t - \sqrt{3}y_{t-1} + y_{t-2}$ será estacionário e $y_t \approx y_{t-12}$, implicando numa *raiz unitária anual.*

O teste inicial era para dados trimestrais e Franses (1990) estendeu o teste HEGY para o caso de dados mensais, tal procedimento é baseado na seguinte regressão auxiliar, estimada por Mínimos Quadrados Ordinários:

$$\begin{aligned} \phi^*(B)y_{8,t} = & \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-1} + \pi_4 y_{3,t-2} + \pi_5 y_{4,t-1} \\ & + \pi_6 y_{4,t-2} + \pi_7 y_{5,t-1} + \pi_8 y_{5,t-2} + \pi_9 y_{6,t-1} + \pi_{10} y_{6,t-2} \\ & + \pi_{11} y_{7,t-1} + \pi_{12} y_{7,t-2} + \mu_t + \epsilon_t \end{aligned} \quad (2.8)$$

em que $\phi^*(B)$ é uma função polinomial de B (o operador de defasagens), μ_t é um componente determinístico, que pode ser uma constante, variáveis *dummies* sazonais e/ou uma tendência e ϵ_t um ruído branco.

As variáveis da equação 2.8 foram criadas de forma a permitir a prova de cada uma das condições, e são descritas pelos seguintes polinômios:

$$\begin{aligned} y_{1,t} &= (1 + B)(1 + B^2)(1 + B^4 + B^8)y_t \\ y_{2,t} &= -(1 - B)(1 + B^2)(1 + B^4 + B^8)y_t \\ y_{3,t} &= -(1 - B^2)(1 + B^4 + B^8)y_t \\ y_{4,t} &= -(1 - B^4)(1 - \sqrt{3}B + B^2)(1 + B^2 + B^4)y_t \end{aligned}$$

$$y_{5,t} = -(1 - B^4)(1 + \sqrt{3}B + B^2)(1 + B^2 + B^4)y_t$$

$$y_{6,t} = -(1 - B^4)(1 - B^2 + B^4)(1 - B + B^2)y_t$$

$$y_{7,t} = -(1 - B^4)(1 - B^2 + B^4)(1 + B + B^2)y_t$$

$$y_{8,t} = (1 - B^{12})y_t$$

Partindo de (2.8) podemos testar as seguintes hipóteses:

- $H_a : \pi_1 = 0$ (tendência estocástica)
- $H_b : \pi_2 = 0$ (raiz bimestral)
- $H_c : \pi_3 = \pi_4 = 0$ (raiz quadrimestral)
- $H_d : \pi_5 = \pi_6 = 0$ (raiz trimestral)
- $H_e : \pi_7 = \pi_8 = 0$ (raiz semestral)
- $H_f : \pi_9 = \pi_{10} = 0$ (raiz de 5 ciclos por ano)
- $H_g : \pi_{11} = \pi_{12}$ (raiz anual)
- $H_h : \pi_1 = \pi_2 = \dots = 0$ (todas as raízes presentes)
- $H_i : \pi_2 = \pi_3 = \dots = 0$ (todas raízes sazonais presentes)

Após a estimação dos parâmetros de (2.8), analisa-se a significância estatística de π_1 e π_2 por meio da estatística "t", e a significância conjunta dos outros parâmetros através da estatística "F", cujos valores críticos são apresentados em Franses (1991) e Franses e Hobjn (1997).

De acordo com Franses (1991), o processo gerador da série y_t terá raiz unitária não-sazonal na frequência zero (tendência estocástica), se $\pi_1 = 0$, e raiz unitária sazonal se qualquer um dos demais π_i 's ($i = 2, 3, \dots, 12$) for igual a zero. Devido ao fato das raízes unitárias complexas serem conjugados, deve-se notar

que essas raízes estão presentes apenas quando os pares de π'_i s são iguais a zero simultaneamente. Por exemplo, as raízes i e $-i$ só estão presentes quando π_3 e π_4 são iguais a zero. Não haverá raiz unitária sazonal se π_2, \dots, π_{12} são simultaneamente diferentes de zero, nesse caso e tendo $\pi_1 = 0$ a sazonalidade pode ser modelada com *dummies* sazonais e a primeira diferença da série. No caso em que π_1, \dots, π_{12} são simultaneamente iguais a zero, teremos sazonalidade estocástica, sendo apropriado aplicar a diferença de ordem 12. A série terá raiz unitária semestral se $\pi_2 = 0$.

Os parâmetros de tendência e os coeficientes das variáveis *dummies*, que podem ser incluídas no modelo (2.8), são testadas com o auxílio do teste t de Student, de acordo com as hipóteses de decisão e as regras correspondentes ao teste bilateral.

Em Beaulieu e Miron (1993) também temos a extensão para dados mensais, mas a sua equação de teste difere da equação (2.8) e é um pouco mais complicada. Além disso, os autores não consideram o teste F para a presença das raízes complexas.

Os testes de raiz unitária sazonal foram implementados em um número reduzido de programas de análise de séries temporais. Neste trabalho, o JmulTi (L'UTKEPOHL; KR'ATZIG, 2004), que tem o teste HEGY implementado, é o programa utilizado.

2.3.8 Estimação da sazonalidade - método de regressão

O métodos de regressão são ótimos para séries que apresentam sazonalidade determinística. Este método consiste em considerar o modelo $Z_t = T_t + S_t + a_t$, em que $T_t = \sum_{j=0}^m \beta_j t^j$ e $S_t = \sum_{j=1}^{12} \alpha_j d_{jt}$, onde d_{jt} são variáveis periódicas

(senos, cossenos ou variáveis *dummies* sazonais).

Supondo sazonalidade constante, α_j não depende de t . Então

$$d_{jt} = \begin{cases} 1, & \text{se o período } t \text{ corresponde ao mês } j, j = 1, 2, \dots, 12. \\ 0, & \text{caso contrário.} \end{cases}$$

De modo que o somatório dos d_{jt} para $j = 1, 2, \dots, 12$ será 1 e a matriz de regressão não será de posto completo. No entanto, se a restrição adicional

$$\sum_{j=1}^{12} \alpha_j = 0 \quad (2.9)$$

for imposta, um modelo de posto completo é obtido:

$$Z_t = \sum_{j=0}^m \beta_j t^j + \sum_{j=1}^{11} \alpha_j D_{jt} + a_t$$

em que

$$D_{jt} = \begin{cases} 1, & \text{se o período } t \text{ corresponde ao mês } j, \\ -1, & \text{se o período } t \text{ corresponde ao mês } 12, \\ 0, & \text{caso contrário, } j = 1, \dots, 11. \end{cases}$$

Desta forma, através do método de mínimos quadrados pode-se obter os estimadores de α_j e β_j , ressaltando que α_{12} será obtido a partir de 2.9. Após obter os estimadores do modelo, a componente sazonal determinística poderá ser eliminada.

2.3.9 Metodologia Box-Jenkins

Segundo Morettin e Tolo (2006), a construção dos modelos de Box e Jenkins (1976) é baseada em um ciclo iterativo, no qual a escolha do modelo é feita com base nos próprios dados. Essa metodologia consiste em ajustar modelos auto-regressivos integrados de médias móveis, ARIMA, a um conjunto de dados.

1. Especificação: uma classe geral de modelos é considerada para análise;
2. Identificação: consiste em descobrir qual dentre as várias versões dos modelos de Box-Jenkins, sejam eles sazonais ou não, descreve o comportamento da série com base na análise de autocorrelações, autocorrelações parciais e outros critérios;
3. Estimação: os parâmetros do modelo identificado são estimados;
4. Verificação: consiste em avaliar se o modelo estimado é adequado para descrever o comportamento dos dados, e é feita através de uma análise de resíduos.

Caso o modelo não seja adequado, repetem-se as etapas voltando-se à fase de identificação.

2.3.10 Modelos lineares estacionários

Existem três casos particulares de processos lineares estacionários. São eles:

- (i) processo auto-regressivo de ordem p , $AR(p)$;
- (ii) processo de médias móveis de ordem q , $MA(q)$;

(iii) processo auto-regressivo e de médias móveis de ordens p e q , ARMA(p, q);

2.3.10.1 Modelos médias móveis de ordem q , MA(q)

O processo de médias móveis de ordem q , MA(q), é aquele onde os valores atuais de uma série temporal Z_t dependem dos erros aleatórios, representados por ruídos brancos passados. Assim,

$$\tilde{Z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad (2.10)$$

e usando o operador diferença teremos

$$\tilde{Z}_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t = \theta(B) a_t \quad (2.11)$$

onde $\theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ é o operador de médias móveis de ordem q .

Este processo possui algumas características especiais. A sua função de autocorrelação (fac), que consiste nas sucessivas correlações entre uma série temporal e ela mesma atrasada k defasagens, vale zero a partir da defasagem q .

Uma condição importante que deve ser respeitada nos processos de séries temporais é a garantia da inversibilidade do processo. Isto significa que é possível identificar um e somente um processo através das suas características de primeira e segunda ordens. A condição de invertibilidade para um modelo MA(q) é que todas as raízes da equação característica, $\theta(B) = 0$, estejam fora do círculo unitário.

2.3.10.2 Modelos auto-regressivos de ordem p , AR(p)

Um processo auto-regressivo de ordem p é aquele onde os valores atuais dependem dos valores passados da própria série, junto com o ruído branco:

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t \quad (2.12)$$

Com a utilização do operador diferença B podemos reescrever a equação 2.12 como

$$\phi(B)\tilde{Z}_t = a_t \quad (2.13)$$

em que $\phi(B) = (1 - \phi_1(B) - \phi_2(B)^2 - \dots - \phi_p(B)^p)$ é o polinômio auto-regressivo de ordem p .

O processo auto-regressivo precisa sofrer uma restrição para garantir a sua estacionariedade em segunda ordem. Para um $AR(p)$ ser estacionário é necessário que todas as raízes do polinômio $\phi(B) = 0$ caiam fora do círculo unitário.

2.3.10.3 Função de autocorrelação parcial

No modelo auto-regressivo é bem mais interessante a avaliação da autocorrelação parcial (facp) do que a autocorrelação (fac), devido à sua natureza recursiva. A autocorrelação parcial pode ser informalmente definida como a contribuição da correlação em uma determinado defasagem dada a ausência dos coeficientes das demais defasagens.

Seja ϕ_{kj} o j -ésimo coeficiente de um modelo $AR(k)$, de tal modo que ϕ_{kk} seja o último coeficiente. Sabe-se que:

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \dots + \phi_{kk}\rho_{j-k}, \quad j = 1, \dots, k,$$

a partir dos quais obtemos as equações de Yule-Walker, e resolvendo estas equações sucessivamente para $k = 1, 2, 3, \dots$ obtemos

$$\phi_{kk} = \frac{|\mathbf{P}_k^*|}{|\mathbf{P}_k|}$$

onde \mathbf{P}_k é a matriz de autocorrelações e \mathbf{P}_k^* é a matriz \mathbf{P}_k com a última coluna substituída pelo vetor de autocorrelações.

Como no processo de médias móveis, caracteristicamente a facp de um $\text{AR}(p)$ também se iguala a zero quando $k > p$.

Considerando que o processo é $\text{AR}(p)$, tem-se:

$$\text{Var}(\hat{\phi}_{kk}) \approx \frac{1}{N}, k \geq p + 1. \quad (2.14)$$

E de acordo com Morettin e Toloí (2006), se o número de observações, N , for suficientemente grande, $\hat{\phi}_{kk}$ tem distribuição aproximada normal, o que permite a construção de intervalos de confiança para $\hat{\phi}_{kk}$.

2.3.11 Modelos ARIMA

Os modelos ARIMA resultam da combinação de três componentes: o componente auto-regressivo (AR), o de integração (I) e o componente de médias móveis (MA). Este modelo considera a tendência da série temporal, tem ordem (p, d, q) e pode ser representado por:

$$\phi(B)(1 - B)^d Z_t = \theta(B)a_t \quad (2.15)$$

sendo

$\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ o operador auto-regressivo de ordem p ;

$\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ o operador de médias móveis de ordem q ;

B o operador translação para o passado, tal que $B^m Z_t = Z_{t-m}$;

d o número de diferenças necessárias para retirar a tendência da série e deixá-la estacionária.

2.3.12 Identificação, estimação e previsão com modelos ARIMA

A identificação do modelo é feita com base nas autocorrelações e autocorrelações parciais estimadas, o seu objetivo é determinar os valores de p , d e q do modelo $ARIMA(p, d, q)$, além das estimativas preliminares dos parâmetros a serem usados no estágio de verificação. Dessa forma, inicialmente serão feitos os correlogramas das séries de níveis médios do mar, com o objetivo de determinar o modelo a ser ajustado.

Na fase de estimação, será necessário utilizar um procedimento iterativo de estimação não-linear de mínimos quadrados sendo que as estimativas preliminares encontradas na fase de identificação serão usadas como valores iniciais neste procedimento. Assim, escolhido os valores iniciais do modelos, estimaremos seus coeficientes.

Após estimar o modelo, é necessário verificar se ele representa os dados de maneira adequada. A verificação pode ser feita analisando os resíduos. Suponha que o modelo ajustado seja

$$\phi(B)W_t = \theta(B)a_t \quad (2.16)$$

em que

$$W_t = \Delta^d Z_t \quad (2.17)$$

Caso o modelo seja verdadeiro, os erros $a_t = \theta^{-1}(B)\phi(B)W_t$ consti-

tirão um ruído branco. Esta verificação pode ser feita pelo teste de autocorrelação residual e pelo teste de Box-Pierce (Morettin e Tolo, 2006). Dessa forma, as séries de níveis do mar, terão um modelo com um bom ajuste se os resíduos forem um ruído branco.

A previsão é um dos objetivos ao se ajustar um modelo. A previsão de Z_{t+h} , para $h = 1, 2, \dots$ é denotada por $\hat{Z}_t(h)$. O erro de previsão é dado por $e_t(h) = Z_{t+h} - \hat{Z}_t(h)$ em que Z_{t+h} é o valor real e $\hat{Z}_t(h)$ é o valor predito. Após obter um modelo com um bom ajuste para as séries, faremos previsões para analisar como está o comportamento do nível médio do mar.

2.3.13 Critérios para seleção do modelo

Vários modelos podem ser identificados para descrever uma série, mas existem critérios para a escolha do melhor modelo. Dentre diversos critérios, têm-se o critério de informação de Akaike (AIC), o erro quadrático médio de previsão (EQMP) e o erro percentual médio absoluto (MAPE), que são analisados a partir do modelo ajustado.

- Akaike (1973) sugere escolher o modelo cujas ordens p e q minimizem o critério. O critério de Akaike, com N fixo, pode ser expresso por:

$$N \log \hat{\sigma}_a^2 + 2(k + l + 2)$$

em que $\hat{\sigma}_a^2$ é o estimador de máxima verossimilhança de σ_a^2 , $0 \leq k \leq p$ e $0 \leq l \leq q$.

Dessa forma, o modelo escolhido deve ter o menor valor para o critério.

- Caso previsão seja um dos objetivos, o melhor modelo será o que apresentar

o menor o erro quadrático médio de previsão. O EQMP com origem em t é dado por:

$$\text{EQMP}_t = \frac{1}{N} \sum_{h=1}^n \left[Z_{t+h} - \widehat{Z}_t(h) \right]^2$$

onde Z_{t+h} é o valor real e $\widehat{Z}_t(h)$ é o valor predito.

- O erro percentual médio absoluto (MAPE) é utilizado para verificar o bom desempenho de ajuste de um modelo, o melhor modelo também será o que tiver menor MAPE. É dado por:

$$\text{MAPE} = \frac{1}{h} \sum_1^h \left| \frac{e_t(h)}{Z_{t+h}} \right| \times 100$$

em que $e_t(h)$ é o erro de previsão, Z_{t+h} é o valor real e h é o número de previsões.

2.3.14 Análise de resíduos

Após estimar o modelo, temos que verificar se ele representa, ou não, adequadamente os dados. Para isso, testamos a autocorrelação residual através das funções de autocorrelação dos resíduos e também através do seguinte teste:

- Teste de Box e Pierce

Uma variação do teste de Box e Pierce (1970) foi sugerida por Ljung e Box.

Dessa forma temos que se o modelo for apropriado a estatística

$$Q(K) = n(n+2) \sum_{j=1}^K \frac{\widehat{r}_j^2}{(n-j)}$$

terá aproximadamente uma distribuição χ^2 com $K-p-q$ graus de liberdade, onde p e q são, respectivamente, o número de parâmetros auto-regressivos e de médias móveis.

A hipótese de ruído branco para o resíduo é rejeitada para valores grandes de $Q(K)$, ou seja, se $Q(K) < \chi_{K-p-q}^2$ então o resíduo é ruído branco.

2.3.15 Análise de intervenção

De acordo com Morettin e Tolo (2006), intervenção é a ocorrência de algum tipo de evento em dado instante de tempo T , conhecido a priori. Esta ocorrência pode se manifestar por um intervalo de tempo subsequente e que afeta temporariamente, ou permanentemente, a série em estudo. O objetivo da análise de intervenção é avaliar o impacto de um determinado evento no comportamento da série.

A análise de intervenção pode alterar a série na sua manifestação sendo, podendo ser abrupta ou gradual, e na sua duração sendo temporária ou permanente. De maneira geral, um modelo desse tipo pode ser representado da seguinte forma:

$$Z_t = \sum_{j=1}^k v(B)X_{j,t} + N_t,$$

onde, no caso do presente trabalho:

- $X_{j,t}$ é uma variável binária do tipo função impulso

$$X_{j,t} = I_t^T = \begin{cases} 0, & t \neq T \\ 1, & t = T \end{cases}$$

- $v_j(B), j = 1, \dots, k$ são funções racionais da forma $\frac{\omega_j(B)B^{b_j}}{\delta_j(B)}$, onde $\omega_j(B) = \omega_{j,0} - \omega_{j,1}B - \dots - \omega_{j,s}B^s$ e $\delta_j(B) = 1 - \delta_{j,1}B - \dots - \delta_{j,r}B^r$ são polinômios em B , b_j é a defasagem no tempo para o início do efeito da j -ésima intervenção e
- N_t é a série temporal livre do efeito de intervenções e é denominada série residual.

Para as séries de níveis médios do mar, procuraremos identificar possíveis intervenções relacionadas aos fenômenos El Niño, testando modelos que incorporem a intervenção nas séries que apresentarem a manifestação deste evento.

2.4 Procedimentos para comparação de séries temporais

Na análise de séries temporais, muitas vezes, é de interesse verificar se duas séries, ou trechos de uma mesma série, estão sendo gerados pelo mesmo processo estocástico. Se tivermos séries de níveis médio do mar geradas pelo mesmo processo, será suficiente colher amostras em apenas um local, reduzindo tempo e custo. Existem também testes multivariados para a comparação de séries, como o proposto por Carmona e Wang (1994) e Dette e Paparoditis(2009), mas estes não serão abordados aqui.

As técnicas de comparação existentes são aplicáveis aos sistemas que são estacionários, ou para aqueles que não são estacionários e podem se tornar estacionários por alguma transformação simples, como a diferenciação (COSTA; SÁ-FADI, 2010). Esses procedimentos são basicamente baseados na comparação entre as estimativas das densidades espectrais de pares de séries temporais, ou da estimativa dos coeficientes de um modelo auto-regressivo $AR(p)$.

Alguns dos estudos mais relevantes são: no domínio do tempo, elaborados por Quenouille (1958) e Pudney et al. (1999), e no domínio da frequência por Coates e Diggle (1986). Outra distinção entre estas técnicas é se elas são paramétricas ou não paramétricas. Os métodos paramétricos são mais poderosos, no sentido de serem capazes de detectar diferenças menores na série. Por outro lado, os métodos não-paramétricos são mais simples, pois não exigem a construção de um modelo paramétrico auxiliar.

Echeverry e Tolo (2000) apresentam a descrição completa dos testes de Igualdade das Funções de Autocorrelação e das Somas Acumuladas, propostos por Quenouille (1958) e Coates e Diggle (1986), respectivamente. Tais testes são univariados e requerem que as séries sejam estacionárias, podendo também ser usadas séries que se tornaram estacionárias por algum processo. Coates e Diggle (1986) discutiram o problema de testar se duas séries vêm de processos estocásticos estacionários com a mesma forma espectral. A técnica utilizada é não paramétrica, e os periodogramas dos processos são usados para produzir um teste. Quenouille (1958) apresenta um teste, realizado no domínio do tempo, para comparar as estruturas de autocorrelação de um par de séries temporais.

2.4.1 Teste das Somas Acumuladas (Coates e Diggle, 1986)

Sejam $I_1(\lambda_i)$ e $I_2(\lambda_i)$ os periodogramas dos processos estacionários $Z_1(t)$ e $Z_2(t)$, respectivamente. Temos pelo Teorema 1, que $I_j(\lambda) \sim \frac{f_j(\lambda)\chi_2^2}{2}$, com $0 < \lambda < \pi$. Sendo $f_j(\lambda)$ o espectro na frequência λ da série temporal.

Definindo as razões espectrais

$$J(\lambda) = \frac{I_1(\lambda)}{I_2(\lambda)} \quad e \quad u(\lambda) = \frac{f_1(\lambda)}{f_2(\lambda)}$$

com $Z_1(t)$ e $Z_2(t)$ independentes temos que, assintoticamente,

$$J(\lambda) = \frac{I_1(\lambda)}{I_2(\lambda)} \sim u(\lambda)F_{2,2}$$

e

$$\ln J(\lambda) \sim \text{logística} \{ \ln u(\lambda), 1 \},$$

onde F é a distribuição de Fisher-Snedecor e logística $(\alpha, 1)$ denota a distribuição logística cuja função de distribuição é definida por:

$$F(x) = 1 + e^{-(x-\alpha)^{-1}}, \quad -\infty < x < \infty.$$

Temos para $Z \sim \text{logística}(\alpha, 1)$, que $E(Z) = \alpha$ e $Var(Z) = \frac{\pi^2}{3}$. Portanto, $\ln(J)$ é um estimador não viesado, mas inconsistente para o log da razão espectral $(\ln u(\lambda))$ e sua variância assintótica independe de λ .

Pode-se demonstrar, ver Coates e Diggle (1986), que os valores,

$$z_i = \ln(1 + J^{-1}(\lambda_i)) \sim u(\lambda_i)exp(1), \quad (2.18)$$

com $\lambda_i = \frac{2\pi_i}{N}$, $i = 1, \dots, m$, e $m = \lfloor \frac{N-1}{2} \rfloor$.

Dessa forma, se $c_j = \sum_{i=1}^j z_i$ então $o_j = \frac{c_j}{c_m}$ é o valor das estatísticas de ordem com distribuição uniforme no intervalo $(0, 1)$.

De acordo com Echeverry e Tolo (2000), o teste proposto por Coates e Diggle (1986), consiste em construir uma estatística baseada no periodograma e usar a estatística de Kolmogorov-Smirnov (CONOVER, 1990) para testar afastamentos da distribuição $U(0, 1)$. Tendo a seguinte hipótese a ser testada:

$$H_0 : f_1(\lambda) = f_2(\lambda), \quad \text{para todo } 0 < \lambda < \pi$$

$$H_1 : f_1(\lambda) \neq f_2(\lambda), \quad \text{para algum } 0 < \lambda < \pi$$

A seguir, temos os passos para a realização do teste, de acordo com Echeverry e Tolo (2000).

1. Calcular os periodogramas $I_1(\lambda_i)$ e $I_2(\lambda_i)$, com $i = 1, \dots, m$ das séries $Z_1(t)$ e $Z_2(t)$, respectivamente;
2. Calcular a razão dos periodogramas, $J(\lambda_i) = \frac{I_1(\lambda_i)}{I_2(\lambda_i)}$, com $i = 1, \dots, m$;
3. Calcular $z_i = \ln(1 + J^{-1}(\lambda_i))$, com $i = 1, \dots, m$;
4. Calcular $c_j = \sum_{i=1}^j z_i$, com $j = 1, \dots, m$;
5. Obter as estatísticas $o_j = \frac{c_j}{c_m}$, com $j = 1, \dots, m$;
6. Comparar a distribuição dos o_j com a distribuição $U(0, 1)$, utilizando o teste de Kolmogorov-Smirnov. Se o valor p for maior que α , não se rejeita a hipótese H_0 , ao nível de significância α .

Coates e Diggle (1986) alertam para o fato de que o teste pode se tornar mais fraco para valores da razão do periodograma que produzem um número maior de valores negativos de $\ln J(\lambda)$.

2.4.2 Teste de Igualdade das Funções de Autocorrelação (Quenouille, 1958)

O teste proposto por Quenouille (1958) visa comparar se duas séries ou duas seções de séries têm a mesma estrutura de correlação, este teste não requer

que as séries tenham o mesmo número de observações, e tem a seguinte hipótese a ser testada:

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

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

onde $\rho_1(j)$ e $\rho_2(j)$ são as funções de autocorrelação das séries $Z_1(t)$ e $Z_2(t)$, respectivamente.

Segundo Echeverry e Toloi (2000), este método para testar as funções de autocorrelação pode ser descrito pelos seguintes passos:

1. Calcular as funções de autocorrelação $\hat{\rho}_1(j)$ e $\hat{\rho}_2(j)$, com $j = 1, 2, \dots, J$, das séries $Z_1(t)$ e $Z_2(t)$, respectivamente;
2. Calcular a função de autocorrelação comum às duas séries, utilizando a fórmula, $\hat{\rho}(j) = \frac{n_1 \hat{\rho}_1(j) + n_2 \hat{\rho}_2(j)}{n_1 + n_2}$, em que n_1 e n_2 são o número de observações das séries $Z_1(t)$ e $Z_2(t)$, respectivamente;
3. Calcular a função de autocorrelação parcial estimada comum ($\hat{\Phi}(k)$), utilizando a função de autocorrelação comum ($\hat{\rho}(j)$);
4. Utiliza-se ($\hat{\Phi}(k)$) para identificar a ordem autorregressiva p ;
5. Resolvendo as equações de Yule-Walker, estimar os p coeficientes do modelo auto-regressivo;
6. Ajustar para as séries $Z_1(t)$ e $Z_2(t)$ o modelo auto-regressivo com os coeficientes do passo 5, obtendo assim as séries residuais \hat{a}_1 e \hat{a}_2 ;

7. Calcular as funções de autocorrelação parcial v_j e v'_j para as séries residuais \hat{a}_1 e \hat{a}_2 , respectivamente;
8. Obter a estatística do teste, dada por:

$$SQ = \sum_{j=1}^J \frac{(v_j - v'_j)^2}{\frac{1}{n_1-j} + \frac{1}{n_2-j}} \quad (2.19)$$

9. Se $SQ > C_\alpha$, em que C_α é tal que $P(\chi_j^2 > C_\alpha) = \alpha$, rejeita-se H_0 ao nível de significância α .

2.5 Modelo de Regressão Linear Múltipla

2.5.1 Introdução

Geralmente, tem-se o interesse em avaliar a relação de uma variável de interesse Y (variável dependente ou variável resposta) em relação a k variáveis x_j (variável independente), $j = 1, 2, \dots, k$. Neste trabalho o interesse é utilizar componentes climáticas para buscar uma associação entre o nível do mar e estas componentes. Dessa forma, um possível modelo para avaliar essa relação pode ser dado pelo modelo de regressão linear múltipla

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \epsilon_i, \quad i = 1, \dots, n. \quad (2.20)$$

em que n é o número de observações, y_i é a observação da variável dependente para o i -ésimo indivíduo, $x_{i1}, x_{i2}, \dots, x_{ik}$ são as observações das variáveis independentes para o i -ésimo indivíduo, $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ são os coeficientes de regressão (parâmetros) e ϵ_i é um componente de erro aleatório.

O modelo 2.20 depende de algumas suposições que descrevem as formas do modelo e também ditam os procedimentos adequados de estimação e inferência.

1. A variável dependente y_i é função linear das variáveis independentes $x_{i1}, x_{i2}, \dots, x_{ik}$;
2. Os valores das variáveis independentes são fixos;
3. $E(\epsilon_i) = 0$;
4. Os erros são homocedásticos, isto é, $E(\epsilon_i^2) = \sigma^2$;
5. Os erros são independentes e têm distribuição normal.

2.5.2 Estimação dos parâmetros do modelo

A estimativa dos parâmetros será realizada pelo método dos mínimos quadrados, na forma explicitada de acordo com Charnet et al. (2008).

O modelo pode ser representado matricialmente por $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, onde:

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}; \mathbf{X} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{k1} \\ 1 & X_{12} & \cdots & X_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & \cdots & X_{kn} \end{bmatrix}; \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}; \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

\mathbf{Y} é um vetor $n \times 1$ de observações aleatórias;

\mathbf{X} é uma matriz $n \times (k + 1)$ de quantidades fixas conhecidas;

$\boldsymbol{\beta}$ é um vetor $(k + 1) \times 1$ de parâmetros desconhecidos;

$\boldsymbol{\epsilon}$ é um vetor aleatório $n \times 1$;

Temos da notação matricial do modelo que $\boldsymbol{\epsilon} = \mathbf{Y} - \mathbf{X}\boldsymbol{\beta}$.

Minimizando a soma de quadrados dos erros por $\boldsymbol{\epsilon}'\boldsymbol{\epsilon}$:

$$\boldsymbol{\epsilon}'\boldsymbol{\epsilon} = [\epsilon_1 \ \epsilon_2 \ \cdots \ \epsilon_n] \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} = \sum_{i=1}^n \epsilon_i^2 \quad (2.21)$$

Seja:

$$Z = \boldsymbol{\epsilon}'\boldsymbol{\epsilon}$$

$$Z = (Y - X\beta)'(Y - X\beta)$$

$$Z = (Y' - X'\beta')(Y - X\beta)$$

$$Z = Y'Y - Y'X\beta - \beta'X'Y + \beta'X'X\beta$$

Sendo as matrizes $Y'X\beta$ e $\beta'X'Y$ de dimensões 1×1 e uma sendo a transposta da outra, temos:

$$Y'X\beta = \beta'X'Y$$

logo:

$$Z = Y'Y - Y'X\beta - 2\beta'X'X\beta$$

O ponto de mínimo da função Z é obtido diferenciando Z em relação à β e igualando a zero.

Diferenciando Z vem:

$$\partial Z = 2(\partial\beta')(X'X\beta - X'Y)$$

Fazendo $\partial Z \equiv \mathbf{0}$, temos:

$$2(\partial\hat{\beta}')(X'X\hat{\beta} - X'Y) \equiv \mathbf{0}, \text{ como } \partial\hat{\beta}' \neq \mathbf{0} \text{ temos:}$$

$$X'X\hat{\beta} - X'Y = \mathbf{0} \Rightarrow X'X\hat{\beta} = X'Y \quad (2.22)$$

onde $\hat{\beta}$ é o vetor de estimativas dos parâmetros.

O sistema 2.22 é chamado de sistema de equações normais, e sua solução nos fornece as estimativas dos parâmetros constituintes do vetor $\hat{\beta}$. Dessa forma,

pré-multiplicando ambos os membros da equação 2.22 por $(X'X)^{-1}$, teremos:

$$\hat{\beta} = (X'X)^{-1}X'Y \quad (2.23)$$

2.5.3 Métodos de seleção

O critério para a adição ou remoção de variáveis é geralmente baseado na estatística F , comparando modelos com e sem as variáveis em questão. No entanto, nem sempre métodos diferentes chegam ao mesmo resultado, porque podem-se ter modelos de ajustes equivalentes. De acordo com Charnet et al. (2008), dentre esses procedimentos apresentam-se os seguintes:

Forward - Esse procedimento parte da suposição de que não há variável no modelo, apenas o intercepto. A idéia do método é adicionar uma variável de cada vez. A primeira variável selecionada é aquela com maior correlação com a variável resposta.

Backward - quando todos os preditores são incluídos de uma só vez na equação, e depois são retirados, um a um, até que se identifiquem os melhores preditores.

Stepwise - assemelha-se ao Forward, e em cada passo todas as variáveis do modelo são previamente verificadas pelas suas estatísticas F parciais. Inicia-se com uma variável: aquela que tiver maior correlação com a variável resposta. A cada variável adicionada no modelo, testa-se novamente a presença de todas as variáveis já adicionadas. Variáveis pouco explicativas são retiradas do modelo. Continua-se o processo até não incluir ou excluir nenhuma variável.

2.5.4 O coeficiente de determinação múltipla

A qualidade do modelo também pode ser avaliada por meio do valor do coeficiente de determinação, R^2 . Tal coeficiente é a quantidade da variação da variável dependente que é explicada conjuntamente pelas variáveis independentes. É fornecido matricialmente pela expressão:

$$R^2 = \frac{\widehat{\beta}'X'Y - C}{Y'Y - C} \quad (2.24)$$

onde $C = \frac{(\sum Y_i)^2}{n}$.

O valor de R^2 pertence ao intervalo $[0; 1]$ e, quanto mais próximo de 1, melhor o ajuste do modelo considerado. É importante notarmos que R^2 está diretamente ligado ao modelo que estamos considerando e não somente aos dados.

Para corrigir esta deficiência, é preferível utilizar o coeficiente de determinação ajustado (R_a^2), que é uma medida ajustada do coeficiente de determinação, sendo penalizada quando incluímos variáveis pouco explicativas. É dado pela seguinte expressão:

$$R_a^2 = 1 - \frac{n-1}{n-(k+1)} (1 - R^2) \quad (2.25)$$

onde $(k+1)$ representa o número de variáveis explicativas mais a constante.

2.5.5 Teste de hipótese em regressão linear múltipla

2.5.5.1 Teste de significância da regressão

Esse teste estuda a probabilidade dos parâmetros da regressão $\widehat{\beta}_0, \widehat{\beta}_1, \dots, \widehat{\beta}_n$ serem iguais a zero, sendo que neste caso não existiria regressão. O teste é efetivado através da distribuição de Fisher-Snedecor. O coeficiente F calculado é

obtido pela expressão matricial:

$$F_c = \frac{\widehat{\beta}'X'Y - C}{(Y'Y - C)(\widehat{\beta}'X'Y - C)} \times \frac{n - k - 1}{k} \quad (2.26)$$

A hipótese nula é aceita, com nível de significância $\alpha = 5\%$, se: $F_c > F$.

2.5.5.2 Teste sobre os coeficiente individuais de regressão

Neste caso, utiliza-se a distribuição t de Student, tendo como hipótese nula a suposição de que os regressores são diferentes de zero ao nível de 5% de significância. Dessa forma, relaciona-se cada regressor com seu desvio padrão:

$$t_i = \frac{\widehat{\beta}_i}{s(\widehat{\beta}_i)} \quad (2.27)$$

Com base nos valores obtidos da distribuição t de Student com $n - k - 1$ graus de liberdade, aceita-se a hipótese nula se: $t_i > t$.

2.5.6 Estudo da adequação do modelo

2.5.6.1 Verificação dos pressupostos e análise dos resíduos

Para que o uso da equação de regressão seja eficaz na predição da variável dependente em estudo, o pesquisador deve examinar previamente os pressupostos da regressão, bem como identificar as conseqüências da sua violação. Entre os pressupostos citados por Graybill e Iyer (1994), estão: a homogeneidade nas variâncias, a normalidade, independência e média zero para os resíduos.

De acordo com Graybill e Iyer (1994), ao se analisar graficamente as distribuições (resíduos versus valor estimado) verifica-se a existência de homocedasticidade, ou seja, a hipótese de variância constante, que é aceita quando não há nenhuma tendência dos resíduos em relação ao valor estimado, neste caso denominado de heterocedasticidade. A heterogeneidade das variâncias, ou violação da

homogeneidade das variâncias, pode ser reduzida por intermédio da transformação de variáveis que não possuem distribuição normal.

2.5.6.2 Teste de Kolmogorv-Smirnov (normalidade dos resíduos)

A condição de normalidade dos resíduos é necessária para a definição de intervalos de confiança e testes de significância. Antes de se aplicar o teste, pode-se comparar graficamente a distribuição dos resíduos obtidos com a distribuição normal.

O teste de Kolmogorv-Smirnov avalia se duas amostras foram extraídas de uma mesma população. Para se aplicar o teste deve-se ordenar as amostras, construir a distribuições de freqüências acumuladas nos intervalos de classe definidos, calcular as diferenças entre estas freqüências (da primeira menos a da segunda amostra) escolhendo-se a maior diferença em valor absoluto, que será comparada com um valor tabelado. A hipótese nula do teste (existência de normalidade) é rejeitada se o valor absoluto for maior que o valor tabelado.

Para maiores detalhes sobre o teste de Kolmogorv-Smirnov, ver Bunchaft (1997).

2.5.6.3 Teste de Durbin-Watson

A estatística Durbin-Watson é um teste estatístico usado para detectar a presença de autocorrelação nos resíduos de uma análise de regressão. Este teste é baseado na suposição de que os resíduos no modelo de regressão são gerados por um processo autoregressivo de primeira ordem. A hipótese nula é de que não existe autocorrelação.

Sua estatística é definida como:

$$d_w = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2} \quad (2.28)$$

em que T é o número de observações, e_t são os resíduos do modelo e $0 \leq d_w \leq 4$. Sejam d_l e d_u os valores tabelados, inferior e superior, respectivamente. Temos que se $d_w > d_u$ rejeitamos a hipótese nula, se $d_w < d_l$ aceitamos a hipótese nula e caso $d_l < d_w < d_u$ o teste é inconclusivo.

Vale ressaltar que o correlograma é uma alternativa à esse teste.

3 MATERIAL E MÉTODOS

Nesta seção, são descritas as séries utilizadas, bem como os procedimentos para as análises.

Todos os dados de níveis do mar foram obtidos do The Global Sea Level Observing System (GLOSS) database - University of Hawaii Sea Level Center. Eles consistem em elevações médias mensais do nível do mar (mm) em:

1. Cananéia, (Brasil) pertencente ao oceano Atlântico, no período de jan./2005 a dez./2006, consituída de 24 dados mensais, cuja estação costeira tem a localização ($25^{\circ}01'S47^{\circ}56'W$);
2. Aburatsu, (Japão) pertencente ao oceano Pacífico, no período de jan./1985 a nov./2011, consituída de 323 dados mensais, cuja estação costeira tem a localização ($31^{\circ}34'N131^{\circ}25'E$);
3. Key West, (Flórida - Estados Unidos) pertencente ao oceano Atlântico, no período de jan./1985 a nov./2011, consituída de 323 dados mensais, cuja estação costeira tem a localização ($24^{\circ}33'N81^{\circ}49'W$);
4. Ko Taphao Noi, (Tailândia) pertencente ao oceano Índico, no período de jan./1985 a nov./2011, consituída de 323 dados mensais, cuja estação costeira tem a localização ($07^{\circ}50'N98^{\circ}26'E$);

Foram consideradas, para ajuste de modelos (nas séries 2, 3 e 4), as observações de janeiro de 1985 a dezembro de 2010; as observações de janeiro de 2011 a novembro de 2011 foram utilizadas para avaliação de previsão.

A escolha destas séries justifica-se por considerar que as séries pertencem a diferentes oceanos. Vale ressaltar que os dados disponíveis não apresentam a

qualidade ideal, existem muitas séries incompletas, faltando em torno de 40% dos dados, e com curto prazo. Desta forma, tais séries foram escolhidas por conter a menor quantidade de observações faltantes, e também por ter pelo menos 25 anos de observações. No entanto, estas séries apresentaram alguns dados faltantes, e foi feita a média móvel para solucionar esse problema. Na Figura 2, pode-se ter uma visualização da localização de cada uma das estações onde as séries foram coletadas.

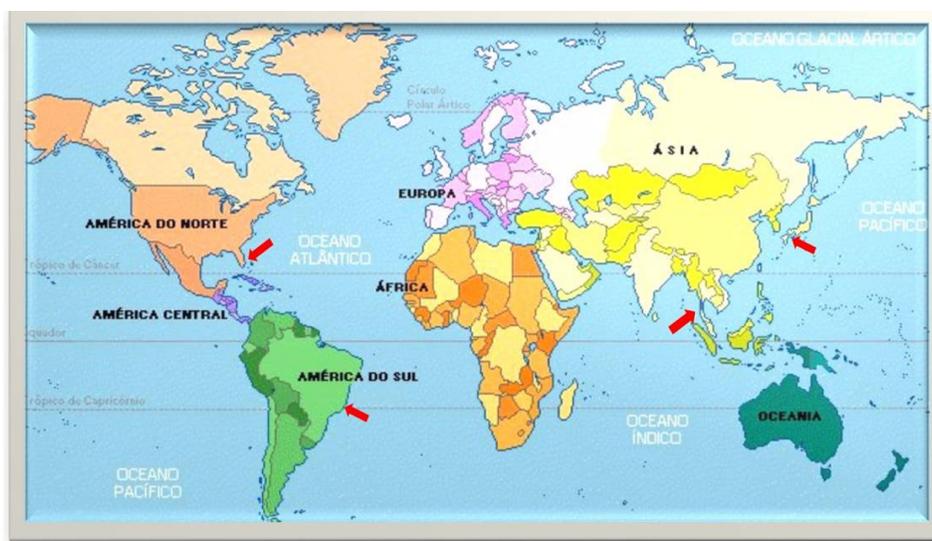


Figura 2 Mapa mundi com destaque nos locais pesquisados (Key West, Cananéia, Ko Taphao Noi e Aburatsu)

Fonte: <http://www.sogeografia.com.br> (2011)

Será considerada inicialmente uma análise descritiva das séries. Em seguida será feita a construção do gráfico da série original para observar indícios de tendência e/ou sazonalidade. Depois, será verificada a necessidade de alguma transformação, por meio do gráfico das médias vs amplitudes.

O teste de Fisher, com a construção do periodograma, será considerado

para verificar a presença de sazonalidade. Caso se confirme a presença dessa componente, será aplicado o teste HEGY para verificar a natureza dessa componente, permitindo o uso do método correto de eliminação da mesma.

O teste de sequência de Wald-Wolfowitz será aplicado para verificar a presença de tendência na série, caso sua presença seja detectada toma-se a diferença na série para eliminar esta componente.

Constroem-se gráficos das funções de autocorrelação (fac) e autocorrelação parcial (facp) da série livre das componentes tendência e sazonalidade para estimar as ordens do modelo.

Em seguida, serão analisadas as intervenções que por ventura ocorreram no período. Depois tem-se o ajuste dos modelos adequados, e a aplicação do teste de Box e Pierce para verificar se os resíduos são ruído branco.

Serão feitas as previsões e a escolha do melhor modelo será feita através do critério de Informação Akaike (AIC), do erro quadrático médio de previsão (EQMP) e do erro percentual médio absoluto (MAPE).

Serão aplicados os métodos para comparação de séries, apenas nas séries de Aburatsu, Key West e Ko Taphao Noi, para verificar se as séries são geradas pelo mesmo processo estocástico.

Em paralelo, será feita análise de regressão tendo o nível médio do mar de Cananéia como variável dependente e, condição climática, direção do vento (graus), velocidade do vento (m/s), referência do barógrafo (mmHg), temperatura psicômetro (°C), umidade relativa (%), referência do higrógrafo (%), referência do termógrafo (°C), temperatura máxima (°C), temperatura mínima (°C), precipitação (mm) e temperatura d'água (°C) como variáveis independentes. Sendo que serão consideradas observações mensais no período de jan./2005 a dez./2006, os dados das variáveis independentes foram obtidos do Instituto Oceanográfico da

Universidade de São Paulo.

Especificamente, será identificado um conjunto de variáveis independentes significativas para comporem o modelo de regressão, para então realizar o diagnóstico do modelo para tomada de decisão quanto a sua aceitação.

As análises serão feitas no software R Development Core Team (2010) e no software JMulTi (L'UTKEPOHL; KR'ATZIG, 2004).

4 RESULTADOS E DISCUSSÃO

Os resultados obtidos, neste trabalho, são apresentados em três etapas. Na primeira, foi feita uma análise inicial com as séries de Aburatsu, Ko Taphao Noi e Key West, de modo a obter séries estacionárias, posteriormente foram ajustados modelos para as três séries, e obtidas as previsões. Na segunda etapa, foram realizados os testes de comparação. Na última etapa, utilizando a série do nível do mar e as variáveis climáticas de Cananéia, ajustou-se um modelo de regressão.

4.1 Análise inicial e ajustes

4.1.1 Aburatsu

Foi feita a análise inicial da série de Aburatsu cuja representação gráfica está apresentada na Figura 3, onde se pode observar a possível da componente sazonal.

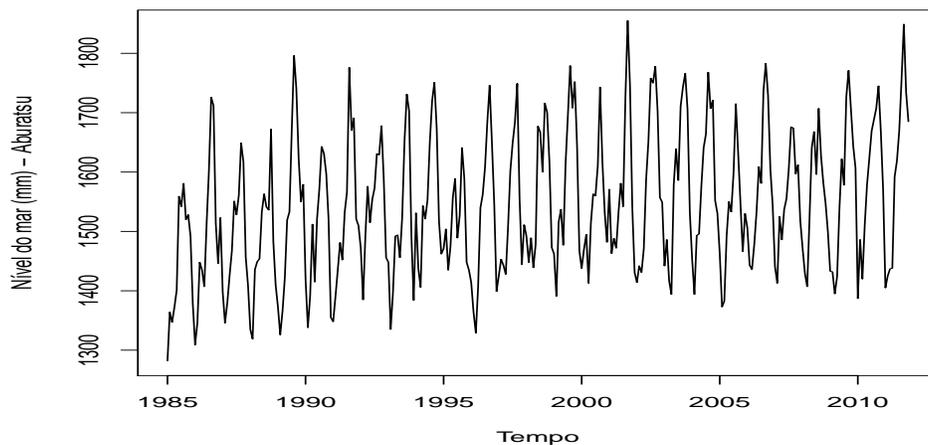


Figura 3 Série mensal de níveis médios do mar de Aburatsu, no período de janeiro de 1985 a dezembro de 2010

Na Figura 4, temos a decomposição da série através do método STL (*Seasonal Decomposition of Time Series by Loess*), descrito em Cleveland et al. (1990):

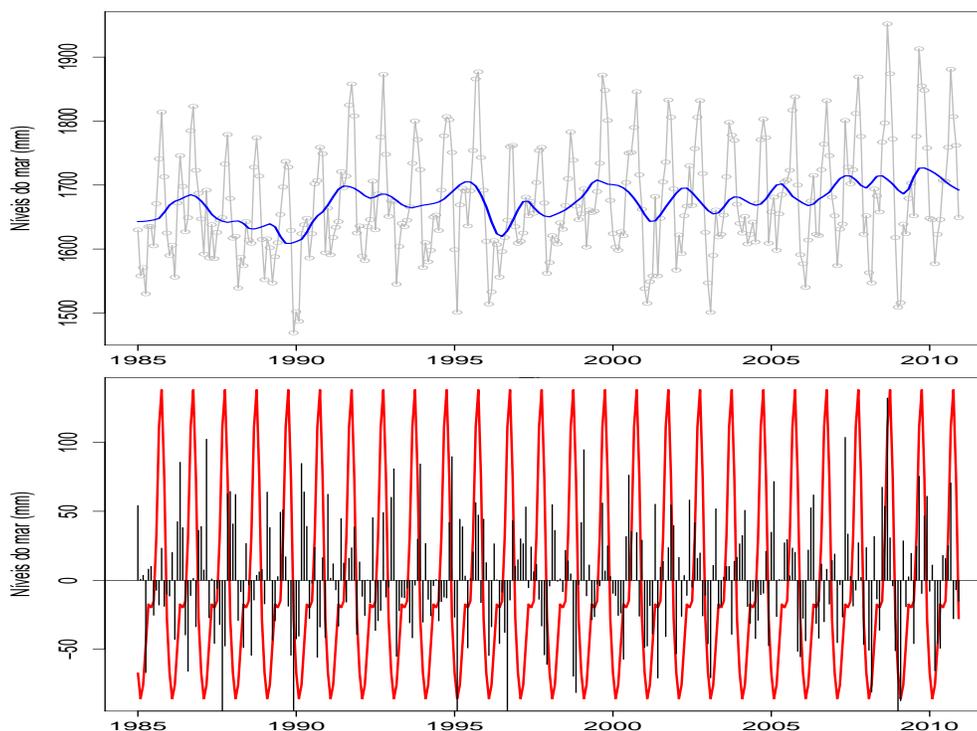


Figura 4 Decomposição sazonal usando STL aplicado à série mensal de níveis médios do mar de Aburatsu, no período de janeiro de 1985 a dezembro de 2010. Os círculos cinzas mostram a série original, a linha azul é a componente tendência, a vermelha mostra a componente sazonal, e as barras pretas são o componente aleatório

A Figura 4 mostra que a tendência está aumentando lentamente, o componente sazonal apresenta picos em torno de fevereiro e setembro de cada ano. A amplitude da componente sazonal não mostra crescimento ao longo da série. O componente aleatório apresenta picos que podem ter sido afetados por diversos fatores influentes no comportamento do nível do mar, como por exemplo, o valor negativo em 1990. Com uma análise dos componentes, aparentemente temos var-

ância constante.

Dessa forma, verificou-se se alguma transformação nos dados, para estabilizar a variância, é necessária. A série foi dividida em grupos de 12 observações, sendo este o período da série, e para cada grupo foram calculadas a média e a amplitude. O resultado pode ser observado na Figura 5, e como os pontos estão espalhados ao redor de uma reta paralela ao eixo das abscissas, nota-se que não há a necessidade de transformação.

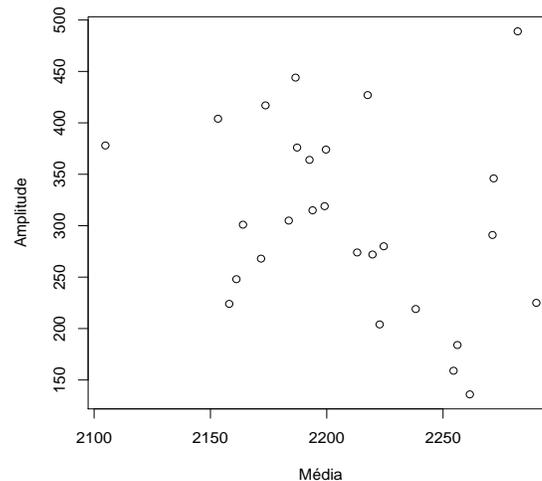


Figura 5 Série mensal de níveis médios do mar de Aburatsu: amplitude vs média

Morettin e Tolo (2006) afirmam que é preciso verificar se a componente sazonal existe, caso ela exista deve ser retirada da série antes de se testar a presença da componente tendência. No periodograma (Figura 6) da série, pode-se verificar a presença de um pico em torno da frequência 0.08, que corresponde ao período 12. Para confirmar a presença do efeito sazonal nesse período aplicou-se teste de Fisher. Como $g = 0.3006 > z_{0,05} = 0.04923$, rejeita-se H_0 ao nível de significância $\alpha = 0,05$, ou seja, a série realmente apresenta sazonalidade no

período analisado.

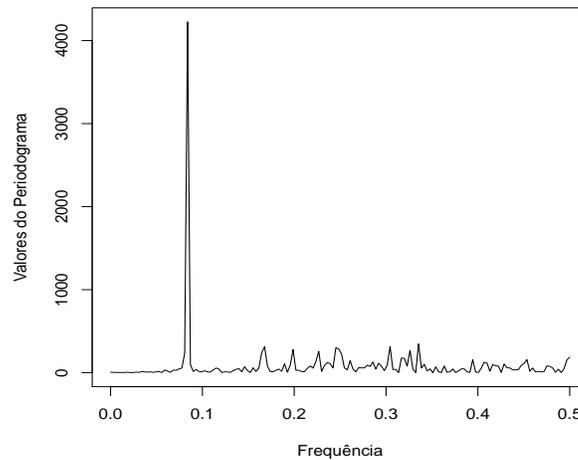


Figura 6 Periodograma da série de níveis médios do mar de Aburatsu

Para avaliar cada subsérie mensal, foi contruída a Figura 7, plotando cada mês em anos sucessivos, e para cada mês a linha horizontal mostra a média dos valores. Através desta figura, podemos observar o padrão sazonal da série, e que existe uma variação não muito grande dentro de cada mês. O padrão global mostra que fevereiro e setembro têm menores e maiores elevações do nível do mar, respectivamente.

Com a constatação de que a série apresenta sazonalidade, o objetivo agora é identificar a natureza das variações sazonais. Para testar se há presença de raiz unitária não-sazonal (ou tendência estocástica) e sazonal, utiliza-se o teste HEGY, para dados mensais. O modelo com uma tendência linear e variáveis *dummies* é estimado. A saída do teste HEGY, realizada no software JMulTi, contém também os valores estimados e testados dos parâmetros do modelo, bem como os valores críticos dos testes correspondentes. Tal software foi utilizado para as análises pois

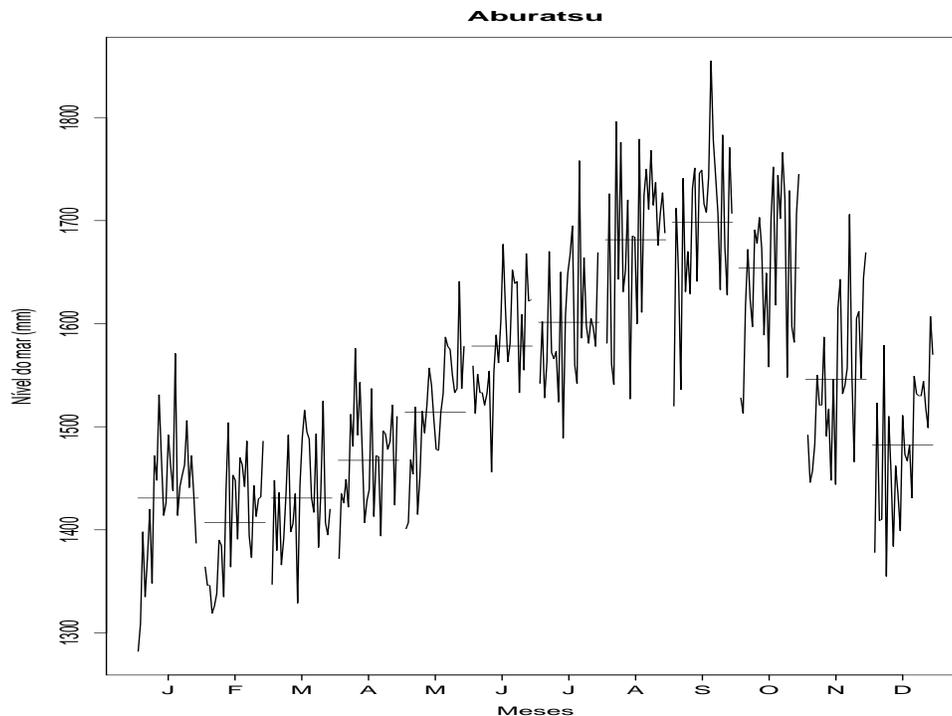


Figura 7 Subséries mensais dos níveis médios do mar de Aburatsu

o pacote do R que fazia tais análises, o "uroot", não está mais disponível pois continha alguns erros.

De acordo com os resultados, temos que as variações sazonais contêm uma parte de natureza determinística. Na Tabela 1 a hipótese nula, testada com o teste t unilateral a esquerda, sobre as raízes π_1 e π_2 é aceita e o modelo tem uma raiz unitária não-sazonal (tendência estocástica) e uma raiz semestral.

Nestas circunstâncias, as variações sazonais são eliminadas através do método de regressão com variáveis *dummies* sazonais. Na Tabela 2 temos o erro padrão e o teste de significância para os coeficientes do modelo ajustado, onde se pode confirmar que os efeitos de tendência e sazonalidade mensal são significantes. E apesar de não termos todos os coeficientes sazonais significantes, a

Tabela 1 Resultado do teste de raiz unitária sazonal (teste HEGY) para a série de nível médio do mar de Aburatsu.

	Valores críticos (5%)	Valores da estatística do teste
$t(\pi_1)$	-3,35	5,37
$t(\pi_2)$	-2,81	5,38
$F(\pi_3 \pi_4)$	6,35	5,70
$F(\pi_5 \pi_6)$	6,48	32,85
$F(\pi_7 \pi_8)$	6,30	28,91
$F(\pi_9 \pi_{10})$	6,40	42,83
$F(\pi_{11} \pi_{12})$	6,46	40,42
$F(\pi_1 - \pi_{12})$	4,44	25,95
$F(\pi_2 - \pi_{12})$	4,58	25,85

regressão é significativa.

Analisando a tabela, nota-se que o coeficiente angular tem o valor de 0,3565, mostrando que o nível médio do mar de Aburatsu apresenta tal crescimento a cada mês (em mm), o que nos leva a um crescimento de 4,478 mm/ano, ou seja um crescimento muito grande, caso esse crescimento permaneça em torno de 100 anos teremos um aumento de meio metro do nível do mar. Os meses de maior aumento do nível do mar são agosto e setembro, e o de menor aumento é fevereiro, estando de acordo com a interpretação da Figura 7.

Após eliminar as variações sazonais com o auxílio do modelo de regressão com *dummies* sazonais, foi então aplicado o teste de sequência (Wald-Wolfowitz) para verificar a existência de tendência. Foi obtido um valor da estatística do teste de $-3,6222$, e o valor $p = 0,0002$, onde se pode concluir que a série apresenta tendência. Dessa forma, com o resultado do teste de sequência e a conclusão de que tem-se raiz unitária não-sazonal, toma-se a primeira diferença da série para torná-la estacionária.

Assim, temos na Figura 8, a série estacionária, e na Figura 9, temos as funções de autocorrelação e autocorrelação parcial da série diferenciada.

Tabela 2 Significância dos parâmetros do modelo de regressão com *dummies* sazonais ajustado para eliminar sazonalidade determinística.

Parâmetros	Estimativas	Erro padrão
β_0	1485,167	187,968
β_1	0,3565	0,1178
α_1	-107,7445	9,7308
α_2	-132,0977	9,7418
α_3	-108,4993	9,7415
α_4	-72,4353	9,7374
α_5	-26,2228	9,7343
α_6	37,5711	9,7336
α_7	60,3347	9,7337
α_8	139,9325	9,7344
α_9	156,0236	9,7388
α_{10}	111,6893	9,7428
α_{11}	3,4889	9,7415
α_{12}	62,0575	9,7415

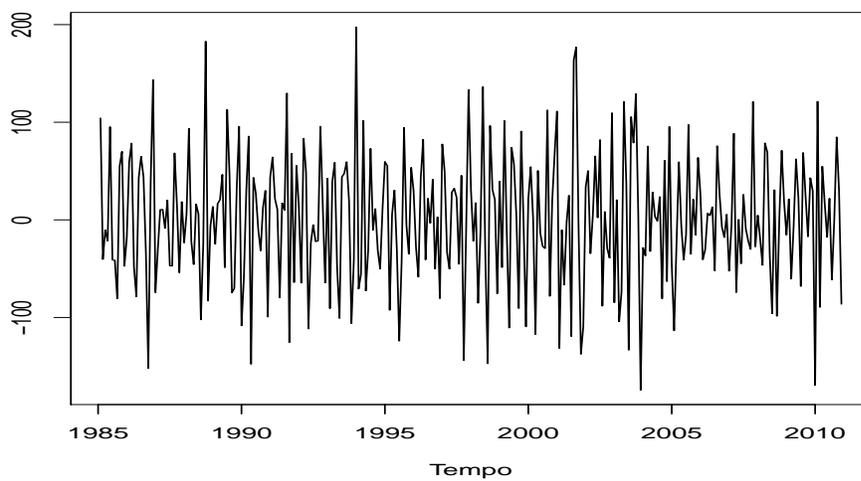


Figura 8 Série diferenciada de níveis médios do mar de Aburatsu

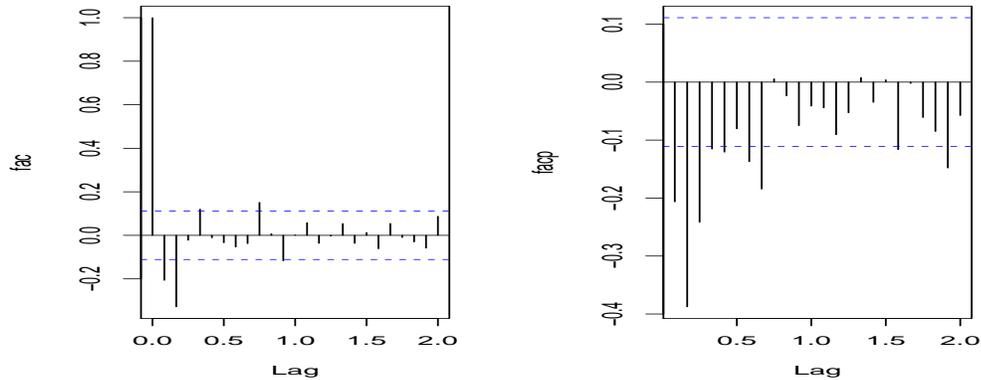


Figura 9 Funções de autocorrelação (a) e autocorrelação parcial (b) da série diferenciada de níveis médios do mar de Aburatsu

Considerando que as ordens dos operadores auto-regressivos são identificadas pela $facp$ e as ordens dos operadores de médias móveis pela fac , o modelo que apresentou melhor ajuste, foi:

ARIMA (3, 1, 1) descrito por:

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3)(1 - B)Z_t = (1 - \theta_1 B)a_t$$

Na Tabela 3, são apresentadas as estimativas dos parâmetros do modelo proposto.

Tabela 3 Estimativas dos parâmetros do modelo ARIMA proposto para a série mensal de níveis médios do mar de Aburatsu, no período de janeiro de 1985 a dezembro de 2010.

Modelo	Parâmetro	Estimativa	Erro Padrão
ARIMA (3, 1, 1)	ϕ_1	0,3990	0,0603
	ϕ_2	-0,2047	0,0604
	ϕ_3	0,1652	0,0602
	θ_1	0,9787	0,0220

As análises das funções de autocorrelação e autocorrelação parcial do resíduo do modelo ARIMA(3, 1, 1), conforme Figura 10, não mostram correlação significativa. Com isso, pode-se afirmar que esse modelo ajusta-se aos dados e que o resíduo é ruído branco.

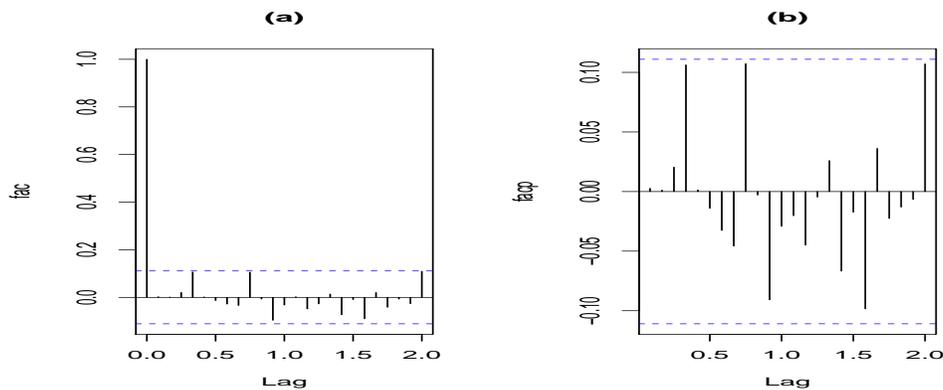


Figura 10 Função de autocorrelação (a) e autocorrelação parcial (b) do resíduo do modelo ajustado para a série de níveis do mar de Aburatsu

A aplicação do teste de Box-Pierce, forneceu o seguinte resultado: $Q(24) = 17.09$, $< \chi_{20,0.05}^2 = 31,41$, o que confirma a qualidade do ajuste.

O modelo substituído por suas estimativas é dado por:

$$(1 + 0,3990B - 0,2047B^2 + 0,1652B^3)(1 - B)Z_t = (1 + 0,9787B)a_t$$

As previsões realizadas, no período de janeiro de 2011 a novembro de 2011, e seus respectivos erros padrão para o modelo ARIMA (3, 1, 1), estão apresentados na Tabela 4. Nesta tabela, também encontram-se os valores observados deste período e os erros de previsão.

Após testar alguns modelos, a escolha deste modelo foi feita por apresentar

Tabela 4 Valores observados (Z_{t+h}), valores preditos com o modelo ARIMA (3, 1, 1) ($\hat{Z}_t(h)$) e erro de previsão ($e_t(h)$) para a série mensal de níveis do mar de Aburatsu, no período de janeiro a novembro de 2011.

Mês	Z_{t+h}	$\hat{Z}_t(h)$	$e_t(h)$
Jan/2011	1405	1460,38	-55,37
Fev/2011	1425	1454,54	-29,54
Mar/2011	1437	1475,61	-38,61
Abr/2011	1438	1506,43	-68,43
Mai/2011	1592	1548,40	43,59
Jun/2011	1619	1618,89	0,10
Jul/2011	1669	1647,42	21,57
Ago/2011	1755	1710,98	44,01
Set/2011	1849	1743,35	105,64
Out/2011	1734	1705,90	28,09
Nov/2011	1685	1606,35	78,64

menor AIC (3388,69) e menor EQMP (2959,52) e também por apresentar um valor de 2,91 para o MAPE, indicando um bom desempenho de ajuste.

Na Figura 11, temos a representação gráfica dos valores de nível médio do mar observados e estimados para o período de janeiro de 1985 a dezembro de 2010, observa-se que as componentes identificadas foram reproduzidas pelos valores estimados, no entanto, os picos da série não foram bem captados pelo modelo.

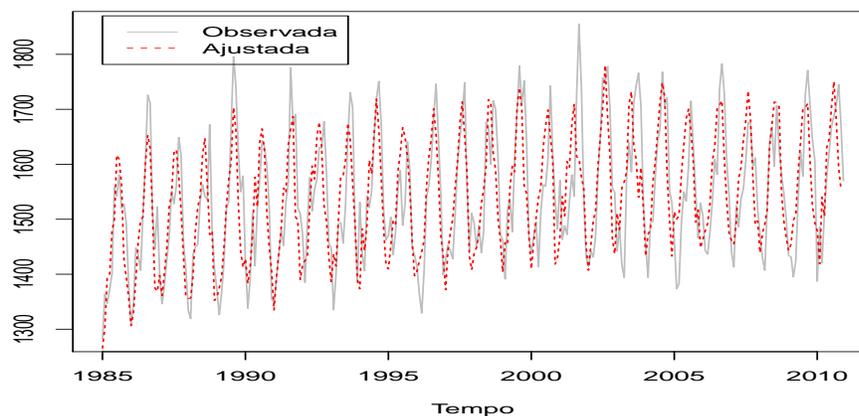


Figura 11 Séries de valores observados e estimados de acordo com o modelo ajustado ARIMA (3,1,1), de nível médio do mar de Aburatsu, no período de janeiro de 1985 a dezembro de 2010

4.1.2 Key West

Apresenta-se na Figura 12, o gráfico da série mensal de níveis do mar da Key West. Observam-se indícios da existência dos componentes tendência e sazonalidade na série.

O procedimento de análise da série para torná-la estacionária é semelhante ao de Aburatsu e pode ser visto no apêndice.

Na Figura 13, temos a série estacionária, e na Figura 14, temos as funções de autocorrelação e autocorrelação parcial da série estacionária.

A representação gráfica da f_{ac} e da f_{pac} da série estacionária, na Figura 14, mostram que os lags 3 e 5 foram significativos, porém o modelo com essas ordens para os operadores auto-regressivos e de médias móveis não apresentou um bom ajuste, bem como outros modelos testados, temos que dentre esses modelos testados, o que apresentou melhor ajuste, através dos critérios de seleção, foi:

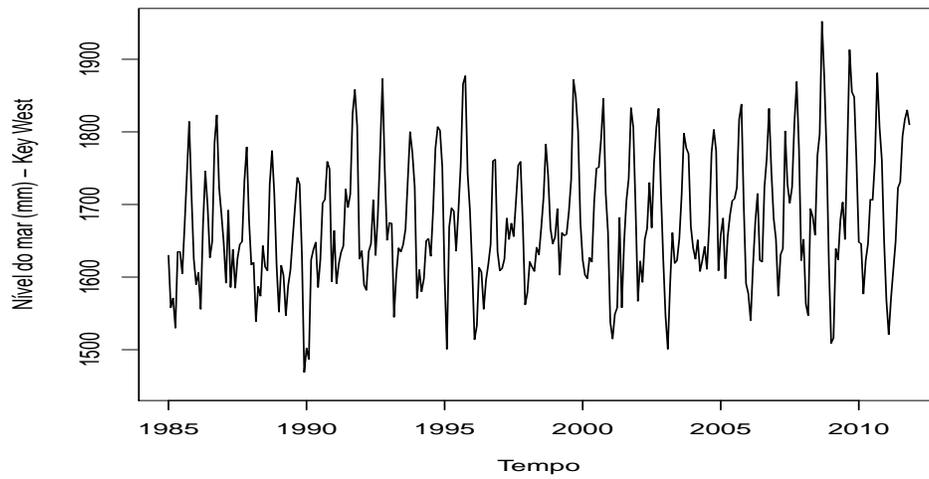


Figura 12 Série mensal de níveis médios do mar de Key West, no período de janeiro de 1985 a dezembro de 2010

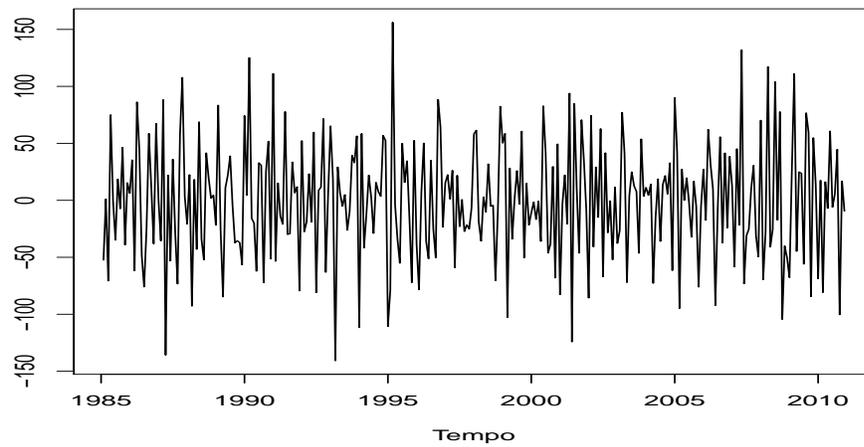


Figura 13 Série diferenciada de níveis médios do mar da Key West

ARIMA (1, 1, 1) descrito por:

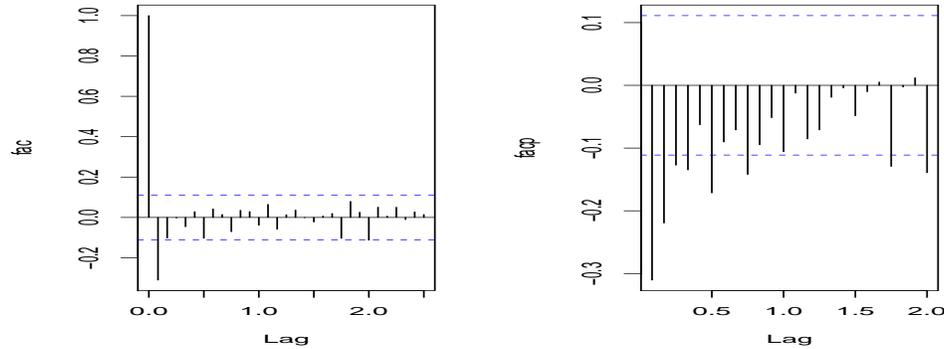


Figura 14 Funções de autocorrelação (a) e autocorrelação parcial (b) da série diferenciada de níveis médios do mar da Key West

$$(1 - \phi_1 B)(1 - B)Z_t = (1 - \theta_1 B)a_t$$

As estimativas dos parâmetros do modelo proposto são apresentadas na Tabela 5 .

Tabela 5 Estimativas dos parâmetros do modelo ARIMA proposto para a série mensal de níveis médios do mar de Key West, no período de janeiro de 1985 a dezembro de 2010

Modelo	Parâmetro	Estimativa	Erro Padrão
ARIMA	ϕ_1	0,4401	0,0512
(1, 1, 1)	θ_1	-1,0000	0.0093

As análises das funções de autocorrelação e autocorrelação parcial do resíduo do modelo ARIMA(1, 1, 1), conforme Figura 15, não mostram correlação significativa. Com isso, pode-se afirmar que esse modelo ajusta-se aos dados e que o resíduo é ruído branco.

A aplicação do teste de Box-Pierce, forneceu o seguinte resultado: $Q(24) = 7,13 < \chi_{21,0.05}^2 = 32,67$, o que confirma que a qualidade do ajuste.

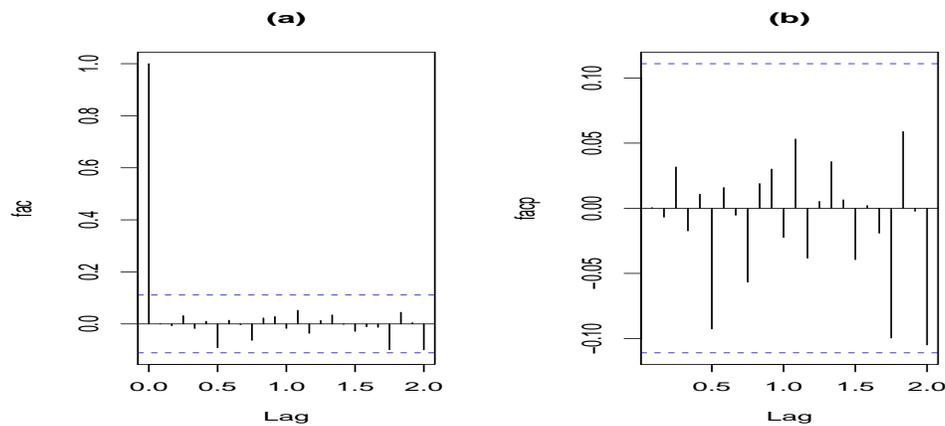


Figura 15 Função de autocorrelação e autocorrelação parcial do resíduo do modelo ajustado para a série de níveis do mar de Key West

O modelo ARIMA (1, 1, 1) com as estimativas dos parâmetros é dado por:

$$(1 - 0,4401B)(1 - B)Z_t = (1 + 1,0000B)a_t$$

As previsões realizadas no período de janeiro de 2011 a novembro de 2011 e seus respectivos erros padrão para o modelo ARIMA (1, 1, 1) estão apresentados na Tabela 6. Nesta tabela, também encontram-se os valores observados deste período e os erros de previsão.

A escolha deste modelo foi feita por apresentar menor AIC (3243,1) e menor EQMP (2519,59), e também por apresentar um valor de 2,49 para o MAPE, que mostra bom desempenho de ajuste.

Os valores do nível médio do mar da Key West observados e estimados para o período de janeiro de 1955 a dezembro de 2010 estão representados na Figura 16. Pode-se notar que a captação dos picos da série não é muito boa com esse modelo.

Tabela 6 Valores observados (Z_{t+h}), valores preditos com o modelo ARIMA (1, 1, 1) ($\hat{Z}_t(h)$) e erro de previsão ($e_t(h)$) para a série mensal de níveis do mar de Key West, no período de janeiro a novembro de 2011.

Mês	Z_{t+h}	$\hat{Z}_t(h)$	$e_t(h)$
Jan/2011	1649	1620,47	-52,47
Fev/2011	1568	1605,95	-84,95
Mar/2011	1521	1620,36	-49,36
Abr/2011	1571	1651,51	-41,51
Mai/2011	1610	1682,27	-33,27
Jun/2011	1649	1682,87	40,12
Jul/2011	1723	1688,10	42,89
Ago/2011	1731	1735,54	58,45
Set/2011	1794	1812,96	4,03
Out/2011	1830	1839,44	-9,44
Nov/2011	1810	1777,60	32,39

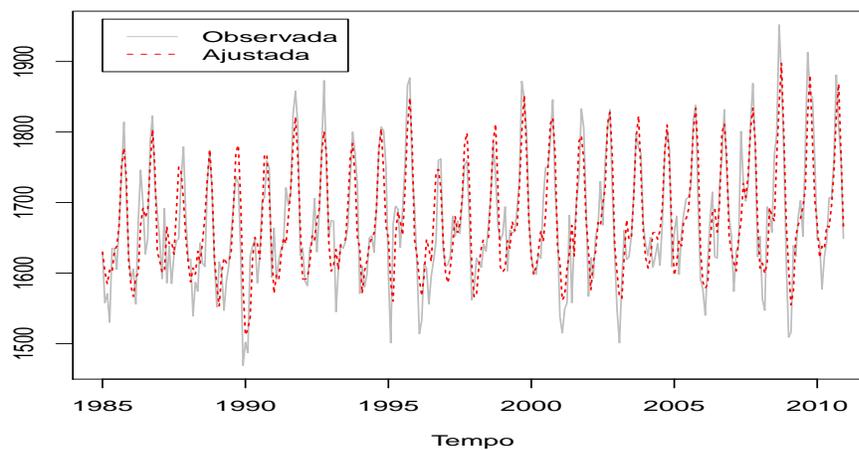


Figura 16 Séries de valores observados e estimados de acordo com o modelo ajustado ARIMA (1,1,2), de nível médio do mar de Key West, no período de janeiro de 1985 a dezembro de 2010

4.1.3 Ko Taphao Noi

A representação da série mensal de nível médio do mar de Ko Taphao Noi pode ser observada na Figura 17. Por meio de uma análise visual, verifica-se a presença de sazonalidade.

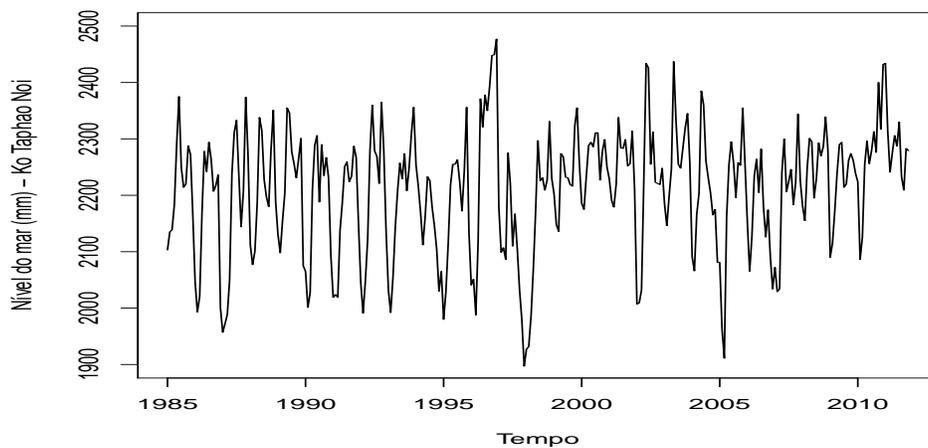


Figura 17 Série mensal de níveis médios do mar de Ko Taphao Noi, no período de janeiro de 1985 a dezembro de 2010

O procedimento de análise da série para torná-la estacionária é semelhante ao de Aburatsu e pode ser visto no apêndice.

Assim, temos na Figura 18, a série em estágio estacionário. Existem indícios de intervenção no período de outubro de 1997 e em janeiro de 2005. O efeito na mudança do nível do mar, refletido nesse período, pode ser explicado pela ocorrência do El Niño mais intenso desde a existência das "observações", que causou anomalia de temperatura da superfície do mar nesse período. Considerando o comportamento da Figura 18, espera-se que o efeito de intervenção seja abrupto temporário. Na Figura 19, temos as funções de autocorrelação e autocorrelação

parcial da série estacionária.

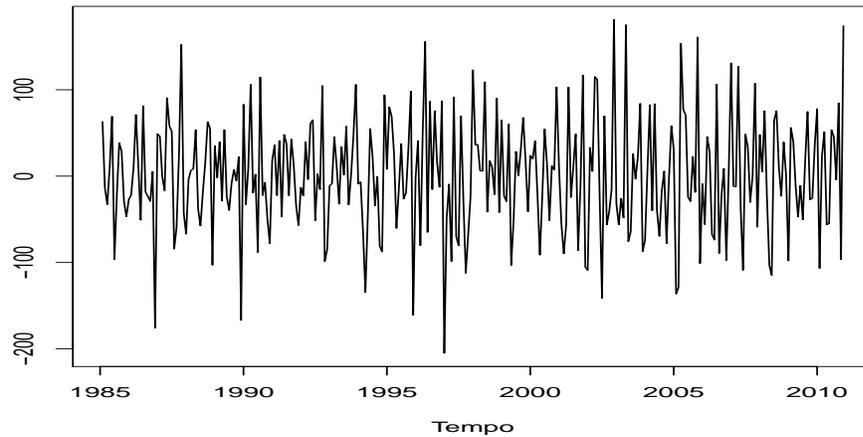


Figura 18 Série diferenciada de níveis médios do mar de Ko Taphao Noi

Na Figura 19, pode-se observar que a fac e $facp$ possuem correlações significativas nos lags, sugerindo um modelo $ARIMA(2,1,3)$, no entanto, esse modelo não apresentou um bom ajuste.

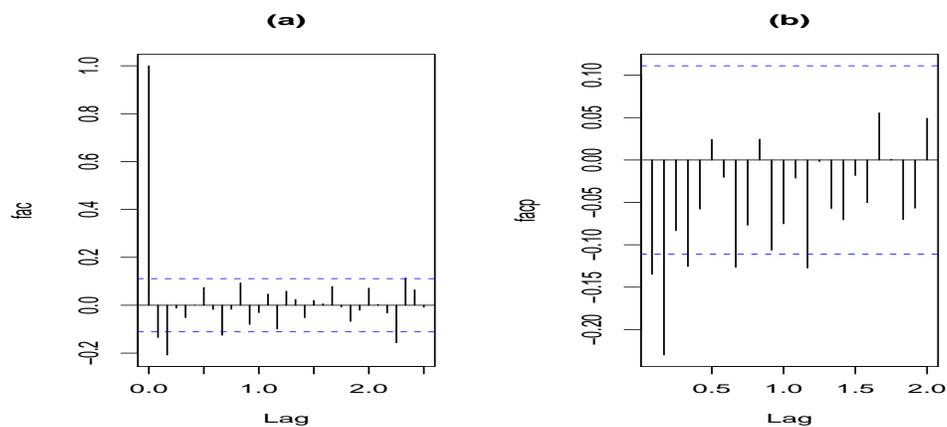


Figura 19 Funções de autocorrelação (a) e autocorrelação parcial (b) da série diferenciada de níveis médios do mar de Ko Taphao Noi

Após verificar vários modelos pelos critérios de seleção, o modelo de melhor ajuste foi:

ARIMA (1, 1, 1) com intervenção em outubro de 1997 descrito por:

$$Z_t = \frac{\omega_1}{1 - \delta_1 B} X_{1,t} + \frac{(1 - \theta_1 B)a_t}{(1 - \phi_1 B)(1 - B)}$$

A intervenção em janeiro de 2005 não foi significativa.

em que a variável de intervenção representada por $X_{1,t}$ é:

$$X_{1,t} \begin{cases} 0, & t = 154 \\ 1, & t \neq 154 \end{cases}$$

Na Tabela 7, são apresentadas as estimativas dos parâmetros do modelos proposto. Observa-se para o modelo $\omega_1 = -180,2$, indicando que no mês de outubro de 1997 houve uma diminuição no nível do mar de 180,2 mm.

Tabela 7 Estimativas dos parâmetros do modelo ARIMA proposto para a série mensal de níveis médios do mar de Ko Taphao Noi, no período de janeiro de 1985 a dezembro de 2010.

Modelo	Parâmetro	Estimativa	Erro Padrão
ARIMA (1, 1, 1) - CI	ϕ_1	0,654	0,0563
	θ_1	0,998	0,0547
	ω_1	-180,2	55,270
	δ_1	0,872	0,056

As análises das funções de autocorrelação e autocorrelação parcial do resíduo do modelo ARIMA(1, 1, 1) - CI, conforme Figura 20, não mostram correlação significativa.

A aplicação do teste de Box-Pierce forneceu o seguinte resultado: $Q(24) = 25,43 < \chi_{22,0.05} = 33,92$. Com isso, pode-se afirmar que esse modelo ajusta-se aos dados e que o resíduo é ruído branco.

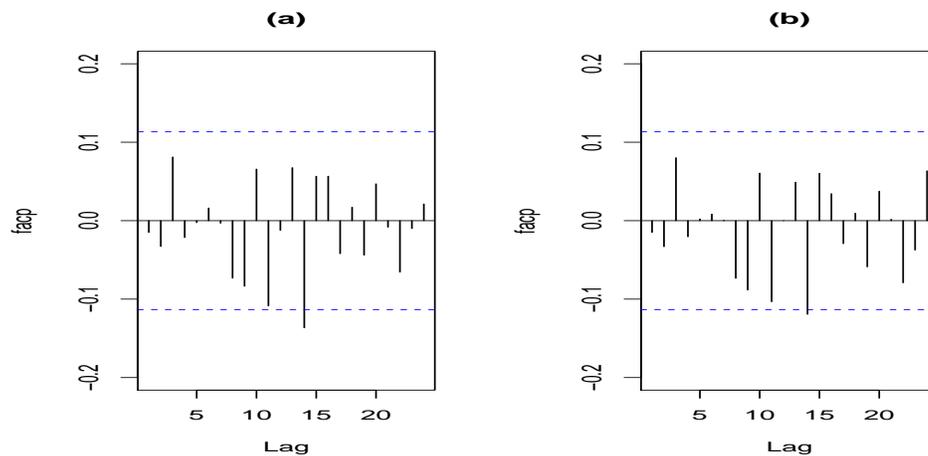


Figura 20 Função de autocorrelação (a) e autocorrelação parcial (b) do resíduo do modelo ajustado para a série de níveis do mar de Ko Taphao Noi

As previsões, realizadas no período de janeiro de 2011 a novembro de 2011 e seus respectivos erros padrão, para o modelo ARIMA (1, 1, 1)-CI, estão apresentados na Tabela 8. Nesta tabela, também encontram-se a previsão para o modelo determinístico, os valores reais deste período e os erros de previsão.

Este modelo foi escolhido pois apresentou menor AIC (3534,3) e menor EQMP (4864,17), e também por apresentar um valor de 2,41 para o MAPE, indicando que teve o melhor desempenho de ajuste.

Na Figura 21, temos a representação gráfica dos valores de nível médio do mar observados e estimados para o período de janeiro de 1985 a dezembro de 2010. Podemos notar que apesar de o modelo incorporar uma intervenção, ele não tem uma boa captação dos picos da série.

Tabela 8 Valores observados (Z_{t+h}), valores preditos com o modelo ARIMA (1, 1, 2)-CI ($\hat{Z}_t(h)$) e erro de previsão ($e_t(h)$) para a série mensal de níveis do mar de Ko Taphao Noi, no período de janeiro a novembro de 2011.

Mês	Z_{t+h}	$\hat{Z}_t(h)$	$e_t(h)$
Jan/2011	2434	2274,19	159,81
Fev/2011	2314	2201,76	112,24
Mar/2011	2241	2192,09	48,91
Abr/2011	2275	2252,22	22,78
Mai/2011	2306	2337,94	-31,94
Jun/2011	2287	2345,25	-58,25
Jul/2011	2330	2310,81	19,19
Ago/2011	2233	2294,93	-61,93
Set/2011	2209	2260,10	-51,10
Out/2011	2283	2297,93	-14,93
Nov/2011	2279	2310,73	-31,73

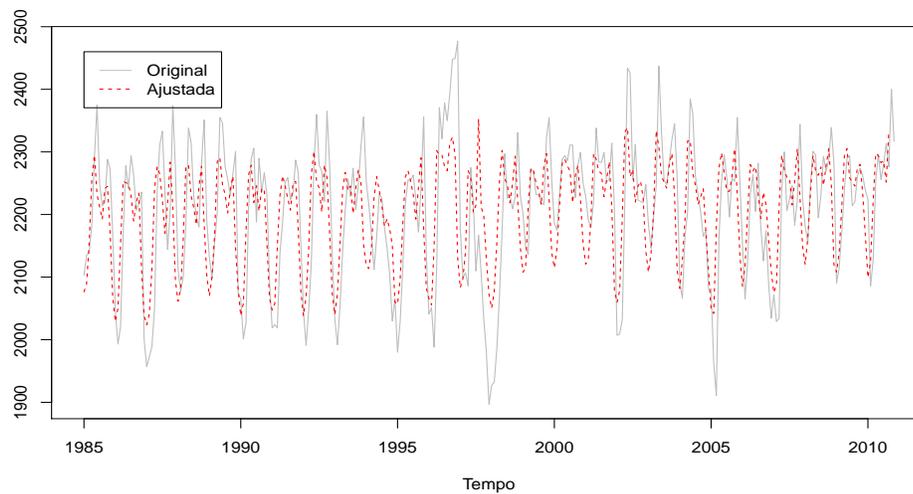


Figura 21 Séries de valores observados e estimados de acordo com o modelo ajustado ARIMA (1,1,1)-CI, de nível médio do mar de Ko Taphao Noi, no período de janeiro de 1985 a dezembro de 2010

4.2 Comparação das séries

Primeiramente, são apresentadas as três séries (Figura 22) para inicial comparação visual.

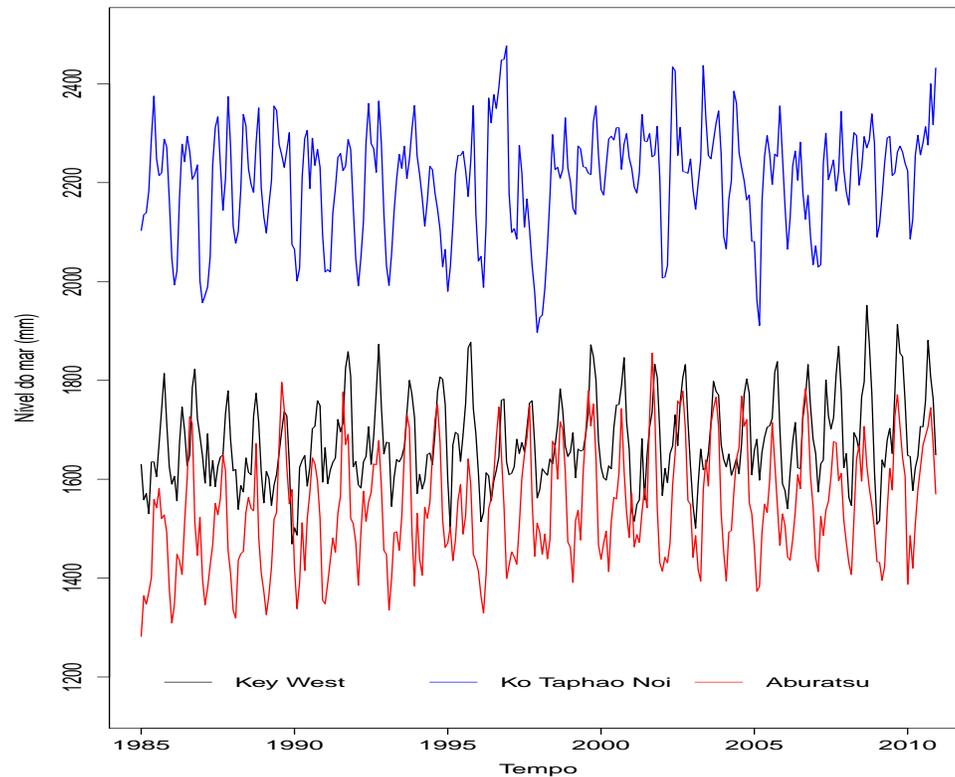


Figura 22 Séries de níveis do mar de Aburatsu, Key West e Ko Taphao Noi, no período de janeiro de 1985 a dezembro de 2010

4.2.1 Teste de Coates e Diggle (1986) aplicado às séries diferenciadas de níveis médios dos mares de Aburatsu e Key West.

Para a realização deste teste utilizaremos o periodograma das séries diferenciadas, no período de janeiro de 1985 a novembro de 2011, que podem ser

observados na Figura 23.

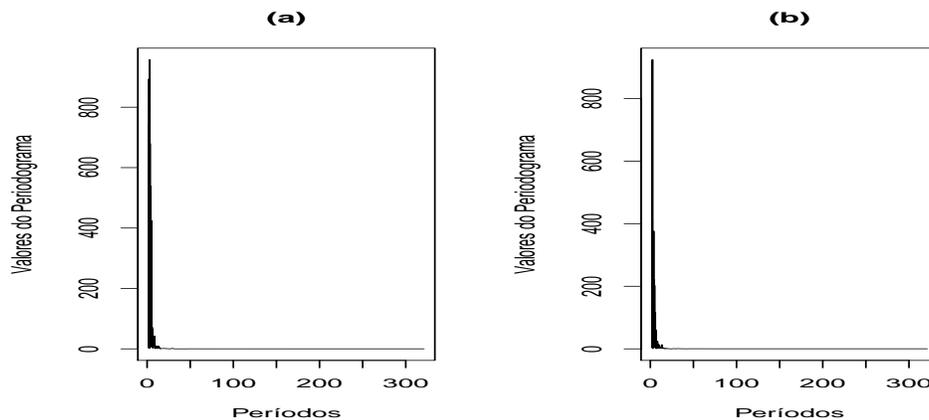


Figura 23 Periodograma da diferença das séries livres de sazonalidade de Aburatsu (a) e Key West (b)

Seguindo os passos para realização do teste, primeiramente calcula-se a razão entre os periodogramas, considerando como numerador os valores do periodograma de Aburatsu e denominador os de Key West, de forma a obter um número menor de valores negativos para o $\ln(J(\lambda))$. Em seguida, calcularam-se os valores de z_i , c_j e o_j .

Para comparar a distribuição dos o_j com a distribuição $U(0, 1)$, utilizou-se o teste de Kolmogorov-Smirnov, e obteve-se um valor $p = 0,3371$, levando a não rejeição da hipótese de igualdade das funções de densidade espectral, ao nível de $\alpha = 0,05$ de significância. Assim, pode-se concluir que as séries são geradas pelo mesmo processo estocástico.

4.2.2 Teste de Coates e Diggle (1986) aplicado às séries diferenciadas de níveis médios dos mares de Aburatsu e Ko Taphao Noi.

O periodograma das séries estacionárias, observados na Figura 24, foram utilizados para a realização do testes. As séries são observadas no período de janeiro de 1985 a novembro de 2011.

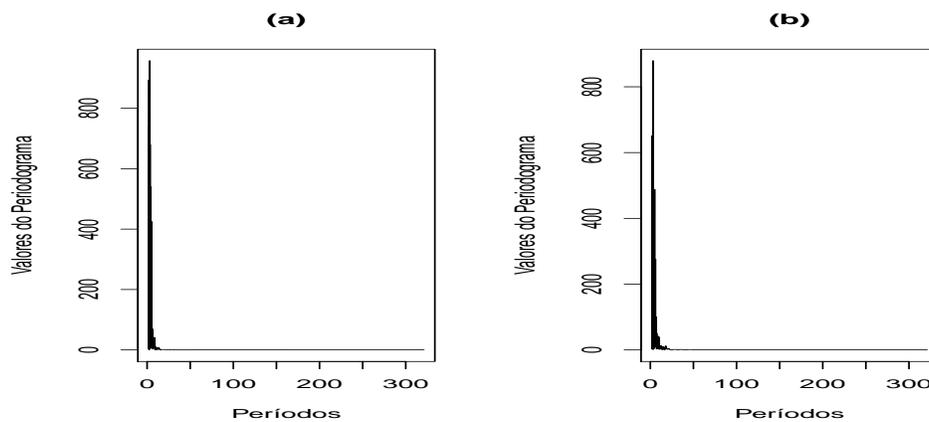


Figura 24 Periodograma das séries diferenciadas de Aburatsu (a) e Ko Taphao Noi (b)

Inicialmente, calcula-se a razão entre os periodogramas, sendo os valores do periodograma de Ko Taphao Noi o numerador, e denominador os de Aburatsu. Foi obtido um valor $p = 1,6 e^{-13}$ para o teste de Kolmogorov, levando a decisão de rejeitar a hipótese de igualdade das funções de densidade espectral, ao nível $\alpha = 0,05$ de significância, de onde se pode concluir que as séries não são geradas pelo mesmo processo estocástico.

4.2.3 Teste de Coates e Diggle (1986) aplicado às séries estacionárias de níveis médios dos mares de Ko Taphao Noi e Key West.

Na realização deste teste, foram utilizadas as séries diferenciadas no período de janeiro de 1985 a novembro de 2011.

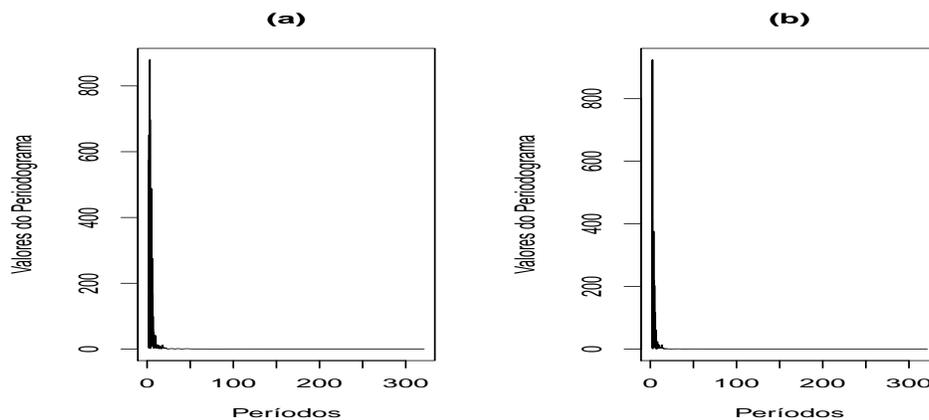


Figura 25 Periodograma das séries diferenciadas de Ko Taphao Noi (a) e Key West (b)

Os periodogramas da série em estudo foram obtidos (Figura 25). Pelo teste de Coates e Diggle (1986) foi observado um valor $p = 0,0258$ no teste de comparação das distribuições. Assim, pode-se concluir que as séries não são geradas pelo mesmo processo estocástico.

4.2.4 Teste de Quenouille (1958) aplicado às séries diferenciadas de níveis médios do mar de Aburatsu e Key West.

Esse teste foi aplicado usando as séries estacionárias de Aburatsu e Key West, no período de janeiro de 1985 a novembro de 2011. Inicialmente, foram obtidas as funções de autocorrelação das séries estacionárias (Figura 26)

A função de autocorrelação parcial estimada comum ($\hat{\Phi}(k)$), que foi obtida

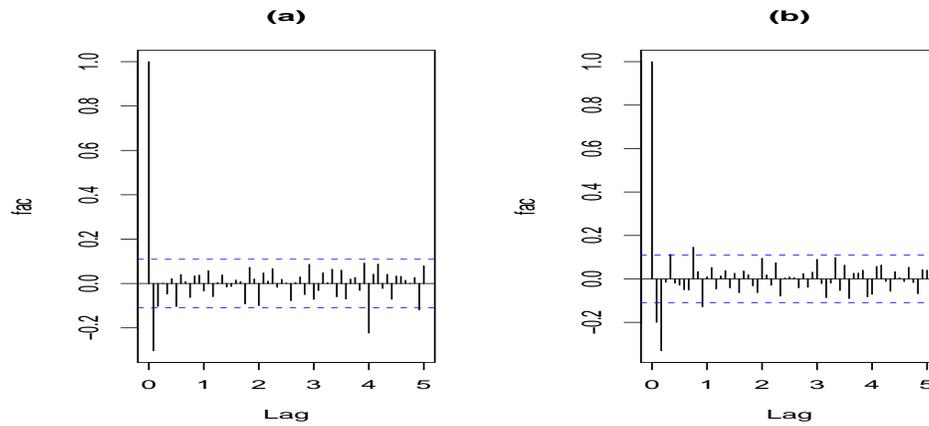


Figura 26 Função de autocorrelação das séries diferenciadas de Aburatsu (a) e Key West (b)

da função de autocorrelação comum $\hat{\rho}(k)$, pode ser observada na Figura 27.

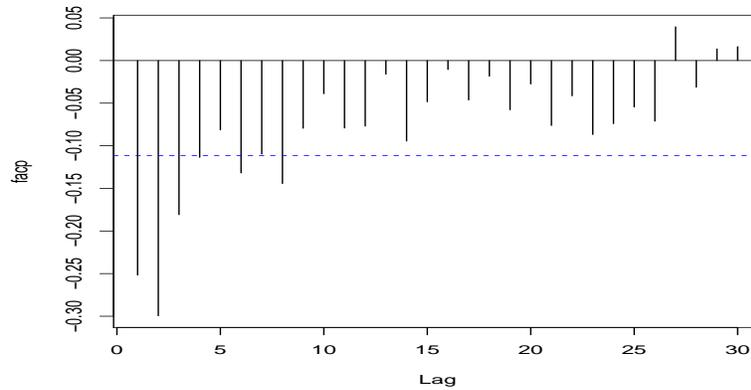


Figura 27 Função de autocorrelação parcial estimada comum ($\hat{\Phi}(k)$)

Analisando a Figura 27 e testando alguns modelos, temos que o modelo autorregressivo que melhor se ajustou foi o de ordem 8.

O modelo ajustado AR(8) pode ser escrito como:

$$Z_t = -0.4523Z_{t-1} - 0.4778Z_{t-2} - 0.3394Z_{t-3} - 0.2672Z_{t-4} - 0.2289Z_{t-5} - 0.2435Z_{t-6} - 0.1728Z_{t-7} - 0.1444Z_{t-8} \quad (4.1)$$

Ajustou-se o modelo 4.1 às séries estacionárias de Aburatsu e Key West, obtendo as séries residuais \hat{a}_1 e \hat{a}_2 . Em seguida, foram obtidas as funções de autocorrelação das séries residuais (Figura 28), onde se pode observar que o ajuste foi bom.

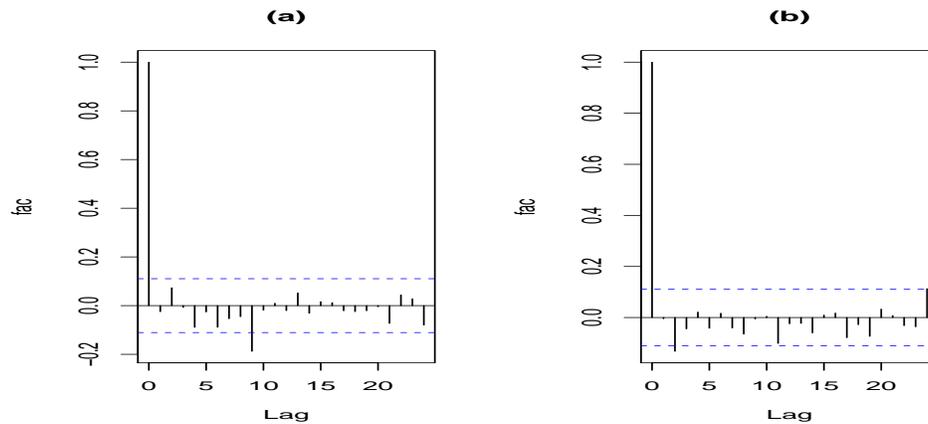


Figura 28 Função de autocorrelação das séries residuais de Aburatsu (a) e Key West (b)

As funções de autocorrelação parcial das séries residuais podem ser observadas na Figura 29. Calculou-se a estatística do teste de Quenouille (1958), obtendo $SQ = 30,20396$ e valor $p = 0,4552$, temos então que a hipótese de igualdade das funções de autocorrelação não pode ser rejeitada. Logo, pode-se concluir que as duas séries são geradas pelo mesmo processo estocástico.

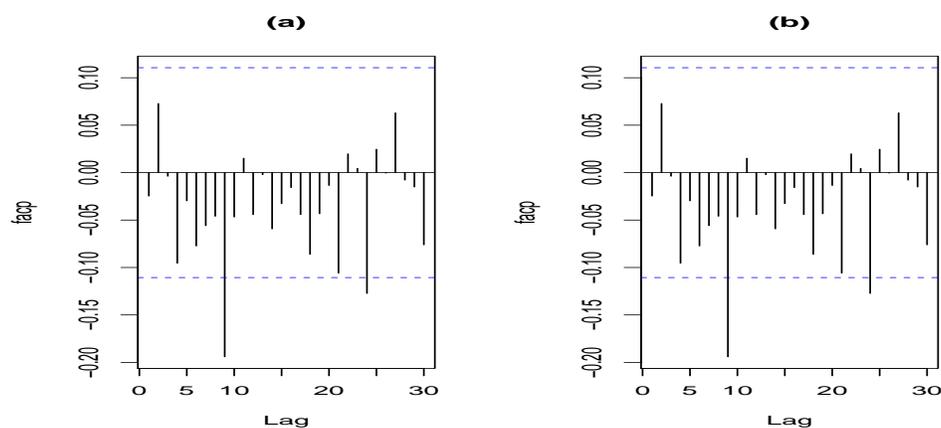


Figura 29 Função de autocorrelação parcial das séries residuais de Aburatsu (a) e Key West (b)

4.2.5 Teste de Quenouille (1958) aplicado às séries diferenciadas de níveis médios do mar de Aburatsu e Ko Taphao Noi.

Para a aplicação deste teste foram usadas as séries diferenciadas de Aburatsu e Ko Taphao Noi, no período de janeiro de 1985 a novembro de 2011. Na Figura 30, temos as funções de autocorrelação das séries estacionárias.

Na Figura 31, é apresentada a função de autocorrelação parcial estimada comum ($\hat{\Phi}(k)$).

Com a análise da Figura 31, e com o teste de alguns modelos, o modelo autorregressivo de ordem 19 foi o que melhor se ajustou.

Este modelo pode ser escrito como:

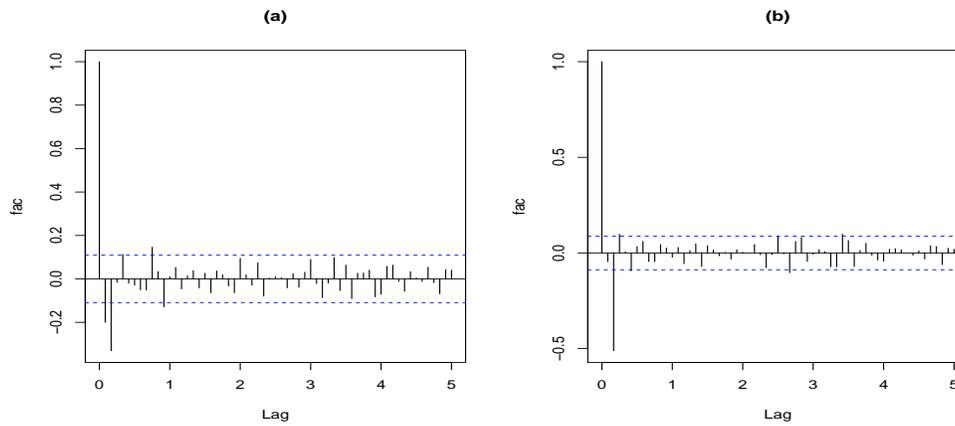


Figura 30 Função de autocorrelação das séries diferenciadas de Aburatsu (a) e Ko Taphao Noi (b)

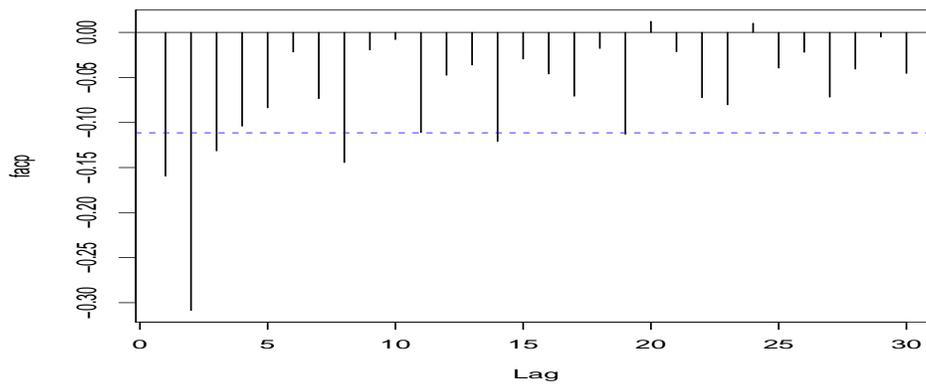


Figura 31 Função de autocorrelação parcial estimada comum ($\hat{\Phi}(k)$)

$$\begin{aligned}
 Z_t = & -0.3124Z_{t-1} - 0.4399Z_{t-2} - 0.2891Z_{t-3} - 0.2358Z_{t-4} - 0.2283Z_{t-5} \\
 & -0.1919Z_{t-6} - 0.2113Z_{t-7} - 0.2569Z_{t-8} - 0.1513Z_{t-9} - 0.1376Z_{t-10} \\
 & -0.2255Z_{t-11} - 0.1647Z_{t-12} - 0.1358Z_{t-13} - 0.1958Z_{t-14} - 0.1043Z_{t-15} \\
 & -0.1071Z_{t-16} - 0.1250Z_{t-17} - 0.0528Z_{t-18} - 0.1129Z_{t-19} \quad (4.2)
 \end{aligned}$$

As séries residuais \hat{a}_1 e \hat{a}_2 foram obtidas com o ajuste do modelo 4.2 às séries estacionárias de Aburatsu e Ko Taphao Noi. Em seguida, foram obtidas as funções de autocorrelação das séries residuais (Figura 32), e as funções de autocorrelação parcial (Figura 33), apresentando um bom ajuste.

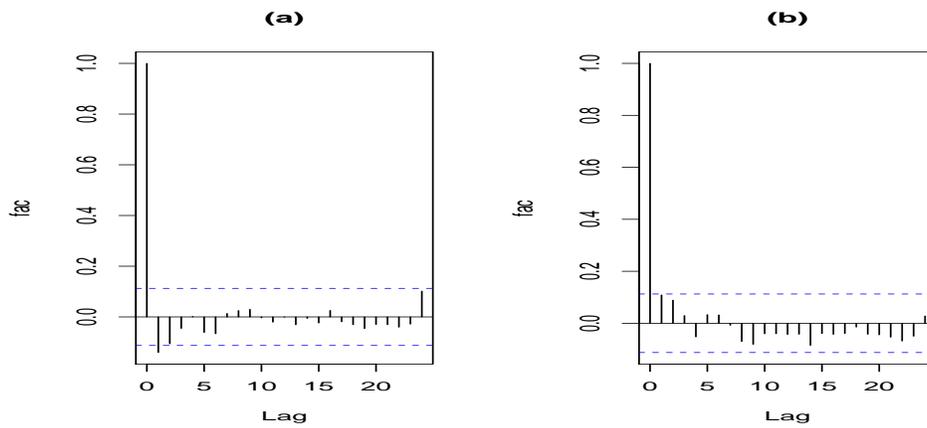


Figura 32 Função de autocorrelação das séries residuais de Aburatsu (a) e Ko Taphao Noi (b)

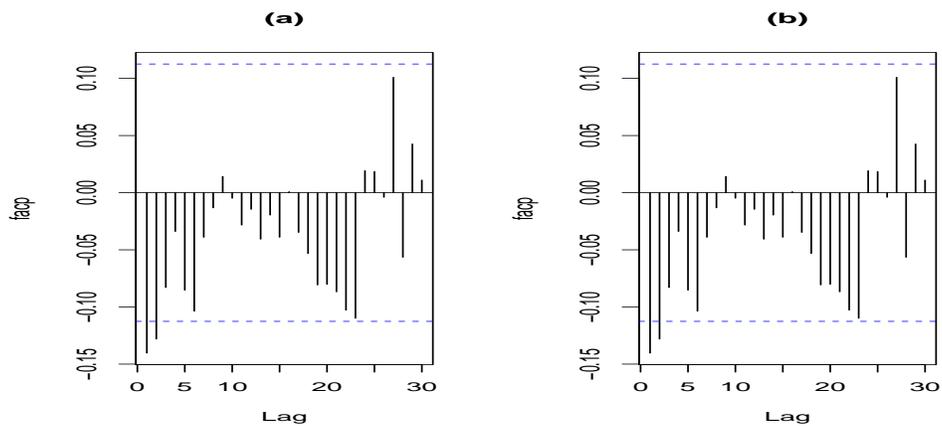


Figura 33 Função de autocorrelação parcial das séries residuais de Aburatsu (a) e Ko Taphao Noi (b)

Calculou-se a estatística do teste obtendo $SQ = 44,07$ e valor $p = 0,0470$, e temos então que a hipótese de igualdade das funções de autocorrelação é rejeitada. Logo, não se pode concluir que as duas séries são geradas pelo mesmo processo estocástico.

4.2.6 Teste de Quenouille (1958) aplicado às séries diferenciadas de níveis médios do mar de Ko Taphao Noi e Key West .

Primeiramente, foram obtidas as funções de autocorrelação das séries estacionárias (Figura 34), no período de janeiro de 1985 a novembro de 2011.

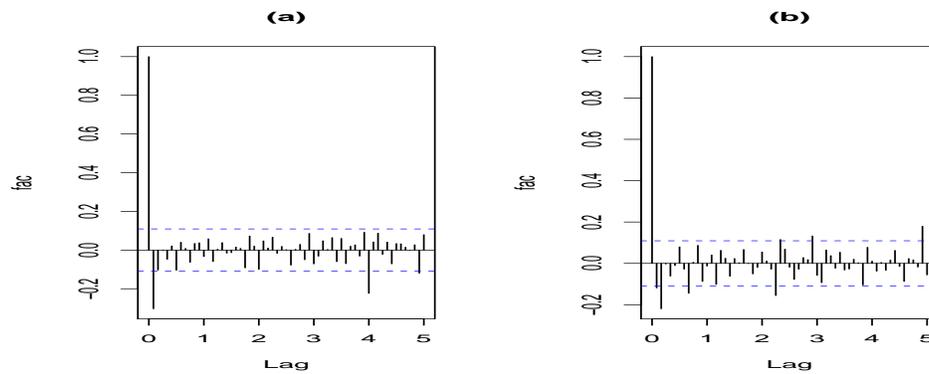


Figura 34 Função de autocorrelação das séries diferenciadas de Key West (a) e Ko Taphao Noi (b)

A função de autocorrelação parcial estimada comum ($\hat{\Phi}(k)$) pode ser observada na Figura 35.

Analisando a Figura 35, e com o teste de alguns modelos, o modelo autorregressivo que melhor se ajustou foi o de ordem 14.

O modelo ajustado AR(14) pode ser escrito como:

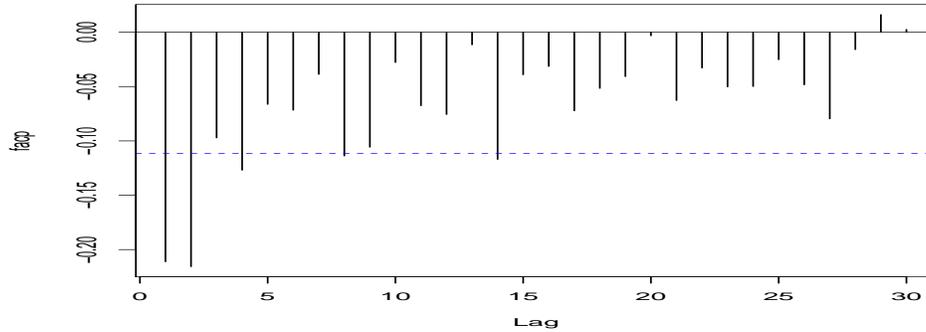


Figura 35 Função de autocorrelação parcial estimada comum ($\hat{\Phi}(k)$)

$$\begin{aligned}
 Z_t = & -0.3348Z_{t-1} - 0.3374Z_{t-2} - 0.2274Z_{t-3} - 0.2396Z_{t-4} - 0.1817Z_{t-5} \\
 & -0.1863Z_{t-6} - 0.1549Z_{t-7} - 0.2061Z_{t-8} - 0.1735Z_{t-9} - 0.1039Z_{t-10} \\
 & -0.1215Z_{t-11} - 0.1180Z_{t-12} - 0.05060Z_{t-13} - 0.1170Z_{t-14} \quad (4.3)
 \end{aligned}$$

Ajustou-se o modelo 4.3 às séries estacionárias de Ko Taphao Noi e Key West, obtendo as séries residuais \hat{a}_1 e \hat{a}_2 . Em seguida, foram obtidas as funções de autocorrelação das séries residuais (Figura 36).

As funções de autocorrelação parcial das séries residuais podem ser observadas na Figura 37. Calculou-se a estatística do teste obtendo $SQ = 39,61$ e valor $p = 0,01127$. Conclui-se que a hipótese de igualdade das funções de autocorrelação deve ser rejeitada, e as séries não são geradas pelo mesmo processo estocástico.

Portanto, os dois testes apresentaram concordância com relação à Aburatsu e Key West serem gerados pelo mesmo processo estocástico. Dessa forma, podemos concluir que, apesar de as duas séries apresentarem médias diferentes, o seu comportamento é semelhante ao longo do tempo.

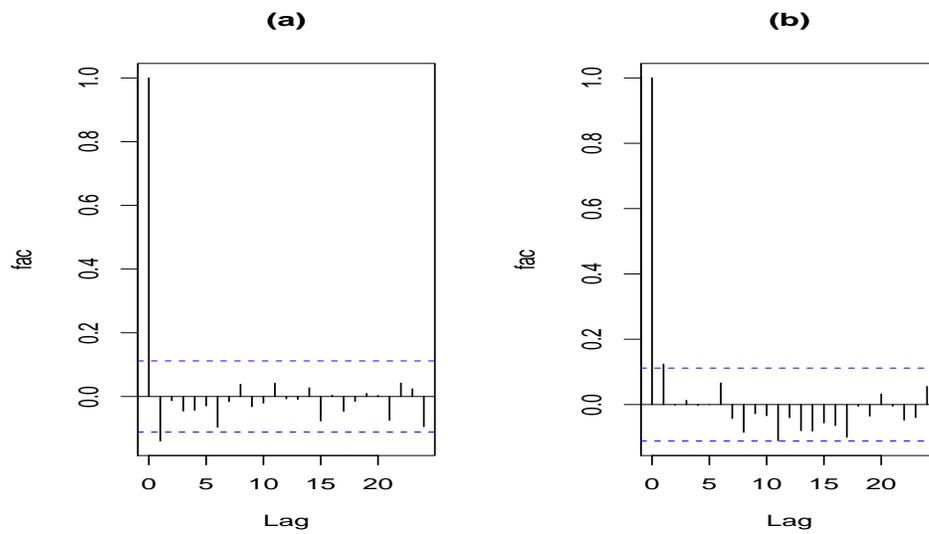


Figura 36 Função de autocorrelação das séries residuais de Key West (a) e Ko Taphao Noi (b)

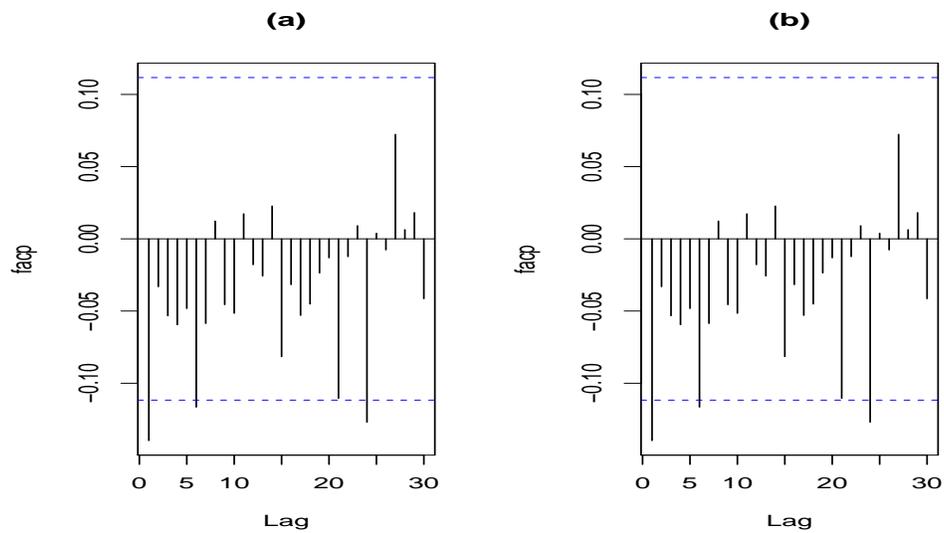


Figura 37 Função de autocorrelação parcial das séries residuais de Key West (a) e Ko Taphao Noi (b)

4.3 Modelo de regressão para a série de níveis do mar de Cananéia

Inicialmente, é apresentada a representação gráfica da série de níveis médios do mar de Cananéia (Figura 38), onde não é observada a presença de tendência e nem de sazonalidade. E através dos gráficos (Figura 39) da função de autocorrelação e autocorrelação parcial da série, pode-se notar que a mesma é um ruído branco.

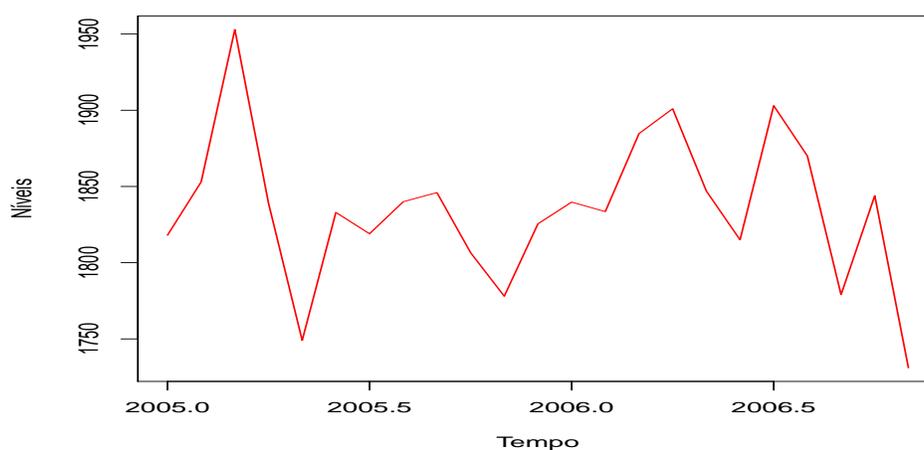


Figura 38 Série mensal de níveis do mar de Cananéia, no período de janeiro de 2005 a dezembro de 2006

O modelo de regressão consistiu na variável dependente como sendo a série de níveis médios do mar de Cananéia. As variáveis independentes são condição climática, direção do vento (graus), velocidade do vento (m/s), referência do barômetro (mmHg), temperatura do psicômetro ($^{\circ}\text{C}$), umidade relativa (%), referência do higrômetro (%), referência do termômetro ($^{\circ}\text{C}$), temperatura máxima ($^{\circ}\text{C}$), temperatura mínima ($^{\circ}\text{C}$), precipitação (mm) e temperatura d'água ($^{\circ}\text{C}$). Todas as variáveis independentes foram analisadas, e dessa forma verificou-se que seguem

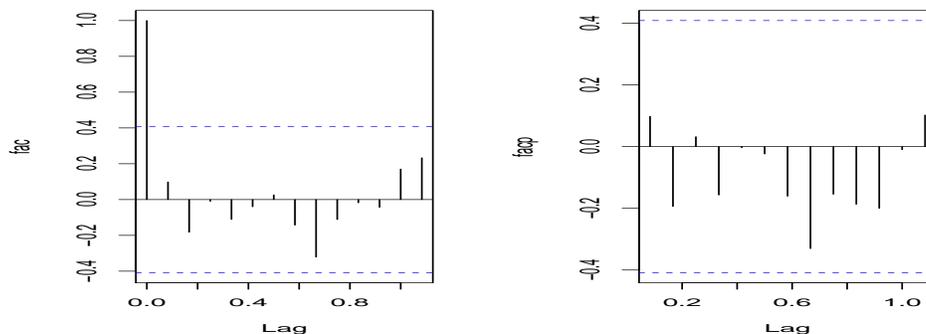


Figura 39 Funções de autocorrelação e autocorrelação parcial (Cananéia)

distribuição normal e atendem aos pressupostos para a regressão.

O três métodos de seleção (stepwise, forward e backward) foram utilizados, no entanto, através do método stepwise não foi encontrado um modelo com estimativas dos parâmetros significativas. Dessa forma, o modelo escolhido foi selecionado através dos métodos forward e backward, que resultaram no mesmo ajuste. Apenas as variáveis temperatura do psicômetro (TP) e temperatura da água (TAG) foram mantidas no modelo ajustado, cuja equação é dada por:

$$\hat{y} = 1950,19 - 110,88 TP + 97,24 TAG$$

A estimativa do parâmetro da temperatura do psicômetro nos mostra que, fixando a temperatura da água, o nível do mar varia em -110,88 mm a cada grau aumentado no psicômetro. Fixando a temperatura do psicômetro, temos que o nível do mar varia 97,24 mm a cada grau de temperatura da água.

O coeficiente de determinação múltipla (R^2) foi de 0,4893, ou seja, 48,93% da variação do nível médio do mar está sendo explicado pelo modelo, e o coeficiente de determinação ajustado foi de (R_a^2) foi de 0,4355.

Na Tabela 9, é apresentada a análise de variância, onde se pode observar

que a hipótese de existência de regressão é aceita.

Tabela 9 Análise de variância do modelo de regressão ajustado para Cananéia.

Fonte	GL	SQ	QM	F
Regressão	2	26420,97	13210,48	9,10*
Resíduo	19	27577,31	1451,43	
Total	21	53998,28		

* significativo ao nível de 5%.

Os regressores foram submetidos ao teste t de Student, e na Tabela 10 tem-se que temperatura da água e temperatura do psicômetro foram estatisticamente diferentes de zero.

Tabela 10 Estimativas dos parâmetros do modelo de regressão, desvio padrão e estatística t de Student.

Parâmetro	Coefficiente	Erro Padrão	t
Intercepto	1950,19	75,91	25,69*
TP	-110,88	26,28	-4,22*
TAG	97,24	23,68	4,11*

* significativo ao nível de 5%.

A análise gráfica da distribuição dos resíduos em função dos valores estimados pelo modelo não revelou a existência de nenhuma tendência, indicando a hipótese de homocedasticidade (Figura 40).

Para independência dos resíduos, foi feito o teste de Durbin-Watson que apresentou os valores $d_w = 1,7675$ com valor $p = 0,2242$, confirmando que os resíduos são independentes, uma alternativa a esse teste é o correlograma dos resíduos (Figura 41), onde pode-se notar a ausência de correlação.

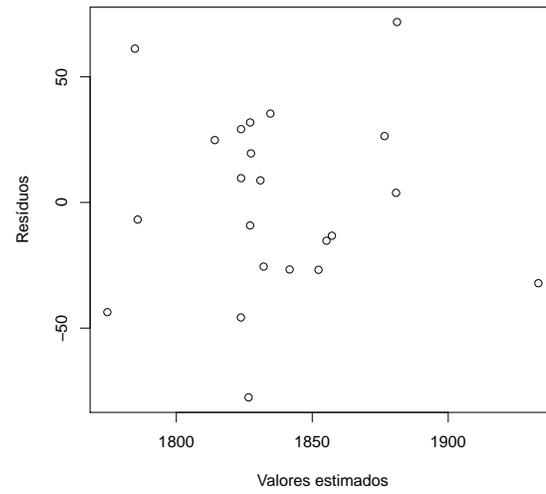


Figura 40 Dispersão dos resíduos em função dos valores estimados pelo modelo

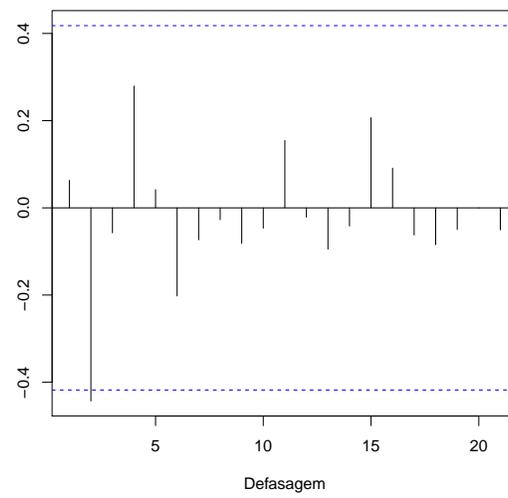


Figura 41 Correlograma dos resíduos do modelo de regressão ajustado

A Figura 42 mostra o gráfico de normalidade dos resíduos, onde se pode verificar que os valores estimados estão dispostos em linha diagonal, indicando uma linearidade satisfatória, a hipótese de normalidade foi aceita com base no teste de Kolmogorov-Smirnov, onde foi obtido um valor $p=0,9979$.

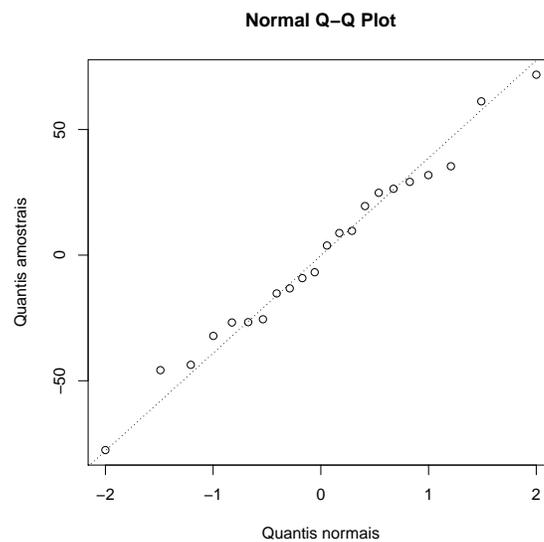


Figura 42 Gráfico da normalidade dos resíduos do modelo de regressão ajustado

Resumindo, de acordo com a análise residual, o modelo proposto não apresenta problemas. As variáveis que apresentaram influência sobre o comportamento do nível do mar foram a temperatura do psicômetro, que avalia a quantidade de vapor de água contido no ar, e temperatura da água. A partir disso, temos que quanto maior for a temperatura do psicômetro menor será o nível do mar, e para a temperatura da água temos o inverso disso. No entanto, é importante ressaltar que no modelo as duas variáveis são consideradas juntas.

5 CONCLUSÃO

Conclui-se que os componentes estruturais de tendência das séries de Aburatsu, Key West e Ko Taphao Noi, apresentaram comportamentos com natureza estocástica, e os componentes sazonais, uma composição com componentes determinísticos. Essas três séries apresentaram tendência crescente, o que confirma o aumento do nível do mar nos últimos anos.

De uma maneira geral, os modelos de Box e Jenkins se ajustaram bem aos dados de níveis médios do mar, somente os picos não foram muito bem captados.

Para a série de Aburatsu, o modelo que melhor se ajustou aos dados foi ARIMA (3,1,1). Na série da Key West tivemos o ARIMA (1,1,1) como modelo de melhor ajuste, e para a série de Ko Taphao Noi o modelo ARIMA (1,1,1) com intervenção em outubro de 1997, foi o de melhor ajuste. Vale ressaltar que no período da intervenção ocorreu o maior pico do fenômeno climático El Niño.

Ao nível de significância $\alpha = 0,05$, os dois procedimentos de comparação de séries aceitaram a hipótese de que apenas as séries de Aburatsu e Key West foram geradas pelo mesmo processo estocástico.

Para a série de Cananéia observou-se que a temperatura do psicômetro e a temperatura da água possuem influência no comportamento do nível do mar deste local.

REFERÊNCIAS

- AKAIKE, H. Maximum likelihood identification of gaussian autorregressive moving average models. **Biometrika**, London, v. 60, n. 21, p. 255-265, 1973.
- ALONSO, J. C.; SEMAÁN, P. Prueba de HEGY en R: Una guía. Tech. Report 23, **Apuntes de Economía**, 2010.
- BEAULIEU, J. J. e MIRON, J. A. Seasonal Unit Roots in Aggregate Data. **Journal of Econometrics**, v.55, p. 305-328, 1993.
- BOX, G. E. P.; JENKINS, G. M.; **Time series analysis:forecasting and control**. San Francisco: Holden Day, 1976.
- BOX, G. E. P.; PIERCE, D. A. Distribution of residual autocorrelations in autorregressive-integrated moving average time series models. **Journal of the American Statistical Association**, New York, v. 65, n. 332, p. 1509-1526, Dec. 1970.
- BUNCHAFT, G. **Estatística sem mistérios**. 4a ed. Petrópolis, RJ: Vozes, 1997.
- CARMONA, R. and WANG, A. Comparison tests for spectral of dependent multivariate time series. **Mathematical models of oceanography**, (eds R. J.Adler, P.Müller and B. L.Rozovskii), p. 69-88. Boston: Birkhäuser, 1994.
- CHARNET, R.; FREIRE, C.A.L.; CHARNET,E.M.R; BONVINO, H. **Análise de modelos de regressão linear com aplicações**. 2a ed. Campinas, SP: Editora da Unicamp, 2008.
- CHURCH, J. A., GREGORY, J. M. , HUYBRECHTS, P., KUHN, M., LAMBECK, K., NHUAN, M. T., QIN, D, WOODWORTH, P. L. Changes in sea level. **Climate Change 2001: The Scientific Basis**, HOUGHTON, J. T. et al., Eds., Cambridge University Press, 639-694, 2001.
- CHURCH, J. A., NEIL J. W., COLEMAN, R., LAMBECK, K., MITROVICA, J.

X. Estimates of the Regional Distribution of Sea Level Rise over the 1950-2000 Period. **J. Climate**, v. 17, 2609-2625, 2004.

CLEVELAND, R. B.; CLEVELAND, W. S.; MCRAE, J.E. and TERPENNING, I. STL: A Seasonal-Trend Decomposition Procedure Based on Loess. **Journal of Official Statistics**, v. 6, 3-73, 1990.

COATES, D.S.; DIGGLE, P. J. Tests for comparing two estimated spectral densities. **Journal of Time Series Analysis**, Clevedon, v. 7, n. 1, p. 7-20, Feb. 1986.

CONOVER, W. J. J. **Practical Nonparametric Statistics**. Nova Iorque: John Wiley, 1990.

COSTA, F. M.; SÁFADI, T. Comparação estatística de duas séries de material particulado (MP_{10}) na cidade de São Paulo. **Revista Brasileira de Biometria** v.28, n.3, p. 23-38, 2010.

DAS, P., RADHAKRISHNA, M. An analysis of Indian tide-gauge records. Sea Level Rise, **International Geophysical Series**, v. 100, Springer India, p. 177-194, 1991.

DETTE, H.; PAPANODITIS, E. Bootstrapping frequency domain tests in multivariate time series with an application to comparing spectral densities. **Journal of the Royal Statistical Society: Series B** 71, p. 831-57, 2009.

ECHEVERRY, G. E.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** v. 61, n.215, p. 51-80, 2000.

FRANSES, P. H. Testing for seasonal unit root in monthly data. Econometric Institute **Report 9032A**. Rotterdam. Erasmus University, 1990.

FRANSES, P. H. Seasonality, non-stationarity and the forecasting of monthly

time series. **International Journal of Forecasting**, v.7, p.199-208, 1991.

FRANSES, P. H.; HOBIN, B. Critical values for unit root tests in seasonal time series. **Journal of Applied Statistics**, v.24, p.25-46, 1997.

GLOSS - The Global Sea Level Observing System. Disponível em <<http://ilikai.soest.hawaii.edu/uhs/c/woce.html>>. Acesso em fev. 2011.

GRAYBILL, F. A.; IYER, H. K. **Regression Analysis. Concepts and Applications**. Duxbury Press, Belmont, 1994.

HARARI, J.; FRANÇA, C. A. S.; CAMARGO, R. Variabilidade de longo termo de componentes de marés e do nível médio do mar na costa brasileira. **Afro-America Gloss News**, v. 8, n. 1, 2004.

HYLLEBERG, S.; ENGLE, R.; GRANGER, C.; YOO, B. Seasonal integration and cointegration. **Journal of Econometrics**, v. 44, p. 215-238, 1990.

L'UTKEPOHL, H.; KR'ATZIG, M. (JMulTi). **Applied Time Series Econometrics**. Cambridge: Cambridge University Press, 2004.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. 2. ed. São Paulo: E. Blucher, 2006. 538p.

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.

PUGH, D., J. Hunter, R. Coleman, and C. Watson, A comparison of historical and recent sea level measurements at Port Arthur, Tasmania, **The International Hydrographic Review**, v. 3, n. 3 (New Series), p. 27 - 46, 2002.

PUDNEY, S. E.; DEADMAN, D. F. and CLARK, F. A simple nonparametric test for the equality of autocorrelation structures of two stochastic processes, with application to experimental data. **Pub. Inst. Stat. Univ**, Paris, v. 43, n. 2- 3, p. 85-101, 1999.

PRIESTLEY, M. B. **Spectral analysis and time series**. London: Academic, 1989. 407p.

R DEVELOPMENT CORE TEAM. **R: a language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2011. Disponível em <<http://www.r-project.org>>. Acesso em: 1 mar. 2011.

SOLOMON, S., QIN, D., MANNING, M., CHEN, Z., MARQUIS, M., AVERYT, K.B., TIGNOR, M., MILLER, H.L. **Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change**, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

TITUS, J.G., NARAYANAN, V. The risk of sea level rise, **Climatic Change**, v. 33, p. 151-212, 1996.

UNESCO. Manual de Medição e Interpretação do Nível do Mar: Volume IV - An Update to 2006. **Comissão Oceanográfica Intergovernamental**, 2006, 88 p.

WÖPPELMANN, G.; LETETREL, C.; SANTAMARIA, A.; BOUIN, M.-N.; COLLILIEUX, X.; ALTAMIMI, Z.; WILLIAMS, S. D. P.; MARTIN MIGUEZ, B. Rates of sea-level change over the past century in a geocentric reference frame, **Geophysical Research Letters**, v.36, L12607, doi:10.1029/2009GL038720, 2009.

APÊNDICE

São apresentados os resultados da análise inicial para os mares de Key West e de Ko Taphao Noi. Para cada série de nível do mar foram utilizados os mesmos critérios estabelecidos na série de Aburatsu.

Key West

Na Figura 43 temos a decomposição da série através do método STL.

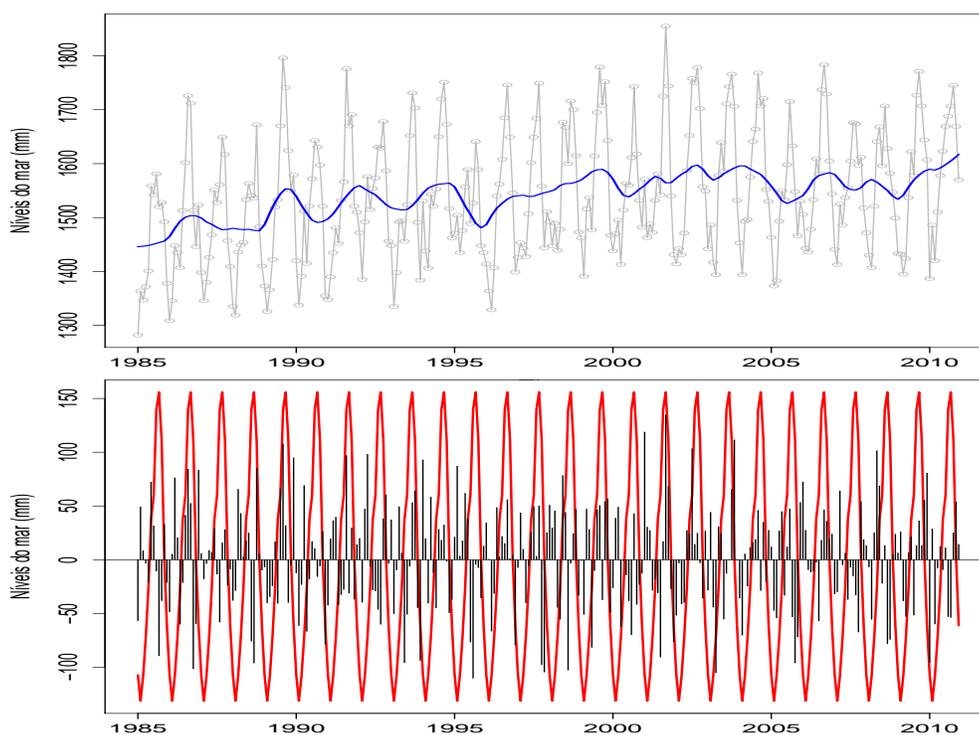


Figura 43 Decomposição sazonal usando STL aplicado à série mensal de níveis médios do mar de Key West, no período de janeiro de 1985 a dezembro de 2010. Os círculos cinzas mostram a série original, a linha azul é a componente tendência, a vermelha mostra a componente sazonal, e as barras pretas são o componente aleatório

Nota-se que a componente tendência apresenta um leve crescimento, e a componente sazonal apresenta-se com poucas mudanças ao longo do tempo, de modo que os anos têm padrão semelhante. O componente aleatório é o que resta na série após a eliminação dos outros componentes.

A série de Key West foi dividida em grupos de 12 observações consecutivas, e para cada grupo foram calculadas a média e a amplitude. Na Figura 44 temos a representação gráfica desse procedimento. O comportamento do gráfico indica que a série não precisa de nenhuma transformação.

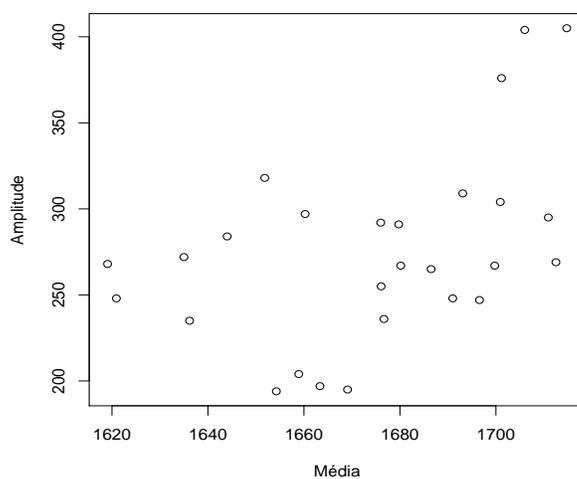


Figura 44 Série mensal de níveis médios do mar de Key West: amplitude vs média

Por meio do periodograma (Figura 45) da série de nível médio do mar é possível observar dois picos, nos períodos 6 e 12 meses. Para testar a existência do efeito sazonal em ambos os períodos aplicou-se o teste de Fisher, em relação ao período 6 temos $g = 0,1992 > z_{0,05} = 0,0494$, a série apresenta sazonalidade de 6 meses, a 5% de significância. O teste referente ao período 12 também

foi realizado, obtendo $g = 0.2102 > z_{0,05} = 0,0492$, constatando portanto a sazonalidade de 12 meses.

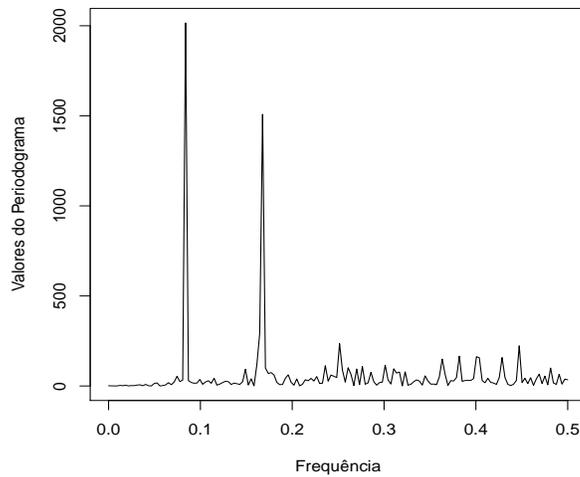


Figura 45 Periodograma da série de níveis médios do mar da Key West

Plotando cada mês em anos sucessivos (Figura 46), podemos observar o padrão sazonal da série, e que existe uma variação não muito grande dentro de cada mês. O padrão global mostra que fevereiro e outubro têm menores e maiores elevações do nível do mar, respectivamente.

Uma vez que foi comprovado que a série apresenta a componente sazonal, tem-se que identificar a natureza da mesma. O modelo com uma tendência linear e variáveis dummies é estimado.

Os resultados mostram que as variáveis dummies e a tendência do modelo foram significativas. Na Tabela 11 pode-se conferir que a hipótese nula sobre as raízes π_1 e π_2 , e para $\pi_3 = \pi_4$ é aceita:

Dessa forma, utilizou-se a regressão com variáveis *dummies* sazonais, para estimar a sazonalidade determinística e, assim obter a série corrigida da vari-

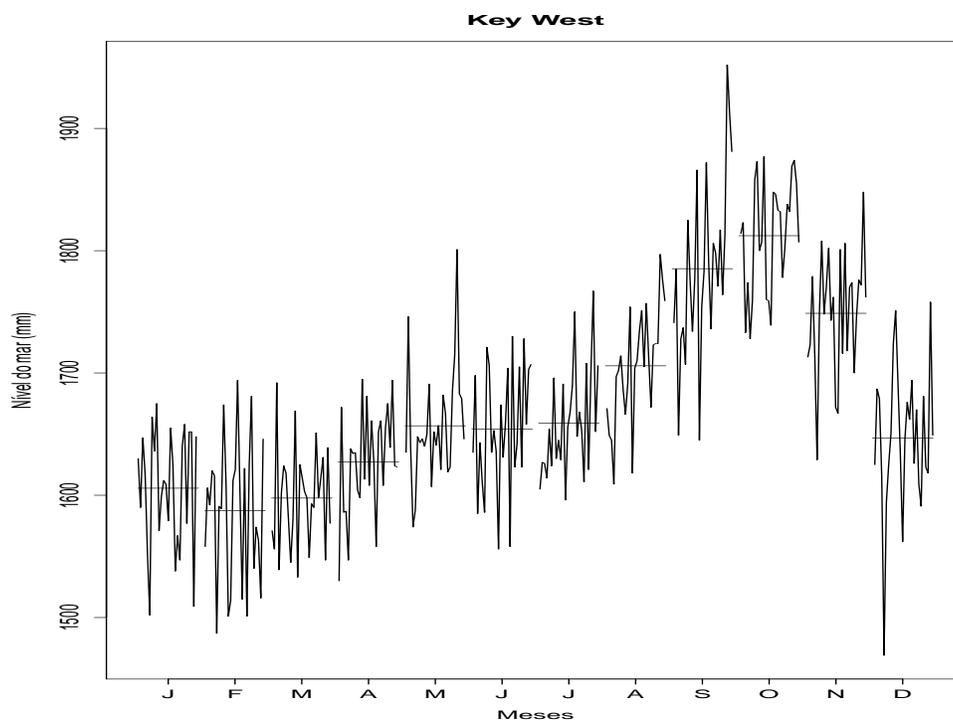


Figura 46 Subséries mensais dos níveis médios do mar de Key West

Tabela 11 Resultado do teste de raiz unitária sazonal (teste HEGY) para a série de nível médio do mar de Key West.

	Valores críticos (5%)	Valores da estatística do teste
$t(\pi_1)$	-3,35	4,34
$t(\pi_2)$	-2,81	4,35
$F(\pi_3 \pi_4)$	6,35	5,94
$F(\pi_5 \pi_6)$	6,48	30,51
$F(\pi_7 \pi_8)$	6,30	17,82
$F(\pi_9 \pi_{10})$	6,40	41,84
$F(\pi_{11} \pi_{12})$	6,46	28,94
$F(\pi_1 - \pi_{12})$	4,44	25,11
$F(\pi_2 - \pi_{12})$	4,58	25,02

ação sazonal. Na Tabela 12 temos o erro padrão e teste de significância para os

coeficientes do modelos ajustado, onde temos que o efeito de tendência e sazonalidade mensal são significantes.

Na Tabela 12 podemos observar que o parâmetro β_1 assume o valor 0,1708, indicando um crescimento de 2,0496 mm/ano. Os meses de maior aumento são setembro e outubro, fevereiro apresenta o menor aumento de todos.

Tabela 12 Significância dos parâmetros do modelo de regressão com *dummies* sazonais ajustado para eliminar sazonalidade determinística.

Parâmetros	Estimativas	Erro padrão
β_0	1647.182	0,00
β_1	0.1708	0.0479
α_1	-67.6719	8.3171
α_2	-86.0377	8.3268
α_3	-75.7608	8.3231
α_4	-46.4636	8.3196
α_5	-16.9233	8.3176
α_6	-19.6953	8.3167
α_7	-14.9305	8.3176
α_8	31.6471	8.3176
α_9	110.9806	8.3195
α_{10}	137.8643	8.3231
α_{11}	74.4385	8.3270
α_{12}	-29.1243	8.3270

O teste de sequência (Wald-Wolfowitz) foi aplicado verificar a existência de tendência na série sem a componente sazonal. Foi obtido um valor da estatística do teste de $-3,7336$ e o valor $p = 0,00018$, levando a concluir que a série apresenta tendência. Como a tendência caracteriza um processo não estacionário, e o ajuste de modelos pressupõe estacionariedade, a série foi diferenciada uma vez para eliminar essa componente.

Ko Taphao Noi

A decomposição da série através do método STL e mostrada na Figura 47. A componente sazonal apresenta três picos predominantes e se mantém constante, a tendência se apresenta oscilatória com pico em 1997 (período analisado com a intervenção). A parte aleatória apresenta alguns picos ao longo tempo.

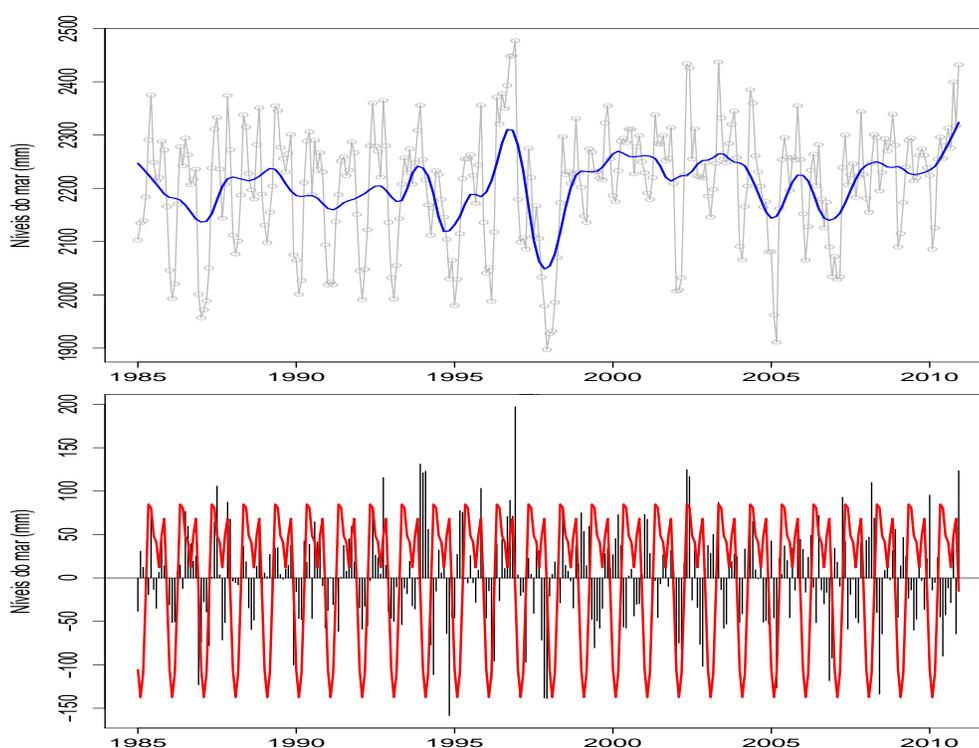


Figura 47 Decomposição sazonal usando STL aplicado à série mensal de níveis médios do mar de Ko Taphao Noi, no período de janeiro de 1985 a dezembro de 2010. Os círculos cinzas mostram a série original, a linha azul é a componente tendência, a vermelha mostra a componente sazonal, e as barras pretas são o componente aleatório

Na Figura 48 é apresentada a representação do procedimento anteriormente descrito para verificar a necessidade de transformação na série. Pode-se

observar que não há a necessidade de uma transformação para estabilizar a variância da série em estudo.

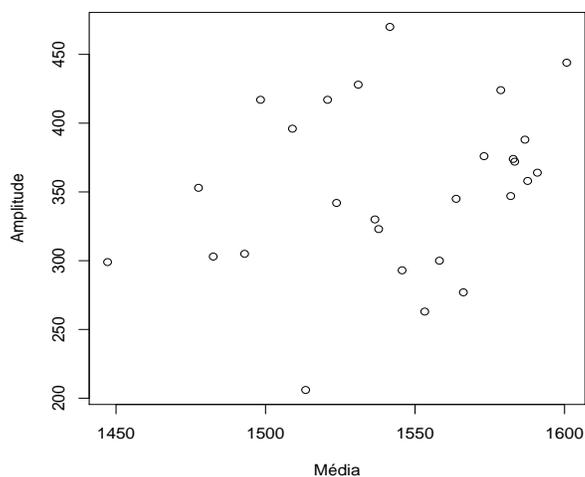


Figura 48 Série mensal de níveis médios do mar da Ko Taphao Noi: amplitude vs média

Através do periodograma da série de níveis médios do mar da Ko Taphao Noi (Figura 49) é possível observar dois picos, nos períodos 6 e 12 meses. Aplicou-se então o teste de Fisher para confirmar a periodicidade, chegou-se às estatísticas $g = 0,1953 > z_{0,05} = 0,0492$ para a periodicidade de 12 meses, e $g = 0,0382 < z_{0,05} = 0,0495$ para a de 6 meses, logo a série apresenta sazonalidade de 12 meses.

Na (Figura 50) pode-se observar uma grande variação ao longo dos anos nos meses de outubro, novembro e dezembro. Os picos superiores ocorrem em maio e novembro, e no mês de fevereiro temos as menores elevações do nível do mar.

Agora será verificada a natureza da componente sazonal através do teste HEGY. O modelo, usado para o teste, com uma tendência linear e variáveis *dummies*

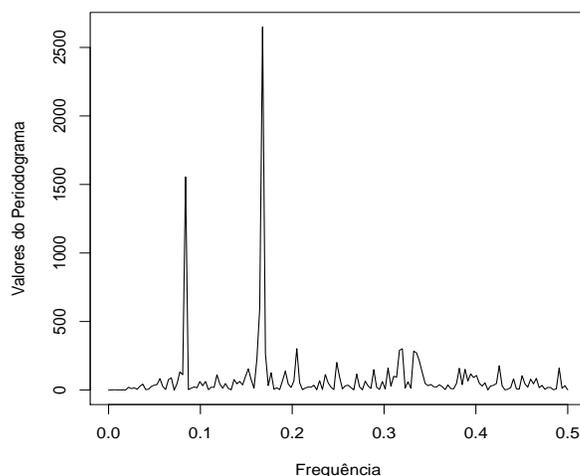


Figura 49 Periodograma da série de níveis médios do mar da Ko Taphao Noi

é estimado.

Os resultados mostram que as variações sazonais contêm uma parte de natureza determinística. Pode-se verificar na Tabela 13 que a hipótese nula sobre as raízes π_1 e π_2 é aceita, o modelo tem uma raiz unitária não-sazonal (tendência estocástica) e uma raiz semestral:

Tabela 13 Valores críticos e da estatística do teste HEGY

	Valores críticos (5%)	Valores da estatística do teste
$t(\pi_1)$	-3,35	4,76
$t(\pi_2)$	-2,81	4,66
$F(\pi_3\pi_4)$	6,35	16,35
$F(\pi_5\pi_6)$	6,48	21,93
$F(\pi_7\pi_9)$	6,30	26,91
$F(\pi_9\pi_{10})$	6,40	32,68
$F(\pi_{11}\pi_{12})$	6,46	25,85
$F(\pi_1 - \pi_{12})$	4,44	26,06
$F(\pi_2 - \pi_{12})$	4,58	24,51

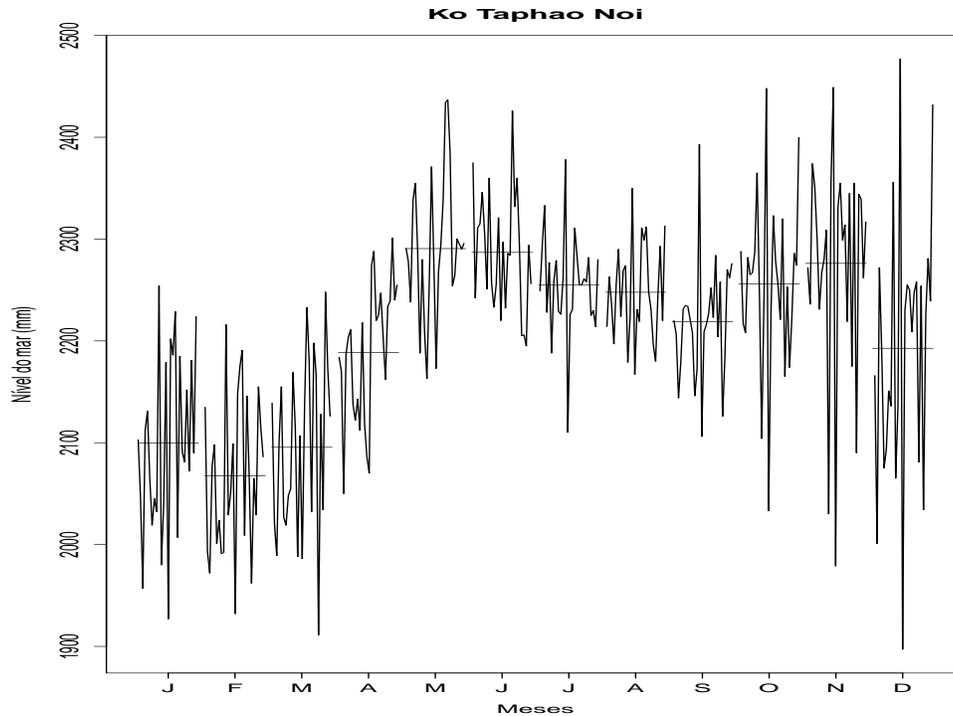


Figura 50 Subséries mensais dos níveis médios do mar de Key West

Numa primeira etapa, elimina-se as variações sazonais por meio do modelo de regressão com variáveis *dummies* sazonais. Na Tabela 14 temos o erro padrão e o teste de significância para os coeficientes do modelo ajustado, onde se pode confirmar que os efeitos de tendência e sazonalidade mensal são significantes.

Com a análise da tabela temos que o mês de menor aumento do nível do mar é fevereiro, e os de maior aumento é maio e junho, havendo certa conformidade com a análise da Figura 50. O parâmetro β_1 que é o coeficiente angular mostra que em Ko Taphao Noi teremos um aumento médio de 0,19 mm/mês, e em um ano isso gera em torno de 2,28mm, sendo também um aumento considerável.

O teste de sequência (Wald-Wolfowitz) foi aplicado para verificar a exis-

Tabela 14 Significância dos parâmetros do modelo de regressão com *dummies* sazonais ajustado para eliminar sazonalidade determinística

Parâmetros	Estimativas	Erro padrão
β_0	2178,89	18,18
β_1	0,19	0,05
α_1	-93,16	22,80
α_2	-124,53	22,80
α_3	-107,38	22,80
α_4	-29,76	22,80
α_5	67,35	22,80
α_6	82,06	22,80
α_7	52,39	22,80
α_8	39,57	22,80
α_9	6,68	22,80
α_{10}	45,71	22,80
α_{11}	59,23	22,80
α_{12}	1,84	22,80

tência de tendência na série corrigida. O valor da estatística do teste foi = -8,3034 e o valor $p = 2,2e^{-16}$, temos então uma série com tendência, e toma-se a primeira diferença para torná-la estacionária.