



Universidade do Estado do Rio de Janeiro

Centro de Tecnologia e Ciências

Faculdade de Engenharia

Hugo Hinostroza Farfán

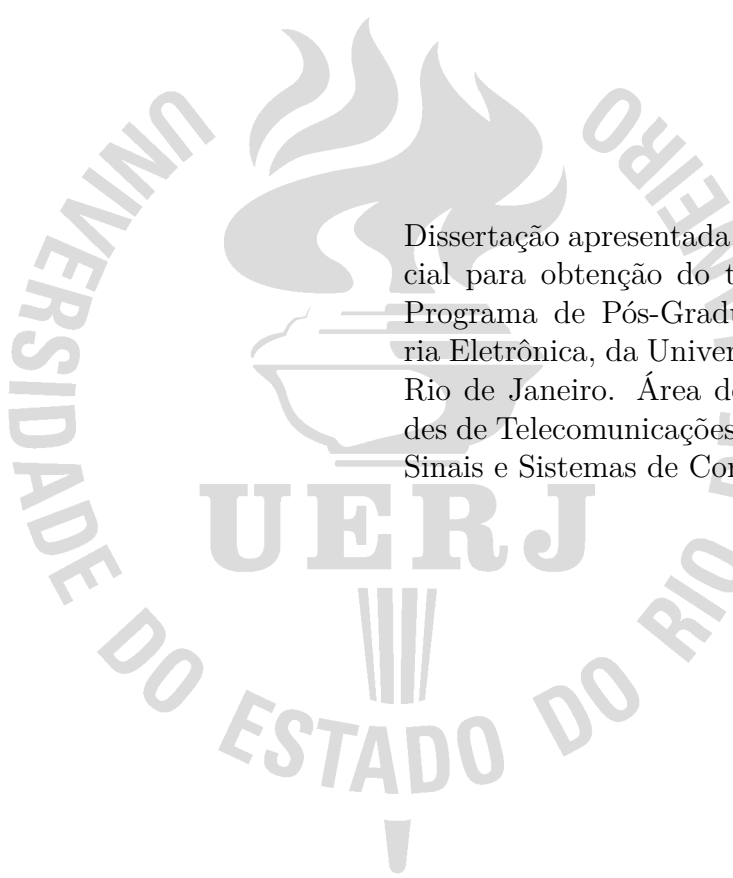
**Filtros Adaptativos no Domínio da Frequência para Correção de
Previsão Climática da Temperatura da Superfície do Mar
Global**

Rio de Janeiro

2023

Hugo Hinostroza Farfán

**Filtros Adaptativos no Domínio da Frequência para Correção de Previsão
Climática da Temperatura da Superfície do Mar Global**



Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Engenharia Eletrônica, da Universidade do Estado do Rio de Janeiro. Área de concentração: Redes de Telecomunicações. Linha de pesquisa: Sinais e Sistemas de Comunicações.

Orientador: Prof. Dr. Michel Pompeu Tcheou

Orientador: Prof. Dr. Lisandro Lovisolo

Rio de Janeiro

2023

CATALOGAÇÃO NA FONTE
UERJ / REDE SIRIUS / BIBLIOTECA CTC/B

F222 Farfán, Hugo Hinostroza.
Filtros adaptativos no domínio da frequência para correção de
previsão climática da temperatura da superfície do mar global /
Hugo Hinostroza Farfán. – 2023.
96 f.

Orientadores: Michel Pompeu Tcheou, Lisandro Lovisolo.
Dissertação (Mestrado) – Universidade do Estado do Rio de
Janeiro, Faculdade de Engenharia.

1. Engenharia eletrônica - Teses. 2. Filtros adaptativos - Teses.
3. Processamento de sinais - Teses. 4. Algoritmos - Teses. 5.
Climatologia - Teses. I. Tcheou, Michel Pompeu. II. Lovisolo,
Lisandro. III. Universidade do Estado do Rio de Janeiro, Faculdade
de Engenharia. V. Título.

CDU 621.391:551.582(26.02)

Bibliotecária: Júlia Vieira – CRB7/6022

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou
parcial desta tese, desde que citada a fonte.

Assinatura

Data

Hugo Hinostroza Farfán

**Filtros Adaptativos no Domínio da Frequência para Correção de Previsão
Climática da Temperatura da Superfície do Mar Global**

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Engenharia Eletrônica, da Universidade do Estado do Rio de Janeiro. Área de concentração: Redes de Telecomunicações. Linha de pesquisa: Sinais e Sistemas de Comunicações.

Aprovada em 19 de maio de 2023.

Banca Examinadora:

Prof. Dr. Michel Pompeu Tcheou (Orientador)
Faculdade de Engenharia – UERJ

Prof. Dr. Lisandro Lovisolo (Orientador)
Faculdade de Engenharia – UERJ

Prof.^a Dra. Chou Sin Chan
Instituto Nacional de Pesquisas Espaciais - INPE

Prof.^a Dra. Zochil González Arenas
Instituto de Matemática e Estatística - UERJ

Rio de Janeiro

2023

DEDICATÓRIA

Primeiramente a minha mãe Filomena Farfán Arango, porque foi minha força, e a todos meus seres queridos que amo, que sempre estiveram ao meu lado me dando total apoio.

AGRADECIMENTOS

Agradeço ao meu orientador Michel P. Tcheou, excelente professor nas aulas e por sua alegria em ensinar. Agradeço ao meu orientador Lisandro Lovisolo, excelente professor e pessoa.

Agradeço à Dra. Chou pela disponibilidade e contribuições durante o desenvolvimento deste trabalho.

Agradeço ao Programa de Pós-Graduação em Engenharia Eletrônica da UERJ.

Agradeço ao Laboratório de Processamento de Sinais, Aplicações Inteligentes e Comunicação (PROSAICO) da UERJ.

Agradeço também pelo apoio financeiro, disponibilização dos dados e pela oportunidade de participar da pesquisa em conjunto com o INPE, CAPES e ANA.

A imaginação é mais importante que o conhecimento.
O conhecimento é limitado. A imaginação envolve o mundo.
Albert Einstein.

RESUMO

HINOSTROZA FARFÁN, HUGO. *Filtros Adaptativos no Domínio da Frequência para Correção de Previsão Climática da Temperatura da Superfície do Mar Global*. 96 f. Dissertação (Mestrado em Engenharia Eletrônica) – Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2023.

Este trabalho de dissertação apresenta resultados da aplicação no domínio da frequência de filtros adaptativos RLS (*Recursive Least Squares*) para a correção de erros de previsão climática global do Modelo Brasileiro do Sistema Terrestre (*Brazilian Earth System Model* – BESM). O BESM é composto por modelos computacionais com o objetivo de gerar cenários de mudanças climáticas, tal como as previsões meteorológicas globais do ERA5 (*ECMWF Reanalysis v5*) produzido e desenvolvido pelo ECMWF (*European Centre for Medium-Range Weather Forecasts*). O desempenho do BESM é medido em referência ao conjunto de dados observacionais do ERA5, a diferença entre as variáveis climáticas em geral aumenta com o tempo. A previsão do BESM é corrigida usando como referência dados observacionais do ERA5. O objetivo central é corrigir os erros de previsão climática global do modelo BESM. Projetam-se filtros usando as sequências de previsões anteriores para uma célula (pequena região do globo onde se aplica a correção) ou nos coeficientes frequenciais obtidos via Transformada de Cosseno Discreta bidimensional (DCT-2D) em uma vizinhança da célula. A DCT permite analisar o conteúdo espectral dos sinais. Os filtros são projetados usando-se o algoritmo adaptativo *Recursive Least Squares* (RLS). Testam-se filtros de diferentes ordens e são consideradas vizinhanças de diferentes tamanhos. A metodologia de correção é testada ainda considerando correções por estações climáticas, separadamente, e anualmente. Avalia-se o desempenho da correção da previsão BESM usando a Raiz do Erro Quadrático Médio (RMSE – *Root Mean Squared Error*) e o Erro Absoluto Máximo (LAE – *Largest Absolute Error*). Aplica-se a proposta na variável prognóstica de Temperatura da superfície do Mar (TSM), para regiões de dimensão $M \times M$, com $M = 8, 16, 32$, e filtros de comprimento $L = 4$. Analisa-se o desempenho da proposta em diversas zonas, como o Oceano Pacífico, o Oceano Atlântico, o Oceano Índico e Oceano Ártico e Antártico, e melhores desempenhos são obtido nas três primeiras zonas, sendo pior nas duas últimas zonas. Consideram-se filtros para correção anual e por estações e os melhores resultados são obtidos para a correção por estações. Em comparação com o método de Correção de Média e Variância (CMV), o uso de filtros RLS gera resultados melhores para o ajuste sazonal. No entanto, em algumas localizações geográficas específicas, o CMV às vezes apresenta um desempenho superior à filtragem adaptativa quando se trata do ajuste anual.

Palavras-chave: Filtros adaptativos. Processamento de sinais. Algoritmo RLS. Modelos climáticos.

ABSTRACT

HINOSTROZA FARFÁN, HUGO. *Adaptive Frequency Domain Filters for the Correction of the Climate Prediction of the Global Sea Surface Temperature*. 2023. 96 f. Dissertação (Mestrado em Engenharia Eletrônica) – Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2023.

This dissertation presents results of the application in the frequency domain of Recursive Least Squares (RLS) adaptive filters for the correction of global climate prediction errors of the Brazilian Earth System Model (BESM). The BESM consists of computational models aimed at generating climate change scenarios, just like the ECMWF (European Centre for Medium-Range Weather Forecasts) which also develops and produces global weather predictions such as ERA5 (ECMWF Reanalysis v5). The performance of BESM is measured with reference to the ERA5 observational dataset, and the difference between climate variables generally increases over time. BESM prediction is corrected using ERA5 observational data as a reference. The main objective is to correct the global climate prediction errors of the BESM model. Filters are designed using the previous prediction sequences for a cell (a small region of the globe where correction is applied) or in the frequency coefficients obtained via two-dimensional Discrete Cosine Transform (DCT-2D) in a cell neighborhood. DCT is a mathematical technique for analyzing periodic signals in the frequency domain with a sum of cosine functions. The filters are designed using the Recursive Least Squares (RLS) adaptive algorithm. Filters of different orders are tested, and neighborhoods of different sizes are considered. The correction methodology is also tested considering corrections seasonally and annually, separately. The performance of BESM prediction correction is evaluated using the Root Mean Square Error (RMSE) and Largest Absolute Error (LAE). The proposal is applied to the prognostic variable of Sea Surface Temperature (TSM) for regions of dimension $M \times M$, with $M = 8, 16, 32$, and filters of length $L = 4$. Among the various zones analyzed, such as the Pacific Ocean, the Atlantic Ocean, the Indian Ocean, and the Arctic and Antarctic Oceans, the best performance of the proposed method is achieved in the first three zones, while the worst is observed in the last two zones. Finally, the best results are obtained for correction by seasons. Compared to the Mean and Variance Correction (CMV) method, the use of RLS filters generates better results for seasonal adjustment. However, in some specific geographic locations, CMV sometimes outperforms adaptive filtering when it comes to annual adjustment.

Keywords: Adaptive filtering. Signal processing. RLS algorithm. Climate models.

LISTA DE FIGURAS

Figura 1	- Séries temporais anuais da Temperatura da Superfície do Mar (TSM) do modelo regional BESM e modelo ERA5 às 12:00h UTC. Latitude = -22.25° (graus a Norte) e longitude = -8.0° (graus a Leste).	20
Figura 2	- Metodologia e ajuste de previsões climáticas do BESM relativamente ao ERA5 com filtros adaptativos no domínio da frequência.	25
Figura 3	- Mapa Mundial indexado por coordenadas de latitude (vertical) e longitude (horizontal).	26
Figura 4	- Erro médio anual (BESM-ERA5) da Temperatura da Superfície do Mar (em graus Celsius) em 2009.	27
Figura 5	- Estrutura típica do conjunto de dados de previsões climáticas. Retirado de [Pydata.org].	28
Figura 6	- Diferença entre os valores da variável climática TSM do BESM e da reanálise ERA5. A superfície do mar é a região onde aplicamos a metodologia proposta, nos continentes, em cinza, a TSM não está definida.	28
Figura 7	- Blocos de 8×8 , 16×16 e 32×32 pixels em torno do pixel alvo (quadrado vermelho). O exemplo considera o pixel em (lat,lon)=(-33.25,-175.0) e as cores representam a Temperatura da Superfície do Mar (TSM) do BESM e do ERA5.	29
Figura 8	- Condições de contorno para um modelo de domínio regional, adaptado de (CALDERÓN, 2015).	30
Figura 9	- DCT-2D de blocos de 8×8 <i>pixels</i> da variável Temperatura da Superfície do Mar (TSM) ao longo da sequência de dias correspondentes a um ano. As cores dos pixels representam o valor da DCT-2D da TSM e o pixel com moldura vermelha é o pixel com a componente DC, o valor médio do bloco.	36
Figura 10	- Diagrama em blocos de um sistema adaptativo – configuração geral de um filtro adaptativo.	39
Figura 11	- Exemplo da seleção de células para aplicação de filtros adaptativos à TSM do BESM visando melhorar sua acurácia relativamente ao ERA5.	46
Figura 12	- Exemplo de um bloco para aplicação da DCT-2D e da célula (lat,lon)=(-10.0,-25.75).	46
Figura 13	- Exemplo de blocos $M=8$, 16, e 32, comparativamente à célula alvo de BESM e ERA5. As cores representam a Temperatura da Superfície do Mar (TSM) e o vermelho é a célula alvo.	48
Figura 14	- Exemplo do cálculo de MSE da TSM ao longo de tempo para BESM e ERA5.	49

Figura 15	- Exemplo de apresentação e avaliação de desempenho usando os índices agregados, indicando as proporções de melhoria ou piora para cada medida de erro (MSE, LAE e PLAE), M é o tamanho da janela empregada para a DCT-2D, quando aplicável.	51
Figura 16	- Evolução dos coeficientes do filtro adaptativo para 2 repetições anuais da série (acima) e para 10 repetições (abaixo). Cada conjunto de gráficos apresenta os $L = 4$ coeficientes do filtro considerando blocos com $M=8, 16$ e 32	53
Figura 17	- Ajuste do BESM em relação ao ERA5 com filtros adaptativos de comprimento $L= 4, 6, 8, 12, 16$ e 32 para o ajuste do BESM na frequência com $M = 32$	54
Figura 18	- Percentuais de redução e aumento do erro com relação ao erro original entre BESM e os dados observacionais de temperatura da superfície do mar, considerando os pixels aleatoriamente escolhidos, resultantes de correção ANUAL, variando-se o comprimento do filtro $L= 4, 12$ e 32 com $M = 8$	55
Figura 19	- Resultados de correção da TSM do BESM para $\lambda= 0.98, 0.99$ e 0.999 por meio de filtragem adaptativa RLS no domínio da frequência, para $M = 32$ e $L = 4$	56
Figura 20	- Comportamento da norma dos coeficientes do filtro adaptativo e sua autocorrelação para $L = 4$ e $M = 8, 16, 32$ para 2 repetições anuais da serie (acima) e para 10 repetições (abaixo) aplicado em BESM . As linhas vermelhas tracejadas indicam os inícios da primeira e segunda repetições anuais. Pixel na posição $lat : -23.5 lon : 75.75$	58
Figura 21	- Comportamento do processo de treinamento do algoritmo RLS aplicados a série temporais de coeficientes DCT-2D, referente a região quadrada 8×8 centrada em $(lat = 29.0; lon = -34.25)$. Para o ajuste anual é considera o primeiro coeficiente da DCT-2D no domínio das frequências.	60
Figura 22	- Comportamento do processo de treinamento do algoritmo RLS aplicados a séries temporais de coeficientes DCT-2D, referente a regiões quadradas de dimensões $8 \times 8, 16 \times 16$ e 32×32 centradas em $(lat = -29.0; lon = -34.25)$	61
Figura 23	- Pixels aleatoriamente selecionados para o empregos dos métodos de correção de viés baseados em RLS e CMV. O pixel vermelho é examinado em detalhe.	62

Figura 24 - Percentuais de redução e aumento do erro com relação ao erro original entre BESM e os dados observacionais de temperatura da superfície do mar, considerando os pixels aleatoriamente escolhidos, resultantes de correção anual.	63
Figura 25 - Percentuais de redução e aumento do erro com relação ao erro original entre BESM e os dados observacionais de temperatura da superfície do mar, considerando os pixels aleatoriamente escolhidos, resultantes de correção de viés por estações.	64
Figura 26 - Correção anual e por estações da temperatura da superfície do mar por meio dos métodos RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Atlântico.	66
Figura 27 - Correção anual e por estações da temperatura da superfície do mar por meio do filtro adaptativo RLS no domínio da frequência (RLS_freq), variando-se o tamanho da região. Célula no Oceano Atlântico.	67
Figura 28 - Erros resultantes da correção da TSM produzida pelo BESM em relação ao dado observacional, corrigidas por ano e por estações. Célula no Oceano Atlântico.	67
Figura 29 - Erros resultantes da correção da TSM produzida pelo BESM em relação ao dado observacional, corrigidas por ano e por estações por meio do método RLS_freq para $M = \{8, 6, 32\}$	68
Figura 30 - Análise a cada três meses da correção por estações da TSM por meio de filtros RLS e CMV nos domínios do tempo e da frequência.	69
Figura 31 - Análise a cada três meses da correção por estações da TSM por meio de filtro RLS no domínio da frequência para regiões de tamanho $M = \{8, 16, 32\}$	70
Figura 32 - Correção anual e por estações da temperatura da superfície do mar por meio dos métodos RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Índico.	72
Figura 33 - Correção anual e por estações da temperatura da superfície do mar por meio do filtro adaptativo RLS no domínio da frequência (RLS_freq), variando-se o tamanho da região. Célula no Oceano Índico.	73
Figura 34 - Erros resultantes da correção da TSM produzida pelo BESM em relação ao dado observacional, corrigidas por ano e por estações. Célula no Oceano Índico.	73
Figura 35 - Análise a cada três meses da correção por estações da TSM por meio de filtros RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Índico.	74

Figura 36 - Análise a cada três meses da correção por estações da TSM por meio de filtro RLS no domínio da frequência para regiões de tamanho $M = \{8, 16, 32\}$	75
Figura 37 - Correção anual e por estações da temperatura da superfície do mar por meio dos métodos RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Pacífico.	77
Figura 38 - Correção anual e por estações da temperatura da superfície do mar por meio do filtro adaptativo RLS no domínio da frequência (RLS_freq), variando-se o tamanho da região. Célula no Oceano Pacífico.	78
Figura 39 - Erros resultantes da correção da TSM produzida pelo BESM em relação ao dado observacional, corrigidas por ano e por estações. Célula no Oceano Pacífico.	78
Figura 40 - Análise a cada três meses da correção por estações da TSM por meio de filtros RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Pacífico.	79
Figura 41 - Análise a cada três meses da correção por estações da TSM por meio de filtro RLS no domínio da frequência para regiões de tamanho $M = \{8, 16, 32\}$	80
Figura 42 - Correção anual e por estações da temperatura da superfície do mar por meio dos métodos RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Glacial Antártico.	82
Figura 43 - Correção anual e por estações da temperatura da superfície do mar por meio do filtro adaptativo RLS no domínio da frequência (RLS_freq), variando-se o tamanho da região. Célula no Oceano Glacial Antártico.	83
Figura 44 - Erros resultantes da correção da TSM produzida pelo BESM em relação ao dado observacional, corrigidas por ano e por estações. Célula no Oceano Glacial Antártico.	83
Figura 45 - Análise a cada três meses da correção por estações da TSM por meio de filtros RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Glacial Antártico.	84
Figura 46 - Análise a cada três meses da correção por estações da TSM por meio de filtro RLS no domínio da frequência para regiões de tamanho $M = \{8, 16, 32\}$	85
Figura 47 - Correção anual e por estações da temperatura da superfície do mar por meio dos métodos RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Glacial Ártico.	87
Figura 48 - Correção anual e por estações da temperatura da superfície do mar por meio do filtro adaptativo RLS no domínio da frequência (RLS_freq), variando-se o tamanho da região. Célula no Oceano Glacial Ártico.	88

Figura 49 - Erros resultantes da correção da TSM produzida pelo BESM em relação ao dado observacional, corrigidas por ano e por estações. Célula no Oceano Glacial Ártico.	88
Figura 50 - Análise a cada três meses da correção por estações da TSM por meio de filtros RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Glacial Ártico.	89
Figura 51 - Análise a cada três meses da correção por estações da TSM por meio de filtro RLS no domínio da frequência para regiões de tamanho $M = \{8, 16, 32\}$	90

LISTA DE TABELAS

Tabela I	- Características principais do ERA5.	34
----------	---	----

LISTA DE ABREVIATURAS E SIGLAS

INPE	Instituto Nacional de Pesquisas Espaciais
ANA	Agência Nacional de Águas
CPTEC	Centro de Previsão de Tempo e Estudos Climáticos
MCGAs	Modelos de Circulação Geral da Atmosfera
NCEP	<i>National Center for Environmental Prediction</i>
DCT - 2D	<i>Two-Dimensional Discrete Cosine Transform</i>
IDCT - 2D	<i>Two-Dimensional Inverse Discrete Cosine Transform</i>
RLS	<i>Recursive Least Squares</i>
NetCDF	<i>Network Common Data Format</i>
MSE	<i>Mean Square Error</i>
BESM	<i>Brazilian Earth System Model</i>
ECMWF	<i>European Centre for Medium-Range Weather Forecasts</i>
TSM	Temperatura da superfície do mar
UNFCCC	<i>United Nations Framework Convention on Climate Change</i>
RCM	Modelos Climáticos Regionais
GFS	<i>Global Forecast System</i>
FIR	<i>Finite Impulse Response</i>
RAMS	<i>Regional Atmospheric Modeling System</i>
C3S	<i>Copernicus Climate Change Service</i>
IFS	Sistema de Previsão Integrado

LISTA DE SÍMBOLOS

σ	Coordenada vertical sigma
p	Pressão da atmosfera
p_S	Pressão na superfície terrestre
p_T	Pressão no topo do domínio do modelo
η	Coordenada vertical eta
z_s	Altitude da superfície em metros
L	Ordem do filtro adaptativo
λ	Fator de esquecimento
δ	Constante de regularização
I	Matriz identidade
$\mathbf{R}_D(n)$	Matriz de correlação determinística do sinal de entrada
$\mathbf{p}_D(n)$	Vetor de correlação cruzada determinístico da sinal de entrada e desejado

SUMÁRIO

	INTRODUÇÃO	18
1	MATERIAIS E MÉTODOS	24
1.1	Metodologia Proposta	24
1.2	Formatos do modelo BESM e ERA5	25
1.2.1	<u>Leitura dos Dados de Modelos Climáticos</u>	26
1.2.2	<u>Região e Variável Prognóstica de Trabalho</u>	27
1.3	Sobre os Modelos Atmosféricos Considerados	29
1.3.1	<u>O BESM</u>	30
1.3.2	<u>O Modelo de Circulação Geral Atmosférico (MCGA)</u>	32
1.3.2.1	Amostragem Espacial e Temporal	32
1.3.2.2	Variáveis Prognósticas	32
1.3.3	<u>Conjunto de Reanálises Climáticas ERA5</u>	33
1.3.3.1	Amostragem Espacial e Temporal	33
1.3.3.2	Metodologia para a Reanálise	35
1.4	Técnicas de Processamento de Sinais	35
1.4.1	<u>Análise Espectral Bidimensional</u>	36
1.4.2	<u>Aplicação da Filtragem Adaptativa Para a Redução dos Erros de Previsão</u>	37
1.4.2.1	Filtragem adaptativa	38
1.4.2.2	Critério de Otimização	39
1.4.2.3	A Solução de Wiener	40
1.4.2.4	O Algoritmo <i>Recursive Least Squares</i> (RLS)	40
1.5	Correção de Média e Variância	42
2	METODOLOGIA EXPERIMENTAL	44
2.1	Variável Prognóstica, Células e Períodos de Trabalho	44
2.1.1	<u>Variável Prognóstica</u>	44
2.1.2	<u>Períodos de Trabalho e Amostragem Temporal</u>	44
2.1.2.1	Ajuste Anual	44
2.1.2.2	Ajuste Estacional ou Sazonal	45
2.1.3	<u>Células Alvo</u>	45
2.1.4	<u>Tamanhos de Blocos para a Análise Espectral</u>	45
2.1.4.1	Definição do Bloco para Análise Espectral em Função da Célula Alvo	47
2.2	Critérios de Avaliação de Desempenho	49
2.2.1	<u>Largest Absolute Error (LAE) e Percentual Largest Absolute Error (PLAE)</u>	50
2.2.2	<u>Mean Squared Error (MSE)</u>	50
2.2.3	<u>Índices Agregados</u>	50

2.3	Configuração do Filtro Adaptativo	52
2.3.1	<u>Configuração do Algoritmo RLS</u>	52
2.3.1.1	Ordem do Filtro (L)	52
2.3.1.2	Fator de Esquecimento	54
2.3.2	<u>O Filtro Ótimo ou Vetor de Coeficientes Ótimos</u>	56
2.3.3	<u>Autocorrelação do Módulo do Filtro</u>	56
3	RESULTADOS	59
3.1	Treinamento do Filtro Adaptativo RLS	59
3.2	Análise por Índices Agregados	61
3.3	Correção da TSM do BESM em Células Específicas em Diferen-	
	tes Oceanos	62
3.3.1	<u>Célula no Oceano Atlântico</u>	65
3.3.2	<u>Célula no Oceano Índico</u>	71
3.3.3	<u>Célula no Oceano Pacífico</u>	76
3.3.4	<u>Célula no Oceano Glacial Antártico</u>	81
3.3.5	<u>Célula no Oceano Glacial Ártico</u>	86
	CONCLUSÃO	91
	REFERÊNCIAS	92

INTRODUÇÃO

Atualmente, o desenvolvimento da tecnologia possibilita diferentes métodos de previsão de fenômenos climáticos tanto para a alta atmosfera como para a previsão de eventos meteorológicos apesar da complexidade computacional. Dada uma região da Terra, entende-se por tempo as condições meteorológicas instantâneas nela vigentes, e por clima o estado médio ao longo do tempo da atmosfera (YNOUE et al., 2017). Hoje, é possível produzir previsões de tempo e clima bastante confiáveis para alguns dias, semanas ou meses (CLARKE; DIAS, 2002).

Dentre os diversos avanços científicos está a capacidade de simular complexos sistemas físicos utilizando modelos numéricos aproximados da realidade. A quantidade de modelos para prever o tempo e o clima aumentou muito nas últimas décadas. Há diferentes modelos para os processos que determinam a evolução do estado da atmosfera. A capacidade de processamento dos computadores permitiu implementar alguns deles. Na década de 1980 e 1990, os modelos matemáticos foram sendo aperfeiçoados a partir de suas implementações viabilizadas pela evolução da tecnologia, em especial, na área de computação de alto desempenho. Vários centros meteorológicos mundiais passaram a produzir previsões numéricas do tempo com alguns dias de antecedência.

O que se chama de um modelo de circulação geral (MCG) visa modelar (e prever) as condições do tempo em diferentes locais do globo e o estado da atmosfera, com alto grau de confiança, para vários dias. Tais modelos podem ser utilizados para prever o clima para a próxima estação ou para simular características mais gerais do clima futuro na escala de décadas ou séculos (BRANKOVIĆ; PALMER; FERRANTI, 1994). O Brasil começou a utilizar um Modelo de Circulação Geral Atmosférico (MCGA) rotineiramente para realizar previsões numéricas de tempo a partir do final de 1994 com a implantação do Centro de Previsão de Tempo e Estudos Climáticos (CPTEC) do Instituto Nacional de Pesquisas Espaciais (INPE) (BITENCOURT et al., 2019).

As variáveis prognósticas de um MCG (como o MCGA) são: o logaritmo da pressão na superfície, a componente vertical do rotacional do vento, a divergência do vento horizontal, a temperatura virtual, umidade específica, dentre outras. Geralmente, as equações do modelo estão definidas no domínio espectral e sua resolução é efetuada por métodos numéricos (BONATTI, 1996). Os modelos de previsão climática decompõem a atmosfera em volumes, simulando numericamente as interações dos processos, e avaliando o estado da atmosfera de forma iterada. Os modelos partem de condições iniciais e, dependendo da abrangência do modelo, podem depender ainda das condições de contorno, que descrevem o estado da atmosfera nas fronteiras/bordas do domínio.

Um Modelo Climático Regional (ou RCM, do inglês *Regional Climate Model*) parte das condições de contorno, os dados nas bordas da região/domínio do modelo, para indu-

zir a previsão no interior do domínio do modelo e assim descrever eventos climatológicos de pequena escala (LAPRISE et al., 2008). Eles têm sido cada vez mais usados para projeções/previsões de mudanças climáticas (OLMOS et al., 2014). No entanto, os desempenhos (da previsão) de tais modelos podem variar conforme a estação do ano, a região, o horário inicial de integração, a resolução do modelo e a quantidade de observações assimiladas pela análise do modelo (SHUKLA, 1998).

Sob outra perspectiva, uma reanálise climática fornece uma descrição numérica do clima recente, produzida pela combinação de modelos com observações. Uma reanálise climática contém estimativas de parâmetros atmosféricos, como a temperatura do ar, a pressão e o vento em diferentes altitudes e outros parâmetros de superfície, como precipitação, umidade do solo, altura das ondas do mar e temperatura da superfície do mar (TSM). Os prognósticos são produzidos para todo o globo e cobrem um longo período de tempo (ECMWF, 2022).

Nesta dissertação, examina-se a discrepância entre o modelo climático *Brazilian Earth System Model* BESM e a reanálise *ECMWF Reanalysis v5* ERA5. Investiga-se se é possível reduzi-la utilizando filtros adaptativos. Projetam-se filtros RLS (*Recursive Least Squares*) para corrigir os erros do BESM em relação ao ERA5.

O modelo atmosférico regional BESM é utilizado pelo Centro de Previsão de Tempo e Estudos Climáticos / Instituto Nacional de Pesquisas Espaciais (INPE/CPETEC) (NOBRE et al., 2013). O BESM é um conjunto de programas computacionais que acopla os componentes de superfície continental, oceano, atmosfera e química globais, com o principal objetivo de gerar cenários de mudanças climáticas para a perspectiva brasileira (VEIGA, 2013). O BESM apresenta como variáveis prognósticas a temperatura do ar, as componentes zonal e meridional do vento, a umidade específica, a pressão na superfície, a energia cinética turbulenta, a umidade e a temperatura do solo, e a água líquida ou gelo das nuvens (CHOU, 1996). O BESM utiliza como condição inicial os dados de análise do modelo *Global Forecast System* (GFS) fornecidos pelo *National Centers for Environmental Prediction* (NCEP) e, como condições de contorno, as previsões do modelo de circulação geral atmosférico do CPTEC (FREITAS, 2015; XUE; SHUKLA, 1998).

O BESM foi empregado para estudos de mudanças climáticas seguindo o protocolo do projeto *Coupled Model Intercomparison Project Phase 5* (ou CMIP5) (TAYLOR; STOUFFER; MEEHL, 2012) e mostrou bom desempenho em reproduzir temperatura, precipitação e padrões atmosféricos comparáveis com os demais modelos de sua categoria (VEIGA et al., 2019; CAPISTRANO et al., 2020). No BESM, a variável Temperatura da Superfície do Mar (TSM) é crucial para a previsibilidade climática (SHUKLA, 1998).

O *European Centre for Medium-Range Weather Forecasts* (ECMWF) produz a reanálise *ECMWF Reanalysis v5* (ERA5). O ERA5 é a última ferramenta de reanálise climática financiada pela União Europeia e representa a quinta geração de reanálise atmosférica do clima global desenvolvida pelo ECMWF (ECMWF, 2022).

Dadas as diferenças inerentes aos procedimentos usados para gerar os dados do BESM e do ERA5, esperam-se diferenças (discrepâncias ou erros) entre os prognósticos do primeiro e os dados observacionais fornecidos pelo segundo, conforme ilustra a Figura 1, já que compara-se o estado observado (ERA5) da atmosfera com o estado previsto (BESM). Tais erros têm diferentes origens, dentre as quais podemos destacar: i) erros nas aproximações físicas do modelo numérico e ii) erros nas condições iniciais, que são amplificados com o tempo de integração do modelo.

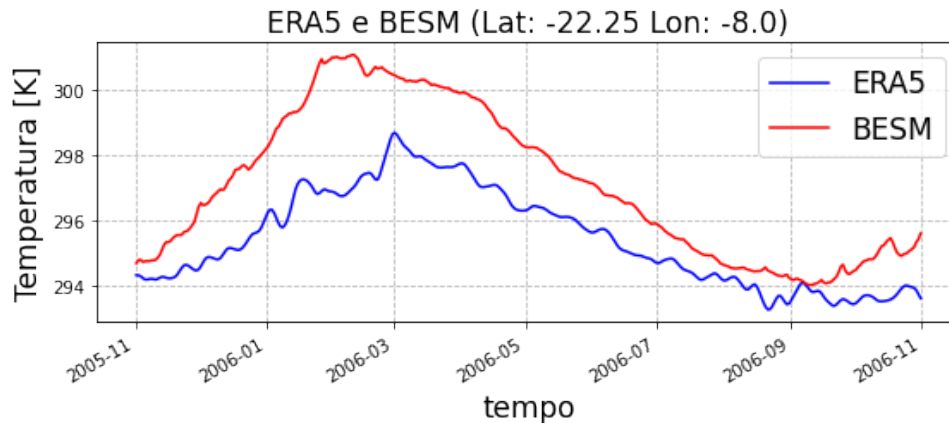


Figura 1 - Séries temporais anuais da Temperatura da Superfície do Mar (TSM) do modelo regional BESM e modelo ERA5 às 12:00h UTC. Latitude = -22.25° (graus a Norte) e longitude = -8.0° (graus a Leste).

De acordo com (LAPRISE et al., 2008), as condições iniciais impactam as integrações de curto a médio prazo, para alguns dias ou semanas. Os erros de previsão geralmente aumentam ao longo do tempo e deterioram a qualidade da previsão de clima. Para previsões sazonais (horizonte de alguns meses) os modelos numéricos apresentam erros nas propriedades estatísticas (CATALDI et al., 2007).

A Figura 1 ilustra a Temperatura da Superfície do Mar (TSM) do BESM e do ERA5, no período de novembro 2005 a agosto 2006 (365 dias) às 00:00 e 12:00h UTC (*Coordinated Universal Time*) na posição de coordenadas latitude = -22.25° (graus a Norte) e longitude = -8.0° (graus a Leste) ou, simplesmente, (lat,lon) = (-22.25,-8.0). Observa-se que há uma discrepância entre as duas séries temporais.

Este trabalho tem por objetivo testar se a filtragem adaptativa pode ser empregada para reduzir o erro da previsão da TSM produzida pelo modelo BESM relativamente ao ERA5, através do pós-processamento da previsão BESM por um filtro projetado a partir de erros passados. Isto é, investiga-se a aplicação de filtros adaptativos para melhorar as previsões climáticas do modelo BESM. Filtram-se as séries temporais de variáveis climáticas provenientes do modelo BESM, visando minimizar seu erro relativamente ao ERA5; pretende-se assim obter uma série mais próxima daquela que seria fornecida futuramente pela reanálise ERA5.

Observa-se que os modelos climáticos são volumétricos e discretos no tempo. Suas células contêm múltiplas variáveis prognósticas temporalmente amostradas e estão espalhadas não só na superfície terrestre mas também na altitude. No entanto, o valor de uma variável prognóstica em uma dada altitude em todo o globo pode ser apresentada como uma imagem cujos pixels estão indexados pelas coordenadas geográficas de latitude e longitude (PINTO et al., 2018). Além disso, os valores de um pixel de uma previsão climática produz uma sequência de números $\{x(n)\}$, tal que n indexa o instante de tempo (OPPENHEIM, 2010; FREITAS et al., 2015).

Um filtro adaptativo é um filtro digital $h(n)$ (que produz na saída $y(n) = x(n) * h(n)$, sendo $x(n)$ a entrada) cujos coeficientes, os valores da sequência $h(n)$, sua resposta ao impulso, são obtidos adaptativamente (DINIZ et al., 1997). A filtragem adaptativa emprega algoritmos iterativos, para o ajuste automático dos coeficientes do filtro, tornando possível o projeto de filtros satisfatórios em ambientes não-estacionários. Os coeficientes do filtro adaptativo são ajustados em função de um sinal de referência; e, conseqüentemente, a aplicação do filtro adaptativo ao sinal de entrada resulta em um sinal de saída filtrado, mais próximo do sinal de referência (DINIZ; WERNER, 2003; FREITAS et al., 2015; PINTO et al., 2015).

Na abordagem investigada nesta dissertação, $\{y(n)\}$ é a sequência observada provida pelo ERA5 e $\{x(n)\}$ é a produzida pelo BESM para um mesmo pixel, i.e., numa posição (lat,lon). A partir disso, aplica-se um algoritmo de filtragem adaptativa para obter o filtro h que, quando recebe o sinal gerado a partir do BESM, produz um sinal na saída que é mais aderente ao provido pelo ERA5 do que o de entrada (BESM). Empregamos o algoritmo de mínimos quadrados recursivo (*Recursive Least Squares* - RLS) (DINIZ; WERNER, 2003; HAYKIN, 2008) para adaptar filtros de resposta finita ao impulso (*Finite Impulse Response* - FIR). Tais filtros são aplicados para reduzir os erros de previsão do BESM relativamente aos dados observacionais do ERA5.

Pixels vizinhos e próximos em uma imagem em geral apresentam características similares, isto é, são correlacionados. Tal correlação pode ser explorada para compactar os valores dos pixels de uma região da imagem através de transformadas (RAO; YIP, 2014; KHAYAM, 2003). Uma das técnicas de transformação de imagens mais utilizadas é a Transformada do Cosseno Discreta (DCT - *Discrete Cosine Transform*). Ela é empregada para descorrelacionar os pixels de imagens digitais naturais e suaves. Cada coeficiente da DCT corresponde à intensidade do padrão espacial correspondente, e está associado a um padrão espacial distinto. A frequência do padrão cresce conforme o índice do coeficiente (a posição) aumenta. Em uma imagem natural, a maior parte da energia está concentrada nos coeficientes de frequências menores (RAO; YIP, 2014; KHAYAM, 2003; PINTO et al., 2018).

Células vizinhas no globo também apresentam algum nível de correlação (PINTO et al., 2015). Conseqüentemente, neste trabalho, investiga-se se a filtragem adaptativa

aplicada a coeficientes da DCT (isto é, no domínio DCT, que corresponde a uma análise frequencial) é uma estratégia razoável para reduzir o erro de previsão do BESM relativamente ao ERA5. Isto é, investiga-se se é possível reduzir erros da previsão a partir de filtros adaptativos no domínio espectral da TSM.

Objetivo

Este trabalho avalia se filtros FIR adaptativos podem aperfeiçoar a previsão climática do modelo BESM. Para tal fim, projetam-se filtros visando minimizar o erro. Com isso, pretende-se concluir se os filtros adaptativos podem ser uma técnica efetiva de pós-processamento para reduzir erros de previsão numérica climática. Utiliza-se a filtragem adaptativa FIR através do algoritmo RLS. Além da correção através de filtros adaptativos no domínio espacial, também investiga-se a aplicação do paradigma usando as componentes espectrais (coeficientes DCT em duas dimensões, DCT-2D).

Organização

Esta dissertação apresenta três capítulos além desta Introdução e da Conclusão.

O Capítulo 1 descreve os materiais e métodos, inicia-se com a metodologia proposta, introduzem-se conceitos sobre modelos atmosféricos e reanálises climáticas que conformam o conjunto de dados empregados, e segue-se apresentando as ferramentas empregadas na metodologia proposta, a análise espectral via DCT-2D e o ajuste de suas componentes através de filtragem adaptativa via algoritmo RLS, e a reconstrução dos blocos de pixels da previsão corrigida via IDCT-2D (*Two-Dimensional Inverse Discrete Cosine Transform*). Discutem-se ainda métodos alternativos aplicando a filtragem no domínio dos pixels e o ajuste via Correção de Média e Variância (CMV).

O Capítulo 2 aborda a metodologia experimental empregada para avaliar desvios de previsão climática e assim medir o desempenho da metodologia proposta, ajustar seus parâmetros, e compará-la a outras abordagens. Discutem-se configurações para a metodologia proposta, as configurações do filtro adaptativo, critérios de avaliação do erro da previsão.

O Capítulo 3 relata os resultados experimentais. Apresentam-se resultados de treinamento dos filtros adaptativos e resultados de ajuste das previsões, apresentando tanto resultados para uma célula considerando os diferentes métodos, como resultados agregados para muitas células. O objetivo é avaliar o comportamento do erro de previsões climáticas ajustadas através de filtragem adaptativa.

Por fim, a Conclusão discute o uso da metodologia proposta – a escolha de parâme-

tros e configurações das abordagens propostas, os resultados obtidos, e caminhos possíveis trabalhos futuros.

1 MATERIAIS E MÉTODOS

Este capítulo apresenta a metodologia proposta. Descrevem-se os conjuntos de dados e os diferentes modelos atmosféricos utilizados que os produzem. Descrevem-se também as técnicas de processamento de sinais usadas – análise espectral via DCT-2D e filtragem adaptativa. Como método comparativo ao proposto, além do treinamento de filtros diretamente no domínio temporal, apresentamos ainda o ajuste de previsões via Correção de Média e Variância (CMV).

1.1 Metodologia Proposta

A Figura 2 ilustra a metodologia empregada para o processamento de uma variável prognóstica de um modelo climático (BESM) via filtragem adaptativa no domínio da frequência visando aumentar sua concordância com a mesma variável gerada por um segundo modelo (ERA5). Obtêm-se os coeficientes da DCT-2D (JAIN, 1989; AHMED; NATARAJAN; RAO, 1974; ZHU; LIU; DAI, 2009) para o bloco de pixels na região em torno da célula/pixel a ser corrigida. Aplica-se a transformação em cada instante de tempo/amostra da variável prognóstica - no caso, as séries empregadas providas pelo BESM e do ERA5 apresentam valores diários. Um filtro FIR para a sequência temporal de coeficientes é obtido por meio da filtragem adaptativa RLS (HAYKIN, 2008; DINIZ; WERNER, 2003) num ano e aplicado à sequência temporal de coeficientes DCT do BESM no ano seguinte para diminuir sua diferença em relação ao dado observacional do ERA5. É importante mencionar de que, de um ano para o próximo, espera-se que a variabilidade climática seja mínima, isto é, supõe-se alguma estacionariedade (ou melhor ainda, correlação) do clima em períodos similares de anos subsequentes.

No “Ano de Treinamento”, o filtro é obtido a partir dos erros da previsão BESM em relação ao ERA5. Usam-se os dados de um ano anterior ao “Ano de Ajuste”. Se usamos especificamente os dados do ano imediatamente anterior, para ajustar a sequência para o ano de 2009, usamos a do ano de 2008 para treinar/adaptar o filtro. No “Ano de Ajuste”, aplicam-se os filtros treinados/adaptados às sequências de coeficientes DCT-2D para corrigir erros de previsão no domínio da frequência. Na sequência, aplica-se a DCT Inversa Bidimensional (IDCT-2D) aos coeficientes corrigidos, retornando ao domínio original da variável prognóstica.

Para avaliar a estratégia proposta, comparam-se os erros entre as sequências temporais da variável prognóstica TSM para o BESM antes e após o emprego da metodologia de correção com os valores fornecidos pela reanálise ERA5. Compara-se ainda a estratégia proposta com a que emprega o filtro adaptativo RLS diretamente na sequência de

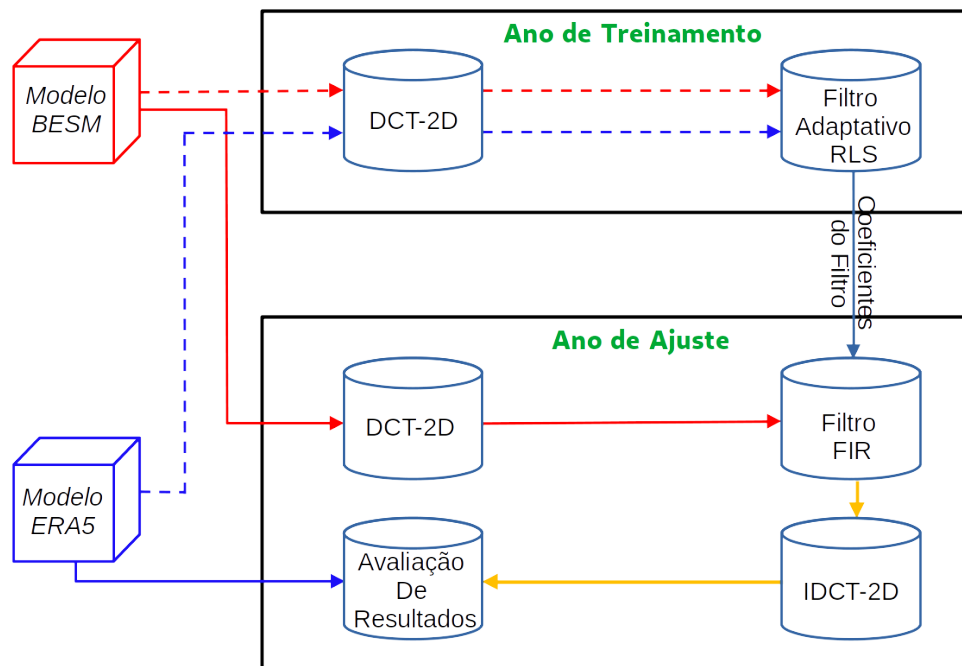


Figura 2 - Metodologia e ajuste de previsões climáticas do BESM relativamente ao ERA5 com filtros adaptativos no domínio da frequência.

valores temporais da variável prognóstica (FREITAS, 2015; PINTO et al., 2018) e com a Correção de Média e Variância (CMV) da sequência da variável prognóstica, em ambos os casos, da sequência BESM relativamente à ERA5. (TCHEOU et al., 2021) apresenta um estudo inicial sobre a aplicação de filtros adaptativos, utilizando o algoritmo RLS (*Recursive Least Square*), para a redução de desvios em previsões climáticas.

1.2 Formatos do modelo BESM e ERA5

A Figura 3 apresenta o mapa global com as referências de latitude e longitude, respectivamente. O modelo BESM e as reanálises ERA5 empregadas nesta dissertação, empregam a mesma amostragem (discretização) espaço-temporal. As coordenadas utilizadas advêm da **Projeção Cilíndrica Equidistante**, também chamada de equirretangular ou projeção de Plate Carré (SERRANO, 2012). Ela é uma projeção equidistante - os paralelos estão igualmente espaçados ao longo dos meridianos.

Tanto os dados provenientes do modelo ERA5 quanto os do BESM estão em formato de “Ponto de Grade” que representam os dados de forma discreta em pontos fixos de uma malha ou grade. Os parâmetros atmosféricos estão disponíveis numa grade regular amostrada com resolução angular de $0,25^\circ \times 0,25^\circ$ em latitude $\times$ longitude (HERSBACH, 2016). Obviamente, a Temperatura Superficial do Mar (TSM) não é definida sobre os continentes.

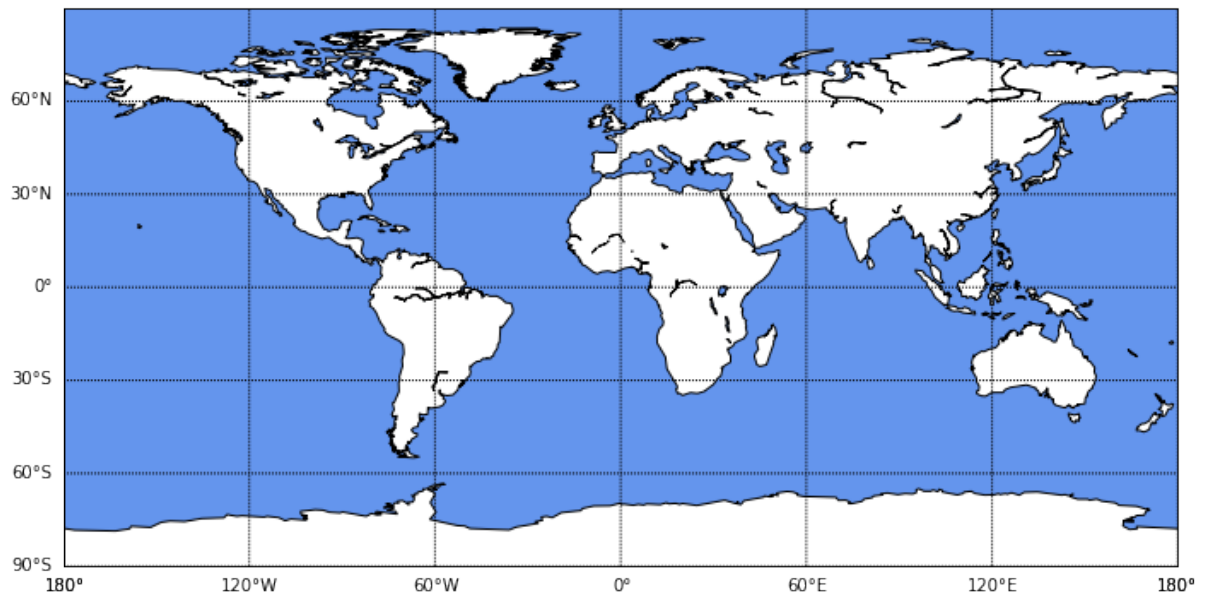


Figura 3 - Mapa Mundial indexado por coordenadas de latitude (vertical) e longitude (horizontal).

1.2.1 Leitura dos Dados de Modelos Climáticos

Os dados de BESM e ERA5 estão disponíveis no formato binário NetCDF. O *Network Common Data Form* (NetCDF) fornece formatos e ferramentas para processar os dados de modelos climáticos. O NetCDF é um conjunto de formatos de dados matriciais e volumétricos que compõem um padrão para compartilhar dados científicos. Há bibliotecas de *software* de suporte ao formato NetCDF para Python, IDL, MATLAB, R, Ruby e Perl (Network Common Data Format (NetCDF), 2022). Neste trabalho, foi utilizada a linguagem Python.

Os dados de BESM e ERA5 estão disponíveis em NetCDF organizadas por meio de quatro dimensões (longitude, latitude, níveis verticais e tempo). Usando a biblioteca *Numpy* e outras, pode-se processar os dados necessários. Por exemplo, a Figura 4 ilustra o erro médio anual (subtraindo-se o ERA5 do BESM) da TSM em 2009.

Os dados BESM e ERA5 são representados no ambiente de Python como matrizes multidimensionais (conhecidas como N-dimensionais, ND, ou “tensores”) (RASCHKA; MIRJALILI, 2019). A Figura 5 exemplifica a estrutura de uma previsão do tempo (XARRAY DOCUMENTATION, 2020). Utiliza-se a estrutura de dados *XarrayDataset* que é uma matriz multidimensional cujas dimensões são rotuladas ou nomeadas. Seus objetos são do tipo *DataArray* que adicionam metadados, como os nomes das dimensões, as coordenadas e os atributos das variáveis, ou também o equivalente multidimensional do *Xarray* de um *DataFrame* (HOFF, 2011). Em outros termos, tem-se algo similar a um dicionário de matrizes rotuladas, objetos do tipo *DataArray*, com dimensões alinhadas.

Python se tornou uma das linguagens de programação mais populares para análise

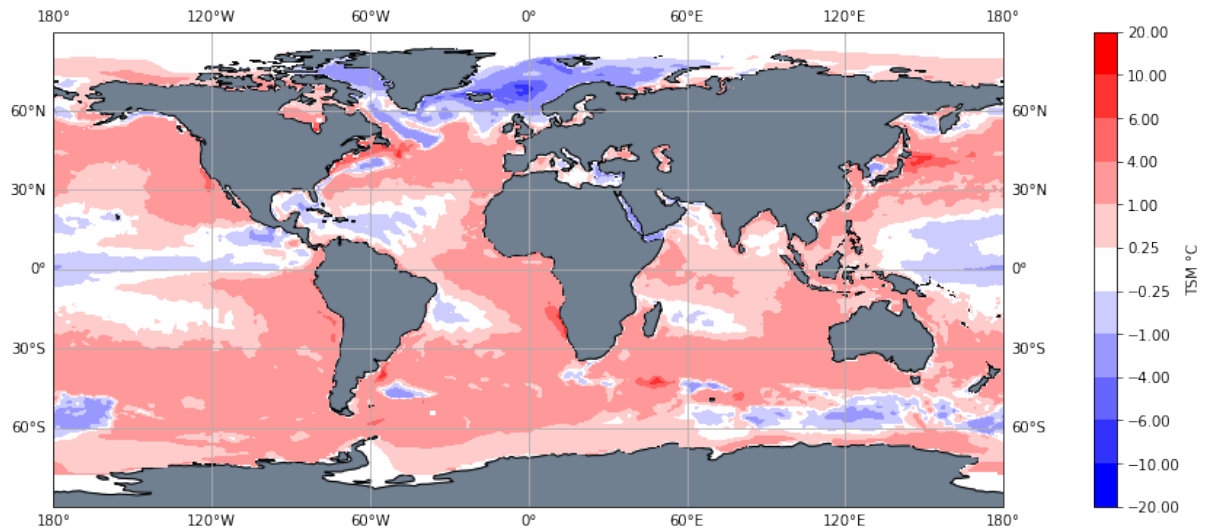


Figura 4 - Erro médio anual (BESM-ERA5) da Temperatura da Superfície do Mar (em graus Celsius) em 2009.

de dados devido à sua simplicidade, versatilidade e à ampla gama de bibliotecas disponíveis. Nesta dissertação, dentre outras, emprega-se bibliotecas do Python para análise de dados, incluindo Numpy, Pandas, Matplotlib, Dataset:

- Numpy: é uma biblioteca essencial para processamento numérico em Python. Ela oferece estruturas de dados para trabalhar com matrizes multidimensionais, além de uma ampla variedade de funções para realizar operações matemáticas.
- Pandas: é uma biblioteca Python projetada para facilitar a análise e manipulação de dados. Ela oferece estruturas de dados, como o DataFrame, que permitem trabalhar com dados tabulados de forma intuitiva.
- Matplotlib: é uma biblioteca de visualização em Python que permite criar gráficos e visualizações dos dados e assim comunicar os resultados.
- Dataset: é uma biblioteca que simplifica o gerenciamento e a exploração de conjuntos de dados. Ela fornece uma interface para baixar, carregar e manipular uma ampla variedade de conjuntos de dados.

1.2.2 Região e Variável Prognóstica de Trabalho

A Figura 6 apresenta o erro da variável TSM (Temperatura da Superfície do Mar) no dia 31/10/2007 entre BESM e ERA5. A variável TSM é quantificada em graus Celsius (°C) ou Kelvin (K). A proposta deste trabalho é aplicada apenas na superfície do mar já que a TSM não está definida nas zonas continentais, em cinza na Figura 6.

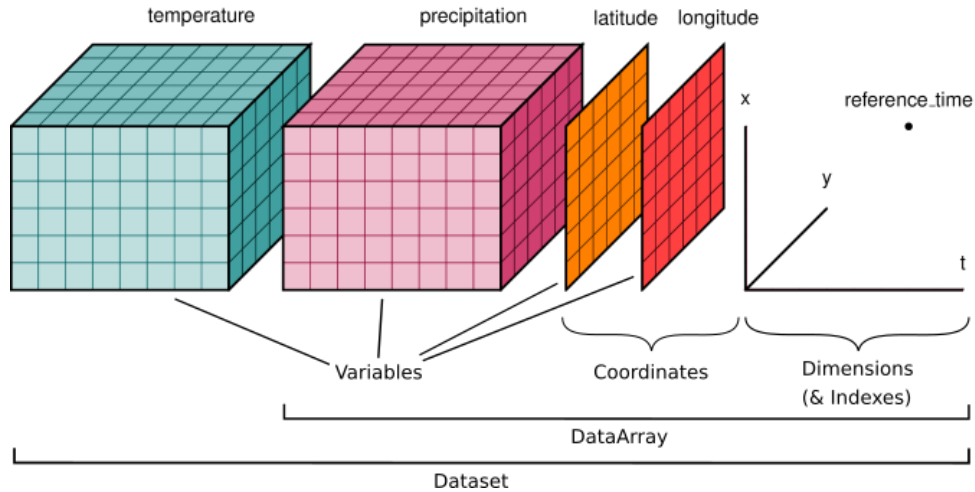


Figura 5 - Estrutura típica do conjunto de dados de previsões climáticas. Retirado de [Pydata.org].

As células (ou pixels, na verdade suas coordenadas) são sorteadas aleatoriamente para aplicar a metodologia proposta, que envolve o projeto de filtros calculados adaptativamente, para correção dos erros da previsão da TSM pelo BESM relativamente ao ERA5, sendo que o BESM e ERA5 têm a mesma amostragem espacial e são indexadas por coordenadas geográficas equivalentes. Regiões em torno das células sorteadas são empregadas para a análise espectral através da DCT-2D. A Figura 7 exemplifica isso, o quadrado vermelho contém a célula alvo. Em torno de cada pixel, se estabelecem vizinhanças (blocos) de dimensões de: 8×8 , 16×16 ou 32×32 pixels para a aplicação da DCT-2D, os coeficientes DCT-2D do BESM são ajustados relativamente aos do ERA5, através de filtros adaptativos.

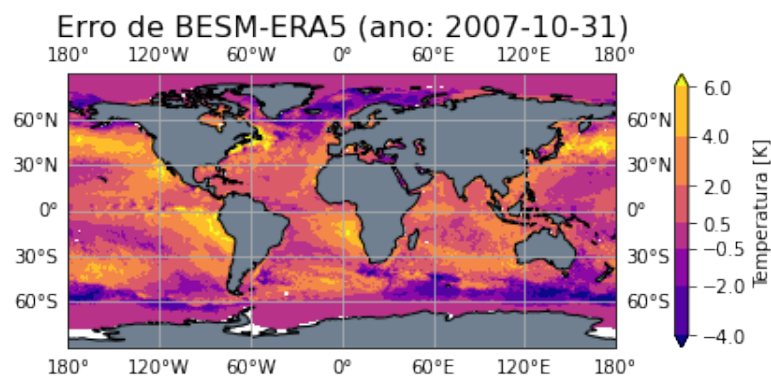


Figura 6 - Diferença entre os valores da variável climática TSM do BESM e da reanálise ERA5. A superfície do mar é a região onde aplicamos a metodologia proposta, nos continentes, em cinza, a TSM não está definida.

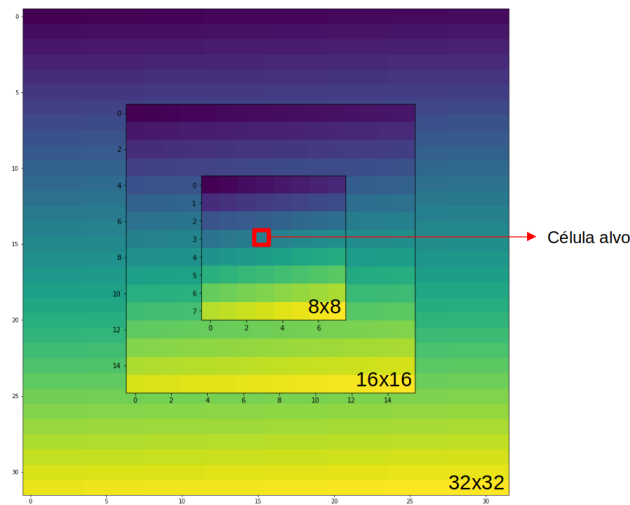


Figura 7 - Blocos de 8×8 , 16×16 e 32×32 pixels em torno do pixel alvo (quadrado vermelho). O exemplo considera o pixel em (lat,lon)=(-33.25,-175.0) e as cores representam a Temperatura da Superfície do Mar (TSM) do BESM e do ERA5.

1.3 Sobre os Modelos Atmosféricos Considerados

Um modelo atmosférico visa reproduzir as condições da atmosfera e sua evolução para utilização de forma diagnóstica ou prognóstica (CAVALCANTI et al., 2002; MARTINS et al., 2015). O termo diagnóstica se refere à avaliação de características climáticas, enquanto prognóstica se refere à previsão de comportamentos futuros do tempo e do clima.

Existem dois modelos atmosféricos muito utilizados na prática: modelos atmosféricos de circulação global (MCGs) e modelos atmosféricos regionais (MCRs). Os MCGs são modelos matemáticos complexos que representam o comportamento da atmosfera em todo o globo terrestre. MCGs são utilizados para prever o clima a longo prazo. Por outro lado, os MCRs são modelos semelhantes que se concentram em uma região geográfica específica e se adaptam as condições climáticas e topográficas únicas da região particular. MCRs são úteis para prever o clima a curto e médio prazo.

Existem duas grandes categorias de modelos atmosféricos, os globais ou os regionais. Um modelo climático global (MCG) reproduz o comportamento atmosférico em toda a superfície terrestre, isto é, em todo o globo terrestre. Já um modelo climático regional (MCR) analisa o comportamento de variáveis atmosféricas dentro de uma determinada região da Terra. Em qualquer caso, o modelo parte de um estado inicial e tenta modelar a evolução das variáveis atmosféricas com o tempo.

As “Condições Iniciais” de um modelo climático são o ponto de partida para a evolução do estado da atmosfera (SUN; GRAHAM, 2001; LAPRISE et al., 2008). Elas representam a atmosfera observada nos pontos da grade do modelo. A partir das Condições Iniciais, o modelo avança no tempo e computa as variáveis atmosféricas.

Um MCR precisa do estado nas bordas da região que modela, além das condições iniciais. As “Condições de Contorno” são as informações climáticas nas bordas da região modelada como ilustra a Figura 8. Pontualmente, as condições iniciais (especificadas no início da simulação) e de contorno (especificadas no entorno da região modelada durante todo o período de simulação) servem para inicializar e balizar os modelos numéricos regionais (BENOIT; MARCOS; BECQ, 1996).

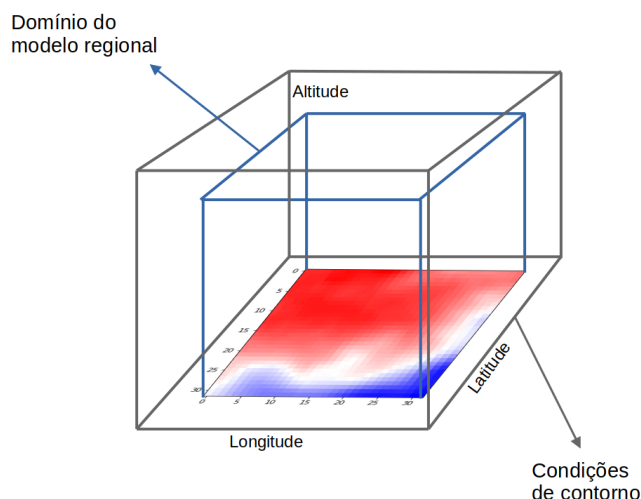


Figura 8 - Condições de contorno para um modelo de domínio regional, adaptado de (CALDERÓN, 2015).

Diferentes modelos climáticos ou atmosféricos podem empregar variáveis de estado (que representam o estado do tempo numa célula), condições iniciais e de contorno distintas e produzir diferentes variáveis prognósticas. Geralmente, um MCG apresenta baixa resolução espacial visando reduzir o custo computacional (ROECKNER et al., 1996; LAPRISE et al., 2008; CHOU; NOBRE, 2015).

1.3.1 O BESM

O Modelo Brasileiro do Sistema Terrestre (BESM – *Brazilian Earth System Model*, na sigla em inglês) é um conjunto de programas computacionais para realizar simulações de cenários climáticos globais futuros (MIGUEL; MONTEIRO, 2015). O BESM é produto dos precursores do INPE/CPTEC em modelagem acoplada oceano-atmosfera, o modelo é capaz de realizar desde previsões estendidas de tempo e clima sazonal, com durações de semanas a meses, até cenários de mudanças climáticas globais com escala temporal de séculos a milênios (MIGUEL; MAHONY; MONTEIRO, 2019).

O BESM é um modelo de ultima geração para simular o sistema terrestre que

inclui a atmosfera, os oceanos a criosfera e biosfera (SAMPAIO; LEITE; SILVA, 2014). O BESM também é utilizado para estudos de diferentes impactos sobre emissões de gases. O modelo é composto por uma série de equações que simulam o comportamento de cada componente do sistema terrestre. O BESM foi desenvolvido para ser adaptado às condições específicas do Brasil. Inclui características regionais da floresta amazônica e do Pantanal. O desenvolvimento do modelo começou em 2005, em parceria do INPE/CPTEC com outras instituições brasileiras, como a Universidade de São Paulo (USP), a Universidade Federal do Rio de Janeiro (UFRJ) e a Universidade Federal do Rio Grande do Sul (UFRGS). Desde a sua criação, o BESM continua sendo atualizado e aprimorado com novas informações e dados, a fim de melhorar a acurácia de suas projeções climáticas.

Neste trabalho, utiliza-se o Modelo Climático Global (MCG) BESM, que fornece as condições iniciais e de contorno para modelos regionais de maior resolução específicas para cada um dos componentes do sistema terrestre. Para o componente da atmosfera, o modelo utiliza as condições iniciais de temperatura, pressão e umidade do ar, e condições de contorno de temperatura, pressão e vento (VERGASTA, 2018). Para o componente de oceanos, o BESM utiliza as condições iniciais de temperatura, salinidade e correntes oceânicas na superfície e em camadas mais profundas do oceano.

As melhorias nas parametrizações físicas do BESM nos modelos componentes (atmosfera, oceano, superfície), entre elas: parametrizações de fluxos atmosféricos da atmosfera para o oceano, implementação de microfísica de nuvens, parametrização de fluxos turbulentos na vertical de calor e momento no modelo oceânico geraram sucesso durante 2013-2015 (progresso do BESM). Dentre as quais, suas principais características e aplicabilidades são:

- Capacidade de gerar duas versões de cenários de mudanças climáticas globais em escala de tempo decenal;
- As versões do BESM utilizam o modelo de processos de superfície continental SSIB (*Simplified Simple Biosphere Model*);
- Utilizam-se formulações para o cálculo das funções de similaridade para os campos de vento, temperatura do ar e umidade específica a 10 m da superfície. A formulação utiliza os valores de vento, umidade e temperatura na superfície na primeira camada vertical sigma (σ) do modelo atmosférico.

Algumas linhas de desenvolvimento do modelo atmosférico são: a) melhorar o ciclo diurno da precipitação tropical, b) melhorar a representação da chuva não convectiva tropical, c) melhorar a representação da precipitação na Amazônia, Zona de Convergência do Atlântico Sul e Sul do Brasil, e d) reduzir excesso de precipitação sobre o Oceano Índico e na ITCZ (*the InterTropical Convergence Zone*). O avanço nessas linhas ajuda a melhorar a previsão de tempo, clima sazonal e intra-sazonal e experimentos de mudanças climáticas usando o BESM (INPE, n.d.).

O *Cloud Resolving Model* (CRM) é usado como uma ferramenta para testes de comparação de parâmetros do modelo com observações, para posteriormente serem usados nos esquemas de parametrização do MCGA.

1.3.2 O Modelo de Circulação Geral Atmosférico (MCGA)

O MCGA do INPE/CPETEC, base do Modelo Brasileiro do Sistema Climático Global (MBSCG), tem sido desenvolvido pelo INPE/CPTEC a partir da sua versão inicial do COLA (*The Center for Ocean-Land-Atmosphere Studies*) (CHOU; NOBRE, 2015). O MCGA tem sua origem no MRF (*Medium Range Forecasting Model*) usado para previsão de tempo de médio prazo pelo NCEP (*National Centers for Environmental Prediction*). Esse modelo combina o código espectral global desenvolvido pelo NCEP com as parametrizações físicas de escala sub-grade desenvolvidas no *Geophysical Fluid Dynamics Laboratory* (GFDL) do *National Oceanic and Atmospheric Administration* NOAA na Universidade de Princeton, EUA (BONATTI, 1996). O MCGA é utilizado pelo INPE/CPTEC para as previsões climáticas.

1.3.2.1 Amostragem Espacial e Temporal

O MCGA tem resolução espacial de $200\text{ km} \times 200\text{ km}$. Ele produz previsões em quatro momentos do dia: 00:00, 06:00, 12:00 e 18:00h UTC. O período de previsão pode ser de até seis meses (FREITAS, 2015; BITENCOURT; SATYAMURTY, 2016).

1.3.2.2 Variáveis Prognósticas

As variáveis prognósticas do MCGA são o logaritmo da pressão à superfície, a vorticidade (componente vertical do rotacional do vento), a divergência do vento horizontal, a temperatura virtual e a umidade específica. Além dessas, existem outras parametrizações incluídas no modelo, tais como a temperatura da superfície, do interior do solo e do interior do dossel (extrato superior de florestas composto pelas árvores mais altas que recebem a intensidade de luz solar) (BONATTI, 1996; ROCHA, 1992). O modelo emprega equações que são escritas no formato espectral resolvidas numericamente (BONATTI, 1996).

1.3.3 Conjunto de Reanálises Climáticas ERA5

O ERA5 é um conjunto de alta confiabilidade que incorpora um registro detalhado da atmosfera global e visa ajudar no mapeamento de mudanças climáticas (HERSBACH et al., 2020). As reanálises climáticas do ERA5 combinam observações passadas com as de modelos para gerar séries temporais consistentes de múltiplas variáveis climáticas.

Dentro do *Copernicus Climate Change Service* (C3S), o *European Centre for Medium-Range Weather Forecasts* (ECMWF) produz a reanálise *ECMWF Reanalysis v5* (ERA5). O ERA5 é a última ferramenta de reanálise climática financiada pela União Europeia e representa a quinta geração de reanálise atmosférica do clima global desenvolvida pelo ECMWF (ECMWF, 2022). É baseado no Sistema de Previsão Integrado (IFS) Cy41r2 que se beneficiou de uma década de desenvolvimentos na modelagem física e dinâmica dos modelos e na assimilação de dados observacionais (HERSBACH et al., 2020). Além disso, o ERA5 incorpora um registro detalhado da atmosfera global, da superfície terrestre e de ondas oceânicas a partir de 1950. Ele fornece uma nova descrição numérica do clima recente e contém estimativas de parâmetros atmosféricos como o estado da superfície terrestre e do mar, em grades regulares de latitude e longitude com resolução de $0,25^\circ \times 0,25^\circ$, juntamente com estimativas de incertezas dos mesmos. O ERA5 cobre a Terra em uma grade de 30 km e divide a atmosfera usando 137 níveis desde a superfície até uma altura de 80 km. O ERA5 fornece dados de hora em hora (frequência de saída de três horas para o conjunto) para muitos parâmetros atmosféricos tanto da superfície terrestre como do estado do mar. Ela é a mais recente reanálise climática produzida pelo ECMWF do C3S (ECMWF, 2023).

A Tabela I destaca algumas das principais características do ERA5. Existem muitas diferenças adicionais, em particular, relativas às mudanças no cálculo de parâmetros atmosféricos individuais.

1.3.3.1 Amostragem Espacial e Temporal

Os parâmetros atmosféricos do ERA5 estão disponíveis em grades regulares de latitude e longitude com resolução de $0,25^\circ \times 0,25^\circ$ (cerca de 30 km) em 37 níveis de pressão (HERSBACH, 2016). O ERA5 gera dados com data inicial em 1950 até o presente. As atualizações estão disponíveis por trimestres (LI et al., 2020).

Além do conjunto referido como “reanálise”, a realização de alta resolução por hora, a cada 31 km, o ERA5 traz um conjunto de dez membros de resolução reduzida (referido como “conjunto” ou “EDA”). Este conjunto (dez membros) é necessário para o procedimento de assimilação de dados 4D-Var (“4D” refere-se à quarta dimensão e “Var” deriva da palavra *variational*). 4D-Var é uma técnica que utiliza a assimilação variacional

Tabela I - Características principais do ERA5.

Período coberto	1950 - presente
Período de produção	Janeiro de 2016 – final de 2017, depois continuou quase em tempo real
Versão do modelo	Ciclo IFS 41r2
Sistema de assimilação	Ciclo IFS 41r2 4D-Var
Resolução espacial	62 km para o <i>Ensemble of Data Assimilations</i> (EDA), 137 níveis para 0,01 hPa e 31 km globalmente.
Frequência de saída (resolução temporal)	Campos de previsão horária, 3 horas para o <i>Ensemble of Data Assimilations</i> (EDA), até 18 horas, com frequência reduzida até 10 dias
Estimativas de incerteza	De um <i>Ensemble of Data Assimilations</i> (EDA) de 10 membros com resolução de 63 km
Entrada do modelo	Adequado para o clima (por exemplo, gases de efeito estufa CMIP5, erupções vulcânicas, TSM e cobertura de gelo marinho)
Esquema de viés variacional	Também dados de ozônio, aeronaves e pressão de superfície
Dados de satélite	RTTOV-11, <i>all-sky</i> para vários componentes
Novos parâmetros	O ERA5 apresenta mais de 240 parâmetros na superfície
Inovações adicionais	Evolução a longo prazo de CO ₂ em RTTOV, correção de pressão de célula SSU, correção de polarização aprimorada para radiossondas, perturbações EDA para cobertura de gelo marinho

Extraída de (ECMWF, 2022).

para melhorar a acurácia e a consistência dos modelos numéricos ao combinar observações e simulações em um contexto de quatro dimensões (espaço e tempo). Geralmente, os dados estão disponíveis em uma frequência sub-diária e mensal e consistem em análises e previsões curtas (num período de 18 horas), inicializadas duas vezes ao dia a partir de análises às 06 e 18 UTC. No entanto, há outros disponibilizados na análise EDA, como, por exemplo, taxas/fluxos e acumulações médias das variáveis fornecidas (HERSBACH et al., 2020).

A frequência de saída da análise climática ERA5 é horária, o que significa que são produzidos dados para cada hora do dia. No entanto, a resolução dos dados difere entre variáveis e zonas geográficas consideradas. Por exemplo, para variáveis atmosféricas como temperatura, umidade e pressão, os dados são produzidos em uma resolução de 0,25 graus (aproximadamente 31 km) em latitude e longitude para todo o mundo. Para outras variáveis, como a radiação solar e a velocidade do vento, a resolução é ligeiramente diferente e depende da localização geográfica (ECMWF, 2023).

É importante ter em conta que os dados são agregados usando valores médios durante períodos de tempo mais longos, como dias, meses e anos. Assim, a frequência de saída dos dados agregados dependerá do período de tempo que estiver sendo considerado.

1.3.3.2 Metodologia para a Reanálise

O ERA5 é produzido usando assimilação de dados 4D-Var e previsões de modelos usando o Ciclo CY41R2 do Sistema Integrado de Previsão (IFS – *Integrated Forecast System*) do ECMWF com 137 níveis híbridos sigma/pressão na vertical e o nível superior em 0,01 hPa (PEREIRA, 2022; DIAS; PALMEIRA; LIBONATI, 2021). Os dados atmosféricos estão disponíveis nesses níveis e são interpolados para os 37 níveis de pressão, 16 de temperatura potencial e um nível de vorticidade potencial por FULL-POS no IFS (PERDOMO; GONZÁLEZ; CAMINO, 2016; HERBACH et al., 2020).

1.4 Técnicas de Processamento de Sinais

O processamento de sinais visa manipular e transformar sinais de forma a obter informações úteis. Nesta seção, discutem-se as técnicas empregadas nesta dissertação.

1.4.1 Análise Espectral Bidimensional

A DCT-2D de uma matriz é a sua representação numa base de matrizes elementares compostas por cossenos bidimensionais separáveis de diferentes frequências. Os coeficientes da DCT-2D medem quanto há das matrizes elementares na matriz original. Cada padrão ou imagem base bidimensional corresponde ao produto de um cosseno na direção horizontal por outro na vertical cujas frequências dependem da posição (índices na matriz) do coeficiente correspondente.

Quando aplicada a um bloco da variável prognóstica, as amplitudes da DCT-2D fazem uma “análise espectral” (no domínio frequência) do bloco. Se o bloco é suave, a maioria da informação significativa estará concentrada em poucos coeficientes correspondentes a imagens-base de baixas frequências.

Se $x_v(l, c)$ é a variável prognóstica na célula de posição (l, c) , os elementos $x(i, j)$ do bloco $\mathbf{x}$ de $M \times M$ pixels em torno da célula (l, c) é dado por

$$x(i, j) = x_v \left(l + i - \left(\left\lceil \frac{M}{2} \right\rceil - 1 \right), c + j - \left(\left\lceil \frac{M}{2} \right\rceil - 1 \right) \right), \text{ onde } i, j = 0 \dots M - 1. \quad (1.4.1)$$

Obtém-se o bloco $\mathbf{x} = [x(i, j)]$, $i, j = 0 \dots M - 1$, para cada dia da série de TSM, tal qual exemplificado na Figura 9.

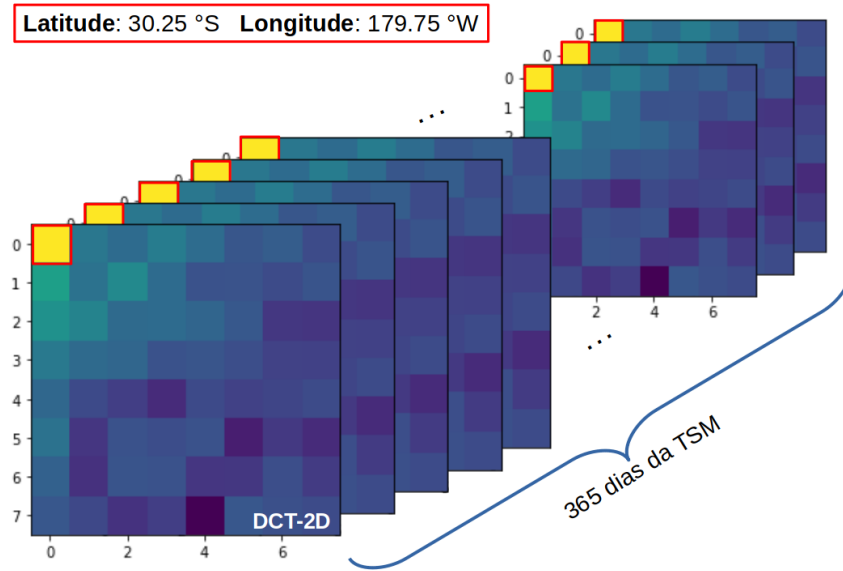


Figura 9 - DCT-2D de blocos de 8×8 pixels da variável Temperatura da Superfície do Mar (TSM) ao longo da sequência de dias correspondentes a um ano. As cores dos pixels representam o valor da DCT-2D da TSM e o pixel com moldura vermelha é o pixel com a componente DC, o valor médio do bloco.

Os elementos da DCT-2D $\mathbf{X}$ de uma matriz $\mathbf{x}$ de dimensões $M \times M$ são (AHMED;

NATARAJAN; RAO, 1974; JAIN, 1989; BELLIFEMINE et al., 1992)

$$X(p, q) = c_p c_q \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} x(i, j) \cos \left[\frac{\pi(2i+1)p}{2M} \right] \cos \left[\frac{\pi(2j+1)q}{2M} \right], \quad 0 \leq p, q \leq M-1, \quad (1.4.2)$$

$$\text{com } c_p = \begin{cases} \frac{1}{\sqrt{M}} & p = 0 \\ \frac{2}{\sqrt{M}} & 1 \leq p \leq M-1 \end{cases} \text{ e } c_q = \begin{cases} \frac{1}{\sqrt{M}} & q = 0 \\ \frac{2}{\sqrt{M}} & 1 \leq q \leq M-1 \end{cases}. \quad (1.4.3)$$

$X(p, q)$ é o coeficiente de índices (p, q) da DCT-2D $\mathbf{X}$ da matriz $\mathbf{x}$. Veja que p e q determinam as frequências dos elementos de $\mathbf{X}$ pois definem os argumentos dos cossenos na Equação (1.4.2). As constantes de normalização c_p e c_q dependem do tamanho da matriz.

A IDCT-2D faz o processo inverso da DCT. Isto é, ela retorna a matriz $\mathbf{x}$ a partir de $\mathbf{X}$. Isto é,

$$x(i, j) = c_i c_j \sum_{p=0}^{M-1} \sum_{q=0}^{M-1} X(p, q) \cos \left[\frac{\pi(2i+1)p}{2M} \right] \cos \left[\frac{\pi(2j+1)q}{2M} \right], \quad 0 \leq i, j \leq M-1, \quad (1.4.4)$$

$$\text{com } c_i = \begin{cases} \frac{1}{\sqrt{M}} & i = 0 \\ \frac{2}{\sqrt{M}} & 1 \leq i \leq M-1 \end{cases} \text{ e } c_j = \begin{cases} \frac{1}{\sqrt{M}} & j = 0 \\ \frac{2}{\sqrt{M}} & 1 \leq j \leq M-1 \end{cases}. \quad (1.4.5)$$

1.4.2 Aplicação da Filtragem Adaptativa Para a Redução dos Erros de Previsão

Para reduzir os erros de previsão do BESM relativamente ao ERA5, aplica-se um filtro FIR aos coeficientes DCT do bloco em torno do pixel alvo; os coeficientes do filtro são ajustados por meio da filtragem adaptativa.

Um sistema ou filtro discreto é constituído pela relação da sequência de entrada $\{x(n)\}$, $n = 0, 1, \dots$ com a sequência de saída $\{y(n)\}$, $n = 0, 1, \dots$, de forma que $\{y(n)\} = H\{x(n)\}$. O operador $H\{.\}$ é o filtro de tempo discreto. Se o sistema é linear e invariante no tempo, temos que $y(n) = x(n) * h(n) = \sum_k h(n-k)x(k)$, e “*” é a notação matemática para a operação de convolução. Daqui em diante vamos representar a sequência $\{y(n)\}$ simplesmente por $y(n)$.

Um filtro digital processa um sinal de entrada para produzir o de saída (DINIZ; SILVA; NETTO, 2010; OPPENHEIM, 2010). Se a saída do sistema depende somente da entrada presente e de um número finito de entradas passadas, então diz-se que o filtro tem Resposta ao Impulso Finita (FIR – *Finite Impulse Response*). Como os valores de saída passados não afetam os valores de saída presentes, este filtro também é dito não-recursivo.

Um filtro linear e invariante no tempo pode ser caracterizado em função de sua resposta $h(n)$ ao impulso $\delta(n)$ ($\delta(n) = 1$ se $n=0$ e $\delta(n) = 0$ se $n \neq 0$). Um filtro FIR é tal que $h(n) = 0$ se $n < 0$ e $n > L$, onde L é o comprimento da resposta ao impulso do filtro, ou simplesmente do filtro – a quantidade de coeficientes que caracterizam o filtro. Os filtros FIR são do tipo *Bounded Input-Bounded Output* (BIBO) estáveis; ou seja, se o sinal de entrada é limitado em amplitude então o sinal de saída também o é.

1.4.2.1 Filtragem adaptativa

A filtragem adaptativa tem sido empregada com êxito em uma grande gama de aplicações, tais como realce de sinais, cancelamento de eco, identificação de sistemas, detecção de novidade, dentre outras (DINIZ; WERNER, 2003; DINIZ; SILVA; NETTO, 2010; HAYKIN, 2008).

Um filtro adaptativo tem os coeficientes alterados em função de um critério de desempenho definido. Isto é, um filtro adaptativo é um sistema variante no tempo, dotado de um mecanismo de ajuste dos coeficientes a partir de um critério específico. Quanto à forma de resolução do problema de filtragem adaptativa, diferentes pontos de vista são empregados. Destacam-se as abordagens baseadas em aproximação estocástica e na consequente minimização de funções-custo determinísticas (DINIZ; WERNER, 2003; HAYKIN, 2008; DINIZ; SILVA; NETTO, 2010).

A configuração geral de um filtro adaptativo poder ser vista na Figura 10. No instante n , a amostra do sinal $x(n)$ é submetida ao filtro adaptativo que produz $y(n)$. Esta é comparada à amostra correspondente do sinal desejado $d(n)$, produzindo o sinal de erro $e(n)$. O sinal de erro é usado para adaptar os coeficientes do filtro na iteração n . Este procedimento é iterado visando reduzir a magnitude do erro $e(n)$ (DINIZ; SILVA; NETTO, 2010).

Para a atualização do filtro, usa-se um vetor $\mathbf{x}(n) = [x(n) \ x(n-1) \ \dots \ x(n-L+1)]^T$ com tantas amostras quanto o comprimento L do filtro. Na Figura 10, $\mathbf{w}(n)$ é o vetor de coeficientes do filtro adaptativo. Por último, $y(n)$ é o sinal de saída do filtro e $e(n) = d(n) - y(n)$ é o sinal de erro – a diferença entre o sinal desejado e a saída do filtro. O algoritmo adaptativo o utiliza para atualizar os coeficientes do filtro.

Neste trabalho, os filtros adaptativos são utilizados num processamento em duas etapas. Primeiramente, os filtros são treinados em regime transitório (atualização de coeficientes) para um determinado ano. E, depois, no regime permanente, os filtros adaptados são aplicados para corrigir a previsão para o ano posterior. No regime transitório, o filtro adaptativo é inicializado com um vetor aleatório e é adaptado/atualizado iterativamente. Após esse período, os coeficientes obtidos são usados para aplicar o filtro em regime permanente.

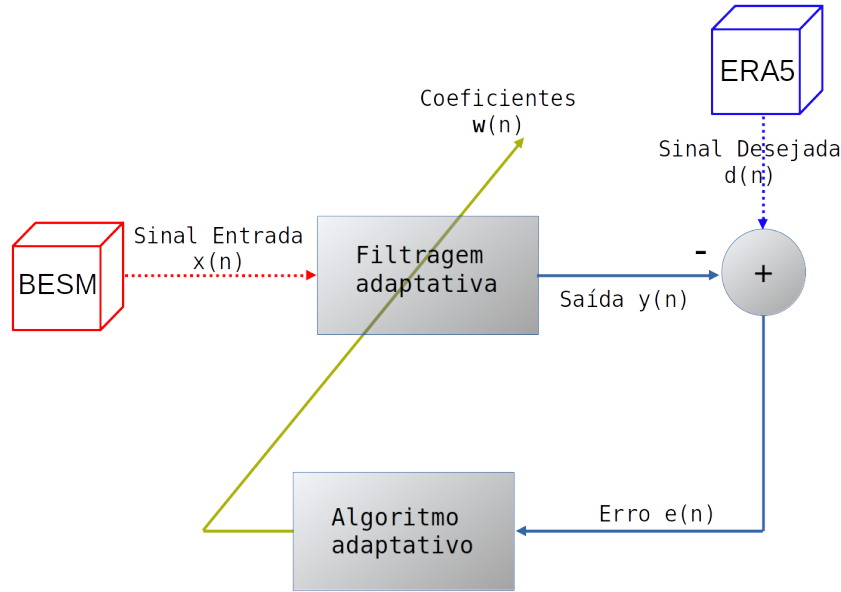


Figura 10 - Diagrama em blocos de um sistema adaptativo – configuração geral de um filtro adaptativo.

Para os propósitos deste trabalho, a sequência $x(n)$ é a sequência temporal dos valores do coeficiente DCT-2D na posição (i, j) , $0 \leq i, j \leq M$, do bloco em torno do pixel a processar do BESM. Por sua vez, $d(n)$ é o equivalente obtido a partir do ERA5. A saída do filtro $y(n)$ de comprimento L é definida por

$$y(n) = \sum_{i=0}^{L-1} \omega_i(n) x(n-i) = \mathbf{w}^T(n) \tilde{\mathbf{x}}(n), \quad (1.4.6)$$

na qual, $L-1$ é a ordem do filtro adaptativo, $\tilde{\mathbf{x}}(n) = [x(n) \ x(n-1) \ x(n-2) \ \dots \ x(n-L+1)]^T$ e $\mathbf{w}(n) = [w_0(n) \ w_1(n) \ \dots \ w_{L-1}(n)]^T$. O símbolo T indica a transposição da matriz.

1.4.2.2 Critério de Otimização

O algoritmo adaptativo é responsável pelo ajuste dos coeficientes do filtro. A filtragem adaptativa FIR geralmente adota o erro quadrático médio (MSE – *Mean Square Error*) como o critério a minimizar (DINIZ, 2013). O MSE é dado por

$$\epsilon(n) = E[e^2(n)] = E[(d(n) - y(n))^2] = E[d^2(n) - 2d(n)y(n) + y^2(n)], \quad (1.4.7)$$

onde $E[\cdot]$ é o operador valor esperado. Se aplicarmos a equação (1.4.6) na Equação (1.4.7), obtém-se

$$\epsilon(n) = E[d^2(n)] - 2E[(d(n)\mathbf{w}^T(n)\tilde{\mathbf{x}}(n))] + E[\mathbf{w}^T(n)\tilde{\mathbf{x}}(n)\tilde{\mathbf{x}}^T(n)\mathbf{w}(n)]. \quad (1.4.8)$$

1.4.2.3 A Solução de Wiener

Assumamos que $\mathbf{w}(n)$ é aproximadamente $\mathbf{w}(n-1)$, isto é, que o vetor de coeficientes é fixo e igual a $\mathbf{w}$, temos que

$$\epsilon(n) = E[d^2(n)] - 2\mathbf{w}^T E[d(n)\mathbf{x}(n)] + \mathbf{w}^T E[\mathbf{x}(n)\mathbf{x}^T(n)]\mathbf{w}. \quad (1.4.9)$$

Sejam $\mathbf{p} = E[d(n)\mathbf{x}(n)]$ e $\mathbf{R} = E[\mathbf{x}(n)\mathbf{x}^T(n)]$, temos que

$$\epsilon(n) = E[d^2(n)] - 2\mathbf{w}^T \mathbf{p} + \mathbf{w}^T \mathbf{R} \mathbf{w}, \quad (1.4.10)$$

na qual, $\mathbf{p}$ e $\mathbf{R}$ são, respectivamente, o vetor de correlação cruzado entre o sinal de referência com o sinal de entrada e a matriz de correlação do sinal de entrada.

O ótimo do vetor $\mathbf{w}$ pode ser obtido a partir do mínimo do MSE em função de $\mathbf{w}$. Para isso, é preciso calcular o vetor gradiente do MSE em relação a $\mathbf{w}$,

$$\mathbf{g} = \frac{\partial \epsilon(n)}{\partial \mathbf{w}} = -2\mathbf{p} + 2\mathbf{R}\mathbf{w}. \quad (1.4.11)$$

O objetivo é a minimização de $\epsilon(n)$, que ocorre para o vetor de coeficientes ótimos. Igualando a zero a Equação (1.4.11) encontra-se que o vetor ótimo

$$\mathbf{w}_0 = \mathbf{R}^{-1}\mathbf{p}. \quad (1.4.12)$$

Esta solução, obtida através de uma perspectiva estatística, é denominada de solução de Wiener. Contudo, estimar $\mathbf{R}$ e $\mathbf{p}$ pode ser inviável (DINIZ; SILVA; NETTO, 2010). A Equação (1.4.12) envolve o cálculo da inversa da matriz $\mathbf{R}$ e, do ponto de vista analítico, na maioria das vezes, tal solução é inviável e, de grande complexidade computacional. Assim, os algoritmos adaptativos visam diminuir as dificuldades da obtenção da solução de Wiener.

1.4.2.4 O Algoritmo *Recursive Least Squares* (RLS)

O algoritmo de Mínimos Quadrados têm como objetivo minimizar a soma dos quadrados da diferença entre o sinal desejado e a saída do filtro do modelo. Quando amostras dos sinais de entrada são recebidas sequencialmente, a solução para o problema dos mínimos quadrados pode ser calculada recursivamente, resultando no algoritmo de Mínimo Quadrados Recursivo (RLS – *Recursive Least Squares*) (GOODWIN, 1977; HAYKIN, 2008).

Os algoritmos adaptativos visam contornar o problema da solução de Wiener. A

escolha do algoritmo pode influenciar a filtragem adaptativa em aspectos como a complexidade computacional e a velocidade de convergência. A velocidade de convergência pode ser descrita como a quantidade de iterações necessárias para que o algoritmo adaptativo, na presença de sinais estacionários, convirja para a solução ótima de Wiener, com o intuito de minimizar o erro médio quadrático. A complexidade computacional é avaliada a partir da quantidade de operações (adições, multiplicações, divisões e outros) requeridas para efetuar uma iteração do algoritmo.

Os algoritmos que usam critérios de mínimos quadrados têm como vantagem a velocidade de convergência, porém, demandam maior complexidade computacional. O algoritmo RLS (*Recursive Least Squares*) busca convergência rápida mesmo quando a dispersão de autovalores da matriz de correlação do sinal de entrada é grande (SILVA et al., 2013). No RLS, os coeficientes do filtro são atualizados iterativamente (DINIZ; WERNER, 2003; HAYKIN, 2008; AMARAL, 2018) com bom desempenho para sistemas sujeitos a rápidas variações na amplitude do sinal. Essas vantagens trazem maior complexidade computacional e alguns problemas de estabilidade (ARDALAN, 1986; CIOFFI, 1987).

Seja a função objetivo determinística

$$\xi(n) = \sum_{m=0}^n \lambda^{(n-m)} \epsilon^2(m), \quad (1.4.13)$$

na qual, $\epsilon(m)$ é o erro de saída *a posteriori* no instante m , e λ é o fator de ponderação exponencial, que deve ser escolhido tal que $0 < \lambda < 1$. Este parâmetro também é chamado de fator de esquecimento e determina o peso de amostras passadas. Vemos que

$$\xi(n) = \sum_{m=0}^n \lambda^{(n-m)} \epsilon^2(m) = \sum_{m=0}^n \lambda^{(n-m)} [d(m) - \mathbf{w}(n)\tilde{\mathbf{x}}(m)]^2, \quad (1.4.14)$$

Com a pretensão de minimizar o erro *a posteriori*, deriva-se $\xi(n)$ em relação a $\mathbf{w}(n)$, obtendo-se

$$\frac{\partial \xi(n)}{\partial \mathbf{w}(n)} = -2 \sum_{m=0}^n \lambda^{(n-m)} \tilde{\mathbf{x}}(m) [d(m) - \tilde{\mathbf{x}}^T(m) \mathbf{w}(n)]. \quad (1.4.15)$$

Ao igualar a zero a Equação (1.4.15), é possível encontrar o vetor ótimo $\mathbf{w}(n)$ que minimiza o MSE. Temos então que

$$-\sum_{m=0}^n \lambda^{(n-m)} \tilde{\mathbf{x}}(m) d(m) + \sum_{m=0}^n \lambda^{(n-m)} \tilde{\mathbf{x}}(m) \tilde{\mathbf{x}}^T(m) \mathbf{w}(n) = [0 \ 0 \ \dots \ 0]^T, \quad (1.4.16)$$

que fornece a expressão de atualização do vetor de coeficientes do filtro $\mathbf{w}(n)$

$$\mathbf{w}(n) = \left[\sum_{m=0}^n \lambda^{(n-m)} \tilde{\mathbf{x}}(m) \tilde{\mathbf{x}}^T(m) \right]^{-1} \sum_{m=0}^n \lambda^{(n-m)} \tilde{\mathbf{x}}(m) d(m). \quad (1.4.17)$$

Com o que, tem-se

$$\mathbf{w}(n) = \mathbf{R}_D^{-1}(n) \mathbf{p}_D(n) \quad (1.4.18)$$

$$\mathbf{R}_D(n) = \sum_{m=0}^n \lambda^{(n-m)} \tilde{\mathbf{x}}(m) \tilde{\mathbf{x}}^T(m) \quad (1.4.19)$$

$$\mathbf{p}_D(n) = \sum_{m=0}^n \lambda^{(n-m)} \tilde{\mathbf{x}}(m) d(m). \quad (1.4.20)$$

Acima, $\mathbf{R}_D(n)$ e $\mathbf{p}_D(n)$ são a matriz de correlação determinística do sinal de entrada e o vetor de correlação cruzada determinística entre o sinal de entrada e o desejado, respectivamente.

No algoritmo RLS convencional, o cálculo da matriz inversa é evitado através do uso do lema de inversão de matrizes

$$[\mathbf{A} + \mathbf{BCD}]^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} [\mathbf{DA}^{-1} \mathbf{B} + \mathbf{C}^{-1}]^{-1} \mathbf{DA}^{-1}. \quad (1.4.21)$$

Com este lema, o inverso da matriz de correlação determinística pode ser calculado para $n \geq 0$ usando

$$\mathbf{R}_D^{-1}(n) = \mathbf{S}_D(n) = \frac{1}{\lambda} \left[\mathbf{S}_D(n-1) - \frac{\mathbf{S}_D(n-1) \tilde{\mathbf{x}}(n) \tilde{\mathbf{x}}^T(n) \mathbf{S}_D(n-1)}{\lambda + \mathbf{S}_D(n-1) \tilde{\mathbf{x}}(n) \tilde{\mathbf{x}}^T(n)} \right] \quad (1.4.22)$$

Inicializa-se o algoritmo RLS convencional com $n < 0$, utilizando

$$\mathbf{S}_D(-1) = \delta \mathbf{I}, \quad (1.4.23)$$

$$\mathbf{p}_D(-1) = \tilde{\mathbf{x}}(-1) = [0, \dots, 0]^T, \quad (1.4.24)$$

onde δ (a constante de regularização) é um valor arbitrariamente pequeno, $\mathbf{I}$ é a matriz identidade (sua dimensão depende do número de coeficientes do filtro). $\mathbf{S}_D(-1)$ é $\mathbf{R}_D^{-1}(n)$ para $n < 0$, e $\mathbf{p}_D(-1)$ e $\tilde{\mathbf{x}}(-1)$ são, respectivamente, o vetor de correlação cruzada determinístico e o vetor do sinal de entrada para $n < 0$.

O Algoritmo 1 apresenta o RLS considerando as derivações anteriormente apresentadas. O valor de δ no algoritmo 1 pode ser o inverso da estimativa de potência do sinal de entrada (DINIZ; WERNER, 2003).

1.5 Correção de Média e Variância

Ajustamos a média e variância da sequência do BESM para que seja a mesma do ERA5 como alternativa de correção, e comparamos os métodos usando filtragem adaptativas à alternativa. Denominamos esta alternativa para ajuste da previsão climática

Algoritmo 1: Algoritmo RLS convencional

Entrada: dados de BESM

```

1 início
2   se  $n < 0$  então
3      $\mathbf{S}_D(-1) = \delta \mathbf{I}$   $\triangleright \delta$  pode ser o inverso de uma estimativa;
4      $\mathbf{x}(-1) = \mathbf{w}(-1) = [0 \ 0 \ \dots \ 0]^T$ 
5   senão
6      $\mathbf{R}_D^{-1}(n) = \mathbf{S}_D(n) = \frac{1}{\lambda} \left[ \mathbf{S}_D(n-1) - \frac{\mathbf{S}_D(n-1)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{S}_D(n-1)}{\lambda + \mathbf{S}_D(n-1)\mathbf{x}(n)\mathbf{x}^T(n)} \right];$ 
7      $\mathbf{p}_D(n) = \lambda \mathbf{p}_D(n-1) + d(n)\mathbf{x}(n);$ 
8      $\mathbf{w}(n) = \mathbf{S}_D(n)\mathbf{p}_D(n);$ 
9      $y(n) = \mathbf{w}^T(n)\mathbf{x}(n);$ 
10     $e(n) = d(n) - y(n)$ 
11  fim
12 fim
```

usando o termo Correção de Media e Variança (CMV).

A média é o valor central de uma distribuição ou de uma lista de valores

$$\mu_x = \frac{x(1) + x(2) + x(3) + \dots + x(n)}{N} = \frac{\sum_{n=0}^{N-1} x(n)}{N}. \quad (1.5.1)$$

A variância (σ^2) é uma medida da dispersão estatística de uma distribuição ou da potência de uma lista de valores

$$\sigma_x^2 = \frac{1}{N} \sum_{n=0}^{N-1} (x(n) - \mu)^2, \quad (1.5.2)$$

Sejam μ_X e σ_X a média e a variância da série $x(n)$, $n = 1 \dots N$, do BESM, e μ_D e σ_D as da série $d(n)$, série correspondente do ERA5, numa mesma célula na posição (l, c) . A Correção de Media e Variância (CMV) da sequência $x(n)$ relativamente a $d(n)$ é

$$\hat{x}(n) = (x(n) - \mu_X) \frac{\sigma_D}{\sigma_X} + \mu_D. \quad (1.5.3)$$

Aplicamos essa correção nos valores das células no tempo e também no domínio transformado. No domínio das frequências, o último caso, obtém-se a DCT-2D, aplica-se o CMV, e aplica-se a IDCT-2D para obter a série ajustada.

No seguinte capítulo, explicamos a metodologia de escolha (sorteio) das células alvo na qual testamos a aplicação dos filtros e algoritmos adaptativos, bem como o método de correção de média e variância. Além disso, detalham-se as configurações dos filtros adaptativos, os diferentes cenários analisados para aplicação da metodologia proposta e os critérios de desempenho empregados na análise do desempenho da proposta de correção de previsões climáticas.

2 METODOLOGIA EXPERIMENTAL

Neste capítulo, são apresentados os critérios de seleção de células alvo e os critérios de avaliação da proposta apresentada. Discute-se ainda a configuração do algoritmo RLS e de outros parâmetros relevantes para a aplicação da filtragem adaptativa no domínio da frequência.

2.1 Variável Prognóstica, Células e Períodos de Trabalho

2.1.1 Variável Prognóstica

O ERA5 e o BESM estão, conforme ilustra a Figura 12, temporal e espacialmente alinhados o que permite corresponder suas células na superfície e avaliar os erros sazonais e anuais entre eles. Assim, aplica-se a metodologia discutida no capítulo anterior para a correção da variável prognóstica Temperatura da Superfície do Mar (TSM) do BESM referente ao ERA5.

2.1.2 Períodos de Trabalho e Amostragem Temporal

Aplicamos a metodologia apresentada para corrigir a variável prognóstica Temperatura da Superfície do Mar (TSM) do BESM tomando o ERA5 como referência, alinhando temporal e espacialmente os períodos de trabalho. Apresentamos os resultados para um período específico, compreendido entre 01 de Novembro de 2006 às 12 horas e 31 de Outubro de 2007 até às 12 horas. Para treinar os filtros, utilizamos o período de 2005-11-01 a 2006-10-31 (um ano); e, em seguida, corrigimos as previsões para o período de 2006-11-01 a 2007-10-31 (ano seguinte ao treinamento). É importante ressaltar que os períodos de trabalho do BESM e ERA5 estão em proporção tanto em termos de tempo quanto de espaço.

2.1.2.1 Ajuste Anual

As previsões do modelo BESM e os dados de reanálise do ERA5 referentes à TSM apresentam um intervalo de saída com valor médio diário. Cada série compreende 365 dias de previsão iniciados em 01 de novembro de um ano e terminados em 30 de outubro do ano subsequente. Para aplicação da metodologia, considera-se neste trabalho, para o

ano de treinamento, o período de 01 de novembro 2005 a 31 de outubro do 2006, e, da mesma forma, para o ano de ajuste o intervalo de 01 de novembro de 2006 a 31 de outubro de 2007.

2.1.2.2 Ajuste Estacional ou Sazonal

Dadas as variações de clima ao longo do ano, consideramos também a correção de previsões por estações climáticas do ano. As avaliações sazonais permitem verificar se a restrição do período para aplicação da metodologia proposta permitem melhorar o desempenho da mesma.

Cada série anual é separada em quatro subséries, cada uma é composta pela previsão para 3 meses. As separações das estações seguem a sequência da previsão, isto é, de três em três meses de Novembro a Outubro; assim, a primeira estação compreende Novembro, Dezembro e Janeiro (do ano seguinte), depois temos o período de Fevereiro a Abril, a estação de Maio a Julho e, por último, a de Agosto a Outubro. Elas contêm entre 90 a 93 dias. A metodologia de correção é aplicada para cada série sazonal, obtendo assim quatro filtros para a correção da previsão anual, um para cada estação.

2.1.3 Células Alvo

As células ou pixels alvo (aos quais aplicamos a proposta) são selecionados de forma aleatória sobre a superfície do mar. Um exemplo do processo de sorteio pode ser visto na Figura 11. Treina-se um filtro para cada célula à qual aplica-se a metodologia de correção. O bloco para a análise espectral é definido para incluir a vizinhança do pixel sorteado; isto é, a partir de suas coordenadas (lat,lon).

2.1.4 Tamanhos de Blocos para a Análise Espectral

Para a análise espectral, computa-se a DCT-2D em torno da célula que deseja-se corrigir usando blocos de 32×32 , 16×16 , e 8×8 *pixels* (Equação 1.4.1). A Figura 12 com a célula (lat,lon)=(-10.0,-25.75) exemplifica um bloco de 32×32 para o cálculo da DCT-2D.

Para cada tamanho de bloco da TSM em torno da célula alvo, obtém-se uma quantidade distinta de séries de coeficientes DCT-2D. Por exemplo, para uma caixa de 8×8 pixels há um total de 64 séries temporais, para 16×16 , 256, e para 32×32 , 1024 séries.

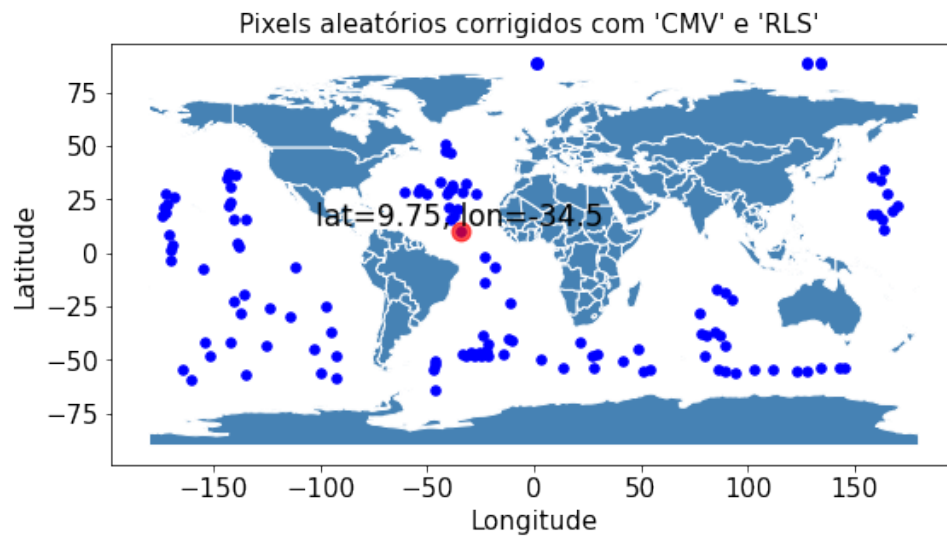


Figura 11 - Exemplo da seleção de células para aplicação de filtros adaptativos à TSM do BESM visando melhorar sua acurácia relativamente ao ERA5.

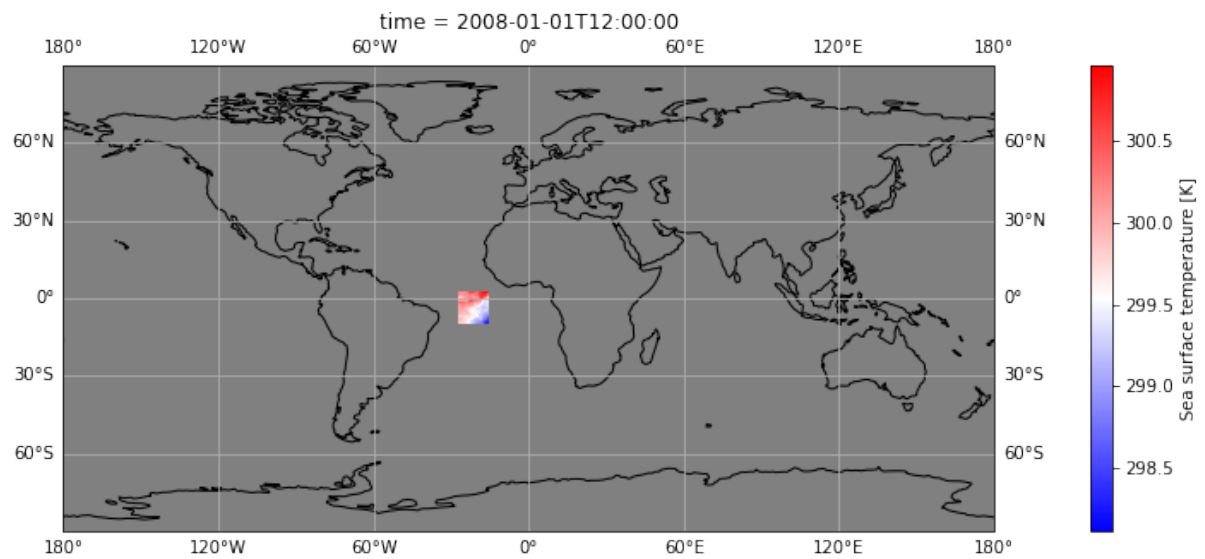


Figura 12 - Exemplo de um bloco para aplicação da DCT-2D e da célula (lat,lon)=(-10.0,-25.75).

2.1.4.1 Definição do Bloco para Análise Espectral em Função da Célula Alvo

Cada bloco de pixels, ao qual aplicamos a análise espectral, refere-se a uma região de dimensões $M \times M$ da TSM, analisada independentemente. O tamanho do bloco afeta a resolução da estimativa espectral. A célula alvo é definida para $M = 8, 16$ ou 32 em ambas coordenadas (latitude e longitude), tal como mostrado na Figura 13. Observa-se a célula alvo (caixa vermelha) da TSM, em dimensões distintas e em diferentes perspectivas. O resultado para valor de M é a princípio distinto. Para a análise no domínio do tempo e da frequência para ambos casos as dimensões (M) permanecem fixas assim como a célula alvo. Para cada M , avalia-se a aplicação do ajuste anual e por estações; ou seja, para a célula alvo e os pixels vizinhos (dentro da região de $M \times M$ pixels) forma-se uma série temporal com os valores da variável para os 365 dias do ano ou com os valores ora 90 a 93 dias. Após o sorteio aleatório da célula alvo, os $M \times M$ pixels em torno dela são usados para aplicar a análise espectral por meio da DCT-2D. A célula alvo é identificada pelo pixel central.

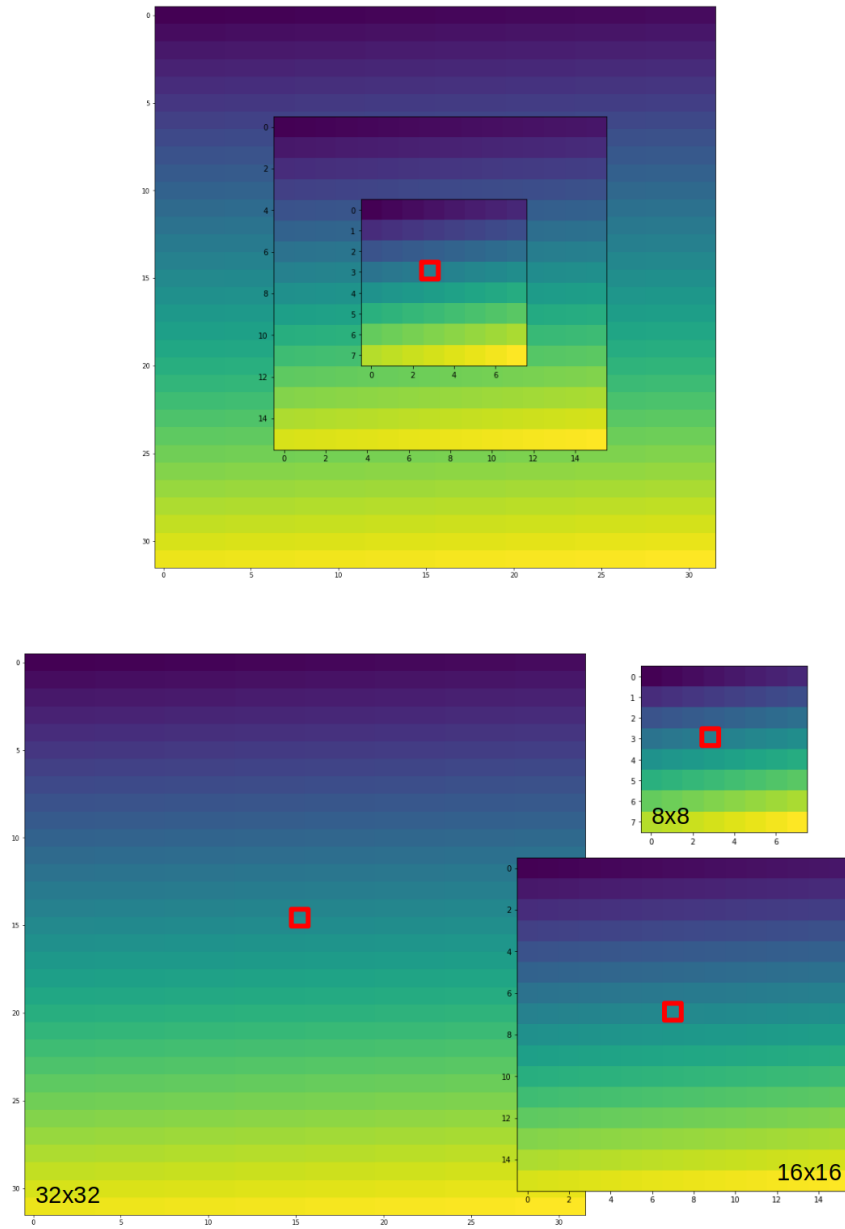


Figura 13 - Exemplo de blocos $M=8, 16$, e 32 , comparativamente à célula alvo de BESM e ERA5. As cores representam a Temperatura da Superfície do Mar (TSM) e o vermelho é a célula alvo.

2.2 Critérios de Avaliação de Desempenho

Empregam-se critérios objetivos para avaliar se os filtros adaptativos ajustam para melhor ou não a série BESM relativamente à ERA5.

A Figura 14 exemplifica a estrutura dos critérios de avaliação. É importante enfatizar que os critérios de avaliação são aplicados no domínio do tempo, o domínio natural dos dados, enquanto propõe-se a aplicação de filtros adaptativos no domínio da frequência.

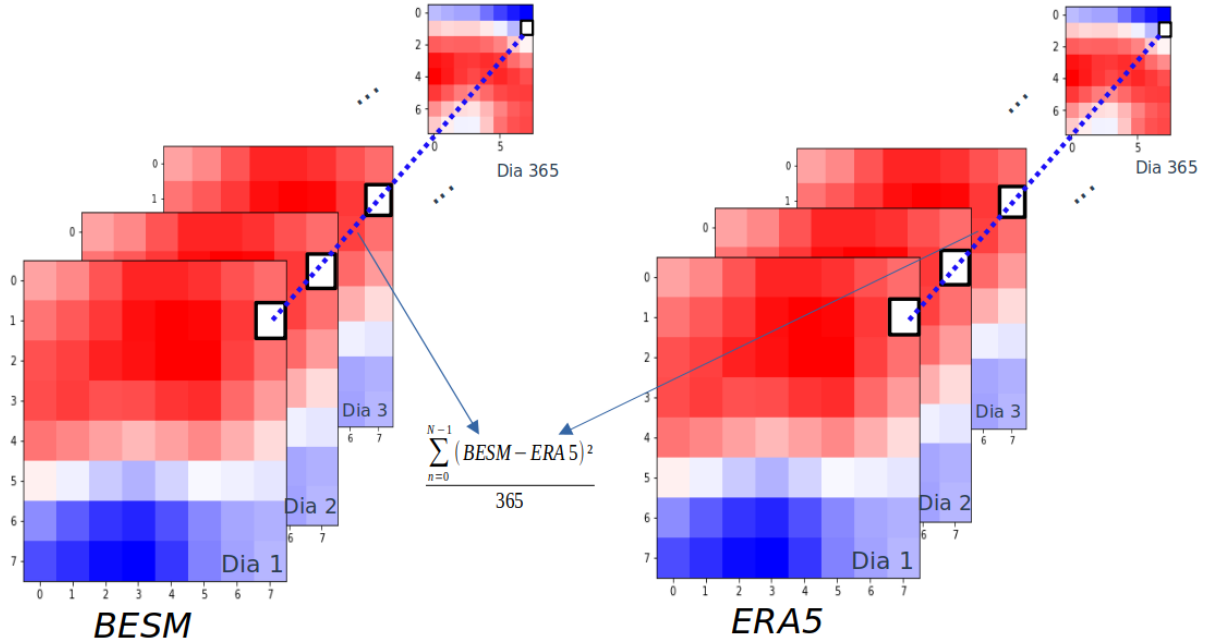


Figura 14 - Exemplo do cálculo de MSE da TSM ao longo de tempo para BESM e ERA5.

Conforme mencionado, avaliamos a melhoria na acurácia da previsão no domínio do tempo (o domínio em que ela é produzida e disponibilizada) e, portanto, em sua confiabilidade. O erro original da previsão numa célula (l, c) no n -ésimo dia é

$$e_n(l, c) = x_n(l, c) - d_n(l, c), \quad (2.2.1)$$

onde $d_n(l, c)$ é o valor da célula advindo do ERA5 e $x_n(l, c)$, o do BESM. Comparamos esse erro com o resultante da correção,

$$\hat{e}_n(l, c) = \hat{x}_n(l, c) - d_n(l, c), \quad (2.2.2)$$

na qual, $\hat{x}_n(l, c)$ é o valor ajustado (filtrado) da variável prognóstica. A partir da definição de erro, definimos alguns índices para avaliar o desempenho do BESM.

2.2.1 Largest Absolute Error (LAE) e Percentual Largest Absolute Error (PLAE)

Para cada célula, obtemos o erro absoluto máximo (LAE)

$$\text{LAE}(l, c) = \max_n |e_n(l, c)|, \quad (2.2.3)$$

e o erro absoluto percentual máximo (PLAE)

$$\text{PLAE}(l, c) = 100\% \max_n \left| \frac{e_n(l, c)}{d_n(l, c)} \right|. \quad (2.2.4)$$

O *LAE* e o *PLAE* são medidas que nos permitem avaliar a qualidade das previsões frente ao ocorrido, e, equivalentemente, da previsão comparativamente ao valor observado (reanálise). O *LAE* retorna a maior magnitude dentre os erros absolutos.

2.2.2 Mean Squared Error (MSE)

Por sua vez, o MSE é dado por

$$\text{MSE}(l, c) = E[e_n(l, c)^2] = \frac{1}{N} \sum_{n=0}^{N-1} e_n(l, c)^2. \quad (2.2.5)$$

O *MSE* avalia a potência (energia do erro); é o critério empregado no projeto dos filtros adaptativos.

2.2.3 Índices Agregados

Para avaliar o desempenho, compara-se o LAE, o PLAE e o MSE antes e após a correção; isto é, para $\hat{e}_n(l, c)$ e para $e_n(l, c)$, de forma a avaliar se há ou não melhorias. Os critérios anteriores são aplicáveis a células. Entretanto, devemos avaliar os métodos de correção globalmente, considerando diferentes células. Médias, máximos e dispersões das medidas de erro anteriores apesar de caracterizarem o desempenho, poderiam mascarar a análise do desempenho global, já que a variável prognóstica pode variar ao redor do globo terrestre.

Consequentemente, definem-se índices que permitem avaliar o desempenho global da correção da previsão. Primeiramente, obtêm-se os índices de desempenho (MSE, LAE ou PLAE) do BESM em relação ao ERA5 antes e após o método de correção aplicado para uma amostra de células. Classificam-se essas células entre as que o método de correção melhorou ou piorou o desempenho conforme cada índice. Finalmente, para cada índice,

obtêm-se as porcentagens de células em que houve melhoria ou piora. Temos assim três índices agregados de desempenho, um para o MSE, um para o LAE, e outro para o PLAE.

A Figura 15 ilustra o resultado da análise conjunta através dos índices agregados. Nela, temos os três índices para quatro métodos de correção. O RLS_{freq} é o ajuste usando a filtragem adaptativa no domínio da frequência proposto nesta dissertação. O RLS_{temp} é o ajuste usando a filtragem adaptativa no domínio do tempo (FREITAS, 2015; PINTO et al., 2018). Os métodos CMV_{freq} e CMV_{temp} se referem aos ajustes de média e variância discutida na Seção 1.5. Na Figura 15, M é o tamanho do bloco (de uma dimensão pois são quadrados) empregada para a DCT-2D, quando aplicável, este não é o caso das correções aplicadas diretamente no domínio temporal.

Devemos perceber que o LAE e o MSE permitem avaliar a precisão do modelo de previsão bem como a correção de viés de sua saída. O MSE mede a potência do erro, é uma medida da distância entre a série de interesse (BESM) e a série alvo ou de referência (ERA5). Por outro lado, o LAE mede o maior erro absoluto, a maior diferença, indicando, assim, os erros extremos. Quanto maior o erro extremo, menor é a coincidência no ponto correspondente. O $PLAE$ relativiza o LAE pelo valor da amostra, já que leva em conta a ordem de grandeza dos valores considerados.

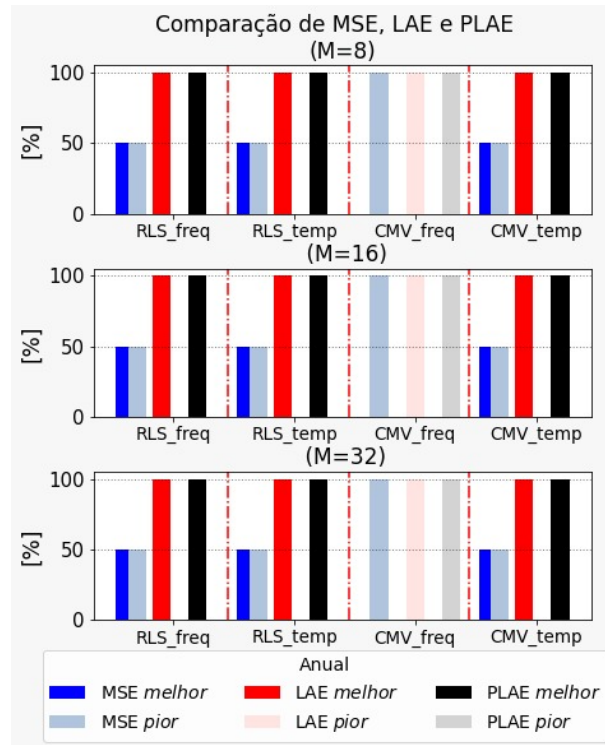


Figura 15 - Exemplo de apresentação e avaliação de desempenho usando os índices agregados, indicando as proporções de melhoria ou piora para cada medida de erro (MSE, LAE e PLAE), M é o tamanho da janela empregada para a DCT-2D, quando aplicável.

Conforme mencionado, o método proposto, RLS na frequência, é comparado com outras três abordagens: RLS no tempo, a correção de média e variância (CMV) no domínio

temporal e no frequencial.

2.3 Configuração do Filtro Adaptativo

Os filtros adaptativos são treinados para obter o filtro para correção das sequências. Na Figura 16, exemplifica-se o comportamento dos coeficientes ao longo das iterações para os diferentes valores de M com $L = 4$ para as repetições das séries do ano de treinamento usadas durante o treinamento. Os coeficientes do filtro são inicializados com um valor aleatório ou zeros. Os filtros usados para o ajuste/correção da previsão são os que resultam ao final do processo de treinamento. Para o ajuste anual (365 dias), com 10 e 2 repetições, temos períodos de 3650 dias e 730 dias respectivamente representados na Figura 16. Os filtros adaptativos serão aplicados para a correção das séries por estações. Vemos que os filtros variam com o valor de M (o tamanho da região em torno do pixel usada para a aplicação da DCT-2D). Os coeficientes dos filtros finais determinam o ajuste da previsão BESM. Esses coeficientes definem o filtro empregado para ajustar as séries do BESM no ano de ajuste no domínio da frequência. A seleção dos “coeficientes ideais”, ou seja, os melhores filtros para o ajuste, depende das características específicas de cada modelo (do ERA5 e do BESM), quer dizer, em cada modelo e seu ano respectivo os modelos variam os coeficientes não serão os mesmos e portanto reduzirá o erro mais ou menos. Em nosso caso, para os ajustes anual, são escolhidos o vetor de coeficientes obtido após o fim das iterações, e por estações posicionadas ao fim de cada três meses, respectivamente.

2.3.1 Configuração do Algoritmo RLS

Os parâmetros relevantes para a aplicação do algoritmo RLS são a ordem do filtro adaptativo (L), o vetor inicial de coeficientes do filtro, a constante de regularização (δ) e o fator de esquecimento ou fator de ponderação exponencial (λ). Nesta seção, discutimos a escolha desses parâmetros na proposta desta dissertação.

O filtro (vetor de coeficientes) $\mathbf{w}(n)$ é inicializado com zeros em todos os casos e fazemos a constante de regularização $\delta = 1$. Nas subseções seguintes, a partir disso, fundamentam-se as escolhas da ordem do filtro e o fator de ponderação exponencial.

2.3.1.1 Ordem do Filtro (L)

Para cada célula foram projetados filtros para $L=4, 6, 8, 12, 16$ e 32 para blocos $M \times M$, com $M= 8, 16$ e 32 . A Figura 17 exemplifica esse procedimento. Observa-se

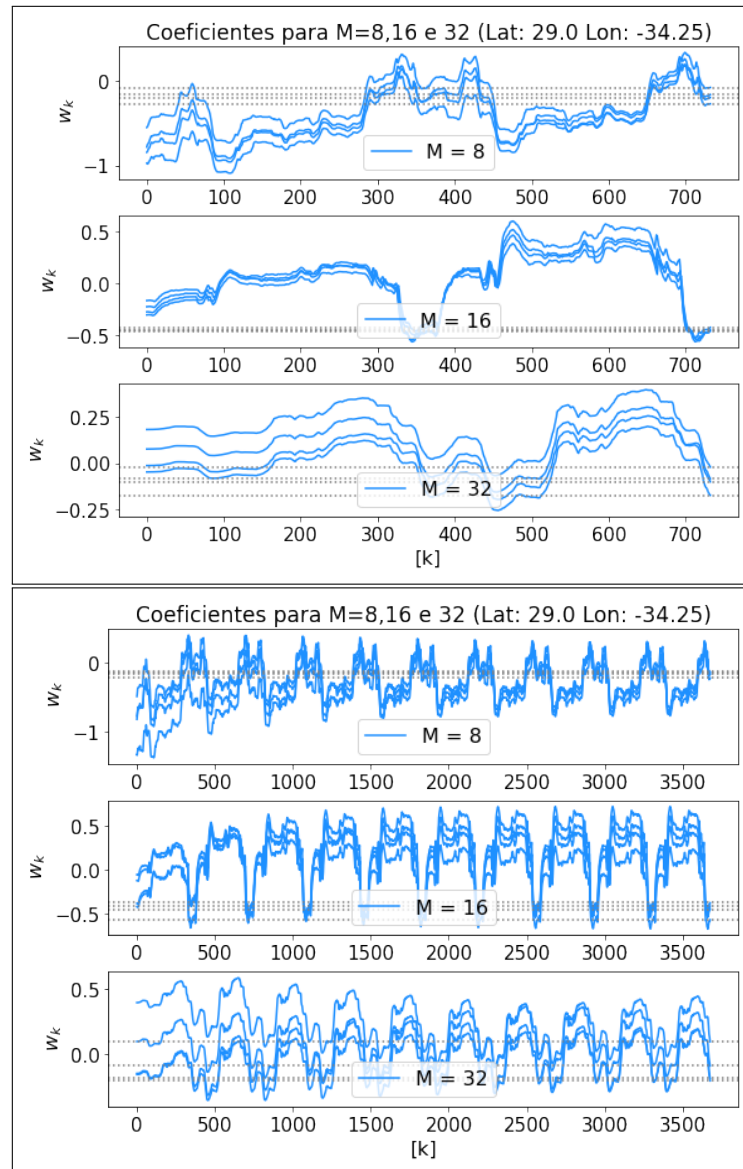


Figura 16 - Evolução dos coeficientes do filtro adaptativo para 2 repetições anuais da série (acima) e para 10 repetições (abaixo). Cada conjunto de gráficos apresenta os $L = 4$ coeficientes do filtro considerando blocos com $M=8, 16$ e 32 .

uma alteração no resultado (a correção da série BESM) em função de L , sendo que para $L = 32$ observa-se um descolamento em relação às demais ordens de filtro.

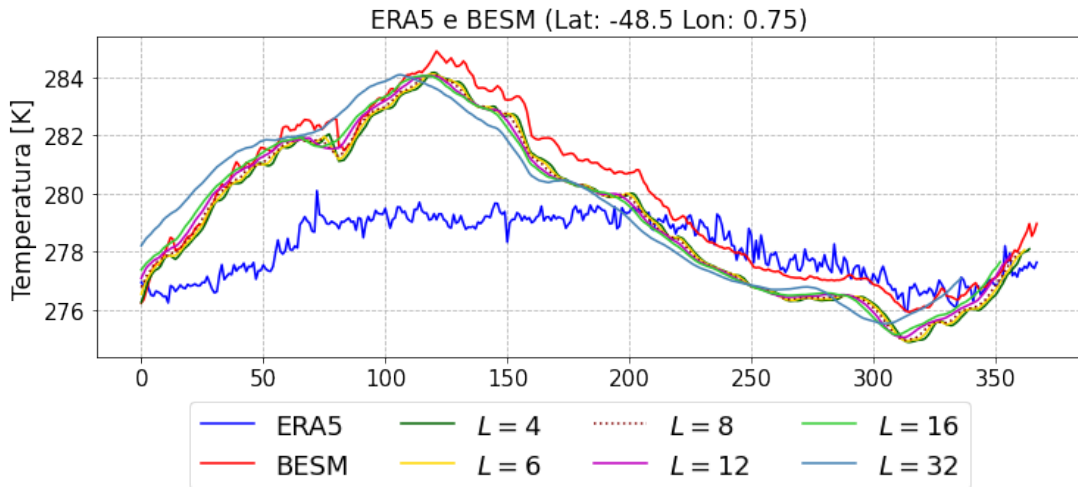


Figura 17 - Ajuste do BESM em relação ao ERA5 com filtros adaptativos de comprimento $L = 4, 6, 8, 12, 16$ e 32 para o ajuste do BESM na frequência com $M = 32$.

Ao observar a figura, é possível notar que a série filtrada apresenta uma maior semelhança visual com o ERA5 em comparação com a série provida pelo BESM. Conforme o comprimento do filtro aumenta, há uma melhora na concordância entre as duas séries. Isso sugere que o uso de filtros adaptativos ajuda a corrigir as discrepâncias entre o BESM e o ERA5, resultando em uma previsão melhor para a temperatura da superfície do mar.

A ordem do filtro L é escolhida a partir dos critérios de avaliação de desempenho. A Figura 18 apresenta os percentuais referentes às ocorrências de redução (“melhor”) ou aumento (“pior”) do erro em termos de MSE, LAE e PLAE, quando comparado aos níveis de erro sistemático entre BESM e ERA5, variando-se o comprimento do filtro para um M fixo. Nota-se, de maneira geral, que os desempenhos dos métodos de correção para $L = 4$ e $L = 12$ são equivalentes; e que para $L = 32$, ocorre um deterioramento no desempenho de correção. Dessa forma, opta-se por $L = 4$, posto que é possível atingir um mesmo nível de correção, com um menor custo computacional.

2.3.1.2 Fator de Esquecimento

O fator de esquecimento ou fator de ponderação exponencial λ dos erros passados é usado para calcular o ajuste a aplicar aos coeficientes do filtro. Neste trabalho, testamos alguns valores λ para todas as combinações de L e M anteriormente apresentadas. A Figura 19 apresenta um exemplo de correção de uma série de TSM para $\lambda = 0,98, 0,99$ e $0,999$ com $M = 32$ e $L = 4$, empregando-se o método de correção proposto RLS_freq. Nesse caso, opta-se por $\lambda = 0,99$ por permitir obter coeficientes que proporcionam de

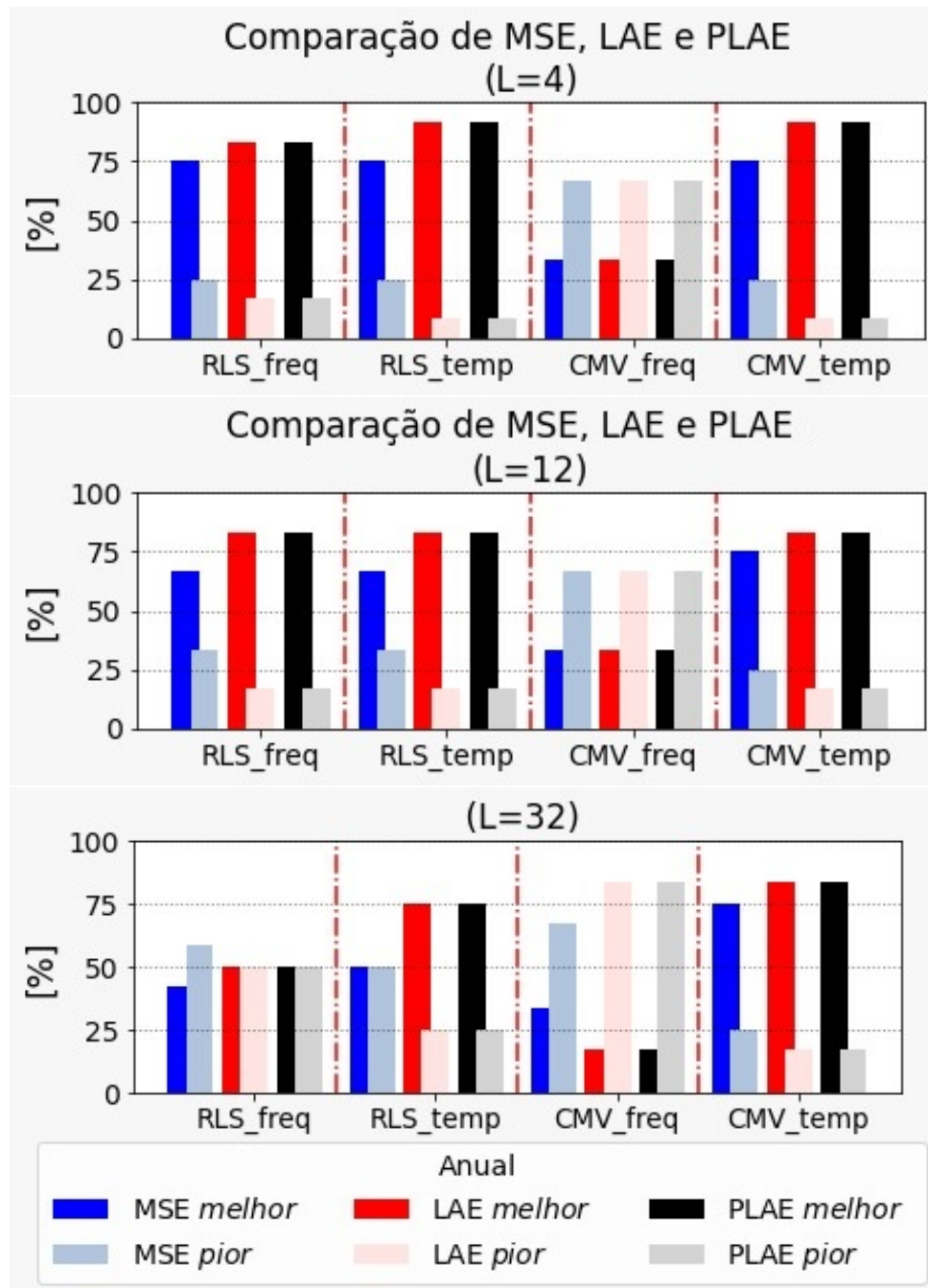


Figura 18 - Percentuais de redução e aumento do erro com relação ao erro original entre BESM e os dados observacionais de temperatura da superfície do mar, considerando os pixels aleatoriamente escolhidos, resultantes de correção ANUAL, variando-se o comprimento do filtro $L=4, 12$ e 32 com $M=8$.

forma geral um maior grau de precisão no resultado da correção de viés.

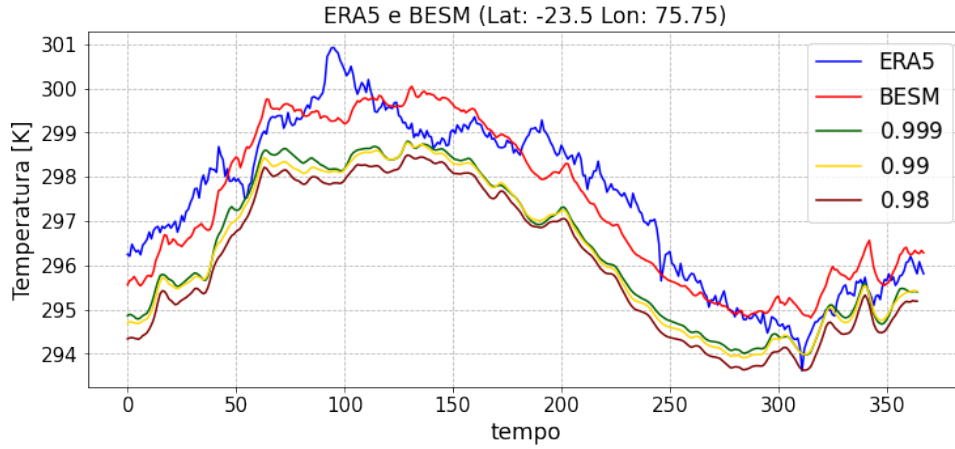


Figura 19 - Resultados de correção da TSM do BESM para $\lambda = 0.98, 0.99$ e 0.999 por meio de filtragem adaptativa RLS no domínio da frequência, para $M = 32$ e $L = 4$.

2.3.2 O Filtro Ótimo ou Vetor de Coeficientes Ótimos

O algoritmo RLS atualiza o vetor de coeficientes $\mathbf{w}(n)$ a cada iteração n . O vetor de coeficientes ótimos ocorre quando a atualização do filtro converge no regime transitório apresentando um regime quase permanente. Isto é, quando o vetor de coeficientes varia muito pouco entre iterações ou apresenta um padrão oscilatório, porém periódico, em função do comportamento não-estacionário da série temporal. Como vimos em exemplos anteriores, nesta proposta, o que ocorre é a segunda situação. Assim, neste trabalho assume-se, como vetor ótimo ou de coeficientes ótimos, aquele obtido na última iteração do algoritmo RLS.

Se o treinamento emprega uma sequência anual, determinamos como filtro ótimo $\mathbf{w}(365)$, independentemente da série considerada. O vetor ótimo obtido para a série do ano de treinamento é aplicado para corrigir a série do ano de ajuste. Se o treinamento emprega uma sequência por estações, os filtros ótimos são: $\mathbf{w}(93)$, $\mathbf{w}(183)$, $\mathbf{w}(275)$ e $\mathbf{w}(365)$.

2.3.3 Autocorrelação do Módulo do Filtro

A autocorrelação refere-se à medida da similitude entre séries temporais. Assim, a variação da autocorrelação pode ser empregada para avaliar a estacionariedade de uma série temporal. Desta forma, analisa-se o módulo do vetor de coeficientes do filtro, $\|\mathbf{w}(n)\|$,

para avaliar ciclos de repetições em intervalos regulares, como dias, semanas ou anos.

Empregamos a autocorrelação $R_{\mathbf{w}}(l)$, com distância “ l ”, do módulo do vetor de coeficientes do filtro $\|\mathbf{w}(n)\|$

$$R_{\mathbf{w}}(l) = \sum_n \|\mathbf{w}(n)\| \|\mathbf{w}(n-l)\| \quad (2.3.1)$$

para encontrar padrões periódicos no processo de treinamento/adaptação do filtro. Como vemos na Figura 20, que traz essa autocorrelação dos coeficientes dos filtros adaptativos com $L=4$ para 2 repetições anuais da serie (acima) e 10 repetições (abaixo) (Figura 17), $M=8, 16$ e 32 , para o ajuste da TSM considerando, apresenta oscilações regulares. Ademais, por meio da função de autocorrelação da norma, é possível identificar também picos intra-anuais (as linhas vermelhas tracejadas indicam o início da primeira e segunda repetições anuais).

No capítulo seguinte, esta autocorrelação é utilizada para identificar padrões nos vetores de coeficientes dos filtros adaptativos durante seus treinamentos, para avaliar a estacionariedade, tendências, ciclos e para uma interpretação melhor do comportamento.

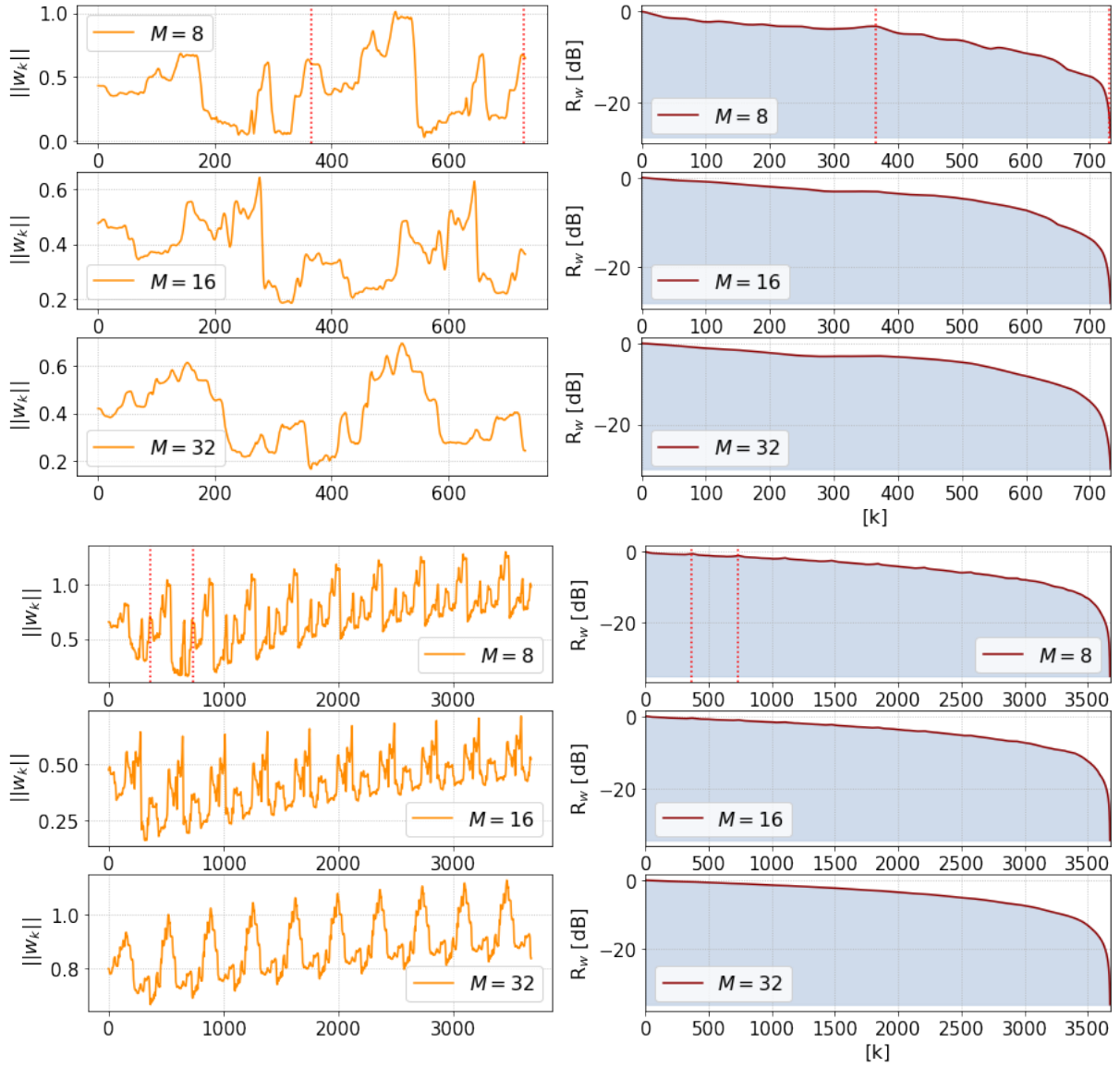


Figura 20 - Comportamento da norma dos coeficientes do filtro adaptativo e sua autocorrelação para $L = 4$ e $M = 8, 16, 32$ para 2 repetições anuais da serie (acima) e para 10 repetições (abaixo) aplicado em BESM . As linhas vermelhas tracejadas indicam os inícios da primeira e segunda repetições anuais. Pixel na posição $lat : -23.5$ $lon : 75.75$.

3 RESULTADOS

Este capítulo apresenta a avaliação dos métodos de correção do BESM relativamente ao ERA5, conforme anteriormente discutido. Apresentam-se os valores finais escolhidos para os parâmetros L , M e λ para o projeto adaptativo dos filtros empregados para a correção. Analisamos as abordagens usando filtros adaptativos RLS e o método CMV aplicados nas séries de valores de uma célula da TSM e usando a análise frequencial a partir do entorno do pixel. Consideramos horizontes de previsão anuais e sazonais. Para avaliar e comparar os diferentes métodos usamos os índices agregados de MSE, LAE e PLAE para amostras de células aleatoriamente escolhidas sobre a superfície do mar em todo o globo terrestre.

3.1 Treinamento do Filtro Adaptativo RLS

Os resultados do processo de treinamento do filtro adaptativo, empregando-se o algoritmo RLS com fator de esquecimento $\lambda = 0.99$ e $L = 4$ coeficientes, para uma célula localizada em latitude e longitude $(-29.0; -34.25)$, são apresentados na Figura 21. Para o ajuste é considerado o primeiro coeficiente da DCT-2D, quer dizer, no espaço das frequências se é $M=8$ dos 64 coeficientes o número 1 é considerado. Observam-se o sinal de erro obtido subtraindo-se o sinal de saída do filtro (BESM corrigido) do sinal de referência (ERA5), os elementos do vetor de coeficientes, o módulo do vetor de coeficientes e a autocorrelação; todos em função do número de iterações. É importante mencionar que neste caso as séries utilizadas no treinamento correspondem ao ano de 2007, ou seja de novembro de 2007 a outubro de 2008.

Nota-se que o erro (no topo da Figura 21) se concentra em valores acima de -50 dB e os coeficientes $\mathbf{w}(k)$ dos filtros não convergem para patamares fixos, dado que o sinal de TSM apresenta um comportamento não estacionário ao longo do ano e das estações. Há uma oscilação bem regular verificada pelo comportamento da norma dos coeficientes $\|\mathbf{w}(k)\|$, em função das 10 repetições das sequências de TSM anuais. As linhas tracejadas verdes indicam uma possível fins de períodos subsequentes de três meses referentes às estações para 10 repetições. Por exemplo para o ajuste sazonal de 365 dias, as linhas verdes ficariam em: 92, 181, 273 e 365 respectivamente. Ademais, por meio da função de autocorrelação da norma, é possível observar uma flutuação lenta intra-anual da energia dos coeficientes. As linhas vermelhas tracejadas indicam o início da primeira e segunda repetições anuais.

Na Figura 22, observam-se os comportamentos de convergência dos $L = 4$ coeficientes ao longo das iterações, para regiões quadradas de diferentes dimensões para

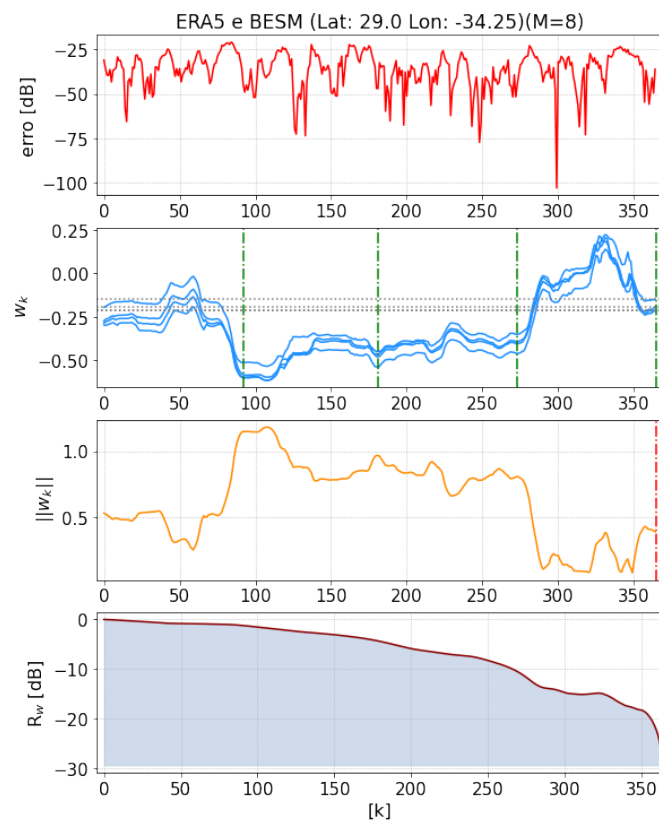


Figura 21 - Comportamento do processo de treinamento do algoritmo RLS aplicados a série temporais de coeficientes DCT-2D, referente a região quadrada 8×8 centrada em $(lat = 29.0; lon = -34.25)$. Para o ajuste anual é considera o primeiro coeficiente da DCT-2D no domínio das frequências.

$M = \{8, 16, 32\}$. Observa-se que quanto maior o valor de M , a convergência se torna mais suave e os ciclos temporais mais bem definidos. Quanto maior a região, maior é a capacidade do algoritmo de ajustar componentes de maiores comprimentos de onda. Pode-se notar também que para cada valor de M , os coeficientes acabam convergindo para diferentes patamares.

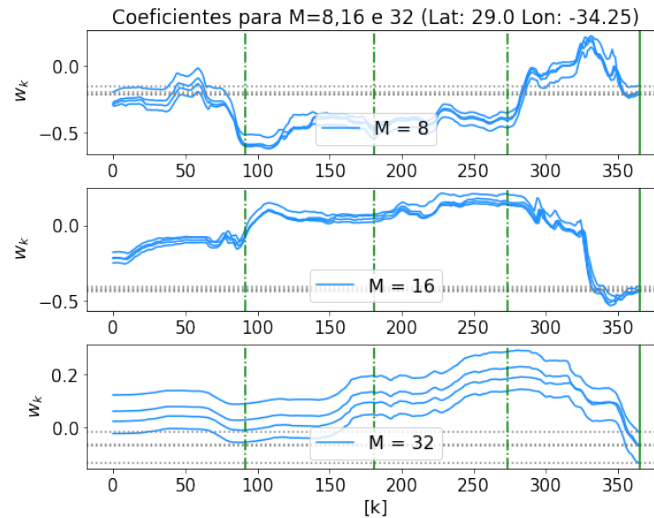


Figura 22 - Comportamento do processo de treinamento do algoritmo RLS aplicados a séries temporais de coeficientes DCT-2D, referente a regiões quadradas de dimensões 8×8 , 16×16 e 32×32 centradas em $(lat = -29.0; lon = -34.25)$.

3.2 Análise por Índices Agregados

Neste trabalho, realiza-se a correção de viés da previsão da TSM do BESM em referência ao ERA5, nos domínios do tempo e da frequência por meio de Filtragem Adaptativa RLS (denominados de RLS_temp e RLS_freq) e do Método de Correção de Média e Variância (CMV_temp e CMV_freq). As previsões de TSM do BESM possuem duração anual com discretização temporal diária, e se iniciam em novembro de um ano específico e terminam em outubro do ano seguinte. Para avaliar a robustez das abordagens de correção de viés sem gerar um alto custo computacional, escolhem-se aleatoriamente T pixels/células ao longo de todo o domínio marítimo do planeta. Nas abordagens de correção no domínio da frequência em que se aplica a DCT bidimensional em uma região quadrada $M \times M$, considera-se o pixel escolhido como o centro dessa região.

A Figura 23 apresenta um mapa com indicações em azul das localizações geográficas de $T = 100$ pixels escolhidos aleatoriamente no domínio marítimo planetário, avaliados nesse trabalho. O pixel com marcação em vermelho, situado em latitude e longitude $(38.0; -35.5)$, será avaliado especificamente em detalhe mais adiante.

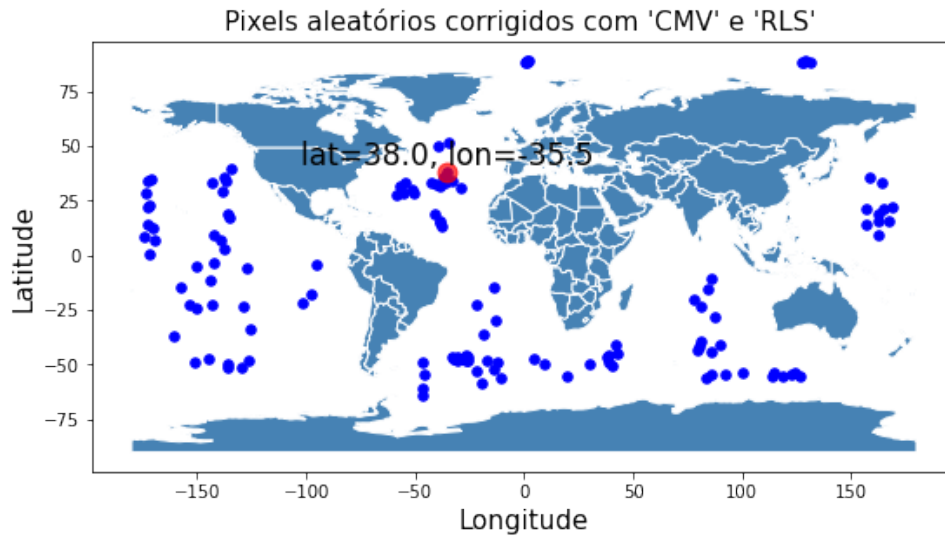


Figura 23 - Pixels aleatoriamente selecionados para o empregos dos métodos de correção de viés baseados em RLS e CMV. O pixel vermelho é examinado em detalhe.

As Figuras 24 e 25 apresentam os percentuais referentes às ocorrências de redução (“melhor”) ou aumento (“pior”) do erro em termos de MSE, LAE e PLAE, quando comparado aos níveis de erro sistemático entre BESM e ERA5. A primeira figura corresponde ao desempenho de correção considerando um ano completo, e a segunda, ao desempenho de correção por estações. Observa-se que em ambas as correções, tanto anual quanto por estações, a estratégia CMV_freq apresenta o pior desempenho, sendo que na maioria dos pixels ocorreu um aumento das medidas de erro em vez de redução.

Já as abordagens baseadas em RLS e o CMV_temp apresentaram desempenhos semelhantes na correção anual, enquanto que na correção por estações, as abordagens baseadas em RLS apresentaram desempenhos superiores ao CMV_temp, com percentuais de redução de erro acima de 75%. Note que nesses casos o percentual de redução de erro foi significativamente maior que o percentual de aumento. Verifica-se que somente o CMV_freq apresentou sensibilidade com relação à variação de M , diminuindo ainda mais o percentual de redução de erro com o aumento de M , tanto na correção anual quanto por estações. Por fim, é possível concluir que as técnicas de correção de viés baseadas em RLS apresentam o mesmo nível de capacidade de redução de erro sistemático entre o modelo BESM e os dados observacionais de TSM.

3.3 Correção da TSM do BESM em Células Específicas em Diferentes Oceanos

Nesta seção, apresentam-se os resultados dos ajustes anuais e sazonais de células nos diferentes oceanos do mapa global. Resultados dos filtros adaptativo RLS e Correção

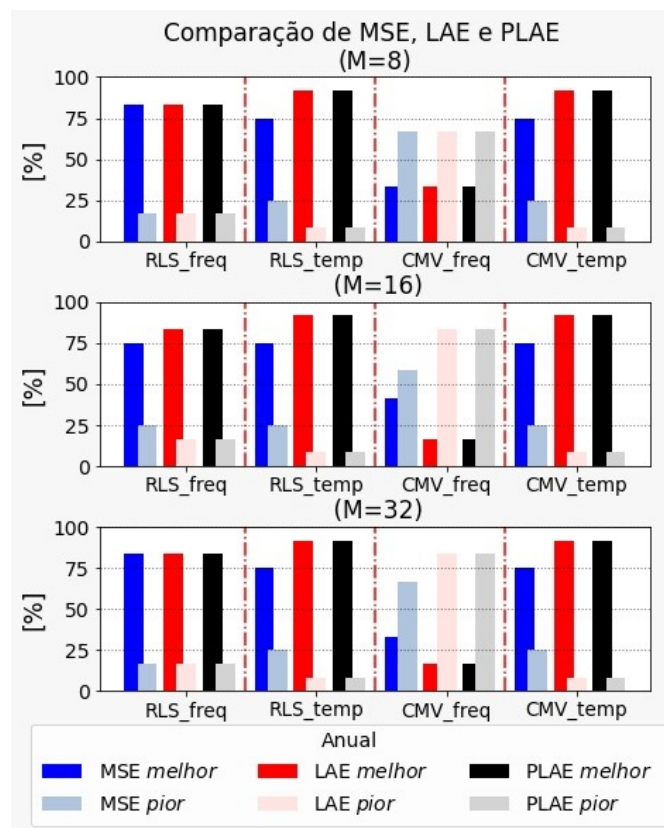


Figura 24 - Percentuais de redução e aumento do erro com relação ao erro original entre BESM e os dados observacionais de temperatura da superfície do mar, considerando os pixels aleatoriamente escolhidos, resultantes de correção anual.

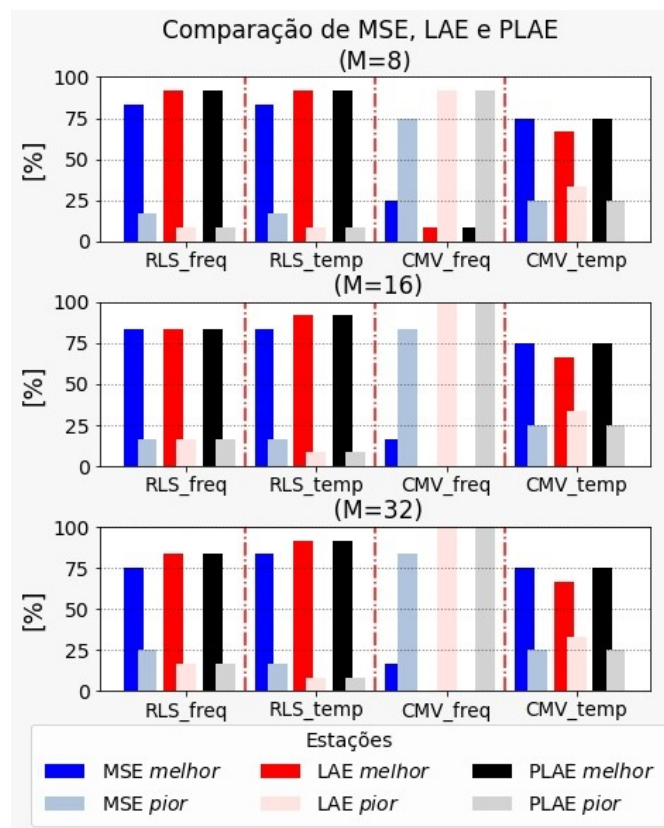


Figura 25 - Percentuais de redução e aumento do erro com relação ao erro original entre BESM e os dados observacionais de temperatura da superfície do mar, considerando os pixels aleatoriamente escolhidos, resultantes de correção de viés por estações.

de Media e Variância (CMV) no domínio do tempo e da frequência para as metodologias utilizadas nos capítulos iniciais.

3.3.1 Célula no Oceano Atlântico

Os resultados apresentados a seguir se referem a correção da série temporal de TSM prevista pelo BESM associado à célula localizada no Oceano Atlântico na coordenada ($lat = 29.0; lon = -34.25$) indicada em vermelho no mapa da Figura 26. Além disso, apresentam-se as séries temporais desta localização produzidas pelo ERA5 e BESM, bem como aqueles decorrentes do ajuste de viés por meio dos algoritmos baseados em RLS e CMV nos domínios do tempo e da frequência. No ajuste anual, observa-se que o CMV_tmp resulta em uma melhor correção a partir do dia 50, já as técnicas baseadas em RLS aproximam melhor a série observacional a partir do dia 200. No ajuste por estações, os métodos RLS_temp, RLS_freq e CMV_temp induzem correções adequadas ao longo de todo o período. Já o CMV_freq não é capaz de produzir uma série que aproxime de forma adequada a série observacional, tanto no ajuste anual quanto no ajuste por estações. Na Figura 27, avalia-se o desempenho de correção do algoritmo RLS_freq para tamanhos de região variados. Nota-se que a variação de M provoca pouco impacto no desempenho da correção de viés.

Na Figura 28, são apresentados os erros iniciais subtraindo-se o sinal de TSM do ERA5 do sinal original ou corrigido do BESM. O gráfico acima se refere ao resultado da correção anual e o de baixo se refere a correção por estações. Nota-se que ao serem empregados, os métodos baseados em RLS promovem erros mais próximo de zero ao longo da série, quando a correção é realizada por estações. O mesmo ocorre para o método CMV_temp, contudo, no caso de CMV_freq o ajuste é deficiente em ambas as correções, anual e por estações.

Na Figura 29, apresentam-se os erros resultantes da correção por meio do método RLS_freq, variando-se o tamanho do região em $M = \{8, 6, 32\}$. Observa-se que neste caso o desempenho de correção é pouco sensível à variação de M tanto para o ajuste anual quanto para o ajuste por estações. Contudo, nota-se que, em torno do dia 200, o erro para blocos menores se afasta de zero, enquanto que, a partir do dia 300, o erro se aproxima de zero.

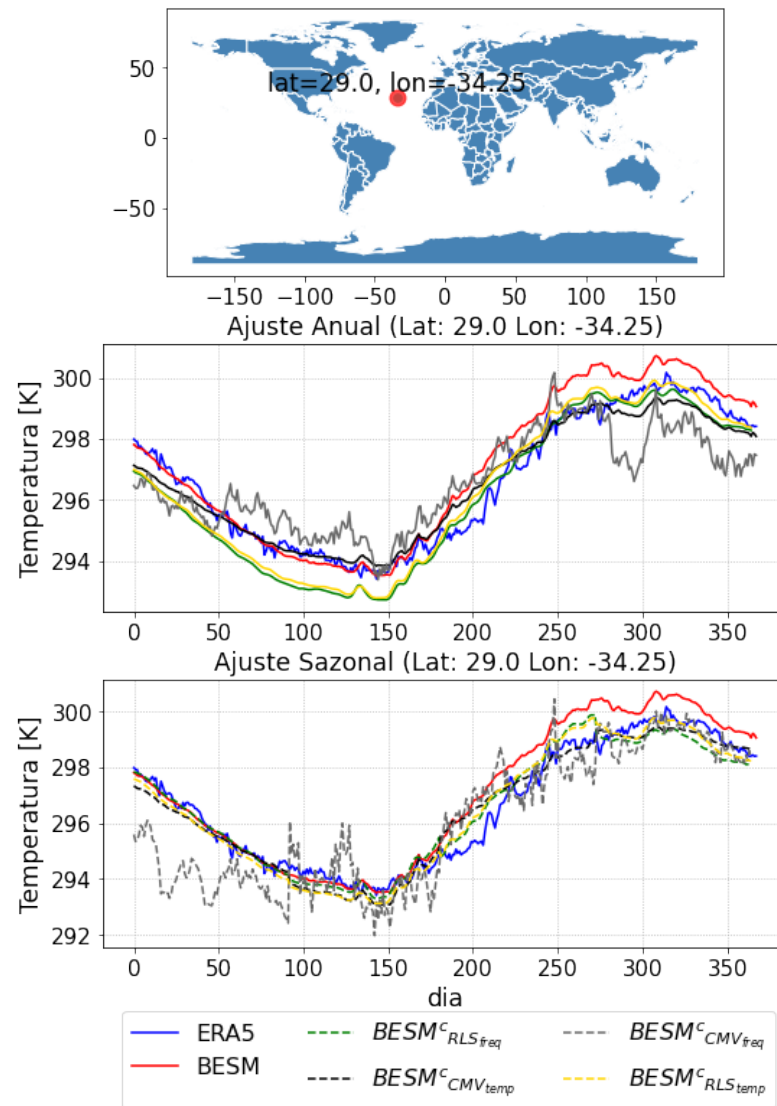


Figura 26 - Correção anual e por estações da temperatura da superfície do mar por meio dos métodos RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Atlântico.

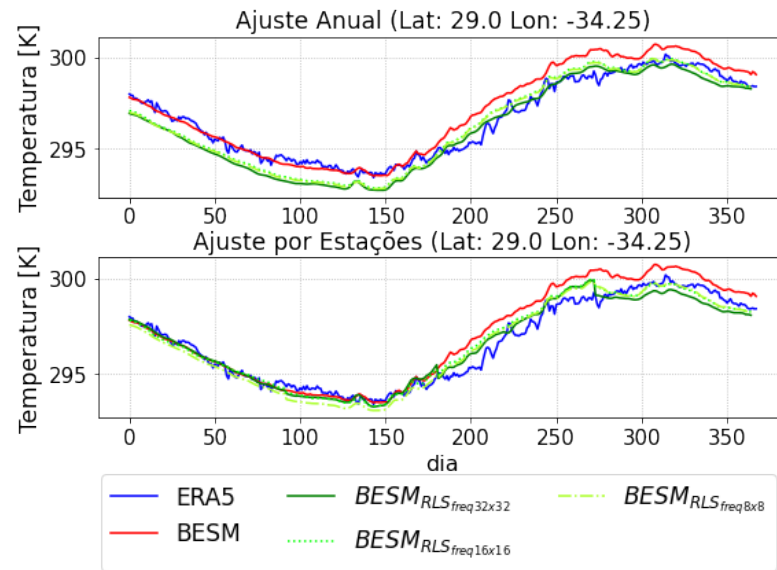


Figura 27 - Correção anual e por estações da temperatura da superfície do mar por meio do filtro adaptativo RLS no domínio da frequência (RLS_freq), variando-se o tamanho da região. Célula no Oceano Atlântico.

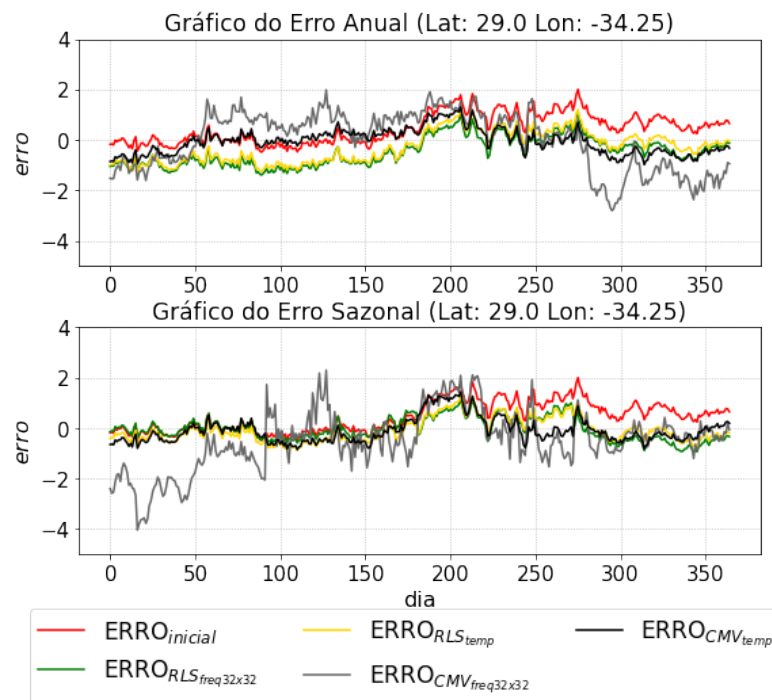


Figura 28 - Erros resultantes da correção da TSM produzida pelo BESM em relação ao dado observacional, corrigidas por ano e por estações. Célula no Oceano Atlântico.

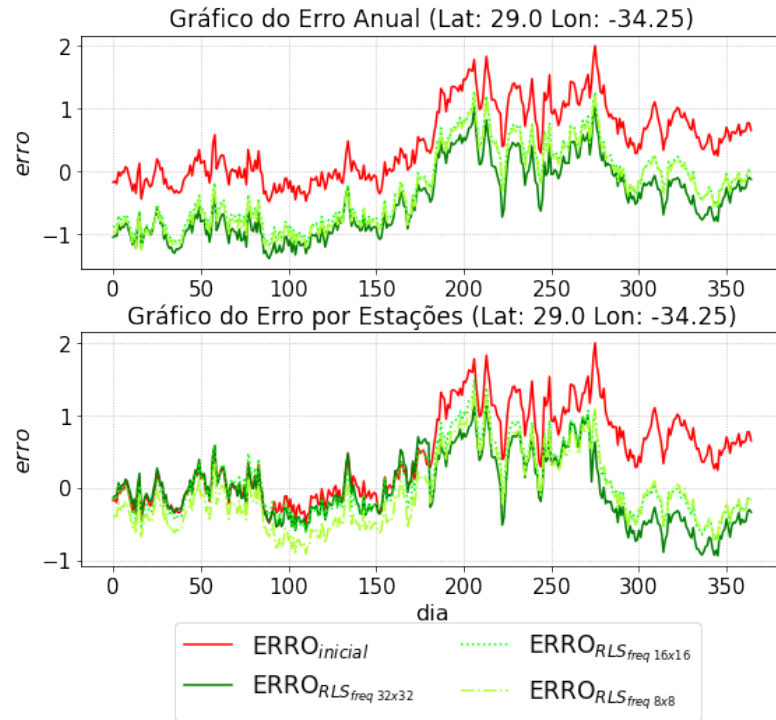


Figura 29 - Erros resultantes da correção da TSM produzida pelo BESM em relação ao dado observacional, corrigidas por ano e por estações por meio do método RLS_freq para $M = \{8, 6, 32\}$.

A Figura 30 apresenta uma visão ampliada a cada três meses do gráfico inferior da Figura 26 referente à correção de viés por estações. É importante lembrar que o modelo produz a previsão iniciando-se em novembro, nesse caso no ano de 2008. Observe que nos 6 primeiros meses, de novembro de 2008 a abril de 2009, o BESM produz uma previsão bastante aderente ao ERA5. De maio em diante, período em que há um descolamento entre BESM e ERA5, os métodos de correção de viés, entre eles, RLS_temp , RLS_freq e CMV_temp , ajustam a previsão do BESM, tornando-a mais aderente ao ERA5. Na Figura 31, analisa-se o desempenho de correção de viés do RLS_freq a cada três meses para diferentes valores de M . Em geral, a variação de M apresenta pouca influência na capacidade de correção do RLS_freq , nota-se apenas um descolamento da série corrigida com $M = 32$ no último trimestre da previsão Ago-Set-Out.

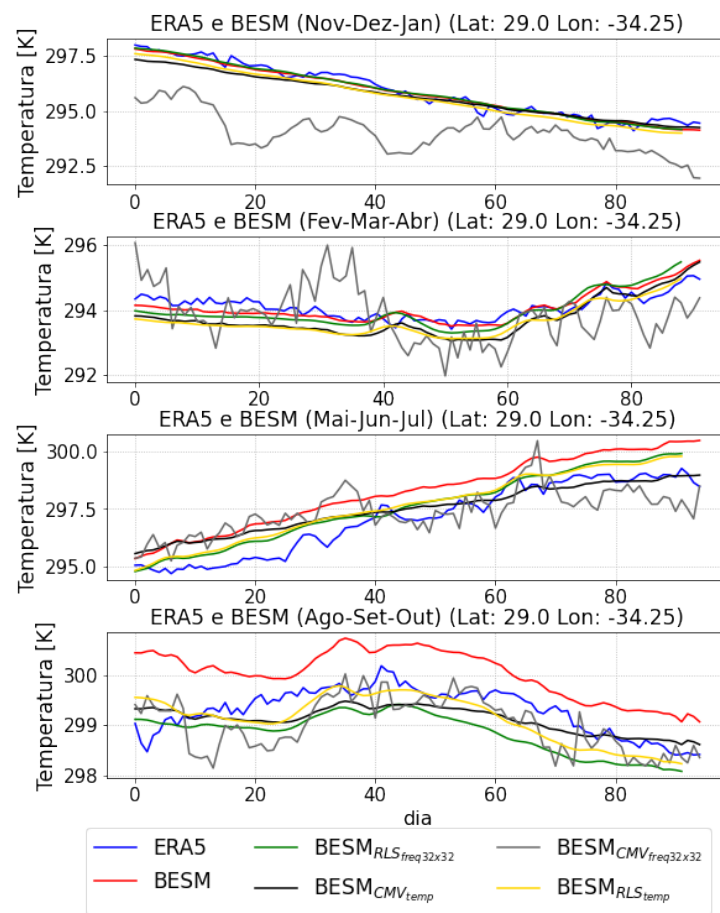


Figura 30 - Análise a cada três meses da correção por estações da TSM por meio de filtros RLS e CMV nos domínios do tempo e da frequência.

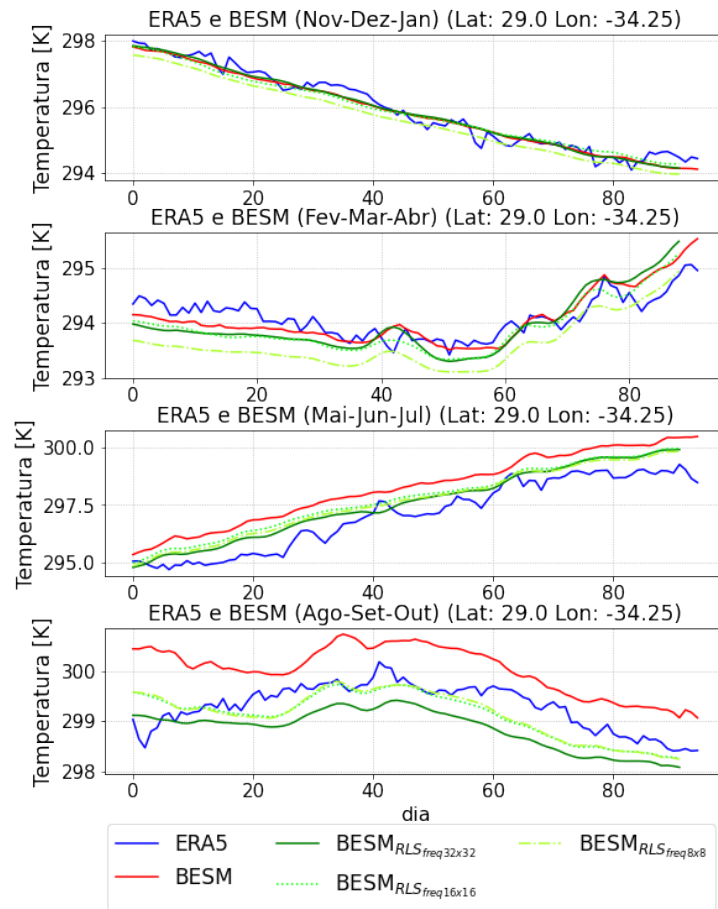


Figura 31 - Análise a cada três meses da correção por estações da TSM por meio de filtro RLS no domínio da frequência para regiões de tamanho $M = \{8, 16, 32\}$.

3.3.2 Célula no Oceano Índico

Os resultados apresentados a seguir se referem a correção da série temporal de TSM prevista pelo BESM associado à célula localizada no Oceano Índico na coordenada ($lat = -23,5; lon = 75,75$) indicada em vermelho no mapa da Figura 32. Além disso, apresentam-se as séries temporais desta localização produzidas pelo ERA5 e BESM, bem como aqueles decorrentes do ajuste de viés por meio dos algoritmos baseados em RLS e CMV nos domínios do tempo e da frequência. Tanto no ajuste anual quanto no sazonal (por estações), observa-se que o CMV_tmp e as técnicas baseadas em RLS apresentam desempenhos de correção semelhantes, aproximando melhor a série observacional em torno do dia 140 e a partir do dia 290. Já o CMV_freq produz séries inaderentes ao ERA5 com elevados desvios em termos de erro absoluto em ambos ajustes, anual e por estações. Na Figura 33, avalia-se o desempenho de correção do algoritmo RLS_freq para tamanhos de região variados. Nota-se que a variação de M provoca pouco impacto no desempenho da correção de viés.

Na Figura 34, são apresentados os erros subtraindo-se o sinal de TSM do ERA5 do sinal original ou corrigido do BESM. O gráfico acima se refere ao resultado da correção anual e o de baixo se refere a correção por estações. Nota-se que ao serem empregados, os métodos baseados em RLS promovem erros mais próximo de zero ao longo da série, quando a correção é realizada por estações. O mesmo ocorre para o método CMV_temp, contudo, no caso de CMV_freq o ajuste é deficiente em ambas as correções, anual e por estações.

A Figura 35 apresenta uma visão ampliada a cada três meses do gráfico inferior da Figura 33 referente a correção de viés por estações. É possível observar que nesse caso o BESM produz uma previsão bastante aderente ao ERA5 ao longo de todo período de simulação. Somente em março de 2009, os métodos RLS_temp e CMV_temp produzem uma correção tornando a série mais aderente à série observacional. Na Figura 36, analisa-se o desempenho de correção de viés do RLS_freq a cada três meses para diferentes valores de M . Em geral, a variação de M apresenta pouca influência na capacidade de correção do RLS_freq, nota-se apenas um descolamento da série corrigida com diferentes valores de M no segundo trimestre da previsão.

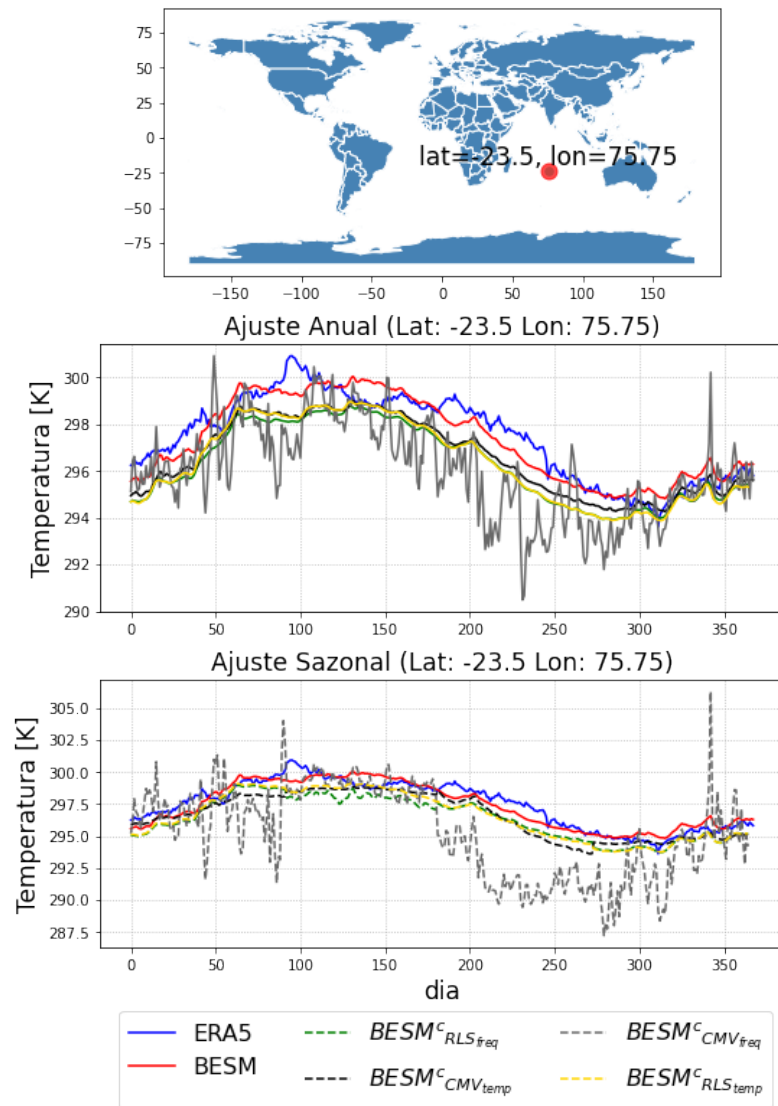


Figura 32 - Correção anual e por estações da temperatura da superfície do mar por meio dos métodos RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Índico.

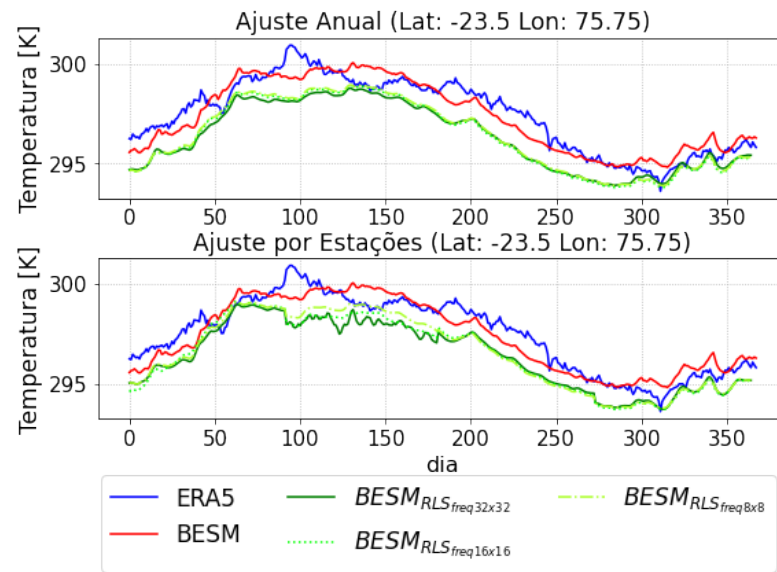


Figura 33 - Correção anual e por estações da temperatura da superfície do mar por meio do filtro adaptativo RLS no domínio da frequência (RLS_freq), variando-se o tamanho da região. Célula no Oceano Índico.

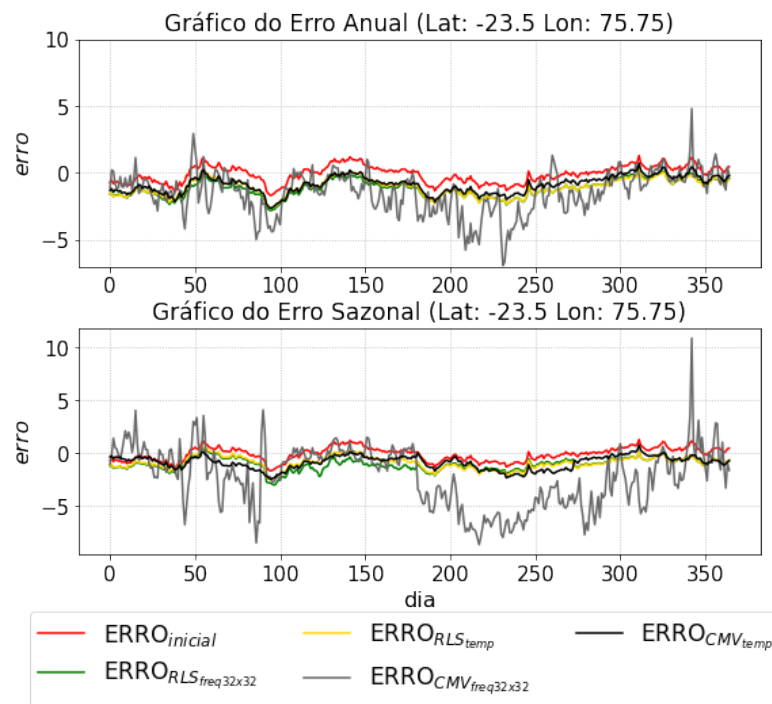


Figura 34 - Erros resultantes da correção da TSM produzida pelo BESM em relação ao dado observacional, corrigidas por ano e por estações. Célula no Oceano Índico.

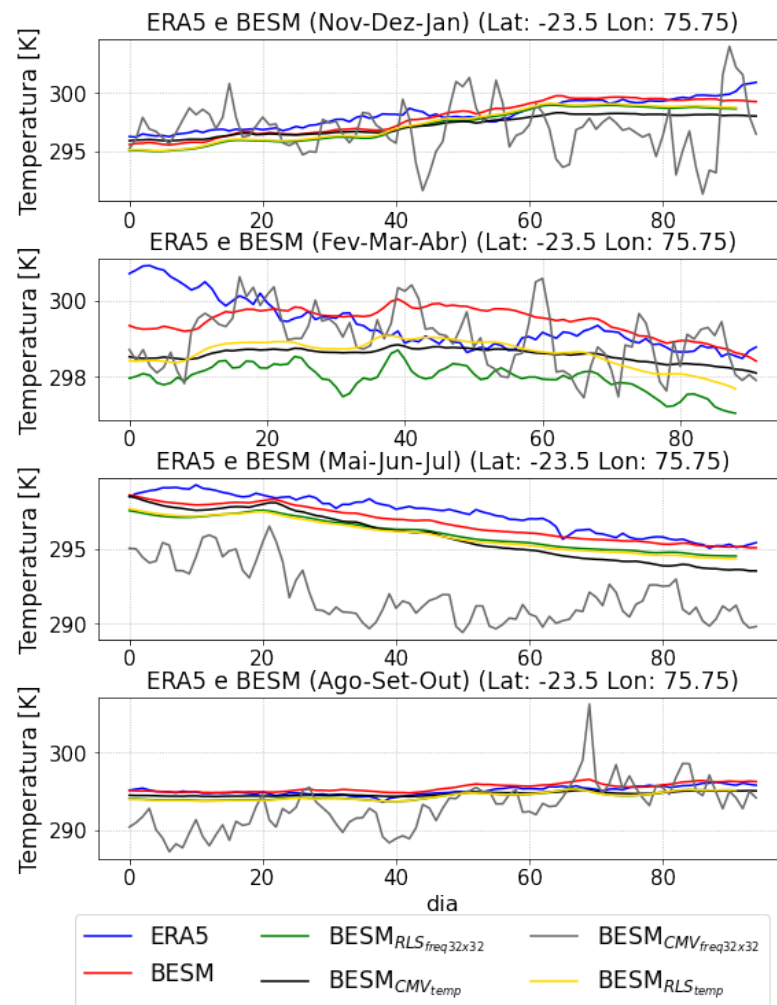


Figura 35 - Análise a cada três meses da correção por estações da TSM por meio de filtros RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Índico.

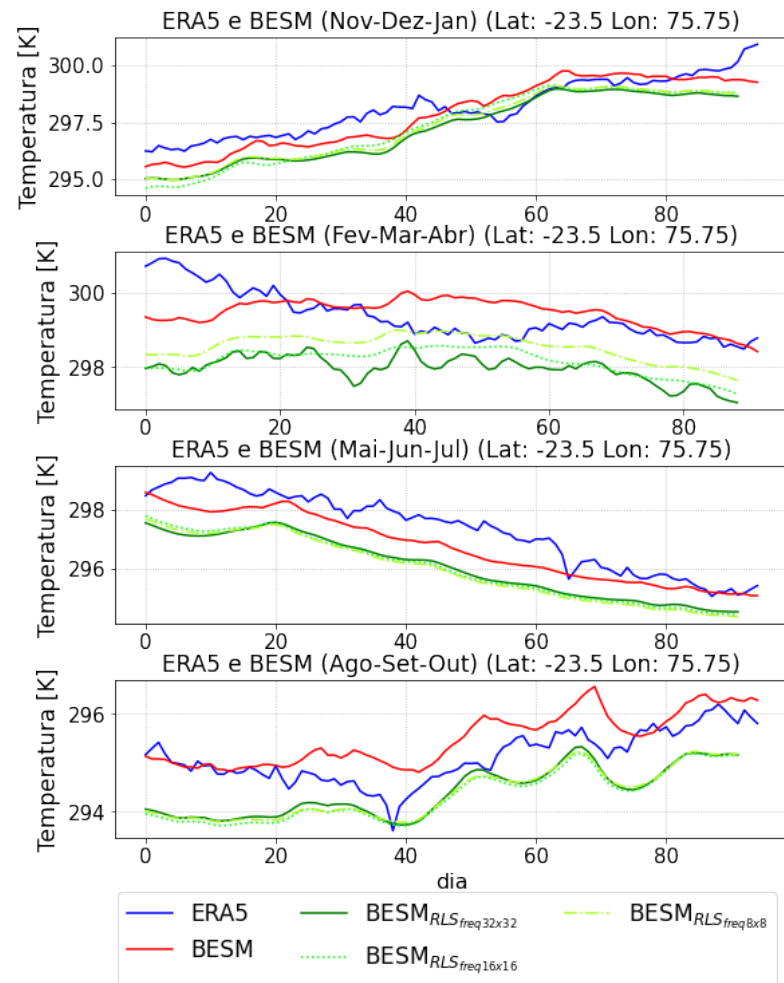


Figura 36 - Análise a cada três meses da correção por estações da TSM por meio de filtro RLS no domínio da frequência para regiões de tamanho $M = \{8, 16, 32\}$.

3.3.3 Célula no Oceano Pacífico

Os resultados apresentados a seguir se referem a correção da série temporal de TSM prevista pelo BESM associado à célula localizada no Oceano Pacífico na coordenada ($lat = -18, 5; lon = -84, 25$) indicada em vermelho no mapa da Figura 37. Além disso, apresentam-se as séries temporais desta localização produzidas pelo ERA5 e BESM, bem como aqueles decorrentes do ajuste de viés por meio dos algoritmos baseados em RLS e CMV nos domínios do tempo e da frequência.

No ajuste anual, observa-se que o CMV_tmp e as técnicas baseadas em RLS apresentam desempenhos de correção semelhantes, aproximando melhor a série observacional a partir do dia 75, sendo que o CMV_freq produz séries inaderentes ao ERA5 com elevados desvios em termos de erro absoluto. No ajuste sazonal, as técnicas baseadas em RLS apresentam desempenho superior às técnicas baseadas em CMV, em especial a partir do dia 80. Observa-se que tanto o CMV_freq quanto o CMV_temp produzem séries inaderentes ao ERA5 com elevados desvios em termos de erro absoluto. Na Figura 38, avalia-se o desempenho de correção do algoritmo RLS_freq para tamanhos de região variados. Nota-se que a variação de M também provoca pouco impacto no desempenho da correção de viés.

Na Figura 39, são apresentados os erros subtraindo-se o sinal de TSM do ERA5 do sinal original ou corrigido do BESM. O gráfico acima se refere ao resultado da correção anual e o de baixo se refere à correção por estações. Nota-se que ao serem empregados, os métodos baseados em RLS promovem erros mais próximos de zero ao longo da série, quando a correção é realizada por estações. O mesmo ocorre para o método CMV_temp no ajuste anual, contudo, no ajuste sazonal, este apresenta um descolamento ao ERA5 em torno do dia 160. Com relação ao CMV_freq, o ajuste é deficiente em ambas as correções, anual e sazonal.

A Figura 40 apresenta uma visão ampliada a cada três meses do gráfico inferior da Figura 38 referente à correção de viés por estações. Observe que nesse caso o BESM produz uma previsão com um descolamento significativo ao ERA5 a menos do período inicial de simulação. Os métodos baseados em RLS produzem uma correção tornando a série mais aderente à série observacional, havendo algum distanciamento no segundo trimestre Fev-Mar-Abr. Na Figura 41, analisa-se o desempenho de correção de viés do RLS_freq a cada três meses para diferentes valores de M . Em geral, a variação de M apresenta novamente pouca influência na capacidade de correção do RLS_freq, nota-se apenas um descolamento da série corrigida para $M = 32$ no segundo e quarto trimestres da previsão.

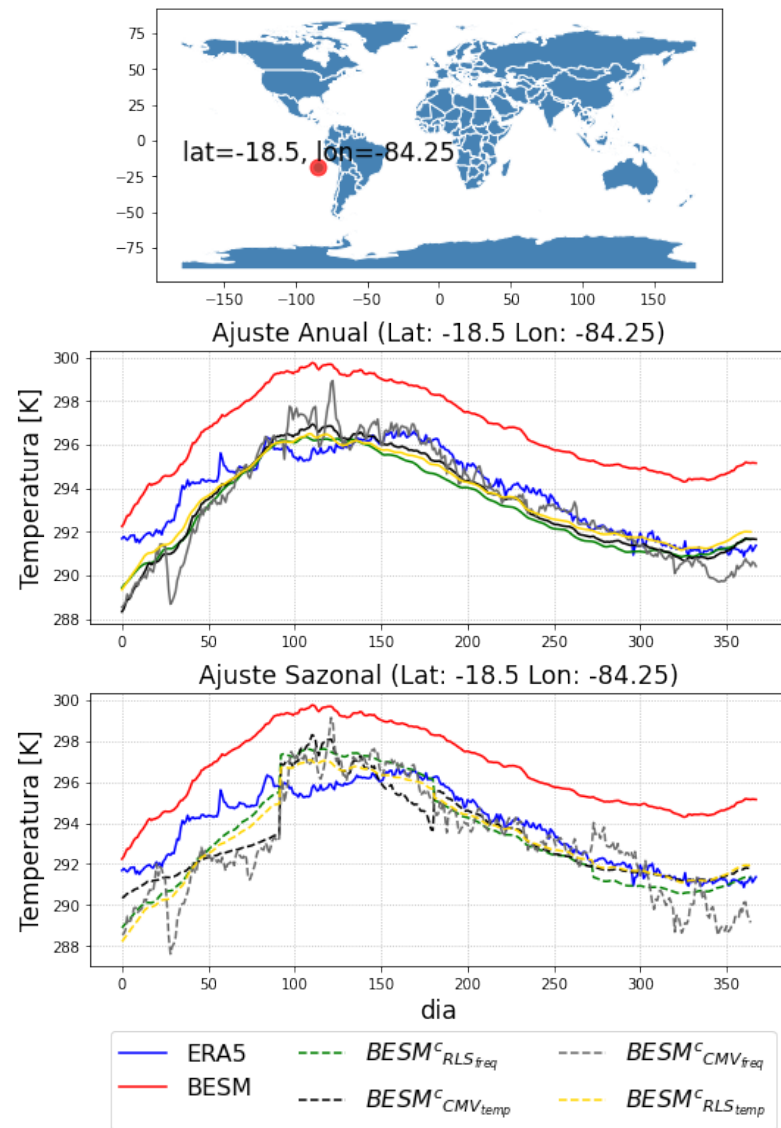


Figura 37 - Correção anual e por estações da temperatura da superfície do mar por meio dos métodos RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Pacífico.

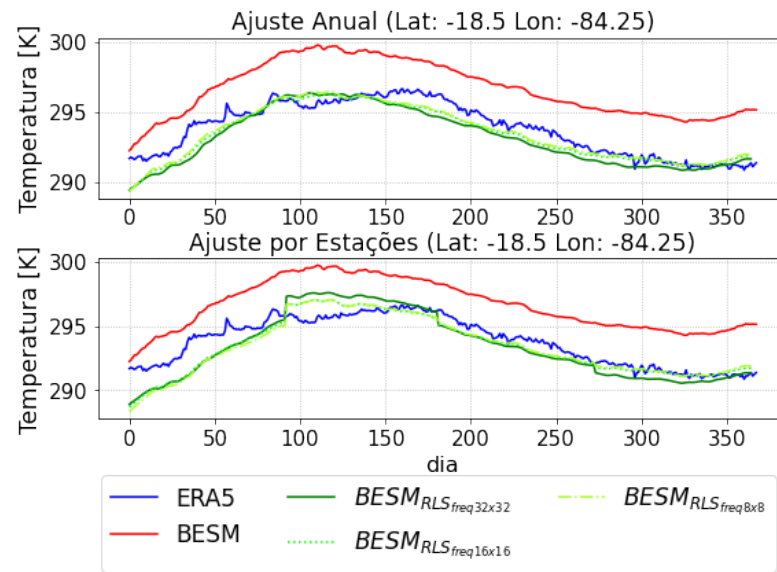


Figura 38 - Correção anual e por estações da temperatura da superfície do mar por meio do filtro adaptativo RLS no domínio da frequência (RLS_freq), variando-se o tamanho da região. Célula no Oceano Pacífico.

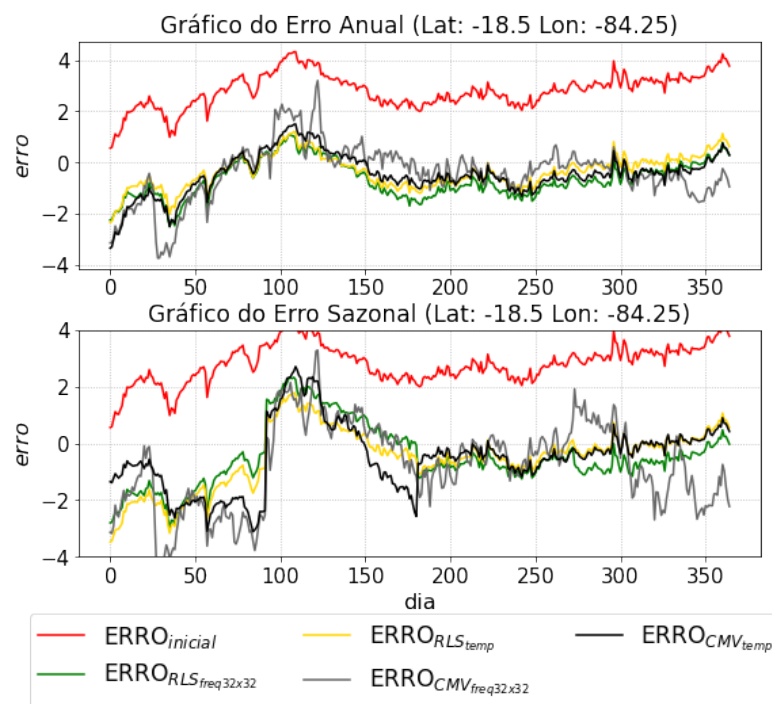


Figura 39 - Erros resultantes da correção da TSM produzida pelo BESM em relação ao dado observacional, corrigidas por ano e por estações. Célula no Oceano Pacífico.

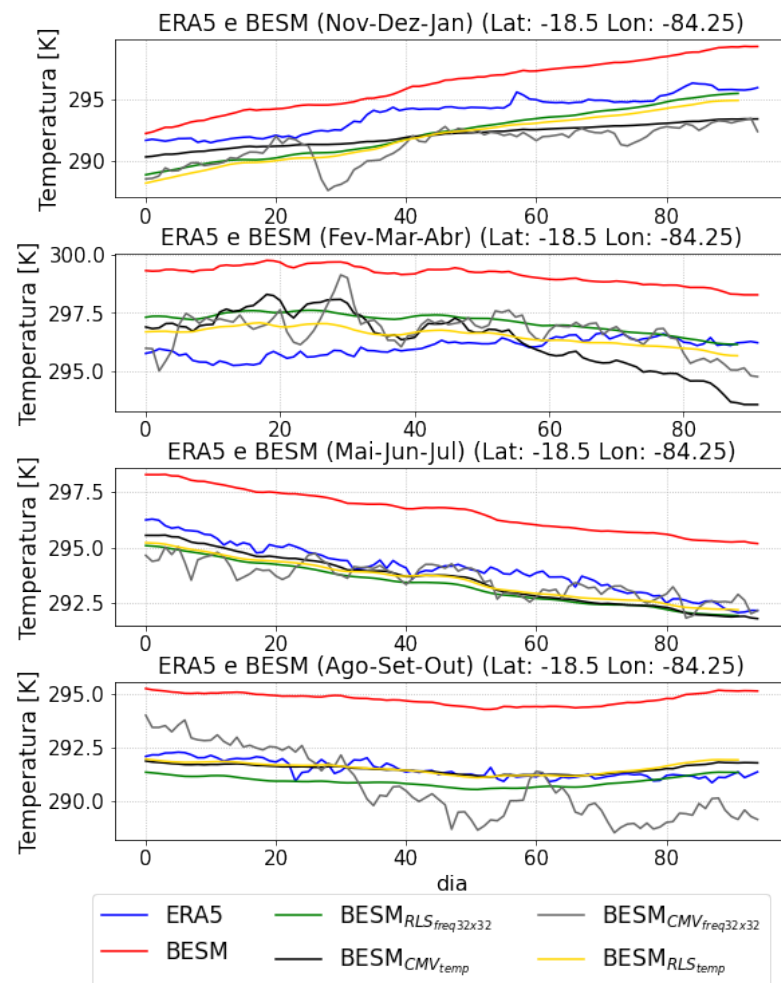


Figura 40 - Análise a cada três meses da correção por estações da TSM por meio de filtros RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Pacífico.

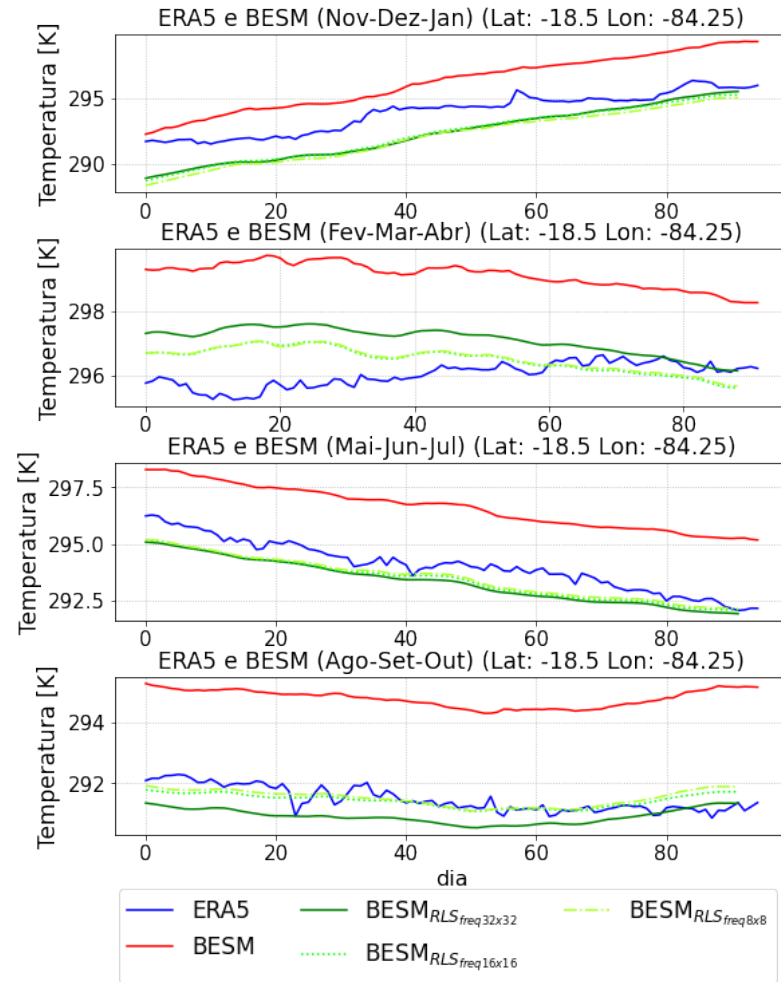


Figura 41 - Análise a cada três meses da correção por estações da TSM por meio de filtro RLS no domínio da frequência para regiões de tamanho $M = \{8, 16, 32\}$.

3.3.4 Célula no Oceano Glacial Antártico

Os resultados apresentados a seguir se referem à correção da série temporal de TSM prevista pelo BESM associado à célula localizada no Oceano Glacial Antártico na coordenada ($lat = -48,5; lon = -0,75$), indicada em vermelho no mapa da Figura 42. Além disso, apresentam-se as séries temporais desta localização produzidas pelo ERA5 e BESM, bem como aqueles decorrentes do ajuste de viés por meio dos algoritmos baseados em RLS e CMV nos domínios do tempo e da frequência. Observa-se no ajuste anual que os métodos baseados em RLS não são capazes de corrigir o BESM até o dia 220, já o CMV_temp produziu uma série corrigida mais aderente ao ERA5 ao longo de todo o período de previsão. No ajuste sazonal, os métodos baseados em RLS lograram êxito na correção do BESM, bem como o CMV_temp, produzindo séries corrigidas mais próximas da série observacional. Já o CMV_freq produz séries inaderentes ao ERA5 com elevados desvios em termos de erro absoluto em ambos ajustes, anual e por estações. Na Figura 43, avalia-se o desempenho de correção do algoritmo RLS_freq para tamanhos de região variados. Nota-se novamente que a variação de M provoca pouco impacto no desempenho da correção de viés.

Na Figura 44, são apresentados os erros subtraindo-se o sinal de TSM do ERA5 do sinal original ou corrigido do BESM. O gráfico acima se refere ao resultado da correção anual e o de baixo se refere à correção por estações. Nota-se que ao serem empregados, os métodos baseados em RLS promovem erros mais próximo de zero ao longo da série, quando a correção é realizada por estações. O mesmo ocorre para o método CMV_temp, contudo, no caso de CMV_freq o ajuste é deficiente em ambas as correções, anual e por estações.

A Figura 45 apresenta uma visão ampliada a cada três meses do gráfico inferior da Figura 43 referente à correção de viés por estações. Observe que nesse caso o BESM produz uma previsão com um distanciamento destacado com relação ao ERA5 ao longo de todo período de simulação. Da mesma forma, os métodos RLS_temp, RLS_freq e CMV_temp produzem uma série corrigida mais aderente à série observacional. Na Figura 46, analisa-se o desempenho de correção de viés do RLS_freq a cada três meses para diferentes valores de M . Em geral, a variação de M apresenta pouca influência na capacidade de correção do RLS_freq, nota-se apenas um descolamento da série corrigida para $M = 8$ no terceiro trimestre da previsão.

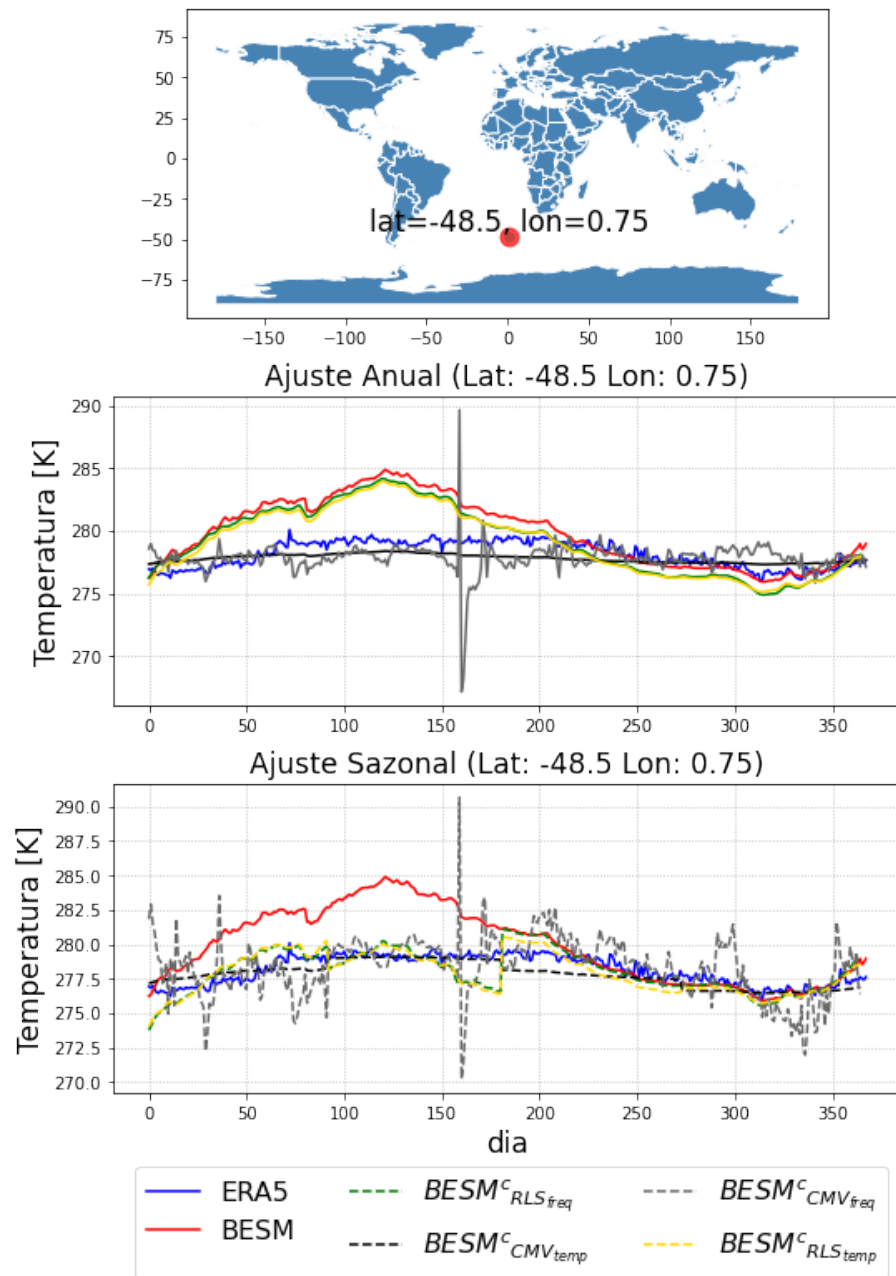


Figura 42 - Correção anual e por estações da temperatura da superfície do mar por meio dos métodos RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Glacial Antártico.

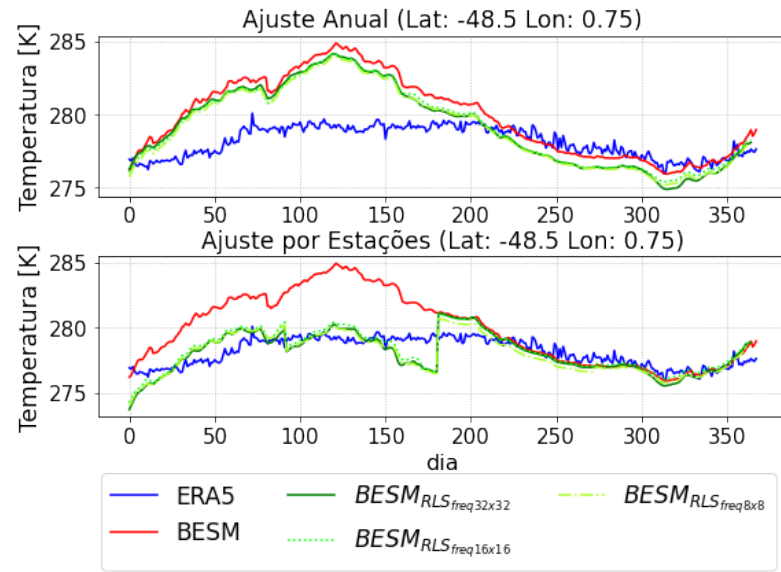


Figura 43 - Correção anual e por estações da temperatura da superfície do mar por meio do filtro adaptativo RLS no domínio da frequência (RLS_freq), variando-se o tamanho da região. Célula no Oceano Glacial Antártico.

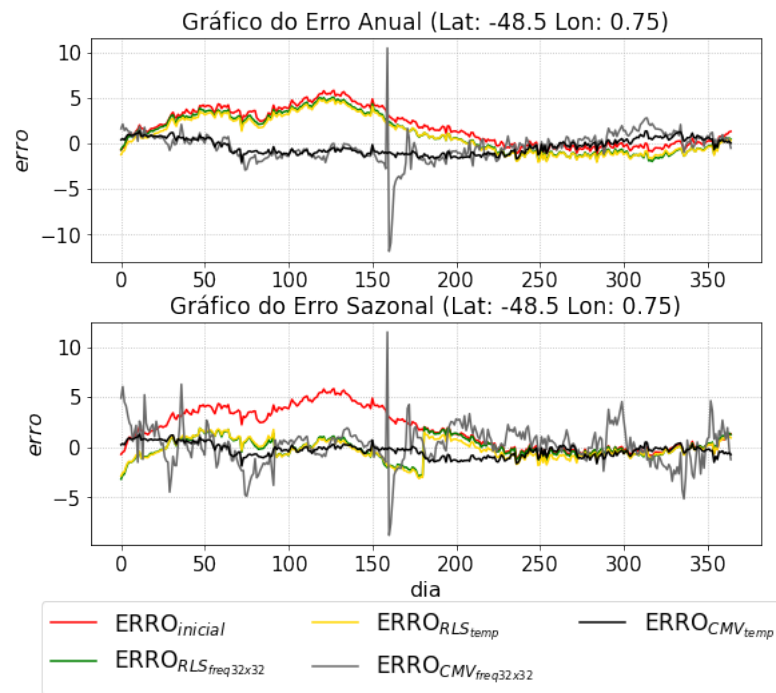


Figura 44 - Erros resultantes da correção da TSM produzida pelo BESM em relação ao dado observacional, corrigidas por ano e por estações. Célula no Oceano Glacial Antártico.

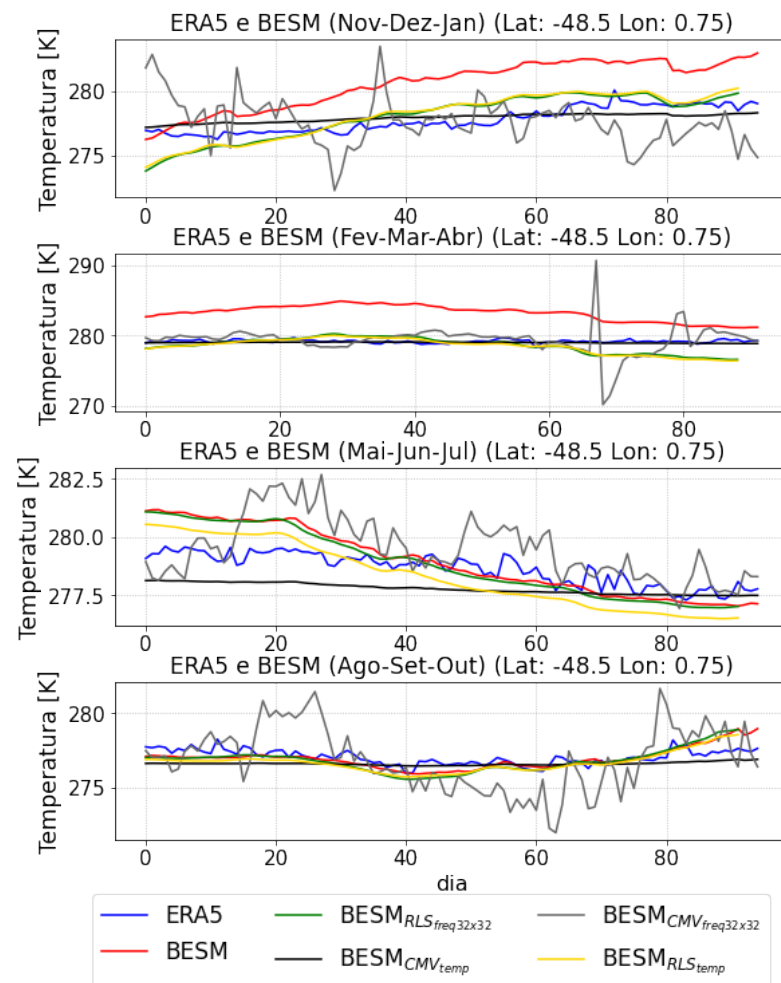


Figura 45 - Análise a cada três meses da correção por estações da TSM por meio de filtros RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Glacial Antártico.

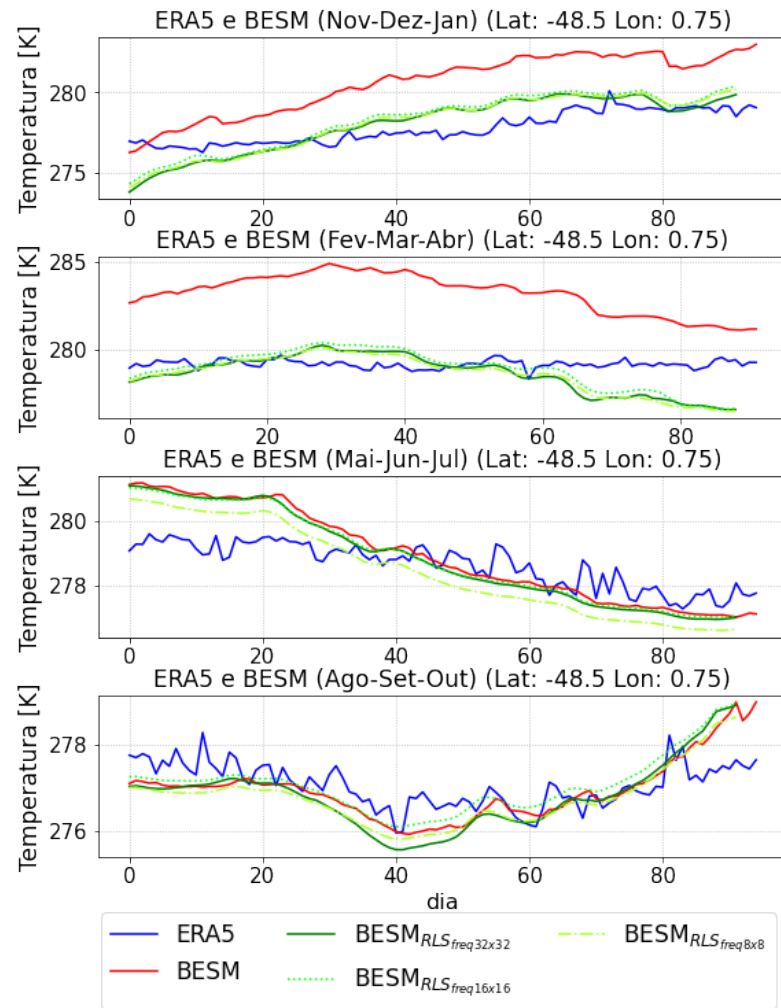


Figura 46 - Análise a cada três meses da correção por estações da TSM por meio de filtro RLS no domínio da frequência para regiões de tamanho $M = \{8, 16, 32\}$.

3.3.5 Célula no Oceano Glacial Ártico

Os resultados apresentados a seguir se referem a correção da série temporal de TSM prevista pelo BESM associado à célula localizada no Oceano Glacial Ártico na coordenada ($lat = 84; lon = 150, 75$) indicada em vermelho no mapa da Figura 47. Além disso, apresentam-se as séries temporais desta localização produzidas pelo ERA5 e BESM, bem como aqueles decorrentes do ajuste de viés por meio dos algoritmos baseados em RLS e CMV nos domínios do tempo e da frequência. Observe que quase não há variação de temperatura nessa localização ao longo do período de previsão. Tanto no ajuste anual quanto no sazonal (por estações), observa-se que o CMV_tmp e o RLS_temp apresentam desempenhos de correção semelhantes, aproximando melhor a série observacional. Já o método RLS_freq apresenta um descolamento da série observacional a partir do dia 260, em ambos os ajustes. O CMV_freq produz séries inaderentes ao ERA5 com elevados desvios em termos de erro absoluto, no ajuste anual, ao longo de todo o período, e no ajuste sazonal, também a partir do dia 260. Na Figura 48, avalia-se o desempenho de correção do algoritmo RLS_freq para tamanhos de região variados. Nota-se uma maior deficiência na correção de viés a partir do dia 260, para $M = 32$.

Na Figura 49, são apresentados os erros subtraindo-se o sinal de TSM do ERA5 do sinal original ou corrigido do BESM. O gráfico acima se refere ao resultado da correção anual e o de baixo se refere a correção por estações. Nota-se que ao serem empregados, os métodos de correção no tempo, RLS_temp e CMV_temp promovem erros mais próximos de zero ao longo da série, tanto na correção anual quanto sazonal. No caso das técnicas de correção no domínio da frequência, o CMV_freq apresenta um ajuste deficiente em ambas as correções, anual e por estações, ao longo de todo o período de simulação, enquanto que a deficiência no ajuste por meio do RLS_freq se destaca a partir do dia 260.

A Figura 50 apresenta uma visão ampliada a cada três meses do gráfico inferior da Figura 48 referente à correção de viés por estações. Observe que no primeiro e no segundo trimestre de simulação, todos os métodos de correção avaliados produziram séries mais aderentes à série de TSM do ERA5. A partir do último terço do terceiro trimestre, surge um descolamento entre as séries corrigidas pelos métodos que atuam no domínio da frequência CMV_freq e RLS_freq, que se torna mais crítico e evidenciado no quarto trimestre. Na Figura 51, analisa-se o desempenho de correção de viés do RLS_freq a cada três meses para diferentes valores de M . Em geral, a variação de M apresenta pouca influência na capacidade de correção do RLS_freq, nos dois primeiros trimestres, e a partir do terceiro trimestre nota-se um descolamento da série corrigida para $M = 32$

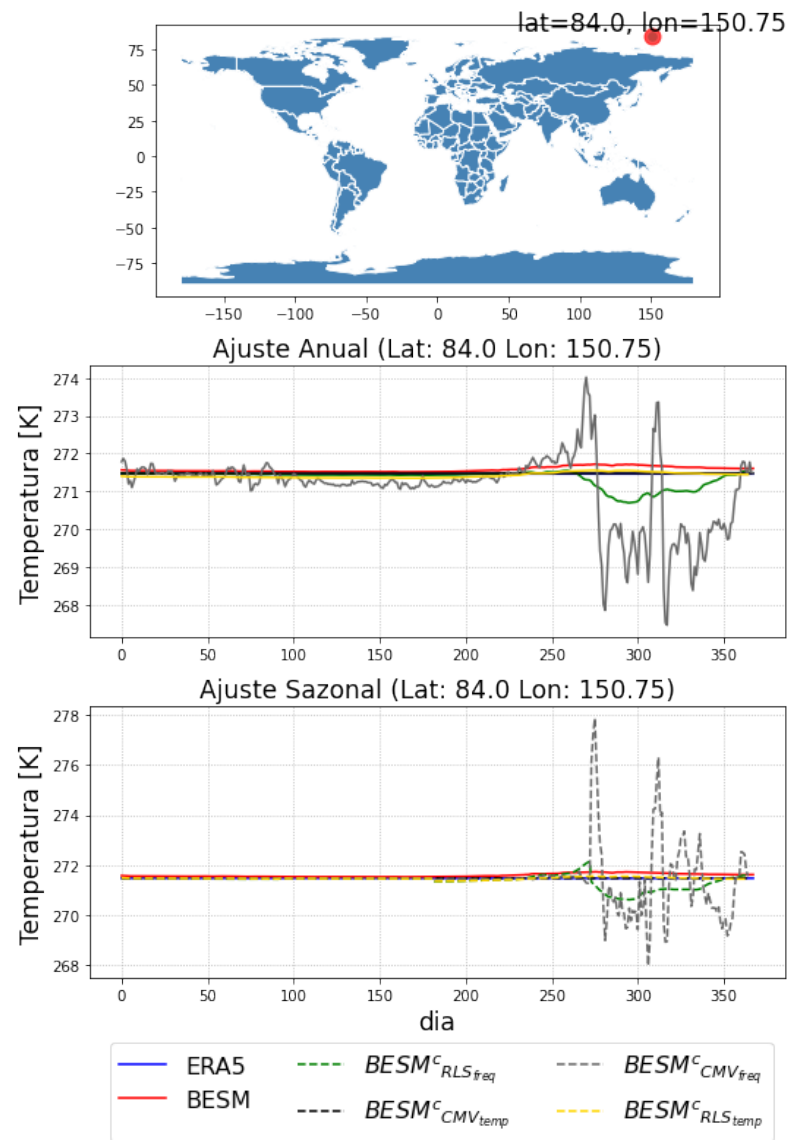


Figura 47 - Correção anual e por estações da temperatura da superfície do mar por meio dos métodos RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Glacial Ártico.

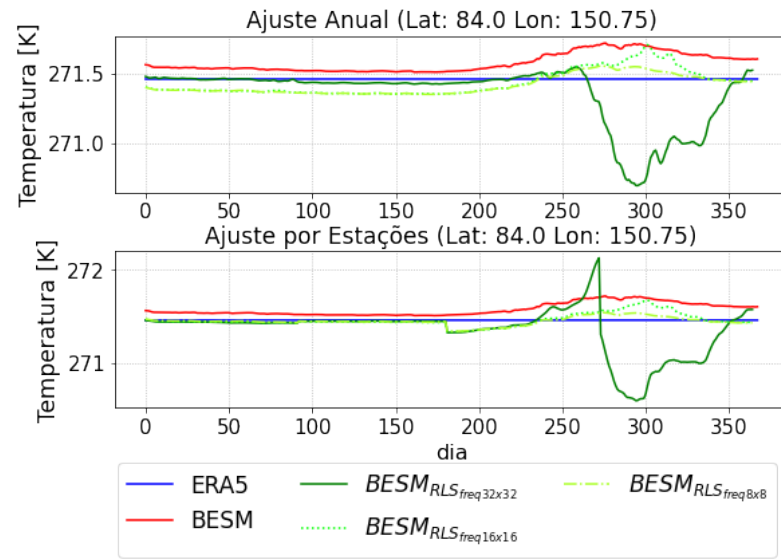


Figura 48 - Correção anual e por estações da temperatura da superfície do mar por meio do filtro adaptativo RLS no domínio da frequência (RLS_freq), variando-se o tamanho da região. Célula no Oceano Glacial Ártico.

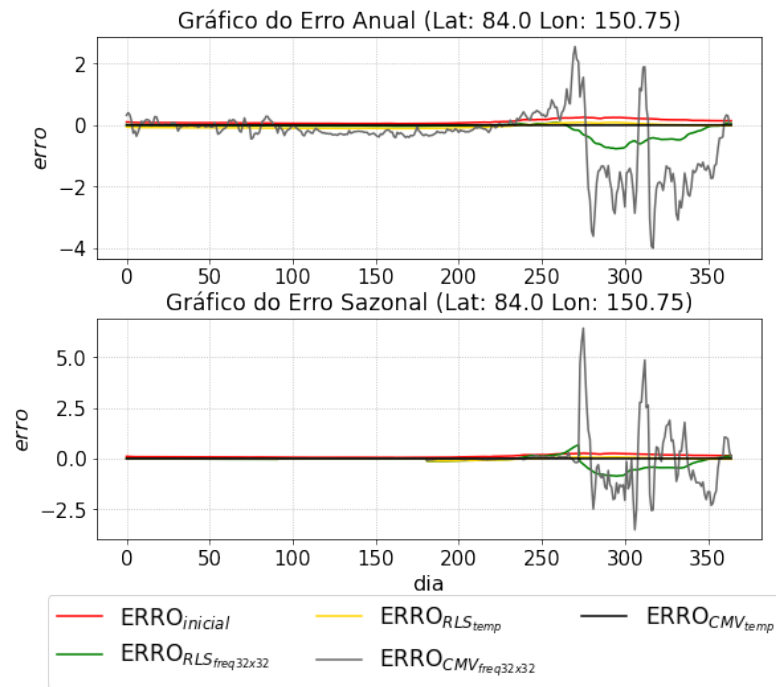


Figura 49 - Erros resultantes da correção da TSM produzida pelo BESM em relação ao dado observacional, corrigidas por ano e por estações. Célula no Oceano Glacial Ártico.

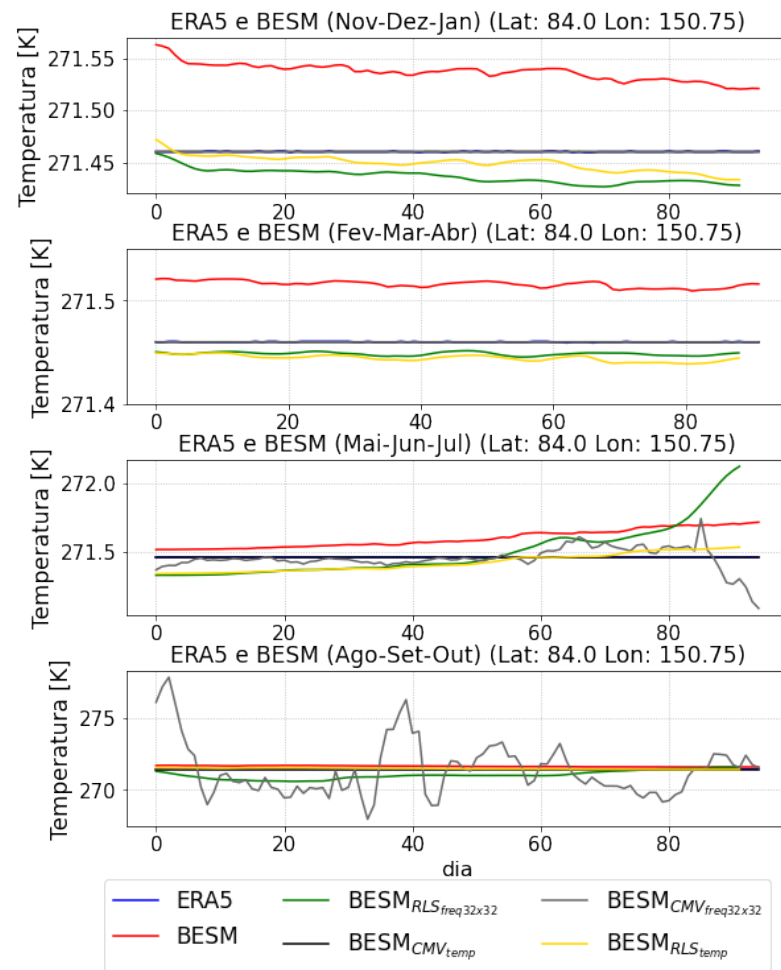


Figura 50 - Análise a cada três meses da correção por estações da TSM por meio de filtros RLS e CMV nos domínios do tempo e da frequência. Célula no Oceano Glacial Ártico.

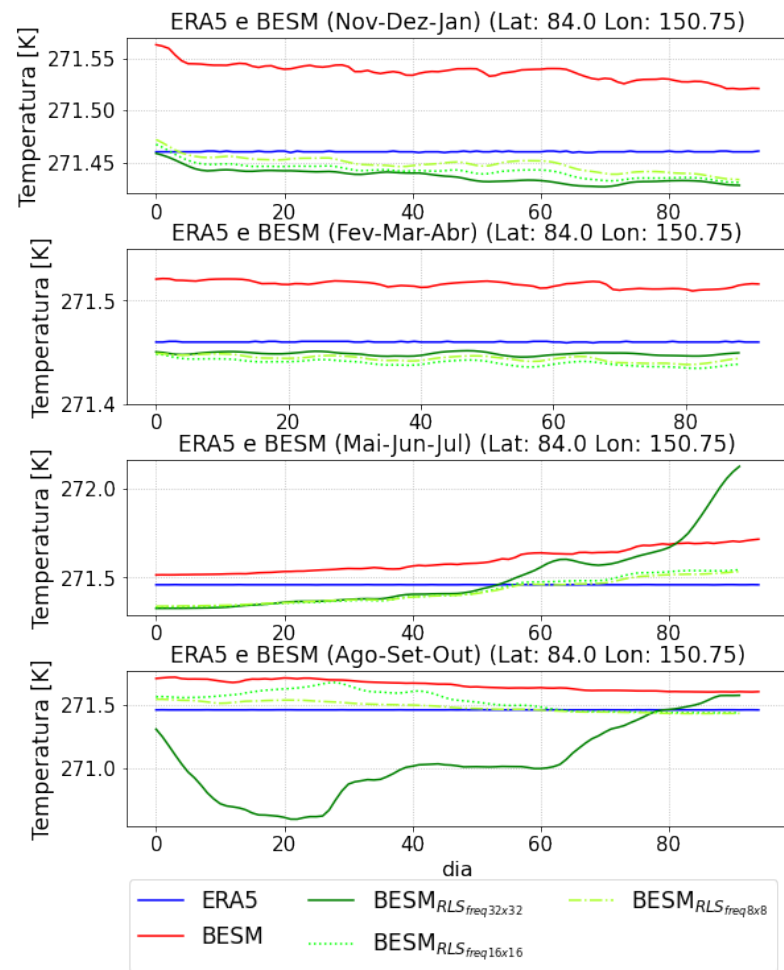


Figura 51 - Análise a cada três meses da correção por estações da TSM por meio de filtro RLS no domínio da frequência para regiões de tamanho $M = \{8, 16, 32\}$.

CONCLUSÃO

Nesta dissertação, propôs-se a correção de viés das previsões de temperatura da superfície do mar do modelo BESM a partir do emprego de filtragem adaptativa RLS. A filtragem é aplicada no domínio da frequência por meio da Transformada Discreta de Cosseno Bidimensional (DCT-2D). A correção de viés proposta foi aplicada nas séries horárias de previsões da temperatura da superfície do mar. As previsões foram produzidas para o horizonte de 1 ano, iniciadas em novembro, dos anos de 2007 e 2008. Comparou-se o desempenho do método de filtragem adaptativa RLS na frequência, ao método RLS no tempo e aos métodos de correção por média e variância (CMV) no tempo e na frequência. De modo geral, verificou-se que o método proposto apresenta uma capacidade de correção equivalente aos do método RLS no tempo e CMV no tempo. Mais especificamente, quanto maior o erro da previsão da TSM em relação aos dados observacionais, melhor é o desempenho do método de correção.

A utilização de um modelo de previsão por estações baseado em filtragem adaptativa mostrou-se mais trabalhosa e eficaz, pois o ajuste da codificação dos filtros não é uma tarefa trivial, sendo necessárias numerosas análises antes de sua implementação.

A filtragem adaptativa FIR com ordem de filtro ($L = 4$) gerou resultados relevantes de redução de erros, conforme indicado pelas métricas de avaliação - erro quadrático médio (MSE), erro absoluto maior (LAE) e erro absoluto maior percentual (PLAE). Para alguns casos, para ordem de filtro maiores do que $L = 4$ são melhores.

É importante considerar a localização geográfica em que o ajuste é empregado. Nos Oceanos Ártico e Antártico, o filtro RLS não consegue ajustar porque a TSM é praticamente constante.

A seleção do coeficiente do filtro “ideal” depende das características específicas de cada modelo (BESM e ERA5), uma escolha cuidadosa do filtro adequado pode melhorar significativamente a precisão do ajuste anual e por estações.

Para trabalhos futuros propõe-se o estudo da utilização de informações climáticas com vieses mais intensos e outras variáveis externas ao modelo de previsão.

REFERÊNCIAS

- AHMED, N.; NATARAJAN, T.; RAO, K. R. Discrete Cosine Transform. *IEEE Transactions on Computers*, IEEE, v. 100, n. 1, p. 90–93, 1974.
- AMARAL, L. F. C. *Um algoritmo não quadrático baseado no RLS estendido*. 88 p. Tese (Mestrado), São Luís, 2018. Programa de Pós-Graduação em Engenharia de Eletricidade/CCET.
- ARDALAN, S. Floating-point error analysis of recursive least-squares and least-mean-squares adaptive filters. *IEEE transactions on circuits and systems*, IEEE, v. 33, n. 12, p. 1192–1208, 1986.
- BELLIFEMINE, F. et al. Statistical analysis of the 2D-DCT coefficients of the differential signal for images. *Signal Processing: Image Communication*, Elsevier, v. 4, n. 6, p. 477–488, 1992. ISSN 0923-5965. Disponível em: <<https://www.sciencedirect.com/science/article/pii/092359659290032B>>.
- BENOIT, M.; MARCOS, F.; BECQ, F. Development of a third generation shallow-water wave model with unstructured spatial meshing. In: *COASTAL ENGINEERING 1996*. [S.l.: s.n.], 1996. p. 465–478.
- BITENCOURT, D.P. et al. The climatology of cold and heat waves in brazil from 1961 to 2016. *International Journal of Climatology*, Wiley Online Library, v. 40, p. 2464–2478, 2019. Disponível em: <<https://doi.org/10.1002/joc.6345>>.
- BITENCOURT, D. P.; SATYAMURTY, P. Demonstração matemática de possíveis diferenças entre a correlação de duas variáveis simples e a correlação de duas variáveis derivadas. *Revista Brasileira de Meteorologia*, Sociedade Brasileira de Meteorologia, v. 31, n. 3, p. 324–329, 2016.
- BONATTI, J. P. Modelo de circulação geral atmosférico do CPTEC: descrição e avaliação. *Climanálise Especial, edição comemorativa*, v. 10, 1996.
- BRANKOVIĆ, Č.; PALMER, T. N.; FERRANTI, L. Predictability of seasonal atmospheric variations. *Journal of Climate*, v. 7, n. 2, p. 217–237, 1994.
- CALDERÓN, M. E. G. *Impacto das condições de contorno químicas na representação do ozônio troposférico na Região Metropolitana de São Paulo com o modelo WRF-Chem*. Tese (Doutorado) — Universidade de São Paulo, 2015.
- CAPISTRANO, V. B. et al. Assessing the performance of climate change simulation results from BESM-OA2.5 compared with a CMIP5 Model Ensemble. *Geoscientific Model Development*, Copernicus GmbH, v. 13, n. 5, p. 2277–2296, 2020.
- CATALDI, M. et al. Análise das previsões de precipitação obtidas com a utilização do modelo eta como insumo para modelos de previsão semanal de vazão natural. *Revista Brasileira de Recursos Hídricos*, v. 12, n. 3, p. 5–12, 2007.
- CAVALCANTI, I. F. et al. Global climatological features in a simulation using the CPTEC–COLA AGCM. *Journal of Climate*, v. 15, n. 21, p. 2965–2988, 2002.

CENTRO EUROPEU DE PREVISÕES METEOROLÓGICAS A MÉDIO PRAZO. *Serviço de mudanças climáticas*. 2022. Disponível em: <<https://climate.copernicus.eu/climate-reanalysis>>. Acesso em: 29 jan. 2022.

CHOU, S. C. Modelo regional ETA. *Climanálise-Boletim de Monitoramento e Análise Climática*, Cachoeira Paulista, SP, Brasil, v. 1, n. especial, 1996.

CHOU, S. C.; NOBRE, P. Avaliação de modelos globais e regionais climáticos. COPPE, v. 1, p. 278–319, 2015.

CIOFFI, J. Limited-precision effects in adaptive filtering. *IEEE Transactions on circuits and systems*, IEEE, v. 34, n. 7, p. 821–833, 1987.

CLARKE, R. T.; DIAS, PL S. As necessidades de observação e monitoramento dos ambientes brasileiros quanto aos recursos hídricos. *Centro de Gestão e Estudos Estratégicos, Secretaria Técnica do Fundo Setorial de Recursos Hídricos*, 2002.

DIAS, E. M.; PALMEIRA, A. C. P. De Almeida; LIBONATI, R. Estimativa dos períodos de retorno da velocidade e rajada do vento e altura significativa das ondas no Atlântico sudoeste. *Sistemas e Gestão*, v. 16, n. 1, p. 84–100, 2021.

DINIZ, P. S. R. *Data Selective Adaptive Filtering*. Boston, MA: Springer US, 2013. 249–304 p. ISBN 978-1-4614-4106-9. Disponível em: <https://doi.org/10.1007/978-1-4614-4106-9_6>.

DINIZ, P. S. R. et al. *Adaptive filtering*. [S.l.]: Springer, 1997. v. 4.

DINIZ, P. S. R.; SILVA, E. Da; NETTO, S. L. *Digital Signal Processing: System analysis and design*. 2nd. ed. Cambridge: Cambridge University Press, 2010.

DINIZ, P. S. R.; WERNER, S. Set-membership binormalized data-reusing LMS algorithms. *IEEE Transactions on Signal Processing*, v. 51, n. 1, p. 124–134, 2003.

FREITAS, A. R. *Filtragem adaptativa aplicada à redução de desvios em séries temporais de previsão numérica climática*. 2015. Dissertação de Mestrado — Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 28 de Setembro, 2015.

FREITAS, A. R. et al. Filtragem adaptativa para a redução de desvios em séries temporais de previsão numérica climática. In: *ANAIS DO XXXIII SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES E PROCESSAMENTO DE SINAIS 2015*. Juiz de Fora, Brasil: Sociedade Brasileira de Telecomunicações (SBTrT), 2015. p. 1–6.

GOODWIN, G. C. Dynamic system identification: experiment design and data analysis. *Mathematics in science and engineering*, Academic press, v. 136, 1977.

HAYKIN, S. S. *Adaptive filter theory*. [S.l.]: Pearson Education India, 2008.

HERSBACH, H. The ERA5 Atmospheric Reanalysis. In: *AGU Fall Meeting Abstracts*. [S.l.: s.n.], 2016. p. NG33D–01.

HERSBACH, H. et al. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, Wiley Online Library, v. 146, n. 730, p. 1999–2049, 2020.

HOFF, P. D. Separable covariance arrays via the tucker product, with applications to multivariate relational data. *Bayesian Anal.*, v. 6, n. 2, p. 179 – 196, 2011.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. *O Modelo Brasileiro do Sistema Terrestre (BESM)*. n.d. Accessed February 18, 2023. Disponível em: <<http://www.inpe.br/besm/relatorios/>>.

JAIN, A. K. *Fundamentals of digital image processing*. 1st. ed. Englewood Cliffs, NJ: Prentice-Hall, 1989.

KHAYAM, S. A. The Discrete Cosine Transform (DCT): theory and application. *Michigan State University, CiteSeer*, v. 114, p. 1–31, 2003.

LAPRISE, R. et al. Challenging some tenets of regional climate modelling. *Meteorology and Atmospheric Physics*, Springer, v. 100, n. 1, p. 3–22, 2008.

LI, J. et al. A refined regional model for estimating pressure, temperature, and water vapor pressure for geodetic applications in China. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, v. 12, n. 11, p. 1713, 2020.

MARTINS, G. et al. Características da circulação atmosférica e precipitação utilizando o modelo acoplado MCGA/IBIS. *Boletim de Geografia*, v. 33, n. 2, p. 100–114, 2015.

MIGUEL, J. C. H.; MAHONY, M.; MONTEIRO, M. S. A. “Geopolítica Infrastructural” del conocimiento climático: el Modelo Brasileño del Sistema Terrestre y la brecha de conocimiento norte-sur. *Sociologias*, SciELO Brasil, v. 21, p. 44–75, 2019.

MIGUEL, J. C. H.; MONTEIRO, M. S. A. Mudanças climáticas, tecnociência e geopolítica: Um modelo do sistema terrestre brasileiro e a soberania na produção de futuros climáticos. *Anais da ReACT-Reunião de Antropologia da Ciência e Tecnologia*, v. 2, n. 2, 2015.

Network Common Data Format (NetCDF). *Network Common Data Form (NetCDF)*. 2022. Accedido en marzo 16, 2022. Disponível em: <<https://www.unidata.ucar.edu/software/netcdf/>>.

NOBRE, P. et al. Climate simulation and change in the brazilian climate model. *Journal of Climate*, American Meteorological Society, v. 26, n. 18, p. 6716–6732, 2013.

OLMOS, G. P. et al. Análisis de sensibilidad en la construcción de ensembles de RCMs de variables meteorológicas. In: *III WORKSHOP EN INVESTIGACIÓN AGROALIMENTARIA: WiA3.14*. Cartagena: [s.n.], 2014. p. 216–219. ISBN 978-84-697-1358-7.

OPPENHEIM, A. V. *Sinais e Sistemas*. 2nd. ed. São Paulo, Brazil: Prentice-Hall, 2010.

PERDOMO, J. A. P.; GONZÁLEZ, J. M. R.; CAMINO, E. R. La nueva generación de modelos climáticos. El proyecto europeo EC-Earth. *Acta de las Jornadas Científicas de la Asociación Meteorológica Española*, n. 34, 2016.

PEREIRA, R. M. *Climatologia e conforto térmico do espaço urbano: um estudo de caso do município de Humaitá-AM, entre 2016 a 2020*. 61 p. Dissertação (Ciências Ambientais) — Universidade Federal do Amazonas, Humaitá (AM), 2022. Dissertação de mestrado.

- PINTO, L. F. G. et al. *Filtragem adaptativa para a redução de desvios em séries temporais de previsão numérica climática*. Dissertação (Mestrado) — Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brasil, 2015.
- PINTO, L. F. G. et al. Redução de desvios de previsão climática usando filtragem adaptativa no domínio da frequência. In: *XXXVI SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES E PROCESSAMENTO DE SINAIS 2018 (SBrT2018)*. Rio de Janeiro: [s.n.], 2018.
- RAO, K. R.; YIP, P. *Discrete cosine transform: algorithms, advantages, applications*. [S.l.]: Academic press, 2014.
- RASCHKA, S.; MIRJALILI, V. *Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2*. [S.l.]: Packt Publishing Ltd, 2019.
- ROCHA, H. R. Modelagem da interação biosfera-atmosfera para áreas de floresta e de pastagem na Amazônia. *Inpe-4391-rpe/593*, INPE, 1992.
- ROECKNER, E. et al. *The Atmospheric General Circulation Model ECHAM-4: Model Description and Simulation of Present-day Climate*. [S.l.], 1996.
- SAMPAIO, G.; LEITE, P.; SILVA, D. Da. Evolução dos modelos climáticos e de previsão de tempo e clima. *Revista USP, São Paulo*, n. 103, p. 41–54, 2014.
- SERRANO, P. C. S. *Projeções e curvas notáveis da esfera*. Tese (Doutorado) — Universidade da Beira Interior, 2012.
- SHUKLA, J. Predictability in the Midst of Chaos: A Scientific Basis for Climate Forecasting. *Science*, American Association for the Advancement of Science, v. 282, n. 5389, p. 728–731, 1998.
- SILVA, C. C. S. et al. *UM ALGORITMO TIPO RLS BASEADO EM SUPERFÍCIES NÃO QUADRÁTICAS*. Tese de Doutorado — Universidade Federal do Maranhão, 2013.
- SUN, L.; GRAHAM, N. Climate simulation over east africa using the ncep regional spectral model. *Submitted to J. Geophys. Res.*, 2001.
- TAYLOR, K. E.; STOUFFER, R. J.; MEEHL, G. A. An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, American Meteorological Society, v. 93, n. 4, p. 485–498, 2012.
- TCHEOU, M. P. et al. Reducing Forecast Errors of a Regional Climate Model Using Adaptive Filters. *Applied Sciences*, MDPI, v. 11, n. 17, p. 8001, 2021.
- THE EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS (ECMWF). *What are the changes from ERA-Interim to ERA5 and ERA5-Land?* 2022. Disponível em: <<https://confluence.ecmwf.int/pages/viewpage.action?pageId=74764925>>. Acesso em: 1 de jul. de 2022.
- _____. *First ERA5-Land dataset to be released this spring*. 2023. Disponível em: <<https://www.ecmwf.int/en/newsletter/159/news/first-era5-land-dataset-be-released-spring>>. Acesso em: 20 Feb. 2023.

VEIGA, A. P. S. A cobertura da apresentação do modelo brasileiro do sistema terrestre pela mídia digital: Análise qualitativa e comparativa. *Razón y Palabra*, Universidad de los Hemisferios, n. 84, 2013.

VEIGA, S. F. et al. The Brazilian Earth System Model ocean–atmosphere (BESMOA) version 2.5: evaluation of its CMIP5 historical simulation. *Geoscientific Model Development*, Copernicus GmbH, v. 12, n. 4, p. 1613–1642, 2019.

VERGASTA, L. A. Avaliação do desempenho do modelo Eta Besm na representação dos componentes do balanço de água na bacia amazônica. *Clima e Ambiente-CLIAMB*, 2018.

XARRAY: N-D LABELED ARRAYS AND DATASETS IN PYTHON. 2020. Disponível em: <<https://xarray.pydata.org/en/stable/index.html>>. Acesso em: 11 de Nov. de 2020.

XUE, Y.; SHUKLA, J. Model simulation of the influence of global sst anomalies on sahel rainfall. *Monthly Weather Review*, American Meteorological Society, v. 126, n. 11, p. 2782–2792, 1998.

YNOUE, R. Y. et al. *Meteorologia: Noções básicas*. [S.l.]: Oficina de Textos, 2017.

ZHU, P.; LIU, J.; DAI, S. Fixed-point IDCT without multiplications based on bg lee’s algorithm. *Digital Signal Processing*, Elsevier, v. 19, n. 4, p. 770–777, 2009.