



AVALIAÇÃO DA EXTENSÃO DA JANELA TEMPORAL DE UM SISTEMA DE
ASSIMILAÇÃO DE DADOS 4D-VAR NA COSTA SUDESTE BRASILEIRA

Gabriel Vieira de Carvalho

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Civil, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Civil.

Orientador(es): Luiz Landau

Luiz Paulo de Freitas Assad

Rio de Janeiro
Setembro de 2019

AVALIAÇÃO DA EXTENSÃO DA JANELA TEMPORAL DE UM SISTEMA DE
ASSIMILAÇÃO DE DADOS 4D-VAR NA COSTA SUDESTE BRASILEIRA

Gabriel Vieira de Carvalho

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE
ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO
GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA CIVIL.

Examinada por:

Prof. Luiz Landau, DSc.

Prof. Luiz Paulo de Freitas Assad, DSc.

Prof. Alexandre Macedo Fernandes, PhD.

Dr^a. Raquel Toste Ferreira dos Santos, DSc.

RIO DE JANEIRO, RJ - BRASIL

SETEMBRO DE 2019

Carvalho, Gabriel Vieira

Avaliação da Extensão da Janela Temporal de um Sistema de Assimilação de Dados 4D-Var na Costa Sudeste Brasileira / Gabriel Vieira de Carvalho. – Rio de Janeiro: UFRJ/COPPE, 2019.

XIII, 92 p.: il.; 29,7 cm.

Orientadores: Luiz Landau

Luiz Paulo de Freitas Assad

Dissertação (mestrado) – UFRJ/ COPPE/ Programa de Engenharia Civil, 2019.

Referências Bibliográficas: p. 79-85.

1. Assimilação de dados. 2. Janela de Assimilação. 3. Não Linearidade. I. Landau, Luiz *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Civil. III. Título.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau Mestre em Ciências (M.Sc.)

AVALIAÇÃO DA EXTENSÃO DA JANELA TEMPORAL DE UM SISTEMA DE ASSIMILAÇÃO DE DADOS 4D-VAR NA COSTA SUDESTE BRASILEIRA

Gabriel Vieira de Carvalho

Setembro/2019

Orientadores: Luiz Landau

Luiz Paulo de Freitas Assad

Programa: Engenharia Civil

Obter a melhor estimativa do estado do mar a partir de dados observados e modelos físico-matemáticos é parte crucial de qualquer sistema de modelagem e previsão oceânica operacional. Como muitos parâmetros de um sistema de assimilação 4DVAR, a escolha do comprimento da janela de assimilação é amplamente baseada em uma abordagem heurística. Neste trabalho, um sistema estabelecido de assimilação de dados 4DVAR para a região oceânica sudeste do Brasil é usado como plataforma de teste para a avaliação objetiva do comprimento da janela de assimilação. Com base na abordagem de experimentos gêmeos, três janelas (4, 7 e 10 dias) são testadas usando observações reais - TSM, ASM e perfis hidrográficos *in situ* - durante um período de 6 meses. Métricas de erros objetivas e análises da minimização da função custo são relacionadas com oscilações da não-linearidade na dinâmica oceânica para subsidiar discussões sobre a validade da premissa linear. O desempenho dos experimentos com janela temporal de 4 e 7 dias foram superiores ao de 10 dias, apresentando menores erros especialmente em períodos com alta energia cinética turbulenta e consequente dificuldade dos sistemas em manter válida a premissa linear.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

ANALYSIS OF THE ASSIMILATION WINDOW LENGTH OF A 4D-VAR
SYSTEM AT THE SOUTHEAST BRAZILIAN OCEAN

Gabriel Vieira de Carvalho

September/2019

Advisors: Luiz Landau

Luiz Paulo de Freitas Assad

Department: Civil Engineering

Obtaining the best estimate of sea state from available observed data and physical-mathematical models is a crucial part of any operational ocean modeling and forecasting system. As many parameters of a 4DVAR assimilation system, the choice of the assimilation window length is widely based on a heuristic approach. In this work, an established 4DVAR data assimilation system for the southeastern oceanic region of Brazil is used as a test platform for the objective assessment of the assimilation window length. Based on a twin experiments approach, three windows (4, 7 and 10 days) are tested using real observations - SST, SSH and *in situ* hydrographic profiles - over a period of 6 months. Objective error metrics and cost function minimization analyzes are related to oscillations of nonlinearity in ocean dynamics to support discussions on the validity of the tangent linear assumption. The performance of the experiments with 4 and 7-days window length were superior to the 10-days and presented smaller errors especially in periods with high turbulent kinetic energy and consequent difficulty of the systems to keep the validity of the linear assumption.

Índice

1	Introdução	1
1.1	Breve histórico da previsão atmosférica e oceânica	2
1.2	Assimilação de Dados	4
1.3	4D-Var Incremental	7
1.4	Sistemas Complexos e não-linearidade	11
1.4.1	Complexidade	12
1.5	Modelagem de fluidos e a mudança de paradigma.....	13
1.6	A não-linearidade nas equações de movimento	15
1.6.1	Turbulencia em escalas “não-resolvíveis”	16
1.6.2	Turbulência em escalas “resolvíveis”	17
1.7	Não-linearidade, efeito borboleta e assimilação de dados 4D-VAR.....	18
1.8	Justificativa e Motivação	20
2	Objetivos	22
2.1	Objetivos Específicos	23
3	METODOLOGIA	23
3.1	Regional Ocean Modeling System (ROMS)	24
3.2	Configuração do NLROMS.....	24
3.3	Observações.....	26
3.4	Configuração do 4DVAR	30
3.5	Área de Estudo	32
3.6	Configuração dos Experimentos.....	34

3.7	Métricas de Avaliação	37
4	resultados e discussão	37
4.1	Comparação com Dados Assimilados	38
4.1.1	Análise da Distribuição dos Erros	38
4.1.2	Observações Satelitais	43
4.1.3	Perfis Hidrográficos.....	48
4.2	Comparação com Dados Independentes.....	56
4.3	Análise da Função Custo	65
4.4	Energia Cinética	70
4.5	Custo Computacional	76
5	Considerações finais	77
6	Referências bibliográficas.....	80
	ANEXO 1 - Mapas Médios Mensais.....	87

Lista de Figuras

Figura 1: Algoritmo do sistema 4DVAR implementado no ROMS.	11
Figura 2: Função custo de três experimentos com diferentes extensões de janelas de assimilação (8, 9 e 9.7). Nota-se que ao aumentar as janelas, aumenta-se o número de “vales” na função custo. Retirado de Pires (1996) (Figura 4a).	19
Figura 3: Domínio do modelo e batimetria utilizada (ETOPO1).	25
Figura 4: Exemplo do dado de TSM utilizado nos experimentos. Fonte: POES AVHRR.	27
Figura 5: Exemplo do dado de ASM utilizado nos experimentos. Fonte: AVISO.	28
Figura 6: Localização dos perfis de temperatura e salinidade entre maio e novembro de 2013, assimilados nos experimentos. Os pontos verdes representam perfis da base de dados EN4, os pontos azuis representam perfis coletados por Gliders do Projeto Azul e os pontos laranja representam dados de perfiladores, também do Projeto Azul.	29
Figura 7: Representação esquemática da estratificação vertical das correntes na região leste e sudeste do Brasil, assim como da região de ocorrência dos vórtices de Vitória (VV), Vórtice do Cabo de São Tomé (VCST) e Vórtice de Cabo Frio (VCF). Na região sobre a plataforma, destacam-se a Corrente Costeira do Brasil (CCB) e a região de ocorrência do fenômeno da ressurgência. Retirado de Godoi (2005).	33
Figura 8: Representação esquemática da configuração dos experimentos com múltiplos ciclos.	35
Figura 9: Configuração da extensão das janelas de assimilação em cada experimento.	35
Figura 10: Número de observações por ciclo ao longo do período do experimento (maio a novembro de 2013). De cima para baixo, experimento 4DIAS, 7DIAS e 10DIAS. As observações <i>in-situ</i> de temperatura e salinidade foram condensadas.	36
Figura 11: Probabilidade acumulada do erro da análise (Ea) dos experimentos com janela de assimilação de 4, 7 e 10 dias para ASM, TSM, e Temperatura e Salinidade <i>in-</i>	

<i>situ</i> , considerando toda a extensão analisada (maio a novembro/2013). Em cada figura os percentis 25, 50, 75 são indicados para os três experimentos comparados.	39
Figura 12: Distribuição de frequência do erro para as variáveis ASM, TSM,, Temperatura e Salinidade in-situ do experimento 4DIAS. A linha preta representa a função de densidade de probabilidade Gaussiana estimada para cada conjunto de amostras.	40
Figura 13: Distribuição de frequência do erro para as variáveis ASM, TSM, Temperatura e Salinidade in-situ do experimento 7DIAS. A linha preta representa a função de densidade de probabilidade Gaussiana estimada para cada conjunto de amostras.	41
Figura 14: Distribuição de frequência do erro para as variáveis ASM, TSM, Temperatura e Salinidade in-situ do experimento 10DIAS. A linha preta representa a função de densidade de probabilidade Gaussiana estimada para cada conjunto de amostras.	42
Figura 15:Série temporal do Jfit para ASM (acima) e TSM (abaixo) para os experimentos 4DIAS (azul), 7DIAS (laranja) e 10DIAS (verde).	45
Figura 16: Gráfico de dispersão do Jfit para as observações de satélites (TSM e ASM) entre os experimentos 4DIAS, 7DIAS e 10DIAS. Acima, Jfit obtido das soluções prévias (ou previsões) e abaixo, Jfit das análises.	46
Figura 17: Jfit médio ao longo dos dias de previsão (acima) e análise (abaixo) para TSM (esquerda) e ASM (direita).	48
Figura 18: Média de temperatura e salinidade na coluna d'água dos três experimentos (4DIAS, 7DIAS e 10DIAS) e do modelo livre (FREE_RUN) nas posições e tempo dos perfis hidrográficos assimilados (OBS). A esquerda, resultados extraídos da solução prévia dos experimentos e a direita, das análises.	50
Figura 19: Desvio padrão (STD) de temperatura e salinidade na coluna d'água dos três experimentos (4DIAS, 7DIAS e 10DIAS) e do modelo livre (FREE_RUN) nas posições e tempo dos perfis hidrográficos assimilados (OBS). A esquerda, resultados extraídos da solução prévia dos experimentos e a direita, das análises.	51

Figura 20: Erro (<i>misfit</i>) médio de temperatura e salinidade na coluna d'água dos três experimentos (4DIAS, 7DIAS e 10DIAS) e do modelo livre (FREE_RUN) nas posições e tempo dos perfis hidrográficos assimilados. A esquerda, resultados extraídos da solução prévia dos experimentos e a direita, das análises.	53
Figura 21: RMSE de temperatura e salinidade na coluna d'água dos três experimentos (4DIAS, 7DIAS e 10DIAS) e do modelo livre (FREE_RUN) nas posições e tempo dos perfis hidrográficos assimilados. A esquerda, resultados extraídos da solução prévia dos experimentos e a direita, das análises.	54
Figura 22: Diagrama de dispersão do Jfit médio para temperatura (azul) e salinidade (laranja) na coluna d'água entre os experimentos 4DIAS, 7DIAS e 10DIAS. Acima, Jfit obtido das soluções prévias (ou previsões) e abaixo, Jfit das análises.	55
Figura 23: Número de derivadores por caixa de 0,5° utilizados na comparação dos resultados dos experimentos. Os pontos no mapa representam as informações individuais dos derivadores na região de estudo.	57
Figura 24: Campo de velocidade média em caixas de 0,5° elaborado a partir dos dados dos derivadores do Projeto Azul para o período de maio a novembro de 2013 e dos resultados dos experimentos analisados (4DIAS, 7DIAS e 10DIAS) e da rodada livre (FREE) amostrados nas posições dos derivadores.	59
Figura 25: Campo de temperatura superficial médio em caixas de 0,5° elaborado a partir dos dados dos derivadores do Projeto Azul para o período de maio a novembro de 2013 e dos resultados dos experimentos analisados (4DIAS, 7DIAS e 10DIAS) e da rodada livre (FREE) amostrados nas posições dos derivadores.	60
Figura 26: Diagrama de dispersão entre os erros (derivadores – experimentos) da componente de corrente meridional (v), zonal (u) dos experimentos 4DIAS, 7DIAS, 10DIAS.	61
Figura 27: Campo de energia cinética média (MKE) em caixas de 0,5° elaborado a partir dos dados dos derivadores do Projeto Azul para o período de maio a novembro de 2013 e dos resultados dos experimentos analisados (4DIAS, 7DIAS e 10DIAS) e da rodada livre (FREE) amostrados nas posições dos derivadores.	64

Figura 28: Campo de energia cinética turbulenta (EKE) em caixas de 0,5° elaborado a partir dos dados dos derivadores do Projeto Azul para o período de maio a novembro de 2013 e dos resultados dos experimentos analisados (4DIAS, 7DIAS e 10DIAS) e da rodada livre (FREE) amostrados nas posições dos derivadores.....	65
Figura 29: Evolução da função custo linear (linha azul) e função custo não linear (pontos laranjas) normalizadas, ao longo de um dos ciclos de assimilação do experimento 4DIAS.....	66
Figura 30: Evolução da função custo linear (linha azul) e função custo não linear (pontos laranjas) normalizadas, ao longo de todos os ciclos de assimilação dos experimentos 4DIAS, 7DIAS e 10DIAS. A linha tracejada indica a redução média da função custo não linear no primeiro e segundo <i>outer-loop</i>	67
Figura 31: Diferença entre a função custo linear ao final dos <i>inner loops</i> e a função custo não linear subsequente para o 1° <i>outer loop</i> (acima) e o 2° <i>outer loop</i> (abaixo)..	68
Figura 32: Evolução da função custo não linear da solução prévia e da solução final e da função custo linear da última iteração dos <i>inner-loops</i> ao longo dos ciclos de assimilação dos experimentos 4DIAS (topo), 7DIAS (centro) e 10DIAS (a baixo).....	69
Figura 33: Mapa da energia cinética média (cm^2/s^2) para os três experimentos e para a rodada livre (FREE-RUN).....	73
Figura 34: Mapa da energia cinética turbulenta média (cm^2/s^2) para os três experimentos e para a rodada livre (FREE-RUN).	74
Figura 35: Série temporal da EKE média em todo o domínio do modelo (acima) e na região de maior concentração de EKE definida entre 29°S-19°S e 49°W-39°W (abaixo).	75
Figura 36: EKE mensal climatológico a partir dos dados ORCA12 (linha sólida preta) e dados de ASM satelital da AVISO (linha pontilhada preta), e tensão de cisalhamento do vento (linha pontilhada vermelha) para a área de 25°S-35°S e 20°W-40°W. O período de maio a outubro é ressaltado pela área em laranja. Adaptado de Rieck, <i>et al.</i> (2015).	75
Figura 37: Campos de corrente média extraídas das soluções posteriores (análises) para o mês de Maio . Da esquerda para a direita, resultados do experimento com janela de	

assimilação de 4, 7 e 10 dias. As linhas tracejadas vermelhas representam as trajetórias dos derivadores do Projeto Azul ao longo do mês de maio..... 88

Figura 38: Campos de corrente média extraídas das soluções posteriores (análises) para o mês de Junho . Da esquerda para a direita, resultados do experimento com janela de assimilação de 4, 7 e 10 dias. As linhas tracejadas vermelhas representam as trajetórias dos derivadores do Projeto Azul ao longo do mês de Junho..... 89

Figura 39: Campos de corrente média extraídas das soluções posteriores (análises) para o mês de Julho . Da esquerda para a direita, resultados do experimento com janela de assimilação de 4, 7 e 10 dias. As linhas tracejadas vermelhas representam as trajetórias dos derivadores do Projeto Azul ao longo do mês de Julho..... 90

Figura 40: Campos de corrente média extraídas das soluções posteriores (análises) para o mês de Agosto . Da esquerda para a direita, resultados do experimento com janela de assimilação de 4, 7 e 10 dias. As linhas tracejadas vermelhas representam as trajetórias dos derivadores do Projeto Azul ao longo do mês de Agosto..... 91

Figura 41: Campos de corrente média extraídas das soluções posteriores (análises) para o mês de Setembro. Da esquerda para a direita, resultados do experimento com janela de assimilação de 4, 7 e 10 dias. As linhas tracejadas vermelhas representam as trajetórias dos derivadores do Projeto Azul ao longo do mês de Setembro. 92

Figura 42: Campos de corrente média extraídas das soluções posteriores (análises) para o mês de Outubro. Da esquerda para a direita, resultados do experimento com janela de assimilação de 4, 7 e 10 dias. As linhas tracejadas vermelhas representam as trajetórias dos derivadores do Projeto Azul ao longo do mês de Outubro. 93

Lista de Tabelas

Tabela 1: Número médio de observações por ciclo de assimilação para os três experimentos e para cada variável assimilada.	36
Tabela 2: Média e desvio padrão (std) do vetor inovação para cada experimento e variável analisada.	40
Tabela 3: Maior diferença (valor absoluto) entre as classes da distribuição de frequência acumulada do vetor inovação e da distribuição gaussiana.	43
Tabela 4: Energia Cinética Turbulenta (EKE), Energia Cinética Média (MKE), RMSE e coeficiente de correlação linear de Pearson (P) para os dados de velocidade (componentes zonal e meridional) e temperatura medidos pelos derivadores e pelos experimentos 4DIAS, 7DIAS, 10DIAS e rodada livre (FREE).	62
Tabela 5: Tempo para a execução de 1 ciclo e para a integração de 140 dias, em horas, para cada experimento analisado.	76

1 Introdução

O trabalho desenvolvido e apresentado neste documento é parte do esforço conjunto de pesquisadores interessados em entender, simular e prever as correntes oceânicas brasileiras. O conhecimento das correntes, da maré e das ondas, foram, e são, fundamentais para o desenvolvimento da sociedade humana. Exemplos vão desde o planejamento de explorações marítimas, a exploração de recursos naturais (p. ex. a pesca, aquicultura e o petróleo), o entendimento do sistema climático global e o controle da poluição antropogênica.

Nesta dissertação, o leitor encontrará uma tentativa de aprimoramento de um sistema de simulação hidrodinâmica com assimilação de dados aplicado ao oceano sudeste brasileiro. O método utilizado consiste em combinar, da melhor forma possível, dados coletados por instrumentos de medição – com alta precisão e limitada cobertura espacial – e resultados de modelos físico-matemáticos – com alta cobertura espacial e limitada precisão. Como toda pesquisa, parte de conhecimentos prévios e se encaixa em um contexto de desenvolvimento específico, buscando contribuir para o assunto estudado com os resultados encontrados a partir das perguntas formuladas.

Espera-se que o leitor possa encontrar no texto elementos para compreender o contexto histórico da pesquisa em assimilação de dados e sua relevância no capítulo 1.1, assim como uma introdução mais específica deste método nos capítulos 1.2 e 1.3. Ainda na introdução, são oferecidos elementos para enriquecer a discussão em torno do assunto central da investigação – a determinação da extensão da janela de assimilação – do capítulo 1.4 ao 1.7. A motivação geral e específica é apresentada no capítulo 1.8 e no capítulo 2 são enumerados os objetivos que este trabalho se propôs a alcançar.

Com o intuito de que se possa compreender as premissas adotadas, as observações utilizadas, e as configurações dos experimentos conduzidos, o capítulo 3 se dedica à apresentação da metodologia. Os resultados dos experimentos são assunto do capítulo 4, que são apresentados conjuntamente a discussões elaboradas à luz dos elementos introduzidos no início do documento. Por fim, as conclusões e considerações finais são apresentadas no capítulo 5, onde busca-se sintetizar o trabalho e elencar as principais contribuições que espera-se sejam úteis para futuras pesquisas correlatas.

1.1 Breve histórico da previsão atmosférica e oceânica

Vilhelm F.K. Bjerknes (1862-1951), cientista norueguês, foi o primeiro a expor claramente que o estado futuro da atmosfera pode ser estimado através da integração das equações de movimento, sendo portanto, em princípio, determinado pelas condições iniciais e de contorno. São de Bjerknes (1904) as palavras:

“If it is true, as every scientist believes, that subsequent atmospheric states develop from the preceding ones according to physical law, then it is apparent that the necessary and sufficient conditions for the rational solution of forecasting problems are the following

1. A sufficiently accurate knowledge of the state of the atmosphere at the initial time.

2. A sufficiently accurate knowledge of the laws according to which one state of the atmosphere develops from another.”

Essas duas condições, enunciadas no início do século 20, resumem os principais desafios enfrentados pelos cientistas dedicados à previsão atmosférica e oceânica (Kalnay, 2003). A primeira condição – conhecer o estado inicial com suficiente acurácia – é dependente de uma rede de observação do sistema natural a ser previsto e consiste no principal objetivo das técnicas de assimilação de dados, apesar deste conceito ter surgido muito tempo depois. A segunda condição consiste no entendimento dos processos físicos que governam os movimentos da atmosfera (e do oceano), uma área em constante desenvolvimento até os dias de hoje.

Se Bjerknes enunciou as bases do problema, coube a Richardson em seu livro *“Weather Prediction by Numerical Process”*, publicado em 1922, a primeira tentativa de estabelecer um método matemático para integrar as equações de movimento. Para testar sua proposta, realizou uma integração numérica das equações primitivas de movimento usando uma grade de 200 Km centrada na Alemanha para prever a pressão atmosférica por 6 horas (Kalnay, 2003). Como dito por Charney (1950), o fato de que a previsão utilizada para testar o método tenha sido um fracasso não representa de maneira nenhuma a medida do valor real de seu trabalho – solidificar os problemas essenciais que teriam que ser resolvidos pelos futuros meteorologistas. Além da

inadequação da condição inicial utilizada, Richardson estava ciente de que a enorme quantidade de cálculos não era compatível com a capacidade computacional da sua época.

“Perhaps some day in the dim future it will be possible to advance the computations faster than the weather advances and at a cost less than the saving to mankind due to the information gained. But that is a dream.” (Richardson, 1922, p.xi)

Desta forma, às duas primeiras condições enunciadas por Bjerknes soma-se agora uma terceira, a capacidade computacional para realizar os cálculos necessários em tempo hábil – ou seja, prever o estado futuro antes que ele aconteça.

Décadas se passaram até que avanços na rede de observação atmosférica e da capacidade computacional permitiram que Charney et al. (1950) realizassem uma histórica previsão do tempo utilizando um dos primeiros computadores eletrônicos desenvolvido por von Neumann no departamento de pesquisa balística do exército americano¹. O trabalho utilizou a ideia de programação armazenada (os *loops* que realizam a mesma operação em diferentes operandos) e um modelo barotrópico simplificado, que filtrava as ondas de gravidade de rápida propagação. Apesar do sucesso da previsão, Charney (1950) sustenta que muito mais poderia ser obtido com a utilização das equações primitivas completas, aponta caminhos para serem desbravados em processos não resolvíveis (sub-grade) e aponta a necessidade de uma análise objetiva de dados meteorológicos para substituir as dispendiosas análises manuais para gerar condições iniciais para os modelos, ao que hoje denomina-se assimilação de dados (Kalnay, 2003).

Desde então, a evolução dos computadores, das representações dos processos físicos, das soluções numéricas e das redes de observação possibilitou grande avanço, abrindo a possibilidade de se resolver modelos de equações primitivas com resolução espacial cada vez maior. Esta evolução tecnológica permite que tenhamos acesso ao

¹ O computador Eniac, utilizado por John von Neumann para pesquisas relacionadas ao desenvolvimento da bomba atômica, no contexto da segunda guerra mundial.

monitoramento e previsão do tempo, do estado do mar e do clima, em uma base operacional.

Nesse contexto, a técnica de assimilação de dados pode ser descrita como o processo pelo qual todas as informações disponíveis são usadas para obter a estimativa mais precisa possível do estado oceânico. As informações disponíveis consistem nas observações (medições *in-situ* ou remotas) e nas leis físicas que governam a evolução do estado oceânico (Talagrand, 1997). Apesar da função originária das técnicas de assimilação de dados ter sido a de produzir campos iniciais para previsões de curto alcance, também pode ser utilizada para a geração de bases de dados de longa duração que representem a melhor estimativa do estado oceânico ou atmosférico naquele período específico.

Retornando à primeira condição de Bjerknes (1904) para a viabilidade de um sistema de previsão “*A sufficiently accurate knowledge of the state of the atmosphere at the initial time.*”, a assimilação de dados é parte crucial de qualquer esforço de se estabelecer um sistema de previsão – atmosférico ou oceânico.

1.2 Assimilação de Dados

Desde os primeiros métodos desenvolvidos para obter a estimativa do estado atmosférico até os dias de hoje a técnica de assimilação de dados evoluiu muito, partindo da simples interpolação manual para métodos estatísticos complexos e automatizados. Em essência, técnicas de assimilação de dados objetivam obter a melhor estimativa de determinada variável (X_a) a partir de duas fontes de dados – solução prévia do modelo ou *first guess* (X_b), e a observação (X_o). A combinação linear objetivada é:

$$X_a = \alpha_1 X_b + \alpha_2 X_o \quad (\text{Eq. 1})$$

Considerando as fontes de informação não enviesadas (*unbiased*), $\alpha_1 + \alpha_2 = 1$, a estimativa da solução posterior pode ser escrita como:

$$X_a = X_b + \alpha_2 (X_o - X_b) \quad (\text{Eq. 2})$$

E a escolha de α_2 que minimiza a covariância do erro de X_a será:

$$X_a = X_b + W(X_o - X_b) \quad (\text{Eq. 3})$$

Como exposto na equação acima, a solução prévia do modelo X_b é atualizada somando-se a diferença entre o valor observado e a própria solução prévia ($X_o - X_b$), ponderada pela função ganho W . A função ganho W é definida a partir das variâncias dos erros das duas fontes de observação:

$$W = \frac{\sigma_b^2}{(\sigma_b^2 + \sigma_o^2)} \quad (\text{Eq. 4})$$

A Eq. 3 representa a Melhor Estimativa Linear Não-Enviesada (*BLUE* - Best Linear Unbiased Estimator), e este conceito se estende para os dois principais métodos estatísticos de assimilação de dados – equivalentes se as premissas forem verdadeiras. As duas grandes famílias de métodos que buscam encontrar esta estimativa (*BLUE*) são os métodos Sequenciais e os métodos Variacionais.

Nos métodos sequenciais, como a Interpolação Ótima (IO) e o *Kalman Filter* (KF), busca-se a atualização de W – função ganho – ao longo dos ciclos de assimilação onde as observações novas em determinado tempo são introduzidas e a estimativa de erro da solução prévia é atualizada.

Dentre os métodos Sequenciais, a obtenção da variância do erro da solução prévia (σ_b^2) e a sua evolução no tempo difere. Por exemplo, na IO, assume-se que a variância do erro da solução prévia aumenta em taxa fixa (a) conforme o modelo é integrado no tempo:

$$\sigma_b^2(t_{i+1}) = a\sigma_b^2(t_i) \quad (\text{Eq. 5})$$

Já no KF, em vez de assumir uma taxa fixa, as matrizes de covariância do erro da solução prévia são computadas usando o próprio modelo em sua forma linearizada (Kalnay, 2003). Desenvolvido para resolver os maiores problemas do KF – o alto custo computacional associado à obtenção e integração no tempo das matrizes de covariância do erro da solução prévia – o *Ensemble Kalman Filter* (EnKF) obtém as matrizes de covariância do erro da solução prévia a partir de um conjunto de soluções do modelo (*ensemble*). Esse conjunto de soluções do modelo é obtido pela integração do modelo não linear n vezes a cada ciclo de assimilação, cada integração possuindo uma pequena perturbação aleatória nas condições iniciais e/ou forçantes atmosféricas. Desta forma, o

espalhamento do conjunto de soluções ao redor da solução média provê as informações necessárias sobre a variabilidade e incerteza do sistema. O EnKf portanto é uma solução atraente pela maior simplicidade de implementação e menor custo computacional, aliado à uma performance equivalente aos métodos mais custosos (Evensen, 2003, Li *et al.* 2017).

Nos métodos variacionais, busca-se diretamente a análise (X_a) que minimiza uma função custo (escalar), J . A obtenção da função custo pode ser vista pela ótica do teorema de Bayes, onde busca-se a distribuição condicional da probabilidade posterior para X_a dado o vetor de observações X_o :

$$p(X_a|X_o) = \frac{p(X_o|X_a)p(X_a)}{p(X_o)} \quad (\text{Eq. 6})$$

Busca-se então maximizar $p(X_a|X_o)$ o que corresponde a minimizar a função custo, que pode ser vista como a combinação das distribuições da observação e da solução prévia. Simplificadamente, a função custo é definida como o quadrado da distância entre a análise e as duas fontes de informação, ponderada pelas estimativas de erro de ambas:

$$J(X_a) = \frac{1}{2} \left[\frac{(X_a - X_b)^2}{\sigma_b^2} + \frac{(X_a - X_o)^2}{\sigma_o^2} \right] \quad (\text{Eq. 7})$$

Portanto, no caso dos métodos variacionais, o foco da minimização não é a variância da análise, mas a própria variável (X_a), problema que é equivalente ao definido nos métodos sequenciais (Lorenz, 1986). As variâncias do erro (σ_b^2 e σ_o^2) são calculadas uma única vez e a princípio não evoluem ao longo dos ciclos de assimilação. Como a solução do modelo depende exclusivamente das condições iniciais, das forçantes atmosféricas e condições de contorno laterais, estes são referidos como vetor controle (z) – busca-se z que leve a minimização da função custo.

No método variacional tridimensional (3DVAR), a análise é feita de uma única vez, considerando todas as observações em um intervalo pequeno de tempo. Já no método variacional tetradimensional (4DVAR), a análise é feita sobre um período de tempo – a janela de assimilação. Desta forma, as observações são utilizadas no tempo em que se encontram, e a análise, ao invés de ser em um tempo único, é feita sobre todo o período da janela de assimilação definida. Apesar de ser mais custoso computacionalmente, este

método tem a vantagem de manter maior coerência com a física do modelo visto que o método não trabalha com as variáveis da solução diretamente, mas sim com o vetor controle, e considera todas as observações contidas na janela de assimilação.

1.3 4D-Var Incremental

Como já enunciado, o objetivo do método 4DVAR é obter a melhor estimativa do estado do mar, denominada solução posterior ou análise x_a , que minimiza a função custo definida como o quadrado da distância entre a análise e as fontes de informação (solução prévia x_b e observação y_o), ponderada pelas estimativas de erro de ambas. O 4DVAR obtém a análise ao longo de um intervalo de tempo finito (janela de assimilação) e opera com base em uma estimativa prévia e observações contidas neste intervalo de tempo. A exposição detalhada deste método, assim como a implementação no sistema de modelagem *ROMS (Regional Ocean Modeling System)* – utilizado neste trabalho - foge ao objetivo desta dissertação e pode ser encontrada em Moore, *et al.* (2011a, 2011b e 2011c), assim como na revisão bibliográfica em português da tese de Faller (2016).

As variáveis prognósticas do ROMS são a temperatura potencial (T), Salinidade (S), velocidade horizontal (u, v) e variação da superfície livre (ζ), que compõe o vetor estado $x(t_i)$ em cada ponto de grade e passo de tempo. O vetor de estado é propagado no tempo através do modelo não-linear, sujeito às condições de contornos laterais $b(t_i)$ e forçantes atmosféricas $f(t_i)$ para fluxos de calor, massa e movimento e a partir das condições iniciais $x(t_0)$. Pela dependência do vetor estado das forçantes de superfície, laterais e condições iniciais, estas são denominadas de vetor controle $z=(x(t_0), f(t_i), b(t_i))$. Portanto, a evolução no tempo do vetor de estado pode ser representada por:

$$x(t_i) = M(t_i, t_0)z \quad (\text{Eq. 8})$$

onde M representa o ROMS não-linear, ou NLROMS.

Para a aplicação da assimilação 4DVAR neste sistema, necessita-se de uma estimativa ou solução prévia para todas as variáveis do vetor controle do modelo ao longo da janela de assimilação (t_0, t_N) , assim como da matriz de covariância de erro para cada variável do vetor controle \mathbf{B}_b , \mathbf{B}_f , e \mathbf{B}_x respectivamente, que representam todas as hipóteses sobre erros e incertezas da solução prévia. Portanto, chamaremos o vetor

controle da solução prévia de $z_b=(x_b(t_0), f_b(t_0, t_N), b_b(t_0, t_N))$ e o conjunto de matrizes de covariância do erro do vetor controle será representado pela matriz $\mathbf{D} = (\mathbf{B}_x, \mathbf{B}_f, \mathbf{B}_b)$. Neste trabalho consideramos que a matriz \mathbf{D} não varia no tempo e utilizamos a formulação de restrição forte (*strong constraint*) que assume que a física do modelo é livre de erros.

Além da solução prévia $x_b(t_0, t_N)$ o sistema inclui as observações contidas na janela de observação, denominados y_o , com suas respectivas matrizes de covariância de erro \mathbf{R} , que analogamente representa todas as hipóteses de erros e incertezas associadas às observações.

Como a solução posterior x_a é definida pelo vetor controle z , o objetivo do sistema é encontrar z que minimiza a função custo \mathbf{J} . De acordo com o teorema de Bayes, a escolha ótima de z que maximiza a probabilidade condicional $p(x_a/y)$ é aquela que minimiza a função \mathbf{J} (agora escrita em relação ao vetor controle z):

$$\mathbf{J}_{NL} = (z - z_b)^T \mathbf{D}^{-1} (z - z_b) + (y_o - H(x))^T \mathbf{R}^{-1} (y_o - H(x)) \quad (\text{Eq. 9})$$

Onde H é o operador da observação, responsável por obter o vetor de estado $x(t)$ no local e tempo das observações utilizando o próprio modelo não-linear. Dessa forma, a função custo \mathbf{J}_{NL} é uma função não-linear, o que significa que sua topografia pode ser complexa e não apresentar um único mínimo global (ver exemplo na Figura 2). Como já introduzido no item 1.7, localizar o mínimo global de funções não lineares pode ser computacionalmente inviável e a maioria dos sistemas de assimilação 4DVAR adota a abordagem *incremental* proposta por Courtier (1994). Nesta, considera-se que a solução posterior $x_a(t)$ pode ser obtida por pequenos incrementos da solução prévia.

$$x_a(t) = x_b(t) + \delta x(t) \quad (\text{Eq. 10})$$

O incremento $\delta x(t)$ é obtido por \mathbf{M} , o operador tangente linear que representa a linearização de M sobre a solução prévia $x_b(t_i)$ e propaga o incremento na condição inicial $\delta x(t_0)$ para frente no tempo, sujeito aos incrementos nas forçantes atmosféricas $\delta f(t)$ e fluxos de contorno laterais $\delta b(t)$:

$$\delta x(t) = \mathbf{M}_{(t_i, t_0)} \delta z \quad (\text{Eq. 11})$$

onde δz é o vetor controle dos incrementos ($\delta z = (\delta x(t_0), \delta f(t), \delta b(t))$) ao longo de toda a janela de assimilação $[t_0, t_N]$. O operador tangente linear \mathbf{M}_b é denominado TLROMS (*Tangent Linear ROMS*) e tem a função de mapear o incremento do vetor controle em incremento no vetor de estado. A já citada premissa linear, portanto, se refere à hipótese de que a solução do modelo não linear a partir do vetor controle atualizado $x_a = \mathcal{M}(z + \delta z)$ não diverge significativamente da solução prévia adicionada a incrementos obtidos pela linearização do modelo:

$$\mathcal{M}(z) + \mathbf{M}(\delta z) \approx \mathcal{M}(z + \delta z) \quad (\text{Eq. 12})$$

Usando a aproximação incremental, a função custo pode ser reescrita como:

$$J = \delta z^T \mathbf{D}^{-1} \delta z + (\mathbf{d} - \mathbf{G} \delta z)^T \mathbf{R}^{-1} (\mathbf{d} - \mathbf{G} \delta z) \quad (\text{Eq. 13})$$

Onde $\mathbf{d} = y_o - H(x)$ é o vetor inovação calculado a cada *outer loop* e \mathbf{G} representa a convolução no tempo de \mathbf{M}_b com \mathbf{H} , e \mathbf{H} é a linearização do operador de observação H . Como J é uma função linear, possui apenas um mínimo global.

Na formulação primal, utilizada neste trabalho, o vetor controle de incrementos δz_a que conduz à solução posterior ótima e que portanto minimiza a função custo linear J , corresponde à solução da equação $\partial J / \partial \delta z = 0$ e é dado por:

$$\delta z = (\mathbf{D}^{-1} + \mathbf{G}^T \mathbf{R}^{-1} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{R}^{-1} \mathbf{d} \quad (\text{Eq. 14})$$

onde \mathbf{G}^T é o adjunto de \mathbf{G} (TLROMS) e é chamado de ADROMS. Enquanto \mathbf{G} é o operador linear responsável por mapear do espaço do vetor controle para o espaço das observações, \mathbf{G}^T é o operador linear que faz o inverso, mapeia do espaço das observações para o espaço do vetor controle. A solução da equação acima passa pela resolução de um sistema de equações lineares usando a formulação de Lanczos pré-condicionada do método do gradiente conjugado descendente (Fisher e Courtier, 1995).

A obtenção de δz_a é feita de forma iterativa em *loops* externos e internos – com número de iterações previamente definida pelo usuário. No primeiro ‘*m*’ *outer-loop*, a solução prévia $x_b(t_0, t_N)$ é obtida a partir do vetor controle $z_b(t_0, t_N)$ pelo modelo não linear (NLROMS), a partir da qual se obtém a função custo não linear J_{NL} e o vetor inovação \mathbf{d} , fornecendo as informações necessárias para a sequência de *inner-loops*. A cada ‘*k*’ *inner-loop*, são obtidos incrementos do vetor controle δz^k no sentido de minimizar J e, após um número previamente definido de *inner-loops* considerado suficiente para que o

mínimo de J tenha sido encontrado, o vetor controle z^k é atualizado a partir do somatório dos incrementos obtidos ao vetor controle prévio:

$$z^k = z_b + \sum_{i=1}^{k-1} \delta z^{k-1} \quad (\text{Eq. 15})$$

A partir do vetor controle atualizado, o modelo não linear é novamente integrado para se obter a solução posterior $x_m(t_0, t_N)$. Caso seja desejado realizar mais de um *outer-loop*, a solução encontrada se torna a solução prévia e dá-se início a um novo ciclo de ' k ' *inner-loops*. Ao final dos ' m ' *outer-loops* e ' k ' *inner-loops*, o vetor controle final z_a é obtido a partir do qual se obtém nova integração do NLROMS para computar a análise ou solução posterior $x_a(t_0, t_N)$. O algoritmo de execução do sistema é esquematizado de forma simplificada na Figura 1.

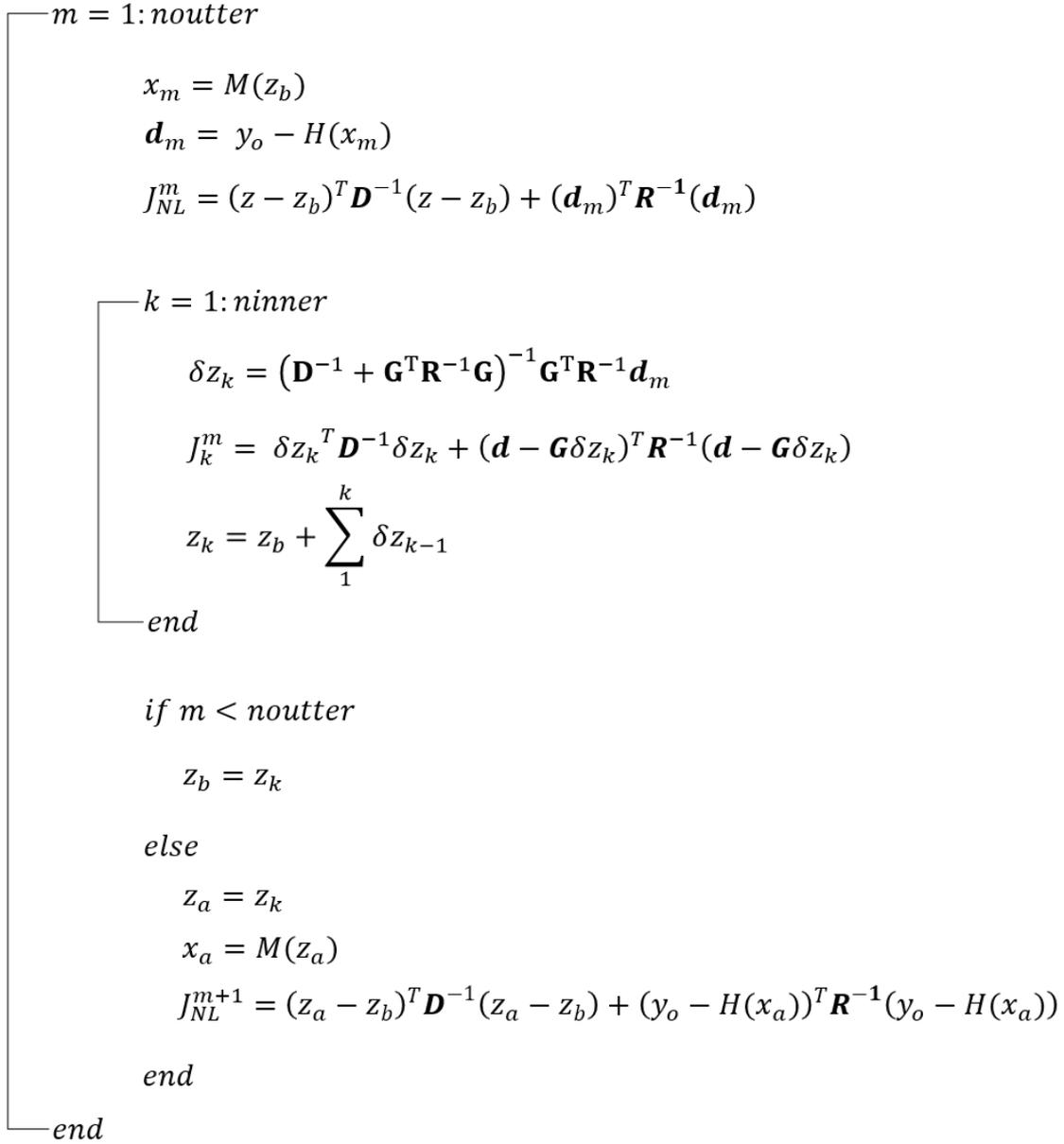


Figura 1: Algoritmo do sistema 4DVAR implementado no ROMS.

1.4 Sistemas Complexos e não-linearidade

Este capítulo explora os conceitos de complexidade e não-linearidade inerentes à sistemas naturais, como o sistema de correntes oceânico. As perguntas fundamentais que levaram a essa investigação são: “De onde vem a não-linearidade? ”; “O que faz um processo ser não-linear? ”; “Onde a não-linearidade se apresenta no sistema de correntes que se quer representar nos modelos hidrodinâmicos?” e “Em que situações a não-linearidade é mais ou menos importante? ”.

Essas e outras perguntas surgiram, pois, a assimilação de dados variacional e incremental utiliza linearizações no processo de minimização da função custo, dando origem a *premissa linear*. Assume-se que o desvio entre a solução linear e a não-linear é suficientemente pequeno dentro da extensão temporal da janela de assimilação. Logo, estes conceitos são importantes para embasar a discussão dos experimentos apresentados nesta dissertação.

1.4.1 Complexidade

A evolução do pensamento ao longo do século XX nos conduz a uma mudança de paradigma. Estamos, atualmente, vivendo esta transição, conscientes do esgotamento do paradigma determinista, mecanicista, cartesiano, que contudo ainda permeia nossa visão e persiste em nossas ações. Tal paradigma, baseado em uma visão racionalista, nos leva a perceber a realidade como uma sequência linear de causas e efeitos, na qual é possível analisar cada parte de um dado sistema separadamente. A “faca analítica” – metáfora utilizada por Robert M. Pirsig em “*Zen e a Arte da Manutenção de Motocicletas*” (1984) – é manejada com intuito de entender a realidade a fundo, separando habilmente um sistema em partes, de partes, de partes... cada vez menores e mais detalhadas. Porém, na busca infinita por especialização, o cientista se vê caminhando em direção oposta, se afastando da verdade almejada de início. Fato é que a aplicação da faca analítica por si só não conduz ao entendimento da realidade - com suficiente tempo se saberia absolutamente tudo sobre absolutamente nada.

Hoje, entende-se que o conhecimento das partes não é suficiente para descrever a realidade de qualquer sistema pois o mesmo é fundamentalmente definido pelas interações das partes entre si mesmas e com o meio. A complexidade decorre do entendimento de que as interações entre as partes de um sistema é tão ou mais importante do que o conhecimento das partes isoladamente para o entendimento do sistema. As interações definem um sistema, não suas partes. As partes são simples reduções cognitivas derivadas do uso da *faca analítica*. A rede de interações mútuas entre as partes de um sistema gera processos de retroalimentação contínuos e torna a definição de causa e efeito linear impossível.

Porém erra quem frente ao entendimento da complexidade dos sistemas, julga inútil o conhecimento específico derivado do antigo paradigma. Inúmeros são os avanços

tecnológicos em todas as áreas das ciências e é inegável a utilidade de se aprender a manejar a *faca analítica* com habilidade para lidar com problemas e situações do mais amplo espectro. Porém, igualmente inúmeras são as mazelas causadas pela utilização desta ferramenta do intelecto humano quando desvinculadas do entendimento das conexões que compõe a nossa realidade e da consciência ética e política no fazer científico. A primeira previsão meteorológica foi realizada no mesmo computador utilizado no desenvolvimento da bomba atômica.

A *busca* pela verdade – frisa-se a *busca* por entender que a verdade é inalcançável - necessita do equilíbrio entre a busca do conhecimento em direção ao micro com a busca do saber em direção ao macro.

1.5 Modelagem de fluidos e a mudança de paradigma

Ao longo do século XX, “*A relatividade eliminou a ilusão newtoniana sobre o espaço e o tempo absolutos; a teoria quântica eliminou o sonho newtoniano de um processo controlável de mensuração; e o caos elimina a fantasia laplaciana da previsibilidade determinista*” (Gleick, 1987). Vê-se agora que as fundações teóricas da previsão meteorológica expressas no início deste trabalho através da citação de Bjerkness (1904) “*If it is true, as every scientist believes, that subsequent atmospheric states develop from the preceding ones according to physical law ...*” são ultrapassadas. Esta afirmação, fruto do paradigma científico do sec. XIX, evidencia a crença determinística do universo na qual o conhecimento exato das leis de movimento e da configuração em determinado instante permite, em princípio, prever toda a sua história. No paradigma emergente do sec. XX, contudo, sabemos que o universo é quântico, ou seja, mesmo conhecendo o estado inicial de um sistema, só poderíamos obter um conjunto de probabilidades para estado futuro. O universo, portanto, é indeterminístico.

Em outra escala, entende-se que a interação entre as partes de um sistema gera processos de retroalimentação que fazem emergir respostas não-lineares e impossibilita a definição clara de causa e efeito. A ação do vento sobre o oceano induz movimento e conseqüente transporte de calor, e a quantidade de calor no oceano induz movimento atmosférico, fechando um ciclo de retroalimentação. A hidrodinâmica em um estuário influencia a distribuição das florestas de mangue que, por sua vez, influencia a

hidrodinâmica ao criar barreiras físicas com suas raízes e caules e modificar a sedimentação do local. Em uma praia, as ondas incidentes induzem a variações na batimetria que, por sua vez, determina a propagação e dissipação da energia das ondas. O entendimento dos sistemas naturais influencia nossa visão sobre o mundo que, por sua vez, molda nosso entendimento dos sistemas naturais. Exemplos de mecanismos de retroalimentação são inúmeros na natureza.

Os modelos físico-matemáticos da dinâmica dos fluidos são determinísticos, isto é, dada uma condição inicial, o estado futuro é sempre o mesmo, e processos não-lineares estão contidos nos termos da equação. Lorenz (1963) demonstra que, mesmo em sistemas de equação simples, a existência de termos não-lineares e processos de retroalimentação levam a uma extrema sensibilidade às condições iniciais. Pequenas alterações na condição inicial levam a grandes diferenças no estado futuro do sistema após determinado tempo de integração. Além disso, as incertezas associadas à escalas menores se propagam para escalas maiores após determinado tempo, contribuindo assim para a impossibilidade da previsão do estado do sistema em um instante suficientemente distante (Lorenz, 1969). O horizonte da previsibilidade é, portanto, dependente da escala do processo e da acurácia das condições iniciais. Quanto menor a escala horizontal simulada, menor o horizonte de previsibilidade – é possível prever a temperatura média na bacia de santos em janeiro daqui a 2 anos com razoável acurácia, mas a temperatura média da água na praia grande em Arraial do Cabo (RJ) não pode ser prevista para além de cerca de uma semana.

No seguinte trecho retirado de Lorenz (1969) fica clara a consciência das limitações de modelos determinísticos e a postura pragmática que o autor acredita que um cientista deve ter frente aos problemas que se propõe estudar. Segundo o autor:

“This is not to imply that fluid dynamicists generally believe that real fluid systems are deterministic. It is a fundamental principle of quantum mechanics, for example, that real systems are indeterministic, and presumably few fluid dynamicists would question the validity of quantum mechanical principles merely because they do not customarily make use of them. More likely, they would simply take it for granted that their equations need to be idealized to some extent, in view of the complexity of most real fluid systems, and that

properties of the exact equations which are not pertinent to the problem under study need not be retained. In many familiar problems the question of determinism or indeterminism is of minor importance, and deterministic equations will yield acceptable results. It is often convenient to look upon an idealized equation as the exact equation for a model of a real system. A model may of course be deterministic by definition.” (Lorenz, 1969)

Lorenz (1969) ainda argumenta que, frente às incertezas na definição do estado inicial do sistema atmosférico (relacionado às limitações nos sistemas de observação) e na formulação das equações governantes, a incerteza associada ao indeterminismo dos sistemas naturais é menos significativa.

Apesar das limitações, a modelagem atmosférica e oceânica continua sendo de crucial importância para diversas atividades humanas e ainda há muito o que se evoluir nas redes de observação e técnicas de assimilação de dados antes de alcançar o limite natural de previsibilidade devido ao indeterminismo da realidade.

1.6 A não-linearidade nas equações de movimento

O termo *não-linear* deriva da representação gráfica cartesiana de um processo cuja resposta não é proporcional ao estímulo oferecido. A raiz dessa falta de proporcionalidade é justamente a existência dos processos de retroalimentação e interação entre as partes do sistema.

Os termos não lineares da equação de movimento emergem do fato de que ao descrever a variação da quantidade de movimento ($\frac{du}{dt}$) em um determinado ponto, é necessário que levemos em conta que uma parcela do fluido, ao se mover para outro ponto, “leva” sua quantidade de movimento para o novo ponto - ao que se chama advecção (Pond e Pickard, 1983). Logo, a descrição da variação da quantidade de movimento em um ponto deve ser dividida em uma componente de variação local e componentes advectivas:

$$\frac{du}{dt} = \underbrace{\frac{\partial u}{\partial t}}_{\text{Componente Local}} + \underbrace{u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z}}_{\text{Componentes advectivas}} = \text{Forçantes}$$

As componentes advectivas são chamadas de não-lineares pois se expressam ou pelo quadrado das componentes da velocidade:

$$u \frac{\partial u}{\partial x} = \left(\frac{1}{2}\right) \left[\frac{\partial (u^2)}{\partial x}\right] \quad (\text{Eq. 16})$$

Ou pelo produto entre uma componente de velocidade e a derivada parcial de outra:

$$v \frac{\partial u}{\partial y} \quad (\text{Eq. 17})$$

Quanto maior for a importância relativa dos termos não-lineares frente aos termos de fricção (que tendem a remover as diferenças de velocidade), menor será o alcance da previsão com acurácia - visto que devido esses termos pequenas perturbações podem se tornar grandes flutuações no campo de velocidade.

1.6.1 Turbulencia em escalas “não-resolvíveis”

Independente da escala de estudo, a dinâmica apresenta turbulência – flutuações das componentes de velocidade em relação ao fluxo médio - em escalas não resolvíveis pelo modelo. A abordagem para solucionar essa questão consiste em separar o fluxo total em fluxo médio – a parcela do fluxo “resolvível” pela escala do modelo – e flutuações – parcela turbulenta, ou não-resolvível na escala de interesse.

Partindo da ideia de que o fluxo total é composto pelo fluxo médio em determinado intervalo de tempo (T) somado à flutuações ($u = \bar{u} + u'$), ao substituir os termos da equação de movimento pelos seus equivalentes “médios + flutuações” e retirarmos a média no intervalo de tempo T , vemos que todos os termos do lado direito (forças por unidade de massa) mantém a mesma forma tanto para o fluxo total quanto para o fluxo médio pois $\overline{u'} = 0$ (Pond e Pickard, 1983). Porém ao aplicar a mesma operação nos termos não-lineares (advectivos), o resultado é a separação em termos advectivos médios e termos advectivos envolvendo as componentes da flutuação de velocidade:

$$\left(\bar{u} \frac{\partial \bar{u}}{\partial x} + \bar{v} \frac{\partial \bar{u}}{\partial y} + \bar{w} \frac{\partial \bar{u}}{\partial z} \right) + \left(\overline{u' \frac{\partial u'}{\partial x}} + \overline{v' \frac{\partial u'}{\partial y}} + \overline{w' \frac{\partial u'}{\partial z}} \right) \quad (\text{Eq. 18})$$

Dessa forma a equação de movimento para \bar{u} passa a ter a mesma forma da equação inicial porém com a velocidade média substituindo a velocidade total, somado aos três novos termos que representam o efeito da turbulência em escalas não resolvíveis pelo modelo utilizado e que possuem o efeito de fricção, tendendo a reduzir diferenças de velocidade. Contudo, essa abordagem adiciona três novas variáveis (u', v' e w') e portanto o novo sistema de equações necessita de um fechamento. Esse problema, chamado de fechamento turbulento é fundamental na mecânica dos fluidos e existem diversos esquemas de fechamento turbulento desenvolvidos, porém todos utilizam parametrizações em algum momento do desenvolvimento matemático e necessitam de comprovação empírica para atestar sua validade. As parametrizações dos esquemas de fechamento turbulento mais utilizados como $K - \epsilon$ ou o de Mellor&Yamada utilizam parametrizações que são função das variáveis resolvíveis, e.g. a energia cinética turbulenta.

Em resumo, pode-se entender que o efeito da turbulência em escalas não resolvíveis pelo modelo são análogas aos efeitos da viscosidade molecular: enquanto a última representa a fricção e troca de momento gerada por colisões moleculares, a viscosidade turbulenta representa a fricção e troca de momento gerada por colisões de parcelas do fluido. Ambas tem o efeito de homogeneizar a energia e a distribuição de massa no fluido porém pelas escalas de interesse usuais em modelos oceânicos, a viscosidade molecular tem efeito desprezível ao contrário da viscosidade turbulenta.

1.6.2 Turbulência em escalas “resolvíveis”

Considerando a parcela da dinâmica resolvida por determinado modelo oceânico – relacionada com a resolução espacial da grade – feições como vórtices e meandros podem ser representados nos resultados das integrações do modelo. Por exemplo, um modelo com resolução de grade de $1/12^\circ$ (~ 9 Km) é capaz de “resolver” vórtices de mesoescala (que apresentam cerca de 100 km de diâmetro) porém é incapaz de “resolver” a sub-mesoescala (que apresentam cerca de 10 km de diâmetro). Dessa forma, tais feições compõe a *turbulência* na escala de discretização espacial do modelo, ou seja, a escala da turbulência que o modelo consegue representar, ou *resolver*.

Analogamente à abordagem de Reynolds, a energia cinética total representada nos resultados de um modelo pode ser dividida em energia cinética média (MKE) e energia cinética turbulenta (EKE), que representam a energia cinética contida no fluxo médio (\bar{u}, \bar{v}) e a energia cinética contida em flutuações do fluxo médio (u', v') .

$$EKE = 0,5 * (u'^2 + v'^2) \quad (\text{Eq. 19})$$

$$MKE = 0,5 * (\bar{u}^2 + \bar{v}^2) \quad (\text{Eq. 20})$$

onde

$$(u', v') = (u - \bar{u}, v - \bar{v}) \quad (\text{Eq. 21})$$

$(u, v) \rightarrow$ *fluxo em determinado instante*

$(\bar{u}, \bar{v}) \rightarrow$ *fluxo médio*

Essa divisão parte do mesmo princípio que $u = \bar{u} + u'$, porém operando no campo de velocidade resolvido pelo modelo. Análises com MKE e EKE podem envolver diferentes tratamentos para o fluxo médio (média anual, mensal, decadal, etc) e para o fluxo turbulento (instantâneo, média em 5 dias, etc), levando à mensuração de diferentes escalas de turbulência representada pelos resultados.

Dessa forma, pode-se afirmar que a energia cinética turbulenta é proporcional à energia contida nas flutuações do fluxo médio – portanto proporcional à energia contida em vórtices e meandramentos e uma medida direta da importância relativa dos termos não-lineares da equação de movimento.

1.7 Não-linearidade, efeito borboleta e assimilação de dados 4D-VAR

Ainda que formuladas a cerca de 50 anos, as questões expostas nos artigos de Lorenz (1963; 1969) ainda são discutidas em artigos recentes no campo da assimilação de dados (Lorenz e Payne, 2007; Pires, 1996).

A assimilação de dados variacional ajusta a solução global do modelo utilizando as observações disponíveis ao longo de todo período de assimilação (*janela de assimilação*). Assegura-se que as informações contidas nas observações são propagadas temporalmente ‘para frente’ e ‘para trás’, ou seja, o estado do modelo em qualquer

instante e em qualquer ponto depende de toda e qualquer observação dentro da janela de assimilação (Pires, 1996).

Como já introduzido, os algoritmos variacionais determinam a melhor estimativa do estado do sistema (análise) através da minimização de uma função escalar que mede o desvio entre as observações e a solução do modelo, a *função custo*. A função custo é definida como o quadrado da distância entre a análise e as duas fontes de informação (observações e soluções do modelo), ponderada pelas estimativas dos erros de cada fonte, que possuem presumidamente uma distribuição Gaussiana (Eq. 5).

A presença de não-linearidades tanto no modelo quanto nas observações potencialmente causam desvios na distribuição Gaussiana dos erros do modelo e das observações criando complexidades na “topologia” da função custo que pode conter múltiplos mínimos. Os experimentos utilizando assimilação variacional no modelo de Lorenz (três equações diferenciais) apresentados em Pires (1996) evidenciam que os múltiplos mínimos na função custo são também função da extensão da janela de assimilação: quanto maior a janela, mais complexa é a topologia da função custo, que apresenta maior número de “vales” (Figura 2). A melhor estimativa do estado do sistema é, teoricamente, a que se encontra no mínimo global.

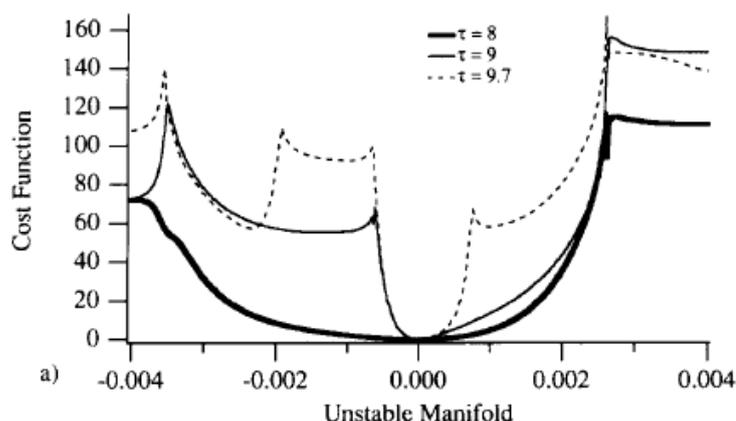


Figura 2: Função custo de três experimentos com diferentes extensões de janelas de assimilação (8, 9 e 9.7). Nota-se que ao aumentar as janelas, aumenta-se o número de “vales” na função custo. Retirado de Pires (1996) (Figura 4a).

Porém, obter o mínimo global da função custo em modelos atmosféricos ou oceânicos realísticos é computacionalmente inviável e por isso simplificações são necessárias (Bonavitta, 2018). A principal delas, proposta por Courtier et al. (1994), é a linearização do modelo e do operador de observação, o que torna a função custo

quadrática e permite a obtenção do mínimo global através do uso de algoritmos de minimização iterativos.

Este método é chamado de variacional-incremental e consiste em minimizar a função custo da solução não-linear através da minimização de uma série de funções custo quadráticas, obtidas a partir de soluções do modelo linearizado, na qual pequenos incrementos no vetor controle são aplicados progressivamente. Logo, quanto maior for a janela temporal de assimilação, maior será a tendência a existirem múltiplos mínimos na topologia da função custo não-linear e, portanto, mais grosseira será sua aproximação linear e conseqüentemente menos válidos serão os incrementos no vetor controle.

Os erros nas soluções dos modelos crescem em uma escala de tempo inversamente proporcional à resolução espacial utilizada (Lorenç e Payne, 2007). As feições turbulentas representadas possuem escala de tempo características (a escala de tempo de formação, translação e dissipação de um vórtice de sub-mesoescala é menor do que a escala de tempo de um vórtice de mesoescala). Portanto, a não-linearidade “age” mais rápido em divergir as soluções quanto maior for a resolução espacial do modelo. Em outras palavras, quanto maior a resolução espacial do modelo, menor será o tempo para que os erros nas condições iniciais se propaguem. Esta constatação é relacionada com o “efeito borboleta”, que descreve um limite finito de previsibilidade devido a existência de uma sucessão de escalas menores.

Como podemos ver, a extensão da janela de assimilação está intimamente relacionada com a não-linearidade da dinâmica local e sua conseqüente influência na topologia da função custo, e com a resolução espacial do modelo, devido à escala temporal de evolução das feições turbulentas representadas.

1.8 Justificativa e Motivação

O território marítimo do Brasil (Zona Econômica Exclusiva) tem 4,4 milhões de quilômetros quadrados, com muitas riquezas marinhas cuja defesa e utilização sustentável é de nossa responsabilidade. Dentre as principais riquezas, pode-se listar em primeiro lugar a gigantesca biodiversidade, a pesca, o petróleo e gás, minerais metálicos no subsolo marinho e o potencial aproveitamento da energia de maré e eólica *offshore*. Segundo o censo de 2010 (IBGE, 2010) o litoral do Brasil tem 7.491 km de extensão e

concentra cerca de 26,6% da população (~50 milhões de brasileiros). Estes dados reforçam a ideia da importância do conhecimento oceanográfico para o Brasil, e o desenvolvimento de sistemas de observação e modelagem oceânica tem um papel importante, sendo demandados para diversos fins. Pode-se destacar, por exemplo, a utilização de modelos hidrodinâmicos para a segurança de navegação e de operações marítimas e o manejo de resíduos lançados ao mar, tanto costeiros quanto oceânicos, planejados ou acidentais.

Na maior parte dos casos, para obter modelos adequados às demandas, ou seja, em alta resolução temporal e espacial e com os processos físicos relevantes representados - maré, fenômenos de meso e larga escala, ondas de plataforma, entre outros - e com boa representação do contorno de fundo, a técnica de aninhamento de modelos é muito utilizada. Desta forma, aninha-se um modelo regional em modelos globais, o que consiste em utilizar os resultados de um modelo global como fonte de informação (condições de contorno e iniciais) para o modelo regional. Assim, objetiva-se manter o modelo regional coerente com os resultados do modelo global, beneficiando-se do seu sistema de assimilação de dados locais, possibilitando o aumento da resolução e da acurácia de informações batimétricas, de linha de costa e de maré locais.

Portanto, acredita-se na importância de que instituições de pesquisa brasileiras se capacitem na aplicação de técnicas de assimilação de dados, tanto para obter autonomia dos produtos globais gerados por instituições européias e norte-americanas, quanto para se inserir no desenvolvimento tecnológico desta área e para ter liberdade de utilizar os dados coletados por redes de observação brasileiras em modelos com assimilação.

Neste contexto, o Projeto Azul figura como uma iniciativa pioneira na área de oceanografia operacional brasileira. O Projeto Azul é um projeto de Oceanografia Operacional iniciado em 2012, baseado em três componentes, uma rede de observação oceanográfica na região da Bacia de Santos, a modelagem computacional e a assimilação de dados (dos Santos, *et al.* 2016). Dentre os dados coletados destacam-se dois anos contínuos de perfis de temperatura e salinidade obtidos por Veículos Autônomos Submarinos, perfiladores do tipo Argos e dados de corrente e temperatura superficiais obtidos por derivadores oceânicos. Na componente modelagem e assimilação de dados, o projeto implementou o primeiro sistema de assimilação

4DVAR (variacional tetradimensional) em um modelo regional nos mares brasileiros (Fragoso *et al.*, 2016).

O trabalho de implementação deste sistema seguiu a orientação de desenvolvedores do 4DVAR no ROMS (Andrew Moore – UCSC e Hernan Arango – *Rutgers University*) e hoje pesquisadores brasileiros envolvidos já participam de desenvolvimentos do sistema (Faller, 2016), além de aprimoramentos do primeiro experimento (Mariano, 2019) e de aplicações do conhecimento adquirido em outras regiões do Brasil contribuindo para fortalecer o sistema de proteção ambiental nacional.

Ainda que a primeira implementação do sistema tenha alcançado bons resultados (Fragoso *et al.* 2016), muitos dos parâmetros e configurações foram determinados a partir da bibliografia e orientações dos desenvolvedores e, portanto, carecem de verificações diretas. Dentre elas, se encontra a determinação da extensão da janela temporal de assimilação que, conforme introduzido, possui grande impacto em sistemas 4D variacionais e é dependente do grau de não linearidade do sistema – portanto variável tanto espacialmente quanto temporalmente. Por um lado, quanto menor for a janela de assimilação, maior a validade da premissa linear e, portanto, mais válida será a solução final do sistema. Por outro, quanto maior a janela de assimilação, maior o número de observações disponíveis para ajustar o sistema. Logo, entende-se que este parâmetro é altamente variável não só de acordo com a região, mas também com a rede de observações disponível. Desta forma, neste trabalho investiga-se objetivamente este parâmetro visando tanto encontrar a configuração mais adequada - dentre os experimentos conduzidos e dadas as métricas de avaliação utilizadas - quanto gerar elementos para discussão e evolução do sistema de assimilação de dados implementado.

2 Objetivos

O objetivo geral desta dissertação é o aprimoramento de um sistema de assimilação de dados do tipo Variacional Tetra-Dimensional Incremental (I4DVAR) implementado para a região sudeste brasileira. O foco da investigação é na definição da janela temporal de assimilação.

2.1 Objetivos Específicos

Os objetivos específicos desta dissertação são:

- I. Apresentar os passos da implementação de um modelo com assimilação 4DVAR
- II. Comparação de experimentos utilizando janelas temporais distintas: 4, 7 e 10 dias.

3 METODOLOGIA

Para alcançar os objetivos, foi utilizada o sistema de assimilação de dados do tipo IS4DVAR (*Incremental Strong-Constraint 4-Dimensional Variational*) disponível no ROMS (*Regional Ocean Modeling System*). A configuração básica do modelo e do sistema de assimilação utilizado parte do trabalho de Fragoso *et al.* (2016). Como será visto, o sistema implementado utilizou parâmetros específicos do 4DVAR obtidos para a região e coerentes segundo a literatura disponível e especialistas da área. Porém, pela alta complexidade e dependência da dinâmica regional, essa primeira implementação – apesar de bem sucedida – é passível de aprimoramentos.

Neste trabalho, a escolha de Fragoso *et al.* (2016) da janela de assimilação de 7 dias é confrontada com dois experimentos partindo das mesmas configurações porém com janelas de 4 e 10 dias. Como a configuração do número de *inner* e *outer loops* e em grande parte o custo computacional envolvido dependem da extensão da janela de assimilação utilizada, a investigação conduzida tem o potencial tanto de melhorar a representação da dinâmica quanto de reduzir os custos do sistema.

A seguir apresenta-se o modelo e o esquema de assimilação de dados utilizado (ROMS IS4DVAR), a configuração básica do modelo não-linear (domínio, fontes de dados de contornos e forçantes), a configuração básica do sistema de assimilação de dados, o experimento realizado nesta dissertação (alteração da janela de assimilação) e as métricas de comparação utilizadas.

3.1 Regional Ocean Modeling System (ROMS)

Amplamente utilizado pela comunidade científica, o ROMS é um modelo oceânico de superfície livre que resolve as equações primitivas de movimento considerando as aproximações hidrostática, de *Boussinesq* e de fluido incompressível pelo método de diferenças finitas. Na horizontal, o sistema utiliza coordenadas curvilíneas (esféricas ou cartesianas) em uma grade do tipo C-Arakawa. Na vertical, as equações primitivas são discretizadas utilizando coordenadas seguidoras de terreno (*sigma*). Tanto na discretização horizontal quanto vertical há flexibilidade de controle da resolução em áreas de interesse (p. ex. aumento da resolução na camada de mistura ou de fundo). As equações de movimento são resolvidas utilizando um esquema de passo de tempo com separação entre a parcela barotrópica (mais rápida) e baroclínica (mais lenta), ou seja, n passos de tempo barotópicos entre passos de tempo baroclínicos. O código do modelo é altamente modular permitindo que o usuário configure a física do modelo selecionando o método de advecção e difusão horizontal e vertical, fechamento turbulento, tipo de condição de contorno e de superfície, entre outros. O código pode ser rodado tanto em série quanto em paralelo, onde a grade 3D do modelo é particionada em *tiles* para serem processados por cada núcleo de processamento. O modelo possui uma ativa rede de usuários e desenvolvedores interagindo em fórum. Uma descrição completa sobre as parametrizações físicas e descrições dos sistemas numéricos podem ser encontradas em Marchesiello *et al.* (2001), Shchepetkin e McWilliams (2003, 2005), Haidvogel *et al.* (2008) assim como mais detalhes e referências podem ser encontrados no *website* www.myroms.org.

3.2 Configuração do NLROMS

Neste capítulo será descrita a configuração do modelo não-linear, que segue a implementação de Frago *et al.* (2016).

O domínio do modelo cobre a região oceânica leste e sudeste brasileira, de 50°W a 25°W e de 15°S a 30°S com uma resolução horizontal de 1/12° (~8 km) e 40 níveis verticais. Pela resolução espacial (~8 Km) adotada, a configuração do modelo livre permite a solução de feições turbulentas de mesoescala na região do domínio de modelagem.

A grade numérica possui 306x181 pontos de grade com três contornos laterais abertos – norte, sul e leste, enquanto o contorno oeste é definido pela linha de costa. A batimetria é oriunda do produto ETOPO1 (Amante&Eakins, 2009) que possui resolução de 1' de grau (cerca de 1,8 km) e cobertura global (Figura 3).

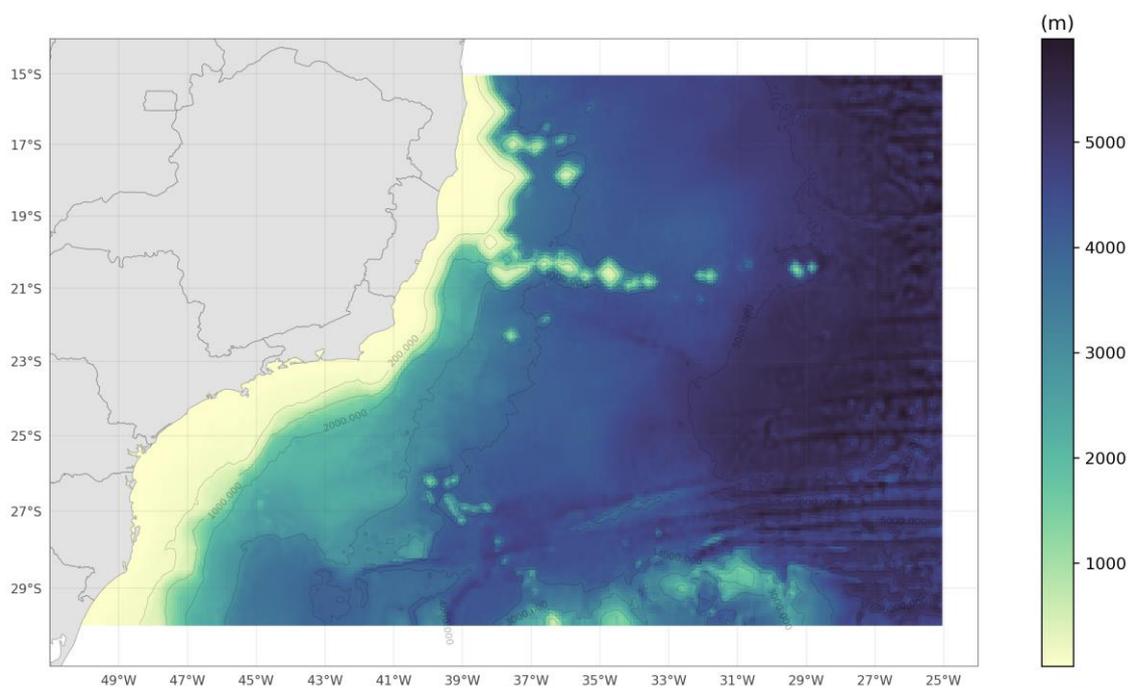


Figura 3: Domínio do modelo e batimetria utilizada (ETOPO1).

O modelo é configurado para utilizar a condição de contorno para superfície livre e velocidade 2D descritas em Chapman (1985) e Flather (1976), respectivamente, e condição de contorno radiacional com nudging para velocidade 3D e traçadores (Marchesiello *et al.* 2001), com escala temporal do *nudging* de 1 e 30 dias para fluxos de entrada e saída, respectivamente.

As condições iniciais e os fluxos laterais do modelo livre são oriundos da reanálise global do modelo oceânico disponibilizado pelo CMEMS (*Copernicus Marine Environment Monitoring System*), especificamente o produto *Global Ocean 1/12° Physics Analysis and Forecast updated daily*, disponível para download pelo site da agência europeia (<http://marine.copernicus.eu/>). A reanálise do CMEMS consiste em médias diárias do campo de temperatura, salinidade, velocidade e elevação da superfície livre em uma grade de 1/12° de resolução.

A condição de contorno atmosférica utiliza a formulação de *Bulk Flux* na qual os fluxos oceano-atmosfera são calculados internamente a partir das variáveis atmosféricas (velocidade do vento a 10m, temperatura a 2m, precipitação e radiação de ondas curtas e longas). As variáveis atmosféricas são oriundas da reanálise global *NCEP_Reanalysis 2* disponibilizadas pela agência americana *NOAA/OAR/ESRL* através do site <https://www.esrl.noaa.gov/psd/> (Kanamitsu et al. 2002). Os dados atmosféricos utilizados possuem resolução espacial de aproximadamente 2,5° e temporal de 6 horas.

3.3 Observações

O conjunto de observações assimiladas inclui dados satelitais de temperatura da superfície do mar (TSM) e de altura da superfície do mar (ASM) e dados *in-situ* de temperatura e salinidade (*in-situ T e S*).

Os dados de TSM são oriundos de sensores AVHRR (*Advanced Very High Resolution Radiometer*) a bordo de satélites do projeto POES (*Polar Operational Environmental Satellites*). Os dados são disponibilizados como compostos diários com resolução espacial de 0,1°. Devido à variabilidade da cobertura de nuvens, os dados nem sempre cobrem todo o domínio (Figura 4). Apesar de existirem produtos de TSM satelital com maior cobertura e resolução espacial, como por exemplo o JPL-MUR, o mesmo não é adequado para assimilação de dados pois a mescla de sensores de infravermelho (que não possuem a limitação de nuvens porém apresentam baixa resolução espacial) e o processo de combinação dos diversos satélites gera um produto espacialmente completo e de alta resolução, porém com maior erro associado.

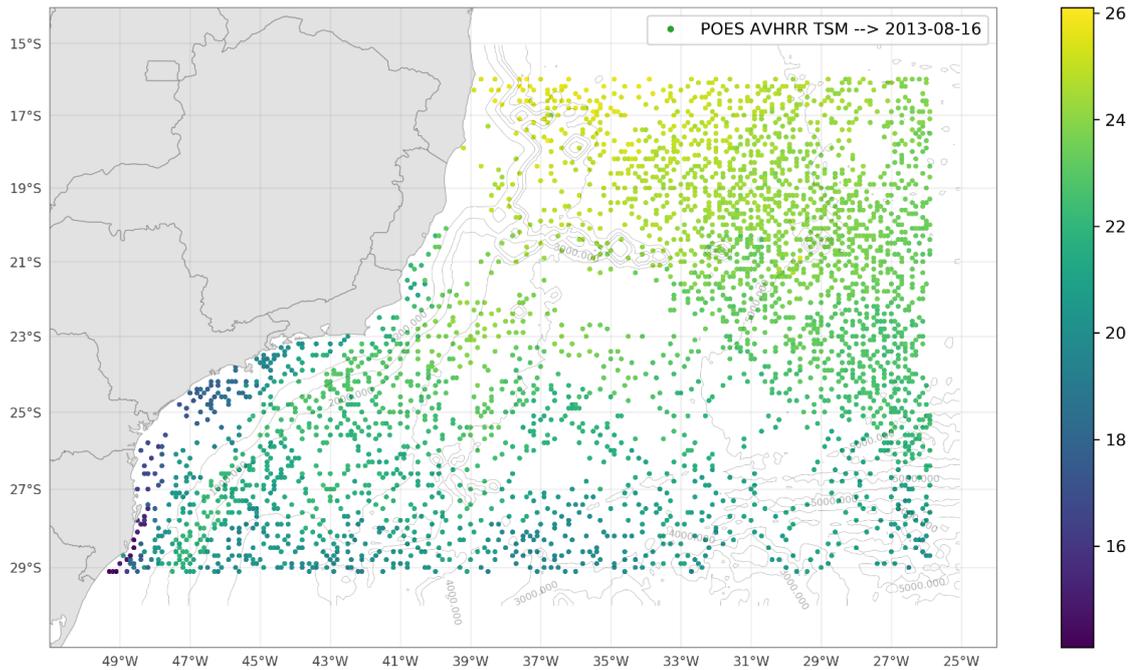


Figura 4: Exemplo do dado de TSM utilizado nos experimentos. Fonte: POES AVHRR.

Os dados de ASM são oriundos do produto gridado de topografia dinâmica distribuída pelo AVISO com aproximadamente $0,3^\circ$ de resolução. Os dados são pré-processados para compatibilizar o nível médio com o nível médio do modelo. A média do dado de ASM é subtraída do dado e a média da ASM do modelo é adicionada. A seguir, foi aplicada uma máscara para evitar a assimilação dos dados entre a costa e a isóbata de 200 metros. Este procedimento é aplicado pois os dados satelitais de ASM apresentam contaminação com a proximidade da costa e também pelo fato de não amostrarem com suficiente resolução espacial e temporal os processos sobre a plataforma continental, terminando por inibir a representação de importantes feições como, por exemplo a passagem de ondas costeiras aprisionadas.

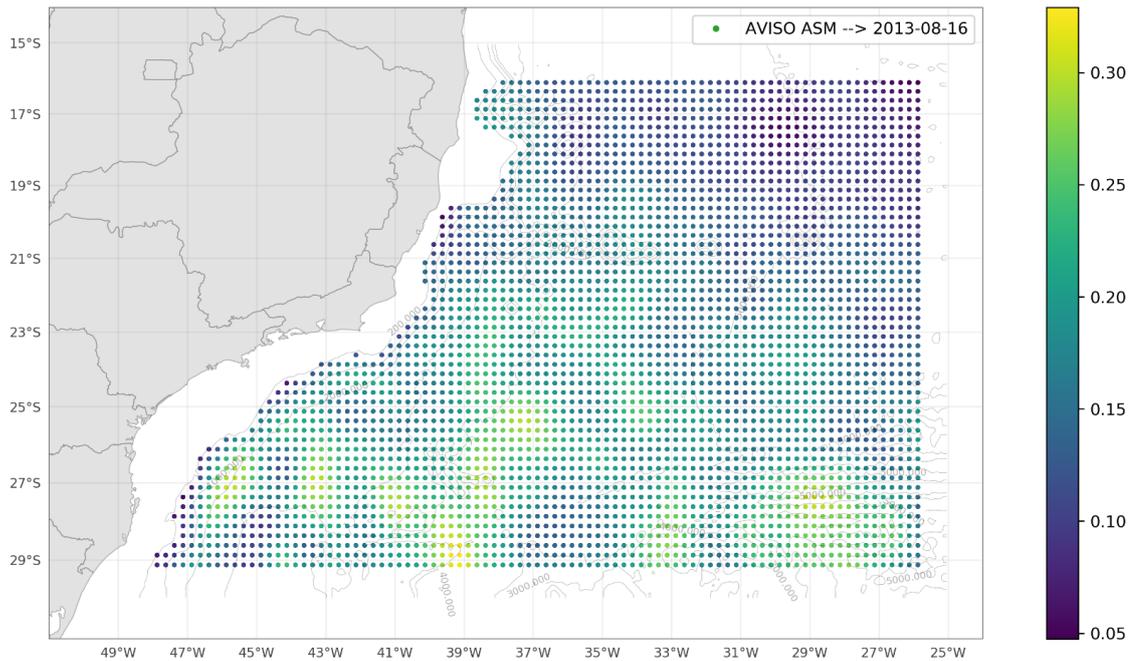


Figura 5: Exemplo do dado de ASM utilizado nos experimentos. Fonte: AVISO.

Os dados *in-situ* de temperatura e salinidade foram obtidos de duas fontes: a base de dados hidrográficos EN4 (Good *et al.* 2013) e os dados coletados no sistema de observação oceânico do Projeto Azul (dos Santos, *et al.* 2015). Os dados do EN4 consistem em um compilado de observações de subsuperfície de temperatura e salinidade com um robusto sistema de controle de qualidade distribuídos gratuitamente pelo *UK Met Office Hadley Center*. Os dados do Projeto Azul são densamente concentrados na região de estudo e foram coletados através da utilização de Gliders (até ~1.000 m) e Perfiladores (~2.000 m) entre 2013 e 2016. Além de dados de subsuperfície, o sistema de observação do Projeto Azul realizou lançamentos periódicos de derivadores de superfície do tipo SVP, que fornecem dados de TSM e velocidade da corrente superficial ao longo das trajetórias. Os dados de derivadores não foram assimilados e portanto puderam ser utilizados como dados independentes para avaliação dos experimentos.

A localização dos perfis hidrográficos assimilados no período dos experimentos (maio a novembro de 2013) é apresentado na Figura 6. Apesar de mais distribuídos no domínio, os dados do EN4 perfazem um número significativamente menor de perfis (622 perfis) quando comparado com os perfis do Projeto Azul (4606 perfis).

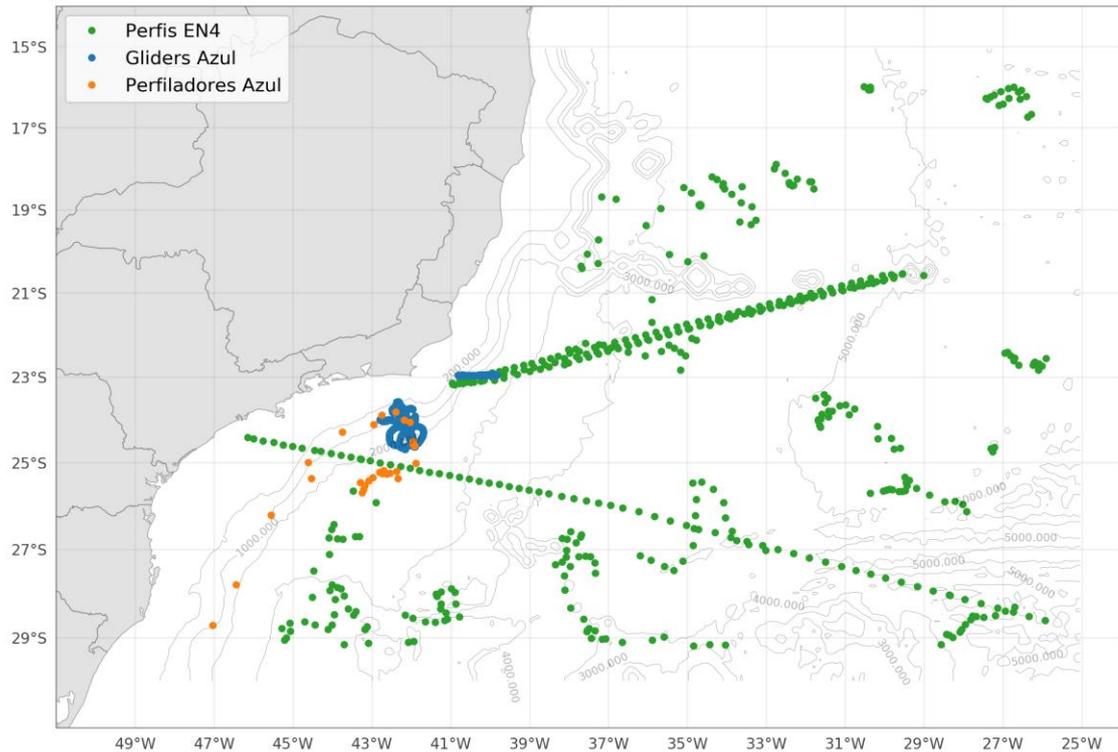


Figura 6: Localização dos perfis de temperatura e salinidade entre maio e novembro de 2013, assimilados nos experimentos. Os pontos verdes representam perfis da base de dados EN4, os pontos azuis representam perfis coletados por Gliders do Projeto Azul e os pontos laranja representam dados de perfiladores, também do Projeto Azul.

As observações são promediadas nas células da grade do modelo de forma que, no caso de existirem mais de uma observação da mesma variável em uma célula de grade em um intervalo definido de tempo, um único valor médio é utilizado – chamado de “super-observação”. O intervalo definido para promediação das super-observações nestes experimentos foi de 12 horas.

Os elementos da matriz de covariância do erro (\mathbf{R}) são derivados de diversas fontes de erros como a estimativa de erro dos sensores, erros de interpolação e erros de representatividade – processos físicos não representados pelo modelo devido à resolução espacial. Em resumo, os erros prescritos para as observações devem representar a medida que o sistema de assimilação de dados é capaz de se ajustar ao dado (Fragoso, *et al.* 2016). Desta forma, foram prescritos os seguintes valores de erros: 0,1°C e 0,4°C para dados *in-situ* e satelitais de temperatura, respectivamente, 2 cm para dados de ASM e 0,05 para dados *in-situ* de salinidade. Esses valores prescritos são modificados no processo de obtenção das super-observações – o valor de erro da super-observação é definida como o maior valor entre o erro prescrito e o desvio padrão do

conjunto de observações contidas na mesma célula da grade e intervalo de tempo definido.

3.4 Configuração do 4DVAR

De um modo geral, a configuração do 4DVAR seguiu o experimento descrito em Frago *et al.* (2016). Os experimentos utilizaram o sistema 4DVAR do tipo primal (realizado no subespaço do modelo), incremental e de restrição forte (considera que o modelo não linear é livre de erros), conforme implementado no ROMS. Considerou-se como vetor controle as condições iniciais $x(t_0)$, os fluxos nos contornos laterais $b(t)$ e os fluxos de superfície $f(t)$.

A prescrição da matriz de covariância dos erros da solução prévia \mathbf{D} é um fator de grande impacto no sistema de assimilação 4DVAR visto que determina como as informações das observações serão propagadas espacialmente e para as variáveis não assimiladas. A matriz diagonal \mathbf{D} é composta pelas matrizes de covariância dos erros de cada elemento do vetor controle (\mathbf{B}_x , \mathbf{B}_b e \mathbf{B}_f) que, por sua vez, são fatoradas segundo de acordo com:

$$\mathbf{B} = \mathbf{K}_b \mathbf{\Sigma} \mathbf{C} \mathbf{\Sigma}^T \mathbf{K}_b^T \quad (\text{Eq. 22})$$

Onde $\mathbf{\Sigma}$ é uma matriz diagonal de desvios padrão e \mathbf{C} é a matriz de correlação e \mathbf{K}_b é a componente de equilíbrio de erros da solução prévia – não utilizada neste experimento.

A matriz de correlação \mathbf{C} é expressa como a solução de uma equação de difusão onde a escala de comprimento de correlação L é aproximadamente a escala de comprimento de correlação natural da variável. Esta escala de correlação controla a extensão da extrapolação dos dados observados e, portanto deve ser relacionada com a escala natural de correlação espacial de cada variável e com o adensamento das observações. Caso L seja muito grande, a influência de cada observação será extrapolada para áreas mais distantes do ponto de observação do que deveriam, se sobrepondo a outras observações da mesma variável, reduzindo a independência de cada observação. Os experimentos conduzido nesta pesquisa seguiram os valores de L utilizados em Frago *et al.* (2016), obtidos a partir dos resultados de uma análise de semi-variograma (utilizando o resultado de uma rodada não-assimilativa), da escala das

principais feições da região de estudo e da resolução do modelo. Desta forma, definiu-se uma escala de comprimento de correlação horizontal de 100 km para altura da superfície livre, temperatura e salinidade e 60 km para as componentes zonal e meridional da corrente. Para a escala de comprimento de correlação vertical, foi utilizado 100 metros para todas as variáveis.

O desvio padrão da solução prévia para as condições iniciais e de contorno foi obtido a partir de uma rodada não assimilativa, refletindo a premissa de que as correções no vetor controle não devem exceder a variabilidade natural. Conforme observado no trabalho de Mariano (2019), o sistema I4DVAR implementado no ROMS apresenta grande sensibilidade às alterações nos valores de desvio padrão utilizados na obtenção da matriz de covariância do erro da solução prévia, sendo sua definição um dos pontos de ativa discussão no meio especializado. Mattern *et al* (2017) sugerem um método iterativo para calibrar as matrizes de covariância de erro (tanto da solução prévia quanto das observações) a partir da análise de consistência descrita em Desroziers *et al.* (2005) em um método que leva em conta mais a consistência matemática dos resultados frente às premissas do sistema 4DVAR do que qualquer pressuposto com sentido físico. O presente trabalho não objetiva propor alterações neste tópico e utiliza valores de desvio padrão mensal obtidos de uma rodada de 1,5 anos do modelo livre. Desta forma, os desvios padrão são espacialmente variáveis e a cada ciclo de assimilação - ao longo dos múltiplos ciclos que compõem os experimentos - o algoritmo de controle prescreve o desvio padrão relativo ao mês no qual o ciclo se encontra.

A quantidade de *inner* e *outer loops* foi definida de forma a obter uma boa estimativa do mínimo global de J_{NL} com um custo computacional viável. Após uma série de testes foram definidos que os experimentos utilizariam 10 *inner loops* e 2 *outer loops*. Esta definição levou em conta testes considerando uma janela de assimilação de 7 dias de extensão e, portanto, podem não ser a melhor escolha para os experimentos aqui apresentados com janelas de assimilação de 4 e 10 dias.

3.5 Área de Estudo

A área de estudo compreende a região oceânica leste e sudeste brasileira, entre as latitudes de -30°S e -15°S e longitudes de -50°W a -25°W (Figura 3). Neste item, será apresentada uma breve descrição da circulação de meso e larga escala na região de estudo.

O sistema de correntes na região de estudo próximo ao talude continental é estratificado e apresenta alta instabilidade baroclínica. Pode ser descrito em três principais camadas— camada superficial dominada pela Corrente do Brasil (CB) com sentido Sul, a Corrente de Contorno Intermediária (CCI) com sentido Norte e a Corrente de Contorno Profunda com sentido Sul (Figura 7).

A circulação superficial é dominada pela Corrente do Brasil, uma corrente de contorno oeste formada a partir da bifurcação do ramo sul da Corrente Sul Equatorial que ocorre entre 5° e 10°S (variação Sazonal). A CB é responsável pelo fechamento oeste do Giro Subtropical do Atlântico Sul, transportando águas quentes e salinas da região equatorial para o Sul (Peterson e Stramma, 1991). A CB transporta as massas d'água Água Tropical (AT) e Água Central do Atlântico Sul (ACAS). A espessura da CB varia desde a sua formação, ganhando em transporte com o aumento de latitude. Na região de formação apresenta cerca de 50 a 100 metros, chegando a aproximadamente 500 metros na região do Rio de Janeiro.

Abaixo da CB, com direção Norte, encontra-se a Corrente de Contorno Intermediária (CCI), que transporta a Água Intermediária Antártica (AIA) para o Norte. A CCI é formada a partir da bifurcação de Santos, em torno de 27°S e ocupa a porção intermediária da coluna d'água, aproximadamente entre 500 e 1300 metros, abaixo da CB e acima da CCP (Silveira, *et al.* 2000). A Corrente de Contorno Profundo flui para Sul entre a CCI e o fundo oceânico, transportando a Água Profunda do Atlântico Norte (APAN) (Silveira, *et al.* 2007).

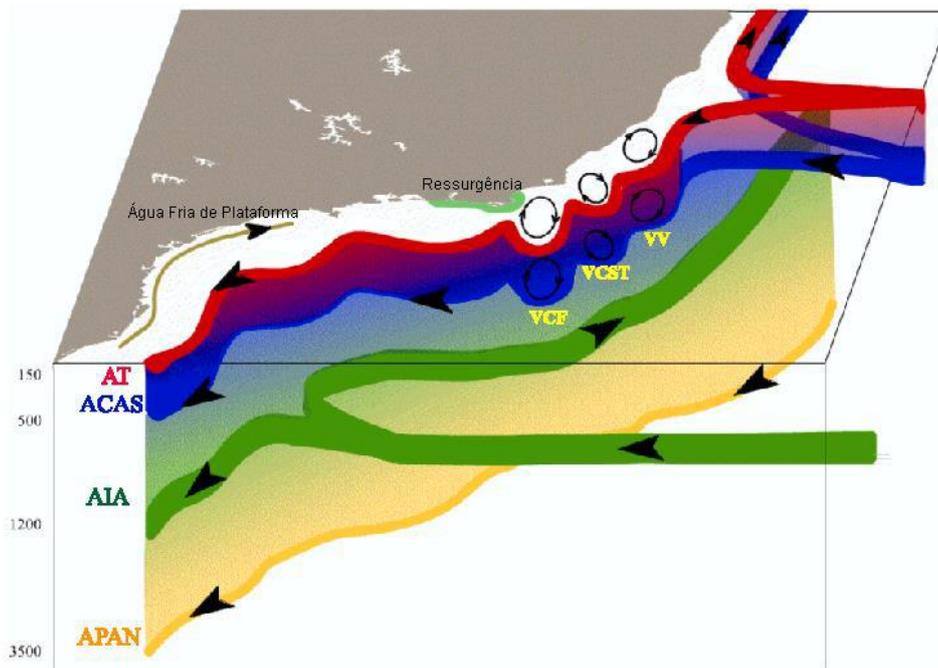


Figura 7: Representação esquemática da estratificação vertical das correntes na região leste e sudeste do Brasil, assim como da região de ocorrência dos vórtices de Vitória (VV), Vórtice do Cabo se São Tomé (VCST) e Vórtice de Cabo Frio (VCF). Na região sobre a plataforma, destacam-se a Corrente Costeira do Brasil (CCB) e a região de ocorrência do fenômeno da ressurgência. Retirado de Godoi (2005).

A partir de aproximadamente 19°S, o sistema de correntes apresenta uma intensa atividade de mesoescala. Meandros e vórtices ciclônicos e anticiclônicos podem ser observados ao longo de todo o ano e apresentam uma escala de comprimento de aproximadamente 100 km de diâmetro. A formação desses vórtices é relacionada com algumas condições tais como à abrupta quebra de orientação das isóbatas no talude causando um aumento na vorticidade potencial e consequente meandramento (Campos *et al.* 1995); ao cisalhamento vertical entre a Corrente do Brasil e a Corrente de Contorno Intermediária que é responsável pela instabilidade baroclínica, “alimentando” a velocidade dos vórtices e meandros (Silveira, 2007). Dentre os vórtices gerados destacam-se os Vórtice de Vitória (ciclônico) formado no embainhamento de Tubarão logo após a passagem da CB pela Cadeia Vitória-Trindade; o Vórtice do Cabo de São Tomé na região de Campos (RJ) e o Vórtice de Cabo Frio (VCF), com giro ciclônico e local de formação em aproximadamente 24°S e 42°W. Durante a ocorrência do VCF, a estratificação entre a CB e a CCI se desfaz e o vórtice pode ser observado até ~1000 metros de profundidade.

Na plataforma continental, as correntes encontram-se em equilíbrio geostrófico com os ventos e, na maior parte do ano, domina a influência do Anticiclone Subtropical do Atlântico Sul com ventos de NE na região gerando correntes resultantes para SW. Na ocorrência da passagem de sistemas frontais, o padrão do ASAS se interrompe e a influência dos ventos do quadrante Sul geram gradientes de pressão barotrópica inverso, forçando a propagação de ondas aprisionadas de plataforma e gerando pulsos de corrente para NE. A maior frequência de passagem de sistemas frontais no período de inverno acaba por gerar um transporte resultante para NE sobre a plataforma continental (Corrente Costeira do Brasil), transportando águas mais frias e ricas em nutrientes até a latitude de 24° aproximadamente (Souza e Robinson, 2004).

3.6 Configuração dos Experimentos

A investigação conduzida neste trabalho consiste em avaliar a influência da extensão da janela de assimilação em um sistema de modelagem oceânica com assimilação de dados 4DVAR com foco na bacia oceânica sudeste brasileira. Para isso, partindo da implementação de Fragoso *et al.* (2016) - descrita nos prévios itens – foram conduzidos experimentos gêmeos com alteração apenas da extensão da janela de assimilação. O experimento base com extensão de 7 dias é comparado com dois experimentos, com 4 e 10 dias.

Os experimentos cobrem o período de maio a novembro de 2013 (6 meses), em uma sucessão de múltiplos ciclos de assimilação nos quais a condição inicial de um ciclo é obtida do último tempo do vetor de estado produzido no ciclo anterior. A partir da condição inicial, forçantes atmosféricas e condições de contorno, o ciclo de assimilação inicia com a obtenção da solução prévia (integração do NLROMS) que será a base para a primeira linearização e sucessão de *inner loops* nos quais serão obtidos incrementos do vetor controle que conduzam à minimização da função custo – sujeita às observações contidas na janela temporal do ciclo atual de assimilação. Ao final dos 10 *inner loops* o vetor controle atualizado é utilizado em uma nova integração do modelo não linear que, por sua vez, será a base para a segunda linearização e sucessão de *inner loops* para a obtenção do vetor controle final e última integração do modelo não linear – a análise. Após o término de um ciclo de assimilação, um script gerenciador move os

resultados e dá sequência ao ciclo seguinte gerando o arquivo de condições iniciais (último tempo da última análise) e disponibilizando o arquivo contendo as observações do ciclo a ser iniciado (Figura 8).

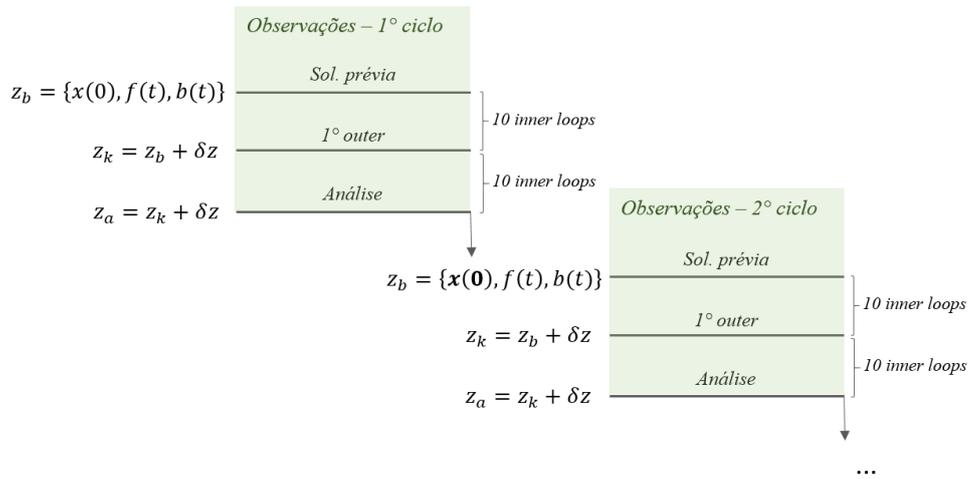


Figura 8: Representação esquemática da configuração dos experimentos com múltiplos ciclos.

A diferença básica entre os experimentos é a extensão da janela de assimilação. Enquanto o experimento 10DIAS irá cobrir o período de 20 dias em 2 ciclos, o experimento 4DIAS irá cobrir em 5 ciclos e o 7DIAS, em 3 (Figura 9). Isso implica diretamente na validade da premissa linear – teoricamente quanto maior a extensão, maior a potencial divergência entre a solução linear e a não linear. Por outro lado, quanto menor a extensão da janela de assimilação, menor é o número de observações disponíveis para ajustar os resultados – lembrando que o sistema 4DVAR utiliza as observações contidas ao longo de todo o período do ciclo de assimilação.

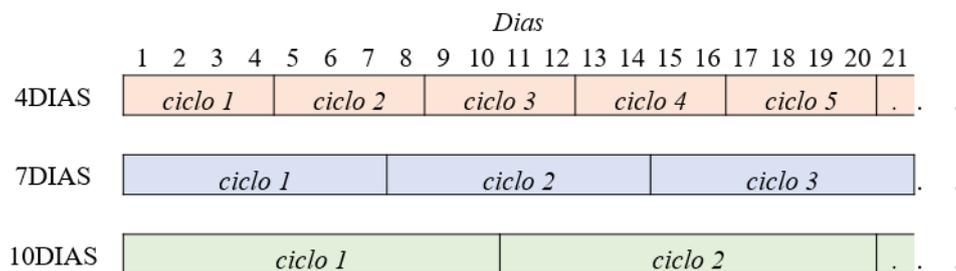


Figura 9: Configuração da extensão das janelas de assimilação em cada experimento.

Nota-se a diferença na distribuição do número de observações de TSM, ASM e de perfis hidrográficos (temperatura e salinidade) ao longo de cada ciclo de assimilação

dos 3 experimentos (Figura 10) e na média do número de observações por ciclo (Tabela 1).

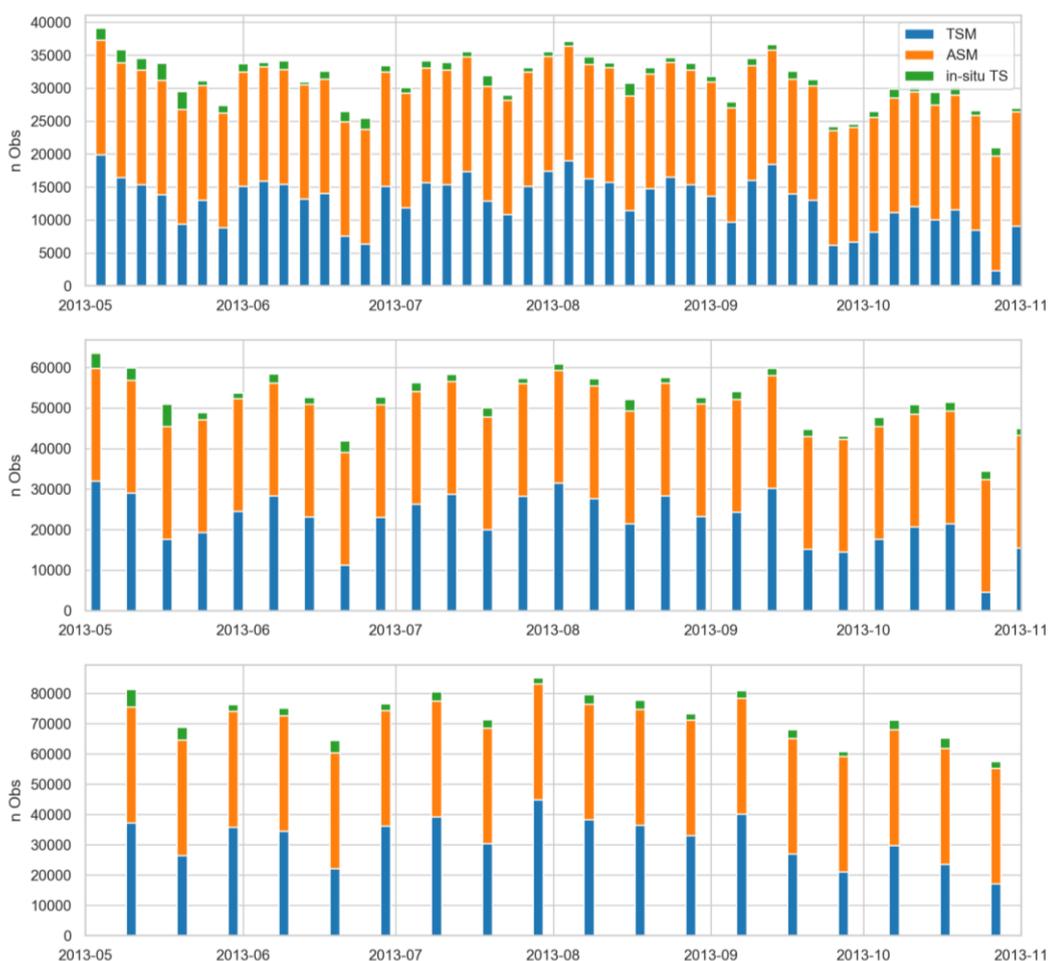


Figura 10: Número de observações por ciclo ao longo do período do experimento (maio a novembro de 2013). De cima para baixo, experimento 4DIAS, 7DIAS e 10DIAS. As observações *in-situ* de temperatura e salinidade foram condensadas.

Tabela 1: Número médio de observações por ciclo de assimilação para os três experimentos e para cada variável assimilada.

	4DIAS	7DIAS	10DIAS
ASM	17380	27808	38236
TSM	12812	22595	32032
<i>in-situ T</i>	690	1196	1864
<i>in-situ S</i>	556	966	1479

O primeiro ciclo de cada experimento utilizou como condição inicial os resultados de Frago *et al.* (2016) e foram iniciados um mês antes do período de análise (abril de 2013) para que o sistema tenha tempo de se ajustar nos primeiros ciclos.

3.7 Métricas de Avaliação

Neste item são apresentadas as métricas de erros e comparação utilizadas ao longo das análises dos resultados. O erro médio (Em) reflete erros sistemáticos no modelo, isto é, indica a tendência geral do erro de super ou subestimação, enquanto o RMSE indica a magnitude média dos erros ponto a ponto. O índice p indica a correlação linear entre o modelo e as observações e varia entre -1 (correlação negativa perfeita) e 1 (correlação positiva perfeita). O índice $Jfit$ (Ngodock *et al.* 2015) indica a relação entre os erros do modelo ($y_i - x_i$) e os erros prescritos para as observações (e_i^o). Logo, valores de $Jfit$ entre 0 e 1 são esperados em um sistema de assimilação de dados funcional.

$$Jfit = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - x_i}{e_i^o} \right| \quad (\text{Eq. 23})$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2} \quad (\text{Eq. 24})$$

$$p = \frac{1}{n\sigma_y\sigma_x} \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x}) \quad (\text{Eq. 25})$$

$$Em = \frac{1}{n} \sum_{i=1}^n (y_i - x_i) \quad (\text{Eq. 26})$$

Onde “ y_i ” são os valores das observações e “ x_i ” são os valores do modelo no ponto e tempo das observações; “ e_i^o ” é o erro associado à observação “ y_i ”; “ \bar{x} ” e \bar{y} são os valores médios das observações e das variáveis correspondentes no modelo; “ σ_y ” e “ σ_x ” são os desvios padrão de “ y ” e “ x ”, respectivamente, e “ n ” é o número de observações. Os valores de “ x ” podem ser tanto da análise dos experimentos quanto da solução prévia – neste caso, análogo à previsão do sistema.

4 resultados e discussão

Após a conclusão dos três experimentos foram realizadas análises com o objetivo de, além de identificar a melhor configuração para o sistema em relação ao parâmetro avaliado, fomentar discussões a cerca da influência da não-linearidade da dinâmica na definição da extensão da janela de assimilação.

De forma a avaliar o desempenho dos experimentos comparativamente, este capítulo é composto por: apresentação de métricas de erro para dados dependentes (assimilados) e independentes (não assimilados); avaliação da minimização da função custo; avaliação da representação da energia cinética média e turbulenta; apresentação de campos de corrente superficiais; análise dos incrementos e análise do custo computacional.

4.1 Comparação com Dados Assimilados

Neste item são apresentadas análises baseadas no o erro “ $y_i - x_i$ ”, isto é, a diferença entre as observações e os valores do modelo amostrados no local e tempo das observações assimiladas. O erro calculado a partir da solução prévia e da análise de cada ciclo de assimilação será denominado \mathbf{E}_b e \mathbf{E}_a , também referidos como *misfit* inicial e final respectivamente.

4.1.1 Análise da Distribuição dos Erros

A comparação da distribuição da frequência acumulada de \mathbf{E}_a (valores absolutos) dos três experimentos evidencia que a diminuição da janela de assimilação favoreceu a redução do vetor inovação das análises de um modo geral (Figura 11). Especificamente para cada conjunto de observações, nota-se que a redução foi maior entre o experimento 10DIAS e os demais porém não houve tanta redução entre os 7DIAS e 4DIAS.

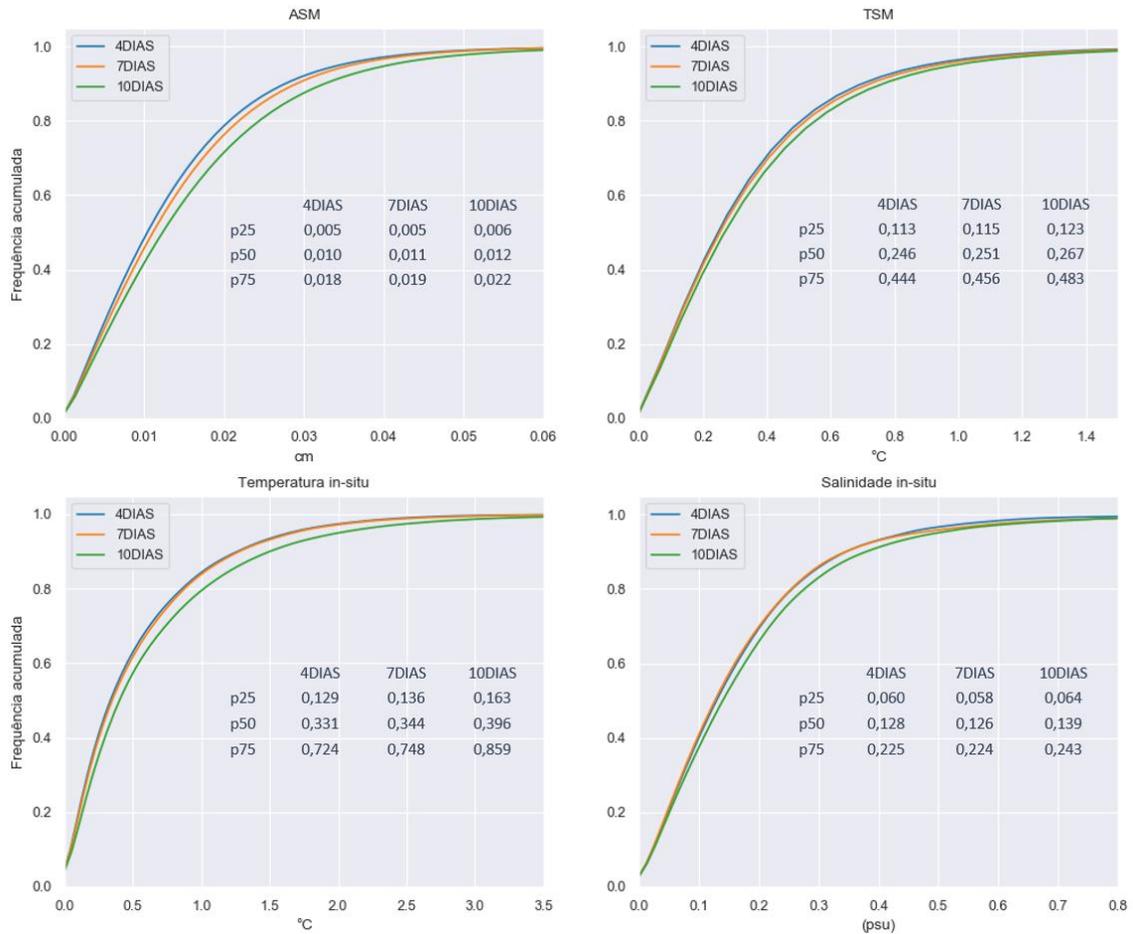


Figura 11: Probabilidade acumulada do erro da análise (**Ea**) dos experimentos com janela de assimilação de 4, 7 e 10 dias para ASM, TSM, e Temperatura e Salinidade *in-situ*, considerando toda a extensão analisada (maio a novembro/2013). Em cada figura os percentis 25, 50, 75 são indicados para os três experimentos comparados.

A seguir, apresenta-se o histograma da distribuição de frequência dos erros para cada variável de cada experimento separadamente. Em cada figura, junto com o histograma foi estimada a função de densidade de probabilidade Gaussiana a partir da média e desvio padrão de cada amostra representada (Figura 12, Figura 13 e Figura 14). Os erros médios e os desvios padrão dos erros encontram-se reunidos na Tabela 2. Os valores de erro médio e desvio padrão do erro de todas as variáveis e experimentos são bem próximos entre si. Contudo, podemos observar que os experimentos de 4DIAS e 7DIAS apresentam erro médio menores que o de 10DIAS e que o 4DIAS apresenta erro médio menor que o 7DIAS para as variáveis ASM e TSM enquanto o 7DIAS é melhor para as variáveis *in-situ*. O desvio padrão dos erros é menor quanto menor é a janela de assimilação.

Tabela 2: Média e desvio padrão (std) do erro da análise para cada experimento e variável analisada.

	4DIAS		7DIAS		10DIAS	
	<i>media</i>	<i>std</i>	<i>media</i>	<i>std</i>	<i>media</i>	<i>std</i>
ASM	0,0001	0,0171	0,0010	0,0179	0,0008	0,0202
TSM	-0,0016	0,4430	0,0037	0,4598	-0,0017	0,4879
T IN-SITU	0,0454	0,7538	0,0011	0,7770	0,0810	0,9221
S IN-SITU	-0,0096	0,2199	-0,0031	0,2344	-0,0049	0,2457

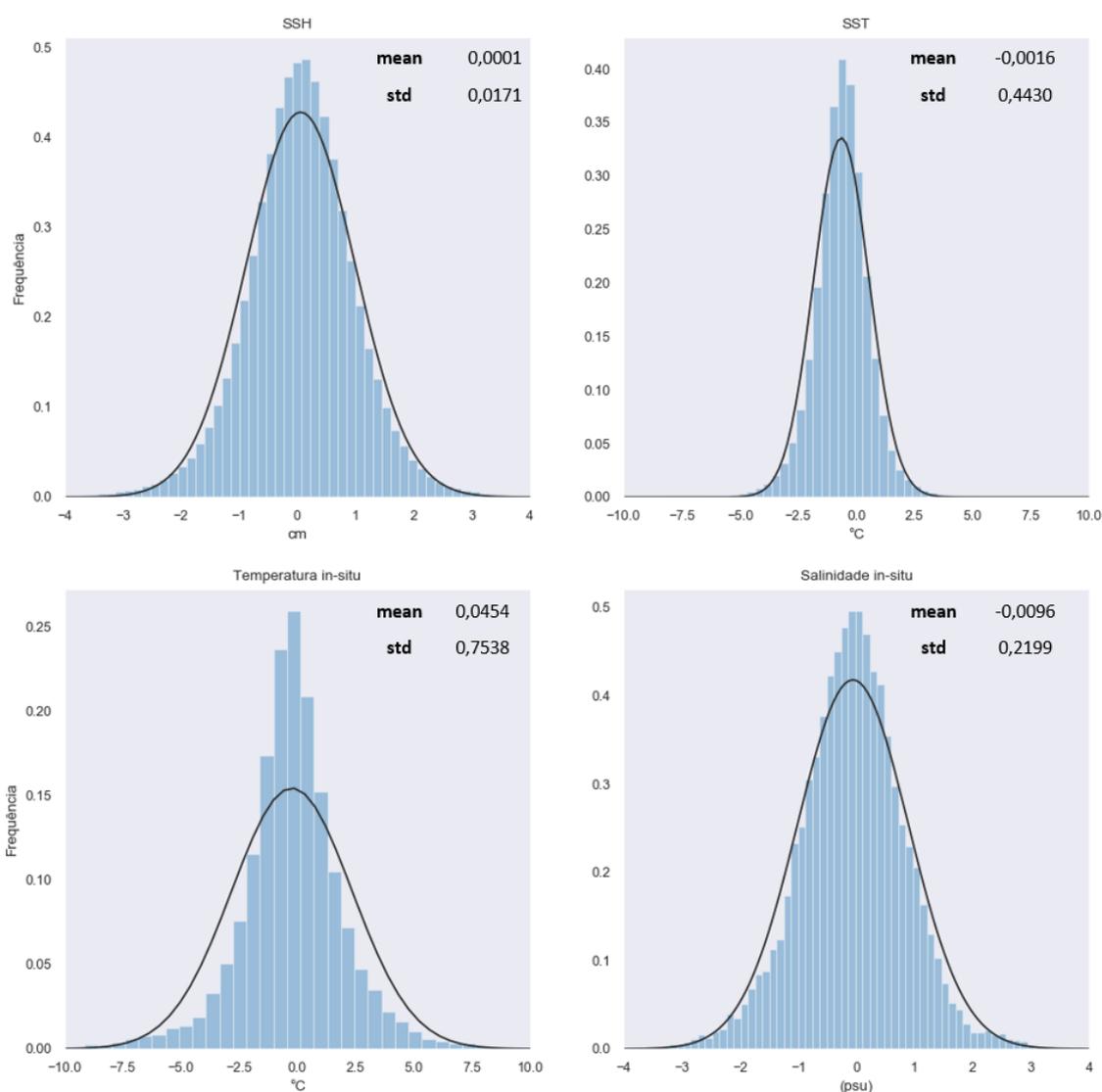


Figura 12: Distribuição de frequência do erro para as variáveis ASM, TSM,, Temperatura e Salinidade in-situ do experimento 4DIAS. A linha preta representa a função de densidade de probabilidade Gaussiana estimada para cada conjunto de amostras.

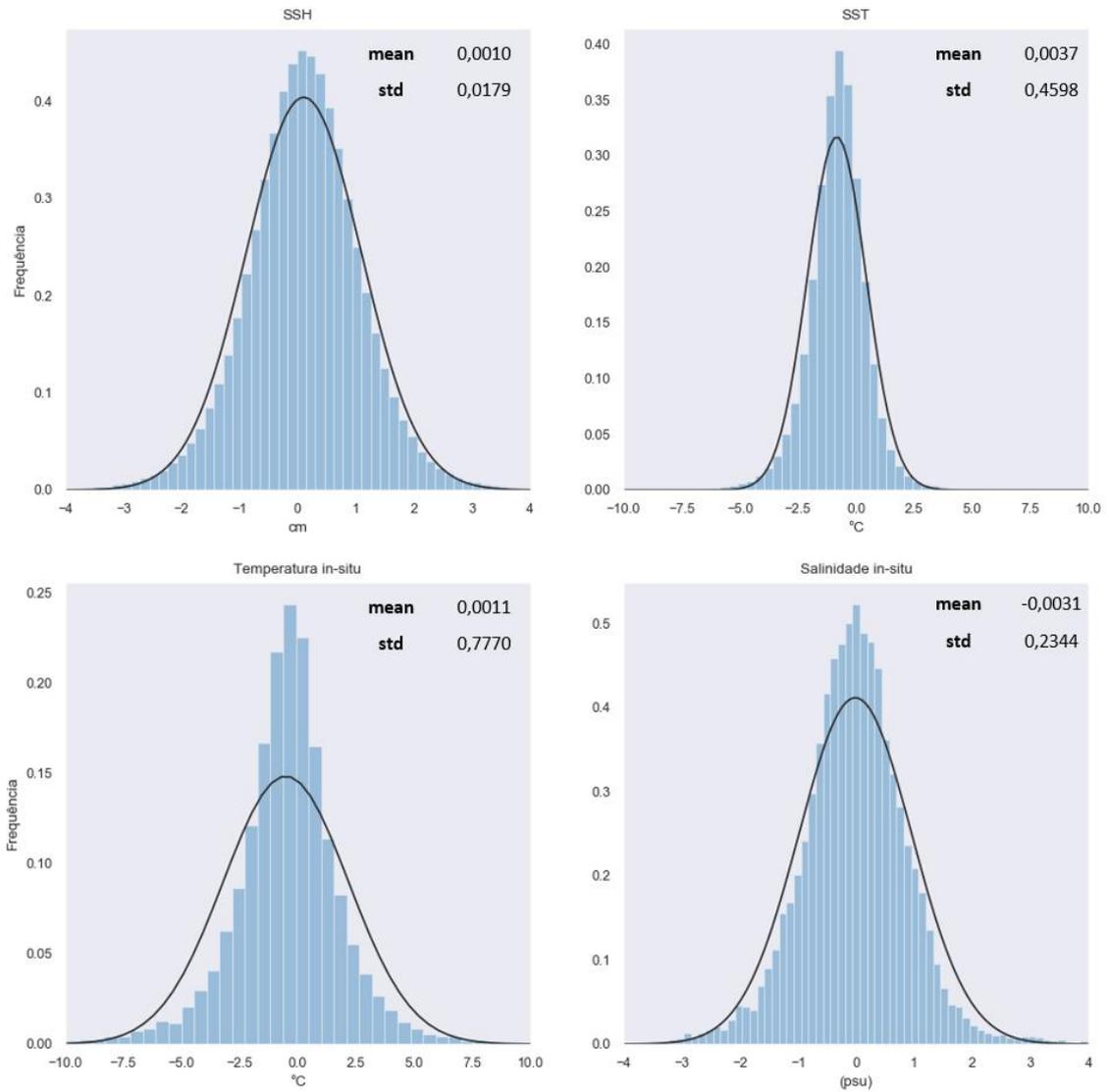


Figura 13: Distribuição de frequência do erro para as variáveis ASM, TSM, Temperatura e Salinidade in-situ do experimento 7DIAS. A linha preta representa a função de densidade de probabilidade Gaussiana estimada para cada conjunto de amostras.

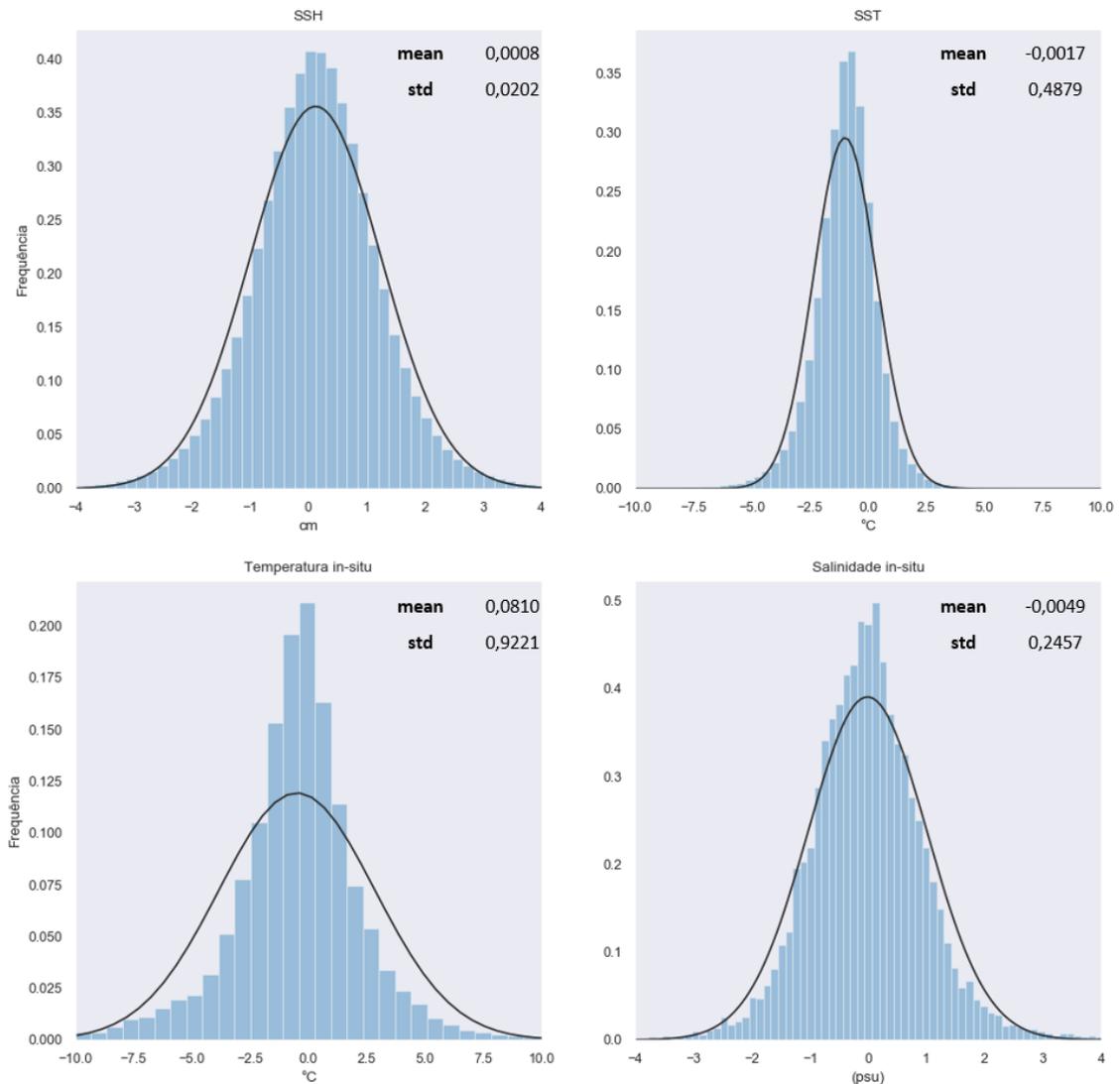


Figura 14: Distribuição de frequência do erro para as variáveis ASM, TSM, Temperatura e Salinidade in-situ do experimento 10DIAS. A linha preta representa a função de densidade de probabilidade Gaussiana estimada para cada conjunto de amostras.

Nota-se nas Figura 12, Figura 13 e Figura 14 que, apesar das distribuições de frequência apresentarem semelhança com a distribuição gaussiana, existem diferenças. Uma das premissas do método variacional é que o vetor inovação seja não tendenciosos (*unbiased*) e apresente distribuição normal (Neveu, 2016). A aplicação do teste de Kolmogorov-Smirnoff (teste KS) rejeitou a hipótese nula de que os valores de inovação são provenientes de uma distribuição normal para todas as variáveis e experimentos. No entanto, o experimento 4DIAS apresentou os menores valores de D (maior desvio da distribuição comparada) para todas as variáveis, exceto SSH, indicando que este experimento apresenta distribuição do vetor inovação mais próxima à distribuição normal (Tabela 3). Segundo Bonnavita *et al* (2018), desvios na pressuposta distribuição

gaussiana dos erros tanto do modelo quanto das observações podem ser potencialmente causadas pela não linearidade da dinâmica.

Tabela 3: Maior diferença (valor absoluto) entre as classes da distribuição de frequência acumulada do vetor inovação e da distribuição gaussiana (D).

Valor de D (Teste KS)				
	4DIAS	7DIAS	10DIAS	Dcrit 10%
<i>SST</i>	0,041	0,047	0,051	0,001
<i>SSH</i>	0,026	0,023	0,026	0,001
<i>in-situ T</i>	0,089	0,092	0,098	0,006
<i>in-situ S</i>	0,032	0,046	0,04	0,007

Ainda que esta análise indique que as premissas do método aplicado não estão 100% satisfeitas, os resultados demonstram que ainda assim o sistema tem sucesso em ajustar o modelo às observações utilizadas. Outros estudos encontram comportamento semelhante nesta mesma análise e relacionam com a presença de *spikes* nas observações (Moore, *et. al* 2013) e a má configuração das matrizes de covariância do erro (Neveu, *et al.* 2016). Segundo o primeiro, possíveis melhorias poderiam ser obtidas pela rejeição de observações distantes $\sim 4(\sigma_b^2 + \sigma_o^2)^{0,5}$ da média (onde σ_b é o desvio padrão do modelo e σ_o , das observações). Já Neveu (*et al* 2016) sugere a calibração dos valores prescritos para compor as matrizes de covariância do erro das observações e do modelo (**B** e **R**). Mattern (*et al.* 2017) sugerem um método iterativo para calibrar **B** e **R** através da metodologia de análise de consistência dos valores prescritos para ambas as matrizes de erro descrita em Desroziers (*et al* 2005) e encontram ajustes significativos nos valores previamente prescritos (de uma maneira geral, de redução dos erros do modelo e aumento dos erros das observações). Mariano (2019) testou uma alternativa para a prescrição de **B**, na qual os valores de σ_b são provenientes de uma rodada assimilativa (usualmente são obtidos de uma rodada livre) e os resultados reforçam a sensibilidade do sistema à matrizes de covariância de erro e indicam melhoria com a aplicação do método proposto. Apesar de não ser o objetivo do presente trabalho propor ajustes na definição de **B** e **R**, este diagnóstico indica um campo para futuras investigações.

4.1.2 Observações Satelitais

A seguir, apresenta-se a série temporal do *Jfit* calculado a partir do *misfit* final de cada experimento e para a rodada livre (Figura 15) e o diagrama de dispersão

comparando os experimentos para os *Jfits* calculados com o *misfit* inicial e final (Figura 16).

Todos os três experimentos reduziram significativamente o *misfit* para as variáveis TSM e ASM em relação à rodada livre (não mostrado na figura), como esperado e já descrito em Fragoso (*et al* 2016). Para ASM, o *Jfit* permaneceu bem abaixo de 1 ao longo de todo o período de análise enquanto que o *Jfit* para TSM se manteve próximo de 1 porém excedendo em alguns momentos. Nota-se na série temporal de *Jfits* dos dados satelitais um indício de sazonalidade, com maiores valores no início e final da série (final do outono e início da primavera). Esta sazonalidade é observada também em Fragoso *et al.* (2016).

O diagrama de dispersão comparativo dos experimentos (Figura 16) demonstra que para ambas as variáveis analisadas, o experimento 4DIAS obteve valores melhores para *Jfit* tanto nas soluções prévias quanto nas soluções finais (análises) e o experimento de 7DIAS foi melhor que o de 10DIAS. A diferença entre o experimento de 4DIAS para o de 7DIAS foi mais significativa na variável ASM, com valores muito próximos para TSM. Nas soluções prévias – que podem ser vistas como a previsão partindo de uma análise – os valores de *Jfit* apresentam maior dispersão, indicando maiores diferenças entre os experimentos quando comparado com as análises, que apresentam valores mais próximos da linha de correlação perfeita (linha pontilhada). O fato de que os experimentos apresentam menores valores de *Jfit* para estas variáveis quanto menor a extensão da janela de assimilação é um indício da influência da não-linearidade nos experimentos realizados.

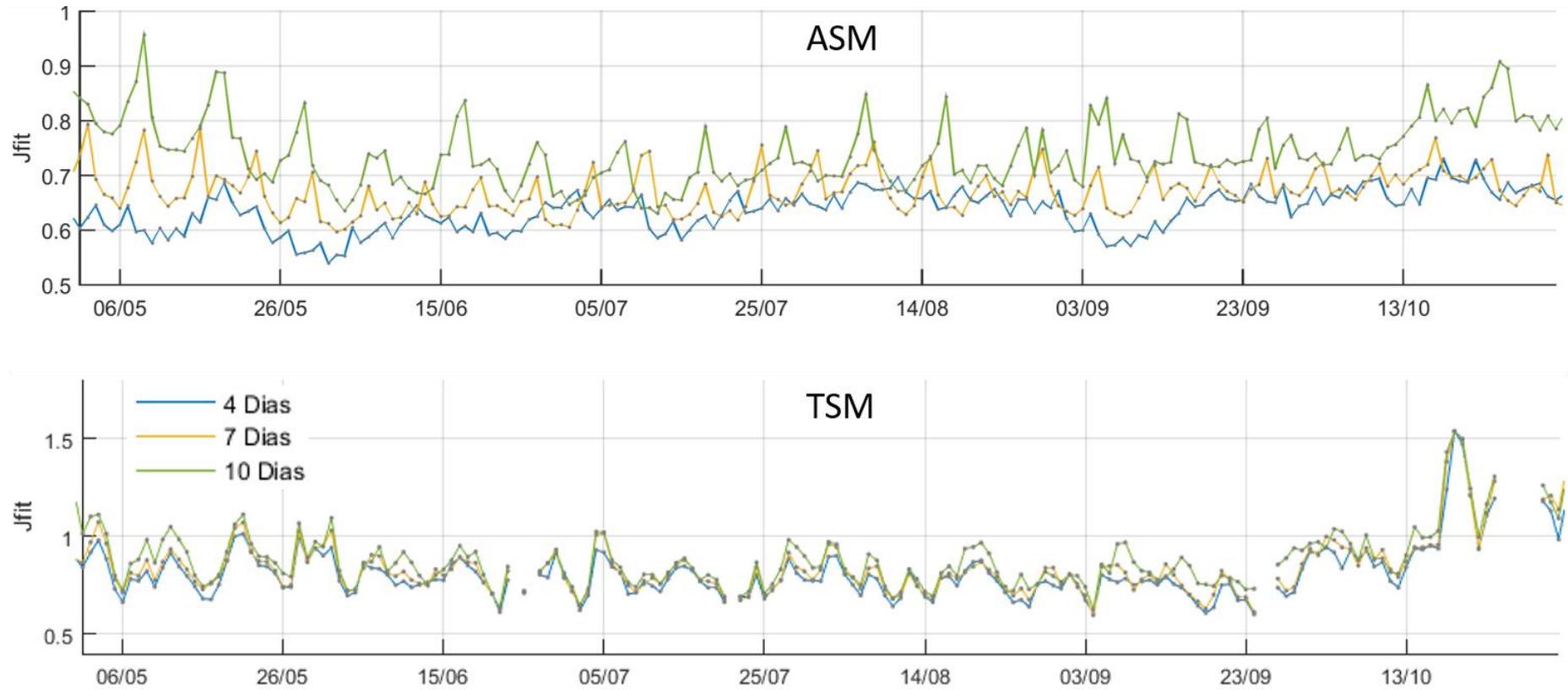
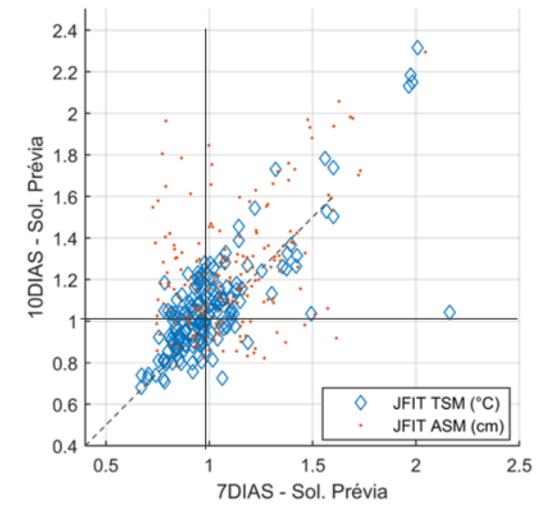
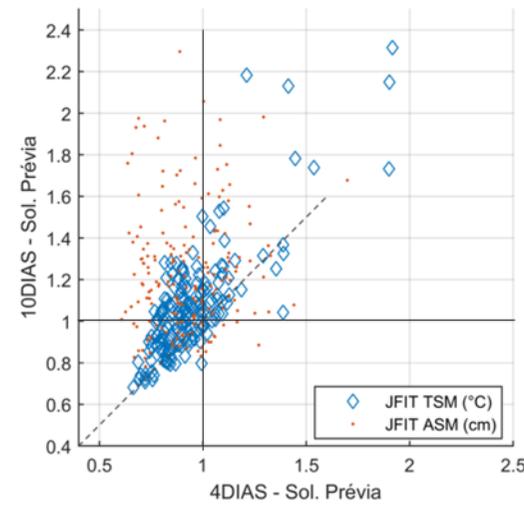
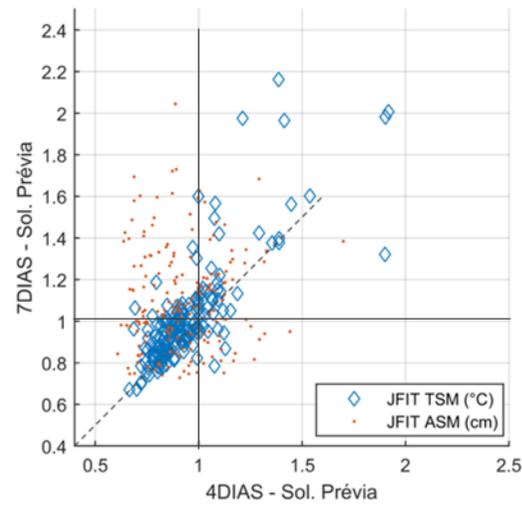


Figura 15: Série temporal do Jfit para ASM (acima) e TSM (abaixo) para os experimentos 4DIAS (azul), 7DIAS (laranja) e 10DIAS (verde).

Sol. Prévia



Análise

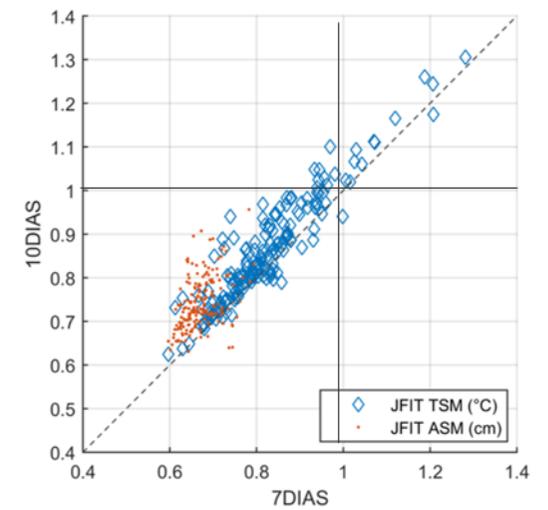
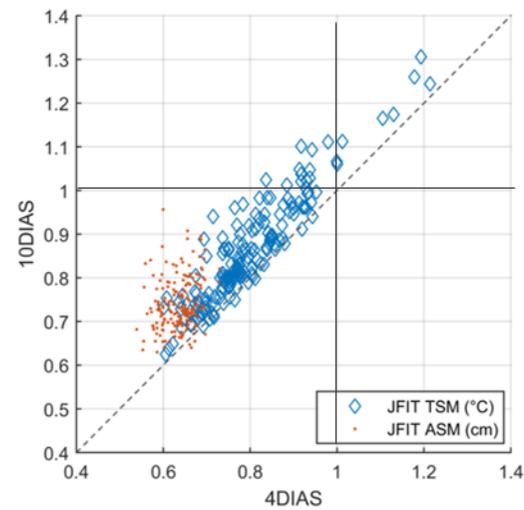
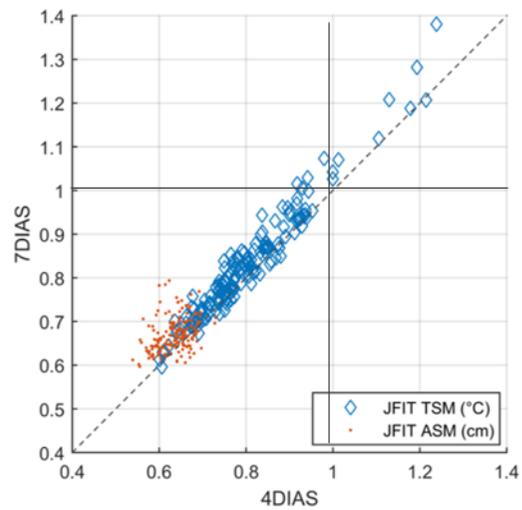


Figura 16: Gráfico de dispersão do Jfit para as observações de satélites (TSM e ASM) entre os experimentos 4DIAS, 7DIAS e 10DIAS. Acima, Jfit obtido das soluções prévias (ou previsões) e abaixo, Jfit das análises.

Para investigar o comportamento do sistema no ajuste aos dados satelitais de TSM e ASM ao longo dos dias dentro de cada ciclo de assimilação, assim como a deterioração da previsão (solução prévia) – que tem como condição inicial o último tempo da análise – são apresentadas médias do $Jfit$ de ASM e TSM do dia 1 ao dia n de cada experimento, sendo n a extensão da janela temporal de cada experimento.

Nota-se que tanto para ASM como para TSM, os três experimentos mantém o $Jfit$ abaixo de 1 ao longo de todos os dias das análises e que o experimento 4DIAS apresenta os menores $Jfit$ para ambas variáveis e que o experimento 10DIAS apresenta os maiores erros também ao longo de toda a janela (Figura 15, abaixo). É possível observar também uma forma côncava nas curvas, com o menor erro ocorrendo aproximadamente no meio da janela temporal dos ciclos de assimilação. Esse efeito é conhecido em sistemas 4DVAR (Pires, *et al.* 1996) sendo decorrente do fato de que os dias centrais na janela temporal de um ciclo de assimilação recebem informação tanto das observações futuras quanto às observações passadas, ao contrário dos dias iniciais e finais. Desta forma, alguns sistemas 4DVAR de análise e previsão oceânica utilizam a técnica de sobreposição entre ciclos de assimilação com o intuito de iniciar um novo ciclo a partir do dia que apresenta em média o menor erro dentro do período de assimilação (Wilkin, *et al.* 2018; Neveu, *et al.* 2016). Apesar desta técnica aumentar o custo computacional do sistema, indica-se como uma investigação futura importante para o avanço deste sistema para previsão operacional.

O aumento do erro ao longo dos dias na solução prévia ocorre em todos os experimentos e variáveis analisadas. Para TSM, os valores são bem próximos entre os experimentos, já para ASM ocorre a já observada melhoria quanto menor a janela de assimilação (Figura 17, acima). Como só foi utilizada a solução prévia nesta avaliação, o tempo de previsão é restrito à extensão da janela, logo só é possível a comparação direta entre os três experimentos nos 4 primeiros dias de previsão. Seria interessante realizar rodadas livres para períodos mais longos a partir de cada experimento apresentado, sendo este outra possível investigação a partir dos resultados aqui apresentados.

Interessante notar também que o $Jfit$ se mantém abaixo de 1 – significando que o erro do modelo está abaixo do erro prescrito para as observações – por um período maior no experimento 4DIAS do que nos demais (aproximadamente 3 dias para TSM e

ASM). Este período pode ser utilizado com um indicativo da extensão da acurácia da previsão do sistema.

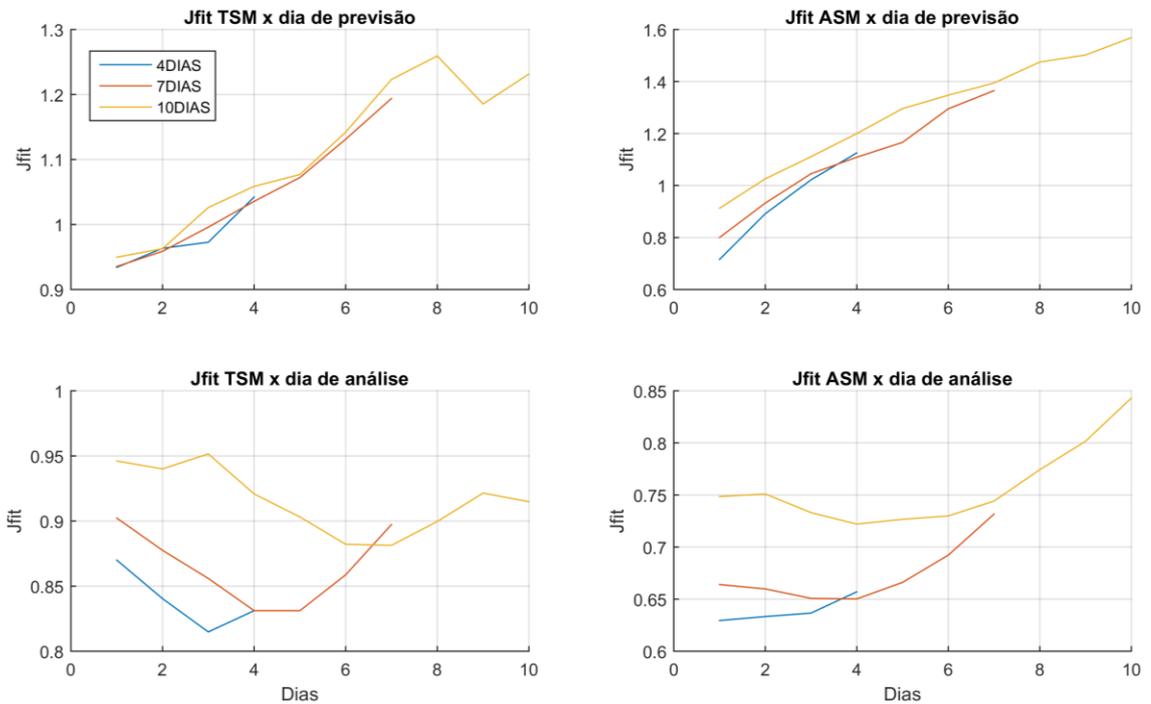


Figura 17: Jfit médio ao longo dos dias de previsão (acima) e análise (abaixo) para TSM (esquerda) e ASM (direita).

4.1.3 Perfis Hidrográficos

Neste item apresenta-se a análise dos perfis hidrográficos *in-situ* assimilados nos três experimentos. Nas Figura 18 e Figura 19, podem ser observados os valores médios e o desvio padrão de temperatura e salinidade na coluna d'água para os três experimentos e para o modelo livre, nas posições dos perfis hidrográficos assimilados. Nota-se que a média dos valores já é bem representada pelo modelo livre e os três experimentos ficam bastante próximos, além de não ser possível neste gráfico observar diferenças entre as soluções prévias e as análises. Porém, quando se observa a representação do desvio padrão das variáveis na coluna d'água, fica evidente a má representação do modelo livre e a melhoria significativa proporcionada pela assimilação dos dados. É possível também observar que as análises representam melhor do que as soluções prévias e que os três experimentos são bastante similares.

Nota-se que a região com maior desvio padrão é em torno de 200 metros, região que também apresenta a melhoria mais sensível proporcionada pela assimilação dos perfis. Isto indica que a variabilidade da posição da termoclina é ajustada nos experimentos.

Em aproximadamente 600 metros – região aproximada da interseção entre a ACAS e AIA - à uma subestimação do desvio padrão de temperatura nas soluções dos experimentos, ainda que apresentem melhoria em relação à solução livre. Já em torno de 900 e 1500 metros, nota-se uma superestimação e subestimação, respectivamente, no desvio padrão de salinidade. Em 900 metros ocorrem os valores mínimos de salinidade na coluna d'água, associadas à massa d'água AIA, onde também nota-se uma superestimação dos valores médios em relação às observações, não corrigido pela assimilação. Isso pode indicar um erro sistemático nos fluxos laterais do modelo livre e uma prescrição de erros da solução prévia para salinidade muito restritiva em camadas abaixo da termoclina, o que merece maiores investigações no sistema visto que existe a recomendação de alguns autores sobre a necessidade de remoção de erros sistemáticos dos modelos globais em sistemas de assimilação de dados (Wilkin, *et al.* 2018).

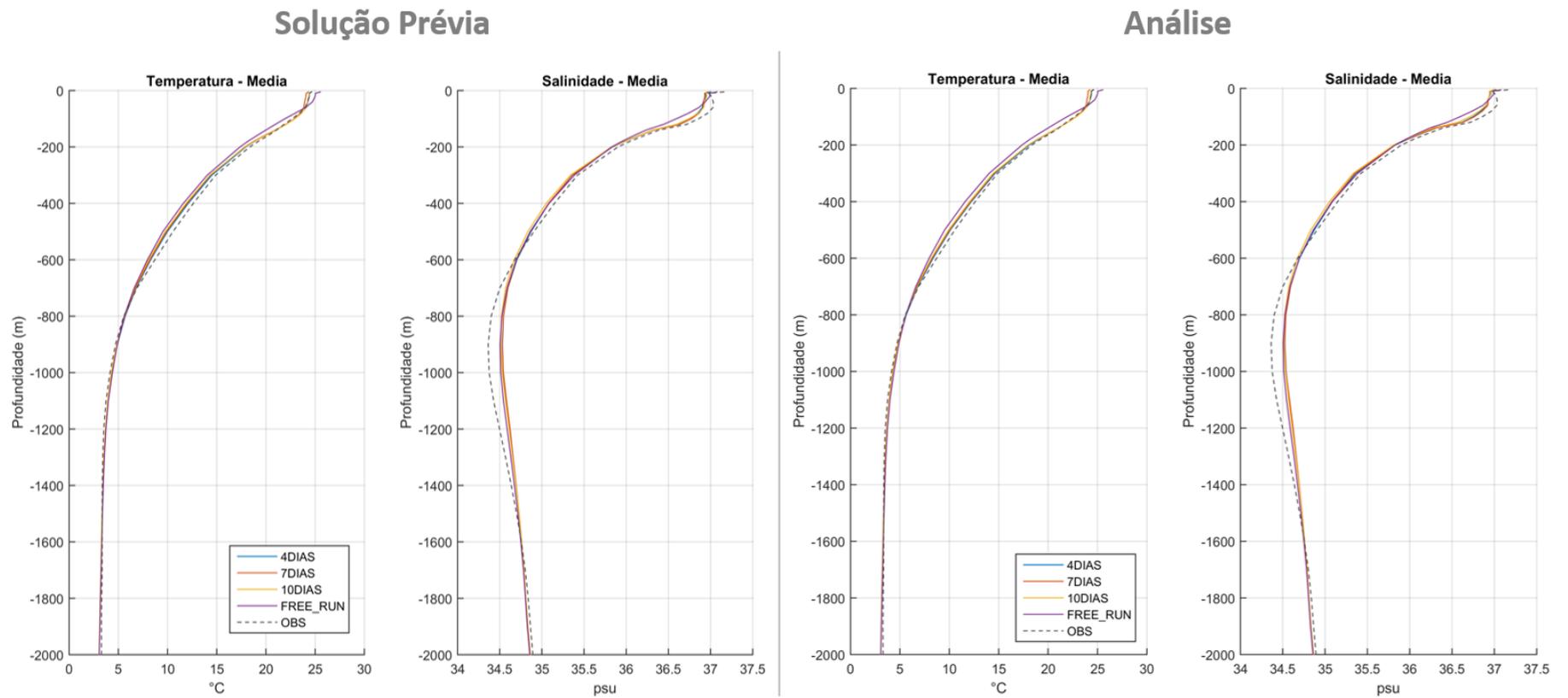


Figura 18: Média de temperatura e salinidade na coluna d'água dos três experimentos (4DIAS, 7DIAS e 10DIAS) e do modelo livre (FREE_RUN) nas posições e tempo dos perfis hidrográficos assimilados (OBS). A esquerda, resultados extraídos da solução prévia dos experimentos e a direita, das análises.

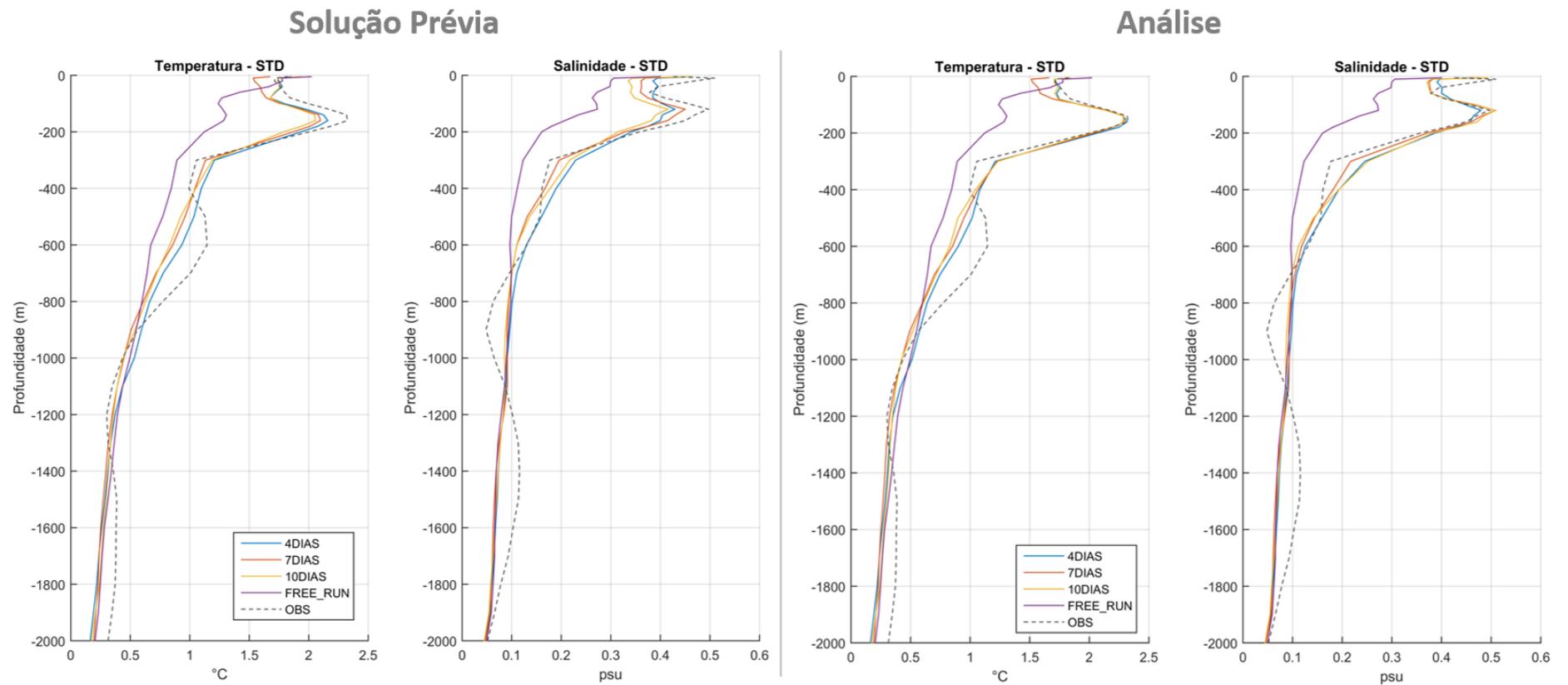


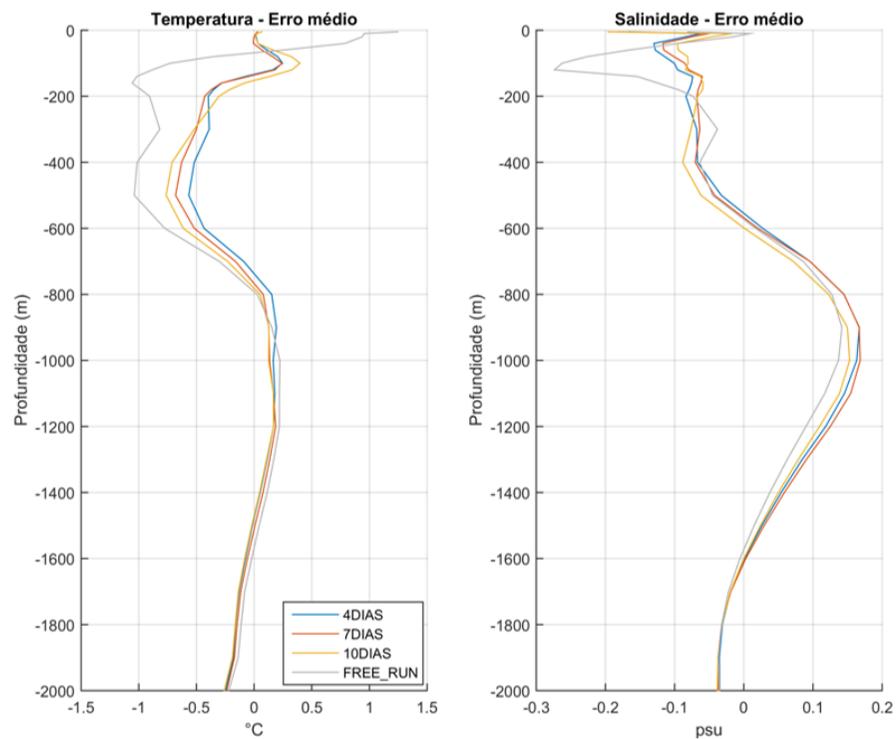
Figura 19: Desvio padrão (STD) de temperatura e salinidade na coluna d'água dos três experimentos (4DIAS, 7DIAS e 10DIAS) e do modelo livre (FREE_RUN) nas posições e tempo dos perfis hidrográficos assimilados (OBS). A esquerda, resultados extraídos da solução prévia dos experimentos e a direita, das análises.

A seguir, apresenta-se a estatística dos erros dos perfis hidrográficos, isto é, a diferença entre os valores de temperatura e salinidade na coluna d'água dos experimentos e do modelo livre nas posições das observações assimiladas. Serão apresentados os erros médios (Figura 20) e o RMSE (Figura 21) na coluna d'água e um diagrama de dispersão para comparação do *Jfit* médio entre os experimentos realizados (Figura 22), para a solução prévia e a análise.

Nestes gráficos pode-se observar mais claramente a redução do erro proporcionada pelos experimentos se dá principalmente na região entre a superfície e 600 metros para temperatura e até 200 metros para salinidade. Nota-se um erro médio negativo de aproximadamente 0,5°C e 0,1 psu nessa região, assim como um erro médio positivo de aproximadamente 0,15 psu em 1000 metros em todos os experimentos. O modelo livre apresenta um RMSE de até 2°C e 0,45 psu na região de 200 metros, reduzido para 1 (1,25) °C e 0,22 (0,3) psu nas análises (soluções prévias) dos experimentos. Os três experimentos apresentam grande similaridade na redução dos erros ao longo de toda a coluna d'água.

O diagrama de dispersão demonstram a redução dos *Jfit* médios da solução prévia para as análises e não evidenciam diferenças significativas entre os experimentos com todos os valores bem próximos da linha de correlação ideal. Nesta análise, o experimento 7DIAS aparece com vantagem entre os demais, seguido pelo 4DIAS e depois o 10DIAS.

Solução Prévia



Análise

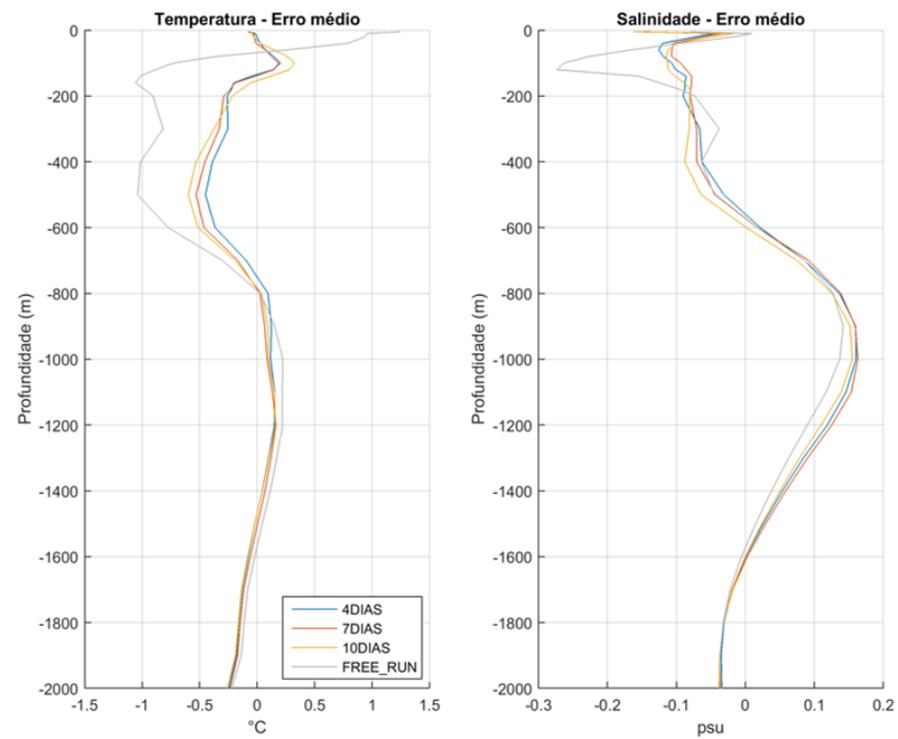
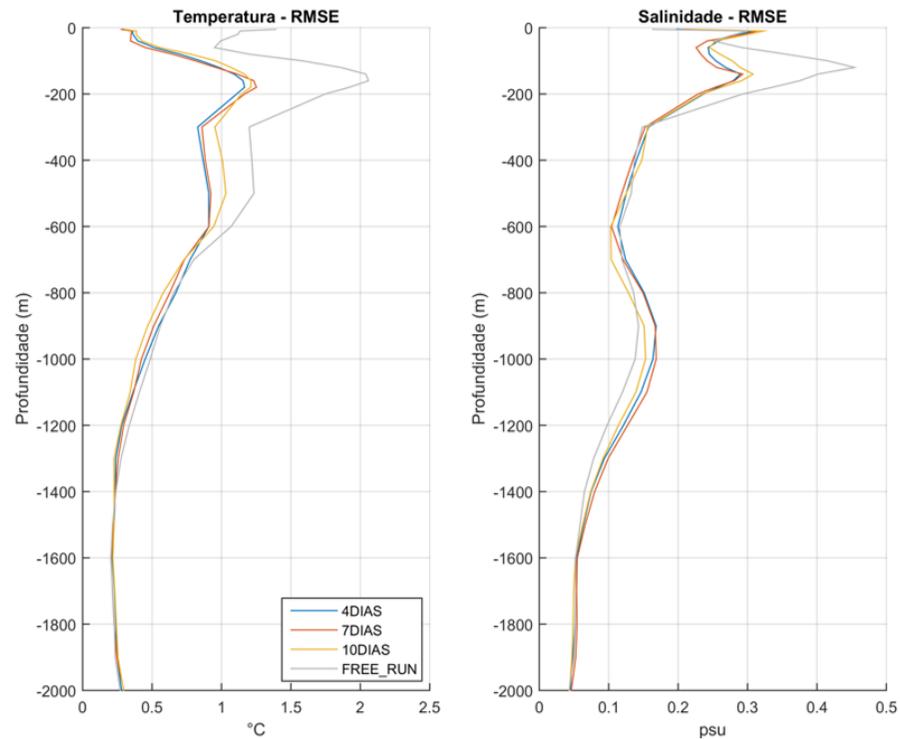


Figura 20: Erro (*misfit*) médio de temperatura e salinidade na coluna d'água dos três experimentos (4DIAS, 7DIAS e 10DIAS) e do modelo livre (FREE_RUN) nas posições e tempo dos perfis hidrográficos assimilados. A esquerda, resultados extraídos da solução prévia dos experimentos e a direita, das análises.

Solução Prévia



Análise

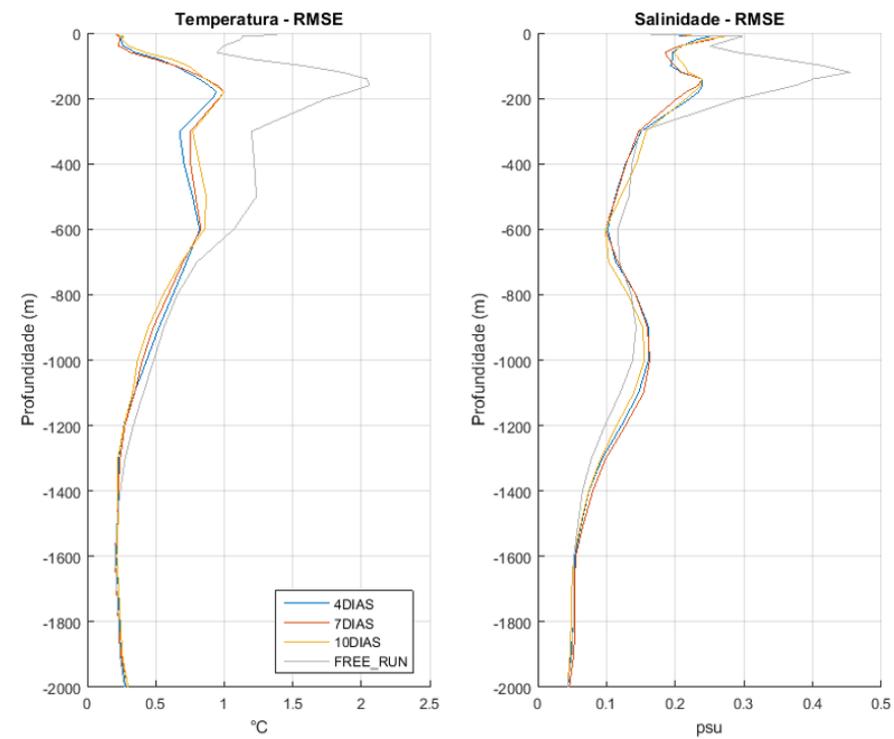


Figura 21: RMSE de temperatura e salinidade na coluna d'água dos três experimentos (4DIAS, 7DIAS e 10DIAS) e do modelo livre (FREE_RUN) nas posições e tempo dos perfis hidrográficos assimilados. A esquerda, resultados extraídos da solução prévia dos experimentos e a direita, das análises.

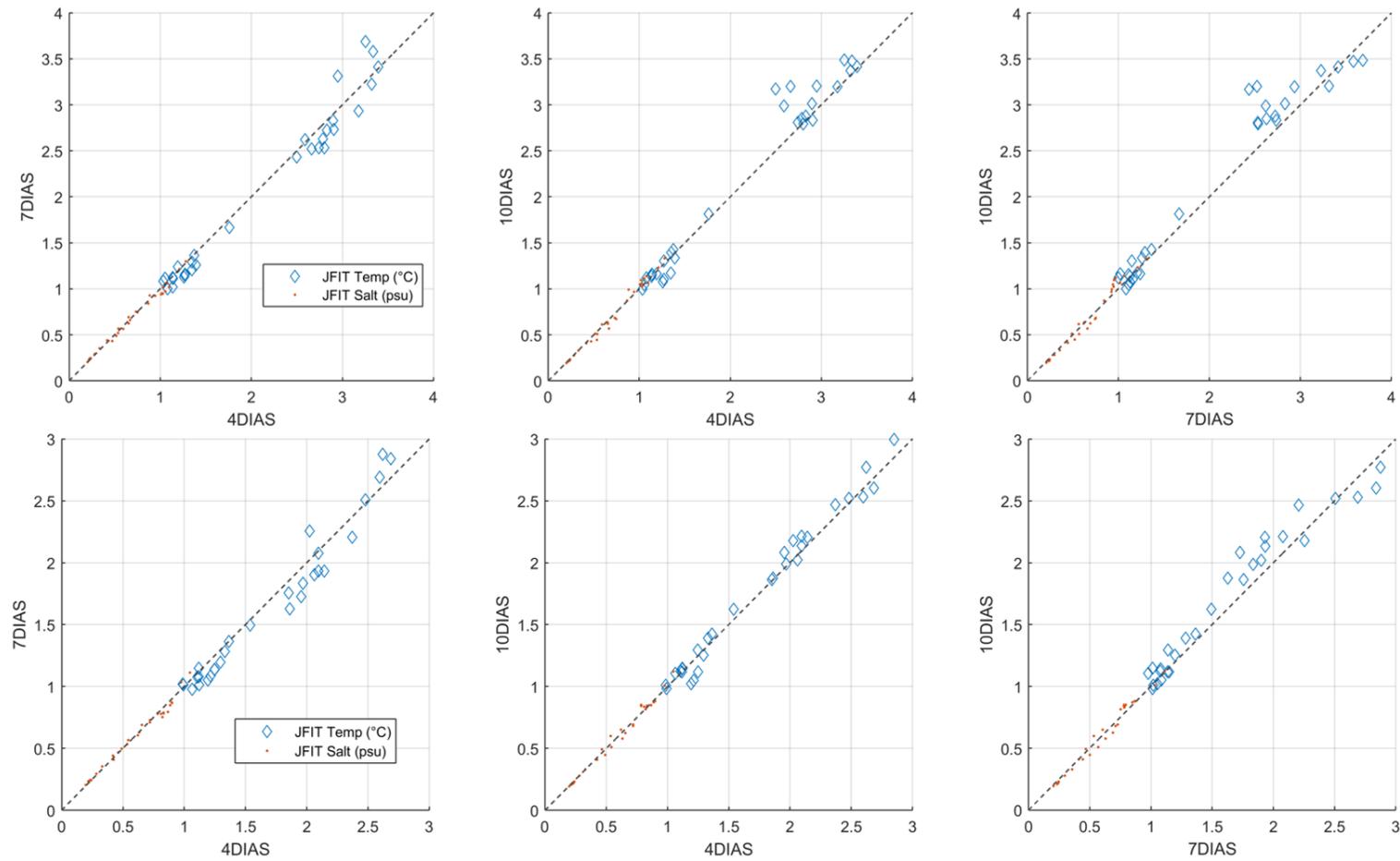


Figura 22: Diagrama de dispersão do Jfit médio para temperatura (azul) e salinidade (laranja) na coluna d'água entre os experimentos 4DIAS, 7DIAS e 10DIAS. Acima, Jfit obtido das soluções prévias (ou previsões) e abaixo, Jfit das análises.

4.2 Comparação com Dados Independentes

A comparação com fontes de dados independentes e de variáveis não assimiladas (velocidade da corrente superficial) é interessante pois reflete a influência da assimilação na dinâmica do modelo e ganhos nesta comparação conferem maior confiabilidade no sistema – para além de ajustar os resultados aos dados assimilados.

Neste capítulo são apresentadas análises com dados de derivadores superficiais oriundos do Projeto Azul lançados mensalmente em aproximadamente -24°S e -42°W a partir de março de 2013. Os derivadores são do tipo SVP e possuem uma vela centrada em ~ 15 metros de profundidade para diminuir a influência do vento e assim representar melhor a deriva forçada apenas pelas correntes superficiais. Cada equipamento provê informações de velocidade e direção da corrente (inferidas pelas posições consecutivas) e temperatura da água a cada 6 horas. Um total de 23 derivadores estão presentes no domínio do modelo no período dos experimentos (maio a novembro de 2013) totalizando 5014 informações de corrente e temperatura.

Para reduzir os erros de representatividade, isto é, provenientes de amostragem de feições em frequência espacial e temporal maior que a capacidade do modelo, e aumentar a confiabilidade estatística, os dados foram binados espacialmente em caixas de $0,5^{\circ}$ e temporalmente em médias diárias. Caixas com menos de 3 observações foram descartadas. A Figura 23 apresenta a quantidade de informação dos derivadores por caixa utilizada e as trajetórias dos 23 derivadores no período de análise (maio a novembro de 2013). Nota-se a grande quantidade de dados na região de lançamento, a influência da Corrente do Brasil na deriva preferencial para SW, recirculações reconduzindo os derivadores para norte a leste da CB e um derivador que entrou na plataforma continental e foi conduzido para NE pela Corrente Costeira do Brasil (este último, analisado em Carvalho, 2017).

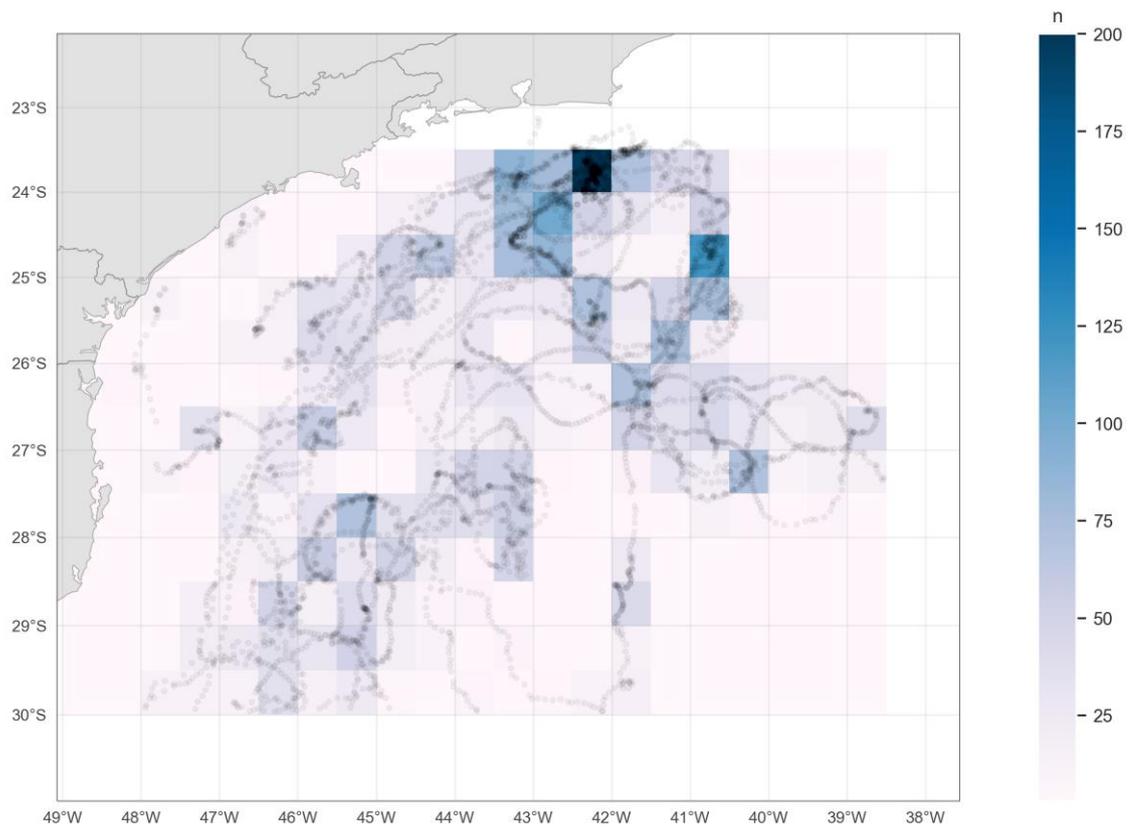


Figura 23: Número de derivadores por caixa de $0,5^\circ$ utilizados na comparação dos resultados dos experimentos. Os pontos no mapa representam as informações individuais dos derivadores na região de estudo.

Os resultados de velocidade e temperatura superficial dos experimentos foram interpolados para as posições espaciais e temporais de cada dado dos derivadores e as mesmas *binagens* em caixas foram aplicadas. Desta forma, a comparação é feita de forma direta e apenas nas posições espaço-temporais amostradas pelos derivadores. Primeiramente, apresenta-se os mapas médios de corrente (Figura 24) e temperatura superficial (Figura 25) elaborados com os dados dos equipamentos e dos resultados dos experimentos.

Nota-se que, ainda que não assimilados dados de corrente, os experimentos com assimilação apresentam grande melhoria na representação do campo de corrente superficial médio para o período analisado quando comparados com o experimento livre (FREE). A posição e intensidade da CB é corrigida (ainda que um pouco subestimada), assim como é possível notar a melhoria na representação de feições de mesoescala para leste da CB. É possível observar um dipolo em aproximadamente 29°S com 46°W e um vórtice anticiclônico em 26°S com 40°W no campo de corrente dos derivadores e a

correta representação pelos experimentos com assimilação de dados, dentre os quais o experimento 4DIAS se destaca. A representação da CCB fluindo para NE sobre a plataforma continental nos instantes de comparação não é representada por nenhum dos experimentos analisados. Carvalho *et al.* (2017) analisa a representação desta feição em sistemas com assimilação de dados e conclui que devido à frequência das forçantes desta feição – ondas de plataforma associadas à passagem de sistemas frontais – os dados de ASM satelitais não são capazes de amostrá-la e portanto, a sua assimilação prejudica sua correta representação. Ainda que os dados de ASM assimilados tenham sido cortados na região da plataforma nestes experimentos, o procedimento não foi suficiente para capturar a CCB no campo de correntes, pelo menos neste instante amostrado pelos derivadores do Projeto Azul.

A representação do campo de temperatura médio também apresenta grande melhoria nos experimentos assimilativos. Além da direta assimilação da TSM satelital, a melhoria na representação da CB pode ter influência na representação da temperatura superficial. Nota-se que, sobre a plataforma, os experimentos assimilativos apresentam temperatura de $\sim 18^{\circ}\text{C}$, consistente com a temperatura média dos derivadores. Apesar da não representação da corrente NE associada à CCB nos campos médios nos instantes de comparação, isto indica que possivelmente esta feição está presente ao longo do período dos experimentos.

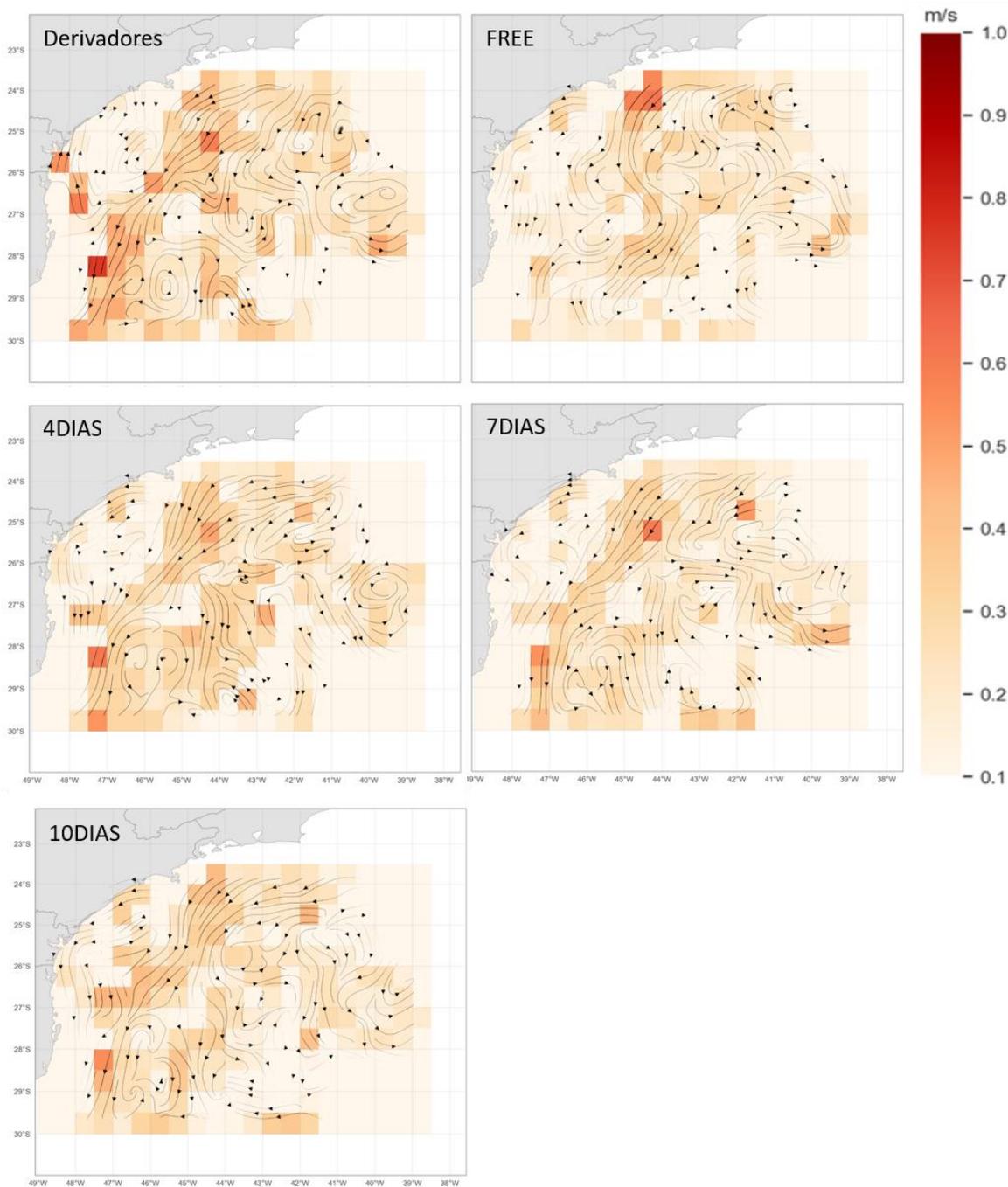


Figura 24: Campo de velocidade média em caixas de $0,5^\circ$ elaborado a partir dos dados dos derivadores do Projeto Azul para o período de maio a novembro de 2013 e dos resultados dos experimentos analisados (4DIAS, 7DIAS e 10DIAS) e da rodada livre (FREE) amostrados nas posições dos derivadores.

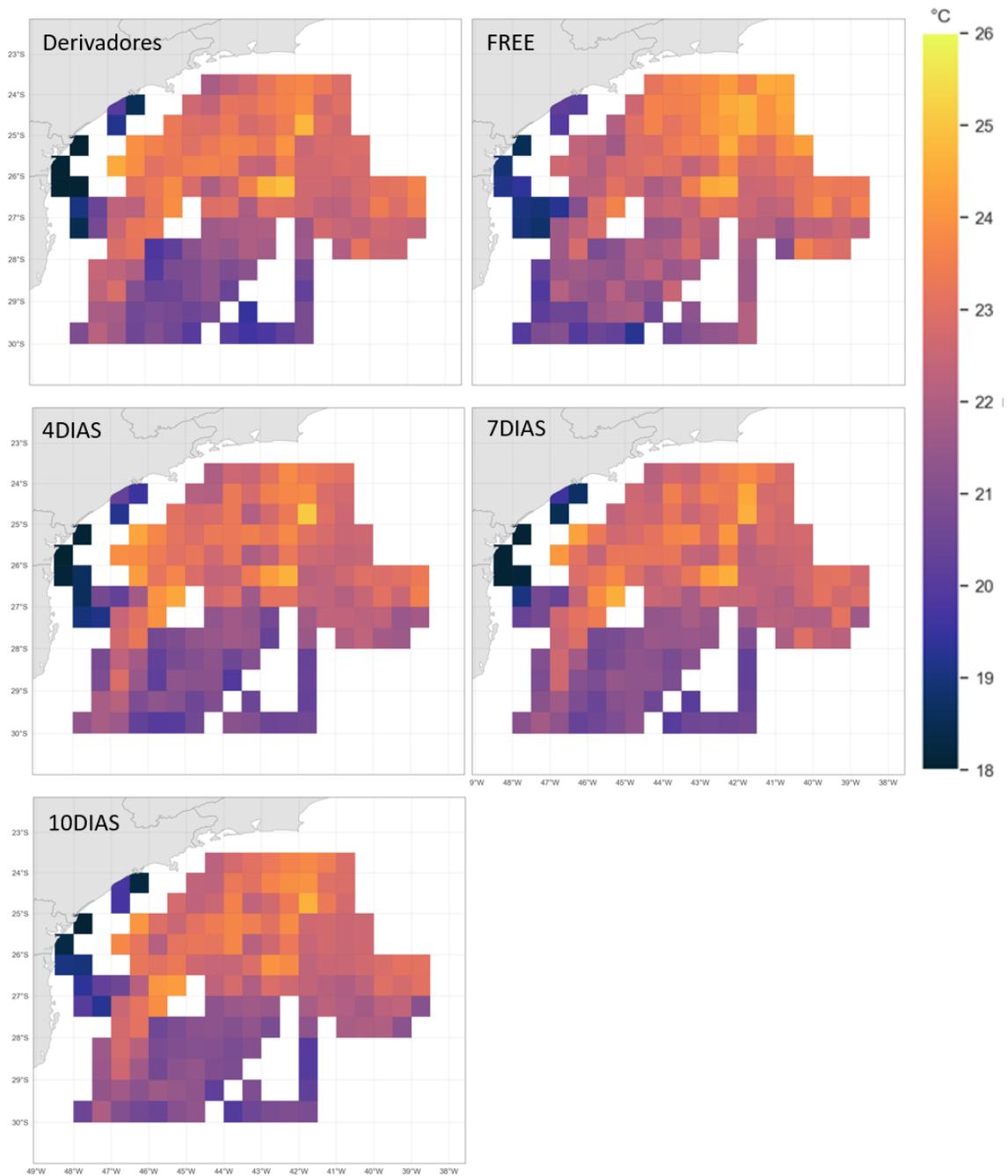


Figura 25: Campo de temperatura superficial médio em caixas de $0,5^\circ$ elaborado a partir dos dados dos derivadores do Projeto Azul para o período de maio a novembro de 2013 e dos resultados dos experimentos analisados (4DIAS, 7DIAS e 10DIAS) e da rodada livre (FREE) amostrados nas posições dos derivadores.

Na Figura 26 apresentam-se diagramas de dispersão dos erros dos experimentos para componente da corrente zonal, meridional e para temperatura. Nota-se que não é possível perceber um destaque dentre os experimentos nesta comparação.

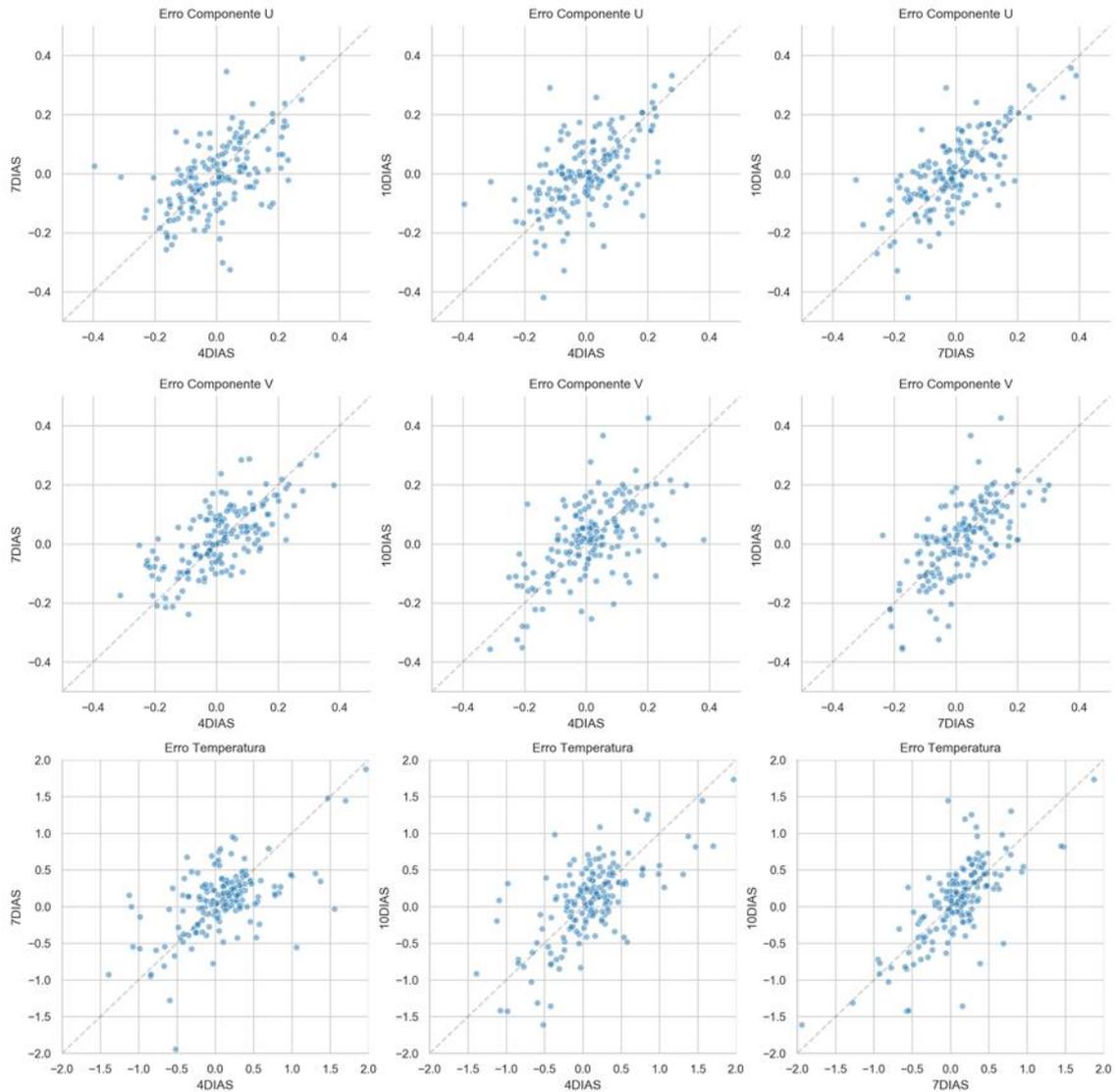


Figura 26: Diagrama de dispersão entre os erros (derivadores – experimentos) da componente de corrente meridional (v), zonal (u) dos experimentos 4DIAS, 7DIAS, 10DIAS.

Na Tabela 4 o erro médio quadrático (rmse) e o coeficiente de correlação linear de Pearson (P) são apresentados para o mesmo conjunto de dados além da rodada livre (FREE). Nota-se que o erro médio quadrático dos experimentos com assimilação é um pouco menor que o da rodada livre, porém a diferença mais marcante é no aumento da correlação linear média entre as séries de dados analisadas. Isso indica que a representação das feições melhorou significativamente – como também pôde ser observado nos mapas de corrente superficial – resultando numa correção de fase entre as séries analisadas. Dentre os experimentos com assimilação, tanto para $RMSE$ quanto para p , os índices nos experimentos 7DIAS são melhores que no de 4DIAS, que por sua vez, são melhores que no de 10DIAS.

Tabela 4: Energia Cinética Turbulenta (*EKE*), Energia Cinética Média (*MKE*), RMSE e coeficiente de correlação linear de Pearson (*P*) para os dados de velocidade (componentes zonal e meridional) e temperatura medidos pelos derivadores e pelos experimentos 4DIAS, 7DIAS, 10DIAS e rodada livre (FREE).

	<i>DRF</i>	<i>4DIAS</i>	<i>7DIAS</i>	<i>10DIAS</i>	<i>FREE</i>
<i>MKE (cm²/s²)</i>	257	235	204	210	183
<i>EKE (cm²/s²)</i>	237	185	137	165	131
<i>rmse U (m/s)</i>	-	0,20	0,19	0,20	0,25
<i>rmse V (m/s)</i>	-	0,19	0,18	0,21	0,23
<i>rmse Temp. (°C)</i>	-	0,60	0,57	0,63	1,01
<i>P (U)</i>	-	0,44	0,46	0,40	0,01
<i>P (V)</i>	-	0,47	0,49	0,35	0,16
<i>P (Temp.)</i>	-	0,90	0,91	0,89	0,75

A energia cinética média (*MKE*) e turbulenta (*EKE*) superficial foram calculadas a partir dos dados dos derivadores e comparadas com os experimentos segundo as equações 19, 20 e 21, onde:

$$(u, v) \rightarrow \text{médias em 60 horas}$$

$$(\bar{u}, \bar{v}) \rightarrow \text{média em todo o período}$$

Desta forma, a *MKE* mensura a parcela da energia contida no fluxo médio e a *EKE* mensura a parcela da energia contida em flutuações do fluxo médio. De forma a filtrar o sinal da maré e de oscilações inerciais, o fluxo momentâneo é representado por médias em caixas de 60 horas, seguindo Cerrone (2010) que apresenta uma extensa análise com derivadores na mesma região de estudo. Posteriormente, serão apresentadas análises na resolução do modelo ($1/12^\circ$) que serão mais adequadas para mensurar a *EKE* contida na totalidade dos vórtices de mesoescala, porém neste item a análise é realizada na resolução espacial adequada para comparação com os derivadores do Projeto Azul (promediados em caixas de $0,5^\circ$).

Apresentam-se mapas espaciais da *MKE* e *EKE* para o período de maio a novembro de 2013 (Figura 28 e Figura 38). Apesar da quantidade limitada de dados dos derivadores, é possível observar que a região com maior *MKE* coincide com a região da CB enquanto que a *EKE* é mais espalhada para leste da CB por provável influência de vórtices. Ao contrário da rodada livre (FREE), os experimentos com assimilação representam bem a variação espacial da *MKE* e *EKE* quando comparados com os

derivadores. A média espacial de ambas é apresentada na Tabela 4 onde nota-se que a assimilação de dados resulta em um aumento da energia cinética média e turbulenta de uma forma geral (~20%) – maior que o observado em Neveu *et al.* (2016) que relata um aumento de 10% da *EKE* em superfície – porém ainda com valores inferiores ao obtido a partir dos derivadores. Segundo Neveu *et al.* (2016) isso indica a capacidade do sistema de assimilação de dados de recuperar a assinatura da *EKE* a partir de dados de ASM.

Chama-se atenção para a região de formação do Vórtice de Cabo Frio (24°S com 42°W) que, apesar de se encontrar no fluxo da CB apresenta *EKE* maior de a *MKE*, o que também é ajustado nos experimentos assimilativos. De fato a *MKE* só começa a se adensar na região da CB para sul de 24°S. No comparativo das médias integradas espacialmente apresentadas na Tabela 4, o experimento 4DIAS apresenta valores mais próximos aos valores oriundos dos dados observados, seguido pelo experimento de 10DIAS e depois pelo de 7DIAS.

A comparação direta com dados dos derivadores representa um importante teste para o sistema, visto que são dados independentes e de variáveis não assimiladas (componentes zonal e meridional da corrente). A melhoria tanto nos campos espaciais de corrente e temperatura quanto no coeficiente de correlação das variáveis e na representação da energia cinética média e turbulenta demonstra a capacidade do sistema de ajustar a dinâmica a partir da assimilação de dados de ASM, TSM e perfis *in-situ* de temperatura e salinidade. Os três experimentos assimilativos analisados apresentaram desempenho bem aproximado para as métricas de erro e correlação utilizadas, com o 7DIAS um pouco superior, seguido pelo 4DIAS, e o experimento 4DIAS foi melhor na representação da *EKE* e *MKE*.

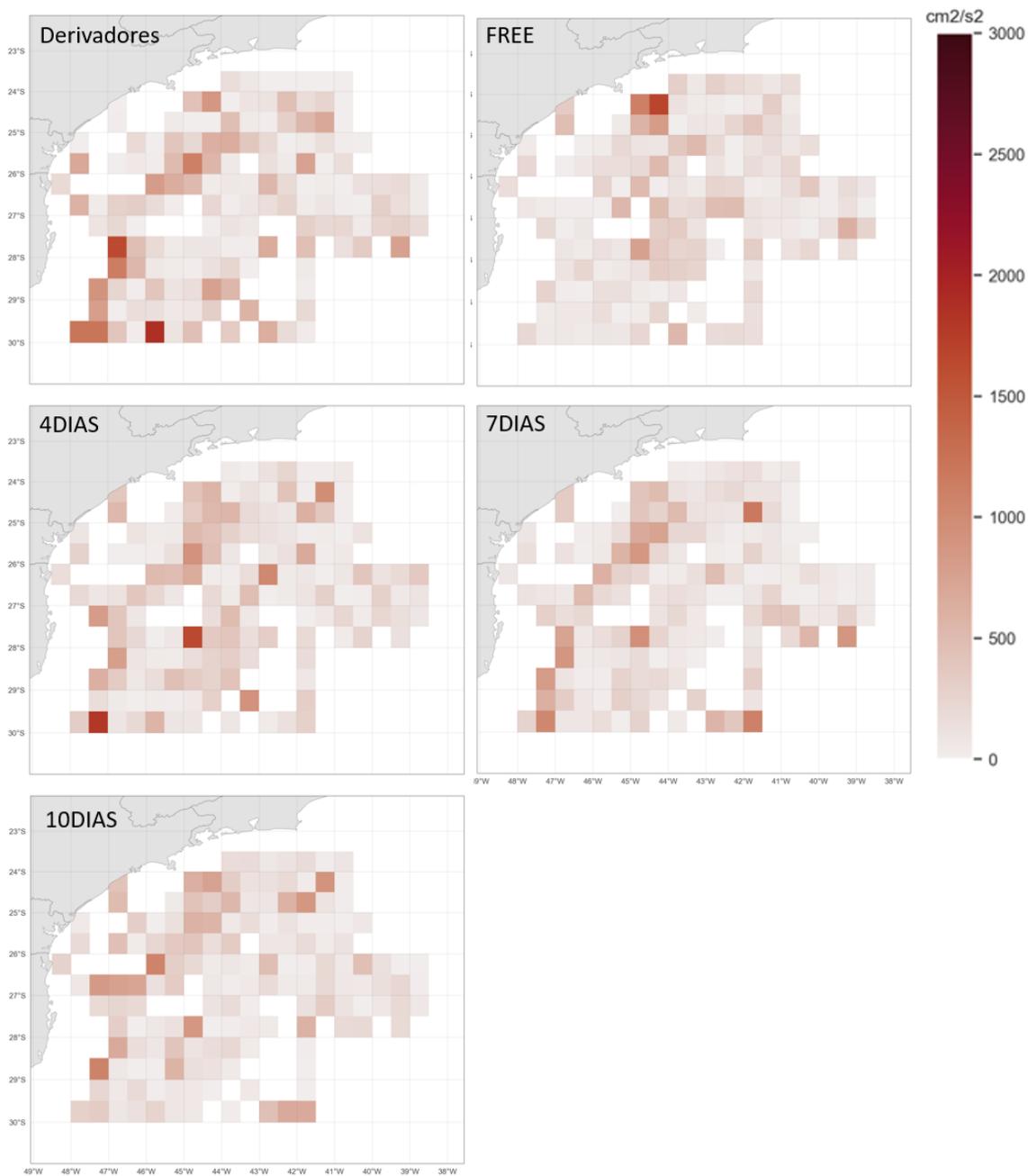


Figura 27: Campo de energia cinética média (MKE) em caixas de $0,5^\circ$ elaborado a partir dos dados dos derivadores do Projeto Azul para o período de maio a novembro de 2013 e dos resultados dos experimentos analisados (4DIAS, 7DIAS e 10DIAS) e da rodada livre (FREE) amostrados nas posições dos derivadores.

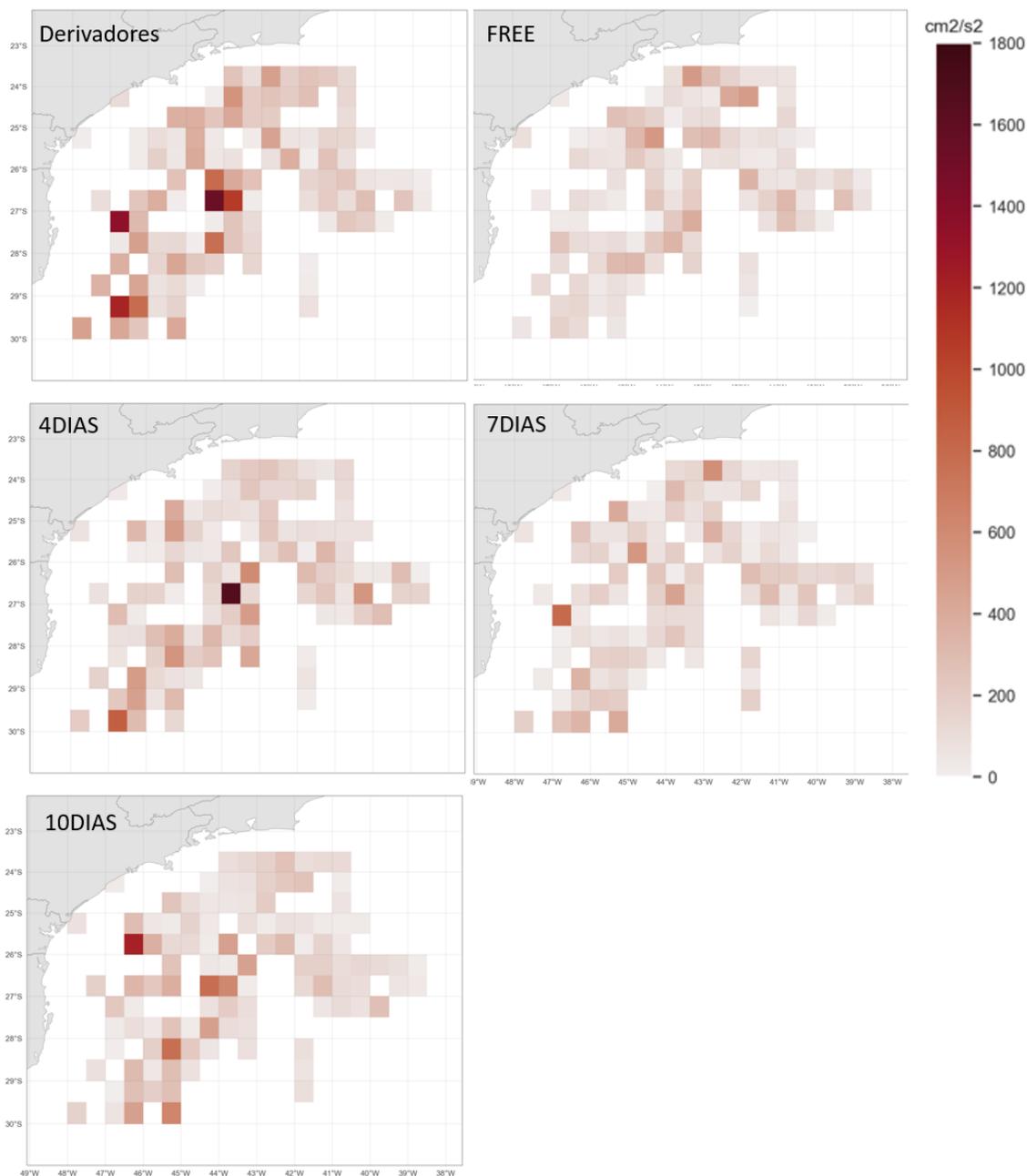


Figura 28: Campo de energia cinética turbulenta (EKE) em caixas de $0,5^\circ$ elaborado a partir dos dados dos derivadores do Projeto Azul para o período de maio a novembro de 2013 e dos resultados dos experimentos analisados (4DIAS, 7DIAS e 10DIAS) e da rodada livre (FREE) amostrados nas posições dos derivadores.

4.3 Análise da Função Custo

Uma vez que o sistema de assimilação variacional busca o vetor controle que minimiza a função custo e desta forma maximiza a probabilidade condicional da melhor

estimativa do vetor estado dadas as observações disponíveis, o monitoramento da função custo é usualmente utilizado para monitorar a performance do sistema. Em cada ciclo de assimilação, o sistema calcula a função custo não linear (JNL) da solução prévia e após cada *outer-loop*, e calcula a função custo linear (JTL) após cada *inner-loop*. Desta forma, para a configuração utilizada de 2 *outer* e 10 *inner loops* utilizada, temos a cada ciclo de assimilação o cálculo de três valores para JNL e 20 valores para JTL. A evolução da função custo (normalizada pelo JNL calculado com a solução prévia) ao longo de um ciclo de assimilação aleatório do experimento 4DIAS é apresentado na Figura 29 de forma a exemplificar este comportamento e auxiliar na explicação das análises que serão realizadas neste item. Pode-se observar os valores da função custo não linear (círculos laranjas) calculados a partir dos resultados da solução prévia, após a primeira atualização do vetor controle (1° *outer-loop*) e após o 2° *outer-loop* – ou seja, o JNL calculado para a solução final (análise). Entre os *outer-loops*, observam-se os valores da função custo calculadas após cada iteração dos *inner-loops*. O comportamento esperado é que o sistema seja capaz de reduzir o JTL a cada iteração e que, ao atualizar o vetor controle e realizar a integração do modelo não linear, o valor de JNL esteja próximo do valor final do JTL. Neste exemplo houve a redução de ~50% da JNL ao final do ciclo de assimilação e pode-se observar que a diferença entre o JNL e o JTL ao final dos *inner-loops* é menor no 2° *outer-loop* do que no 1°.

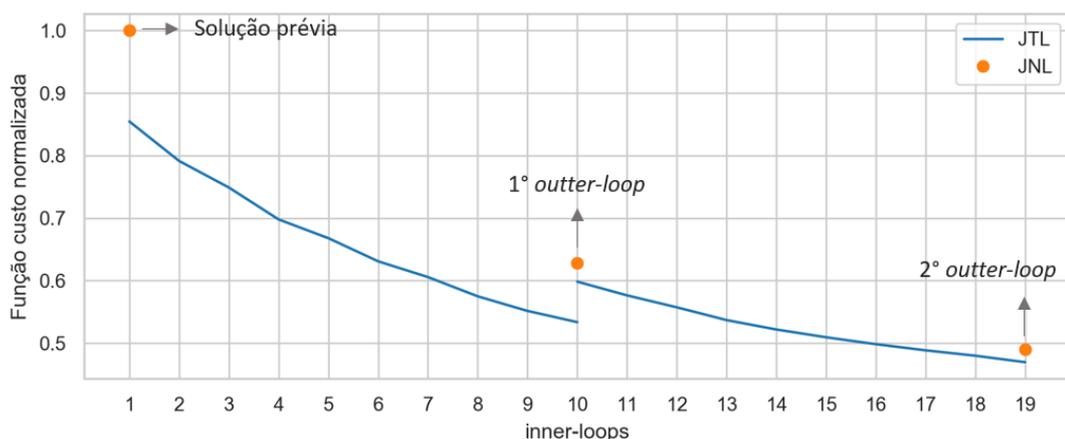


Figura 29: Evolução da função custo linear (linha azul) e função custo não linear (pontos laranjas) normalizadas, ao longo de um dos ciclos de assimilação do experimento 4DIAS.

Na Figura 30 apresenta-se o mesmo gráfico, porém com todos os ciclos de assimilação ao longo do período analisado e para os três experimentos. Nota-se a similaridade no comportamento entre os experimentos e que a redução da função custo

inicial (JNL da solução prévia) é de aproximadamente 30% no 1º *outer-loop*, chegando a ~50% na solução final. O experimento de 4DIAS apresenta redução ligeiramente maior (1%) até o 1º *outer loop* e ligeiramente menor (4%) no 2º *outer loop*.

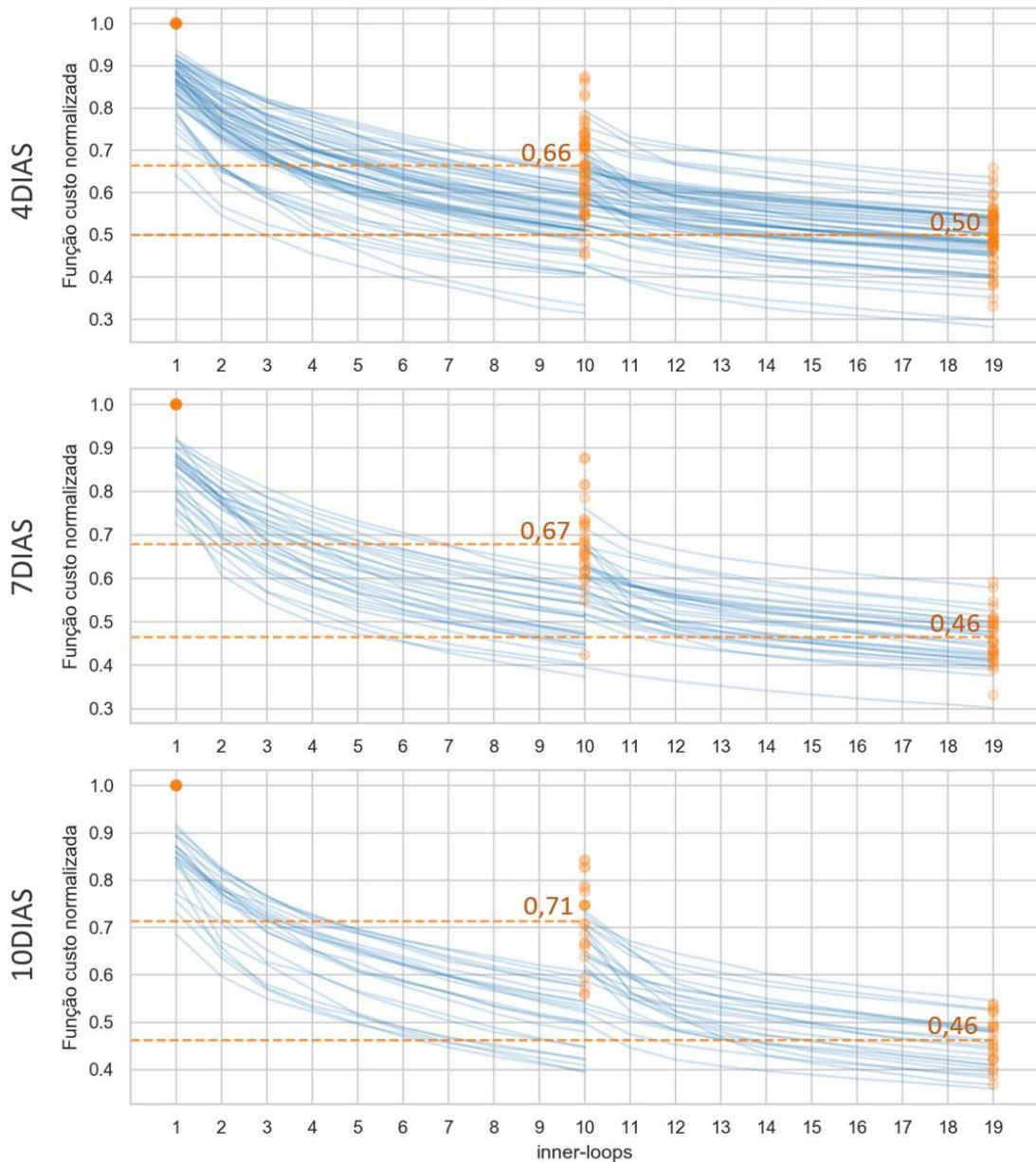


Figura 30: Evolução da função custo linear (linha azul) e função custo não linear (pontos laranjas) normalizadas, ao longo de todos os ciclos de assimilação dos experimentos 4DIAS, 7DIAS e 10DIAS. A linha tracejada indica a redução média da função custo não linear no primeiro e segundo *outer-loop*.

Nota-se que para os três experimentos, existe um “salto” maior entre o último JTL da primeira rodada de *inner loops* e o JNL subsequente quando comparados com o salto do JNL final. Apresenta-se um gráfico com a evolução ao longo dos ciclos destes

“saltos” entre os últimos valores de JTL e o JNL subsequente para cada *outer loop*, normalizados pelo primeiro valor de JNL de cada ciclo para permitir a comparação entre os experimentos (Figura 31). Nota-se que, apesar de bastante similares, o experimento de 4DIAS apresenta menor diferença no 1º *outer loop* e o experimento 7DIAS, no 2º *outer loop*. A maior redução do JNL e menor diferença entre o JNL e o JTL no primeiro *outer loop* quanto menor é a janela de assimilação é um indicativo da influência de reduzir a janela temporal para reduzir os efeitos da não linearidade.

As diferença entre JNL e JTL são causadas pela não-linearidade do sistema e a utilização de mais de 1 *outer-loop* é indicada para controlar seus efeitos em sistemas 4D-Var (Bonnavita, *et al.* 2018). Fica claro que para o número de *inner loops* escolhido, a utilização de 2 *outer loops* é importante para todos os três experimentos conduzidos e que ao final dos ciclos os três experimentos apresentam comportamento muito similar.

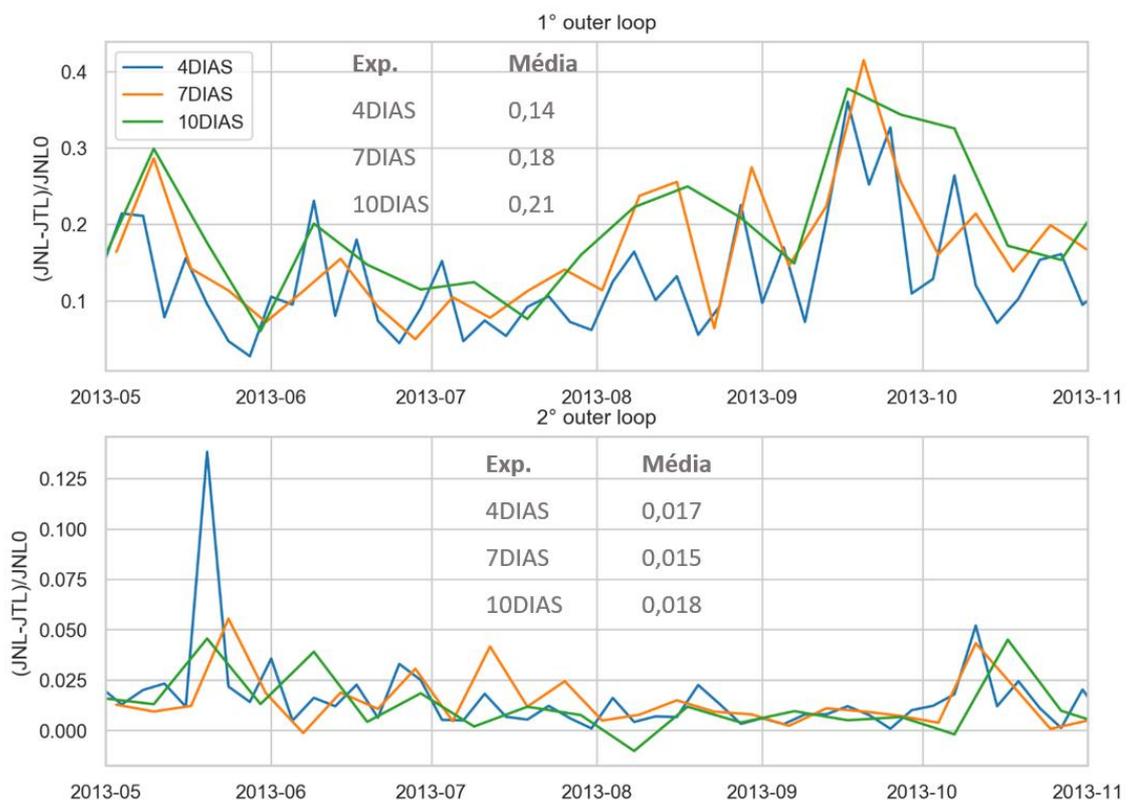


Figura 31: Diferença entre a função custo linear ao final dos *inner loops* e a função custo não linear subsequente para o 1º *outer loop* (acima) e o 2º *outer loop* (abaixo).

Por fim, uma análise usualmente utilizada para monitorar a minimização da função custo é a evolução ao longo dos ciclos do JNL da solução prévia, do último valor de JTL e do JNL da solução final (Mariano, 2018; Neveu, *et al.* 2016; Faller, 2016;

Fragoso, *et al.* 2016; Moore, *et al.* 2013). Conforme pode ser visto na Figura 32, os três experimentos apresentam o comportamento esperado ao longo de todo o período analisado com a redução da função custo em todos os ciclos e com os valores de JNL da solução final bem próximo do JTL ao final dos *inner-loops*. Essa análise indica que a premissa de que é possível obter o vetor controle que minimize a função custo não linear através de sucessivos pequenos incrementos utilizando operadores lineares foi satisfeita para as três extensões da janela de assimilação.

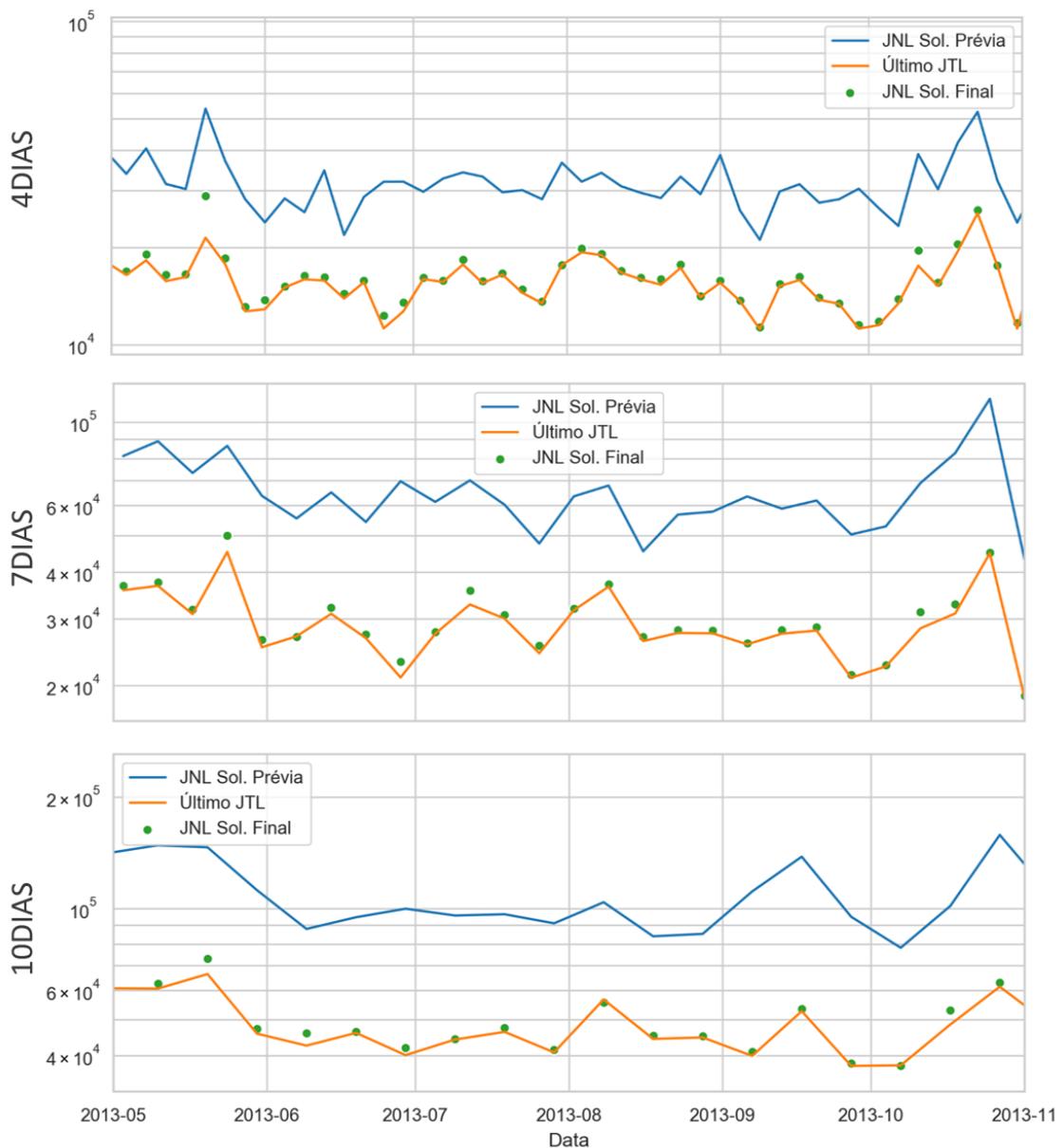


Figura 32: Evolução da função custo não linear da solução prévia e da solução final e da função custo linear da última iteração dos *inner-loops* ao longo dos ciclos de assimilação dos experimentos 4DIAS (topo), 7DIAS (centro) e 10DIAS (a baixo).

Um ponto interessante é observar o comportamento das séries apresentadas acima e na diferença entre JNL e JTL ao longo do período dos experimentos. Nota-se claramente valores mais altos em todas as séries de JNL e JTL nos meses de maio e outubro, assim como maior diferença entre JNL e JTL nos mesmos. Isso é um indicativo de que nestes períodos os processos não lineares foram mais significativos na dinâmica oceânica, o que leva a uma divergência maior da premissa linear e, conseqüentemente, ao ajuste do estado do mar às observações. Como pode ser visto na evolução do ajuste aos dados satelitais (Figura 15) esses dois meses são justamente os que apresentam os piores índices e que, principalmente para o ajuste aos dados de ASM, o experimento 10DIAS apresenta o *Jfit* mais elevado, seguido do 7DIAS e por último, 4DIAS. Esta análise corrobora a hipótese de que a janela de assimilação reduzida atende melhor à premissa linear. No item 4.4 é apresentada uma análise da energia cinética turbulenta (EKE) contida nos resultados dos experimentos com o objetivo de testar a hipótese de que há relação entre o comportamento exposto acima e a não linearidade do sistema.

4.4 Energia Cinética

Neste item é realizada uma análise da evolução da energia cinética turbulenta ao longo do período dos experimentos. O método para o cálculo da energia cinética turbulenta (EKE) e média (MKE) é o mesmo utilizado no item 4.2 porém utilizou-se médias em caixas de 5 dias para a obtenção das flutuações de velocidade, seguindo metodologia descrita em Rieck *et al.* (2018). O cálculo foi realizado apenas para a camada superficial dos modelos.

De início, apresentam-se mapas da MKE e EKE média para o período de análise para os três experimentos assimilativos e para a rodada livre (Figura 33 e Figura 34). Como esperado, a MKE é concentrada na Corrente do Brasil, se intensificando a partir de 19°S. Os experimentos assimilativos apresentam diferença em relação à rodada livre a partir de 26°S, o que já se mostrou um avanço na estimativa da corrente superficial na comparação com dados dos derivadores (item 4.2). A região no entorno de 24°S com

42°W se mostra uma exceção, com valores baixos de MKE e altos de EKE, pois é uma região dominada por atividade de mesoescala com meandros e vórtices, como é amplamente conhecido (Mano, 2009). Os mapas da Figura 34 demonstram que a EKE é mais espalhada por toda a bacia oceânica, pois como é também conhecido, fora das correntes de contorno oeste a energia é dominada por vórtices de mesoescala (Ferrari&Wunsch, 2008).

Devido à assimilação de dados de ASM, pode-se notar que os resultados dos experimentos ajustam a escala das feições de EKE e também alteram significativamente as regiões de maior concentração de energia (indicada pela área retangular na Figura 34). Esta distribuição espacial está de acordo com o trabalho de Rieck *et al.* (2015) sobre a EKE nas bacias oceânicas e sua variabilidade interanual realizadas com dados de ASM satelital.

Na Figura 35 apresenta-se a série temporal da EKE média integrada espacialmente em todo o domínio (acima) e na região de maior concentração definida entre 29°S-19°S e 49°W-39°W (abaixo). Nota-se claramente uma oscilação na EKE diminuindo de maio até julho onde encontra seu valor mínimo e voltando a subir até outubro onde apresenta os valores máximos. Esse comportamento reflete a sazonalidade climatológica da EKE na região como reportado por Rieck *et al.* (2015). Note na Figura 36 que a EKE calculada pelos dados da AVISO para a região 25°S-35°S com 20°W-40°W apresenta o mesmo padrão para o período do ano entre maio e outubro. A comparação direta com os resultados de Rieck *et al.* (2015) não foi possível pois a área analisada está fora do domínio de modelagem. A EKE média na região entre 29°S-19°S e 49°W-39°W apresenta menor sinal desta sazonalidade e, apesar de apresentar máximos em outubro e valores altos em maio, a energia não diminui tanto nos demais meses do ano.

É notório que o experimento 4DIAS apresenta maior EKE que os demais experimentos em ambos os casos (todo o domínio e na área de adensamento de EKE). O experimento de 7DIAS apresenta valores maiores que o 10DIAS nos meses de junho a agosto na média para todo o domínio e valores similares no restante do período e em todo o período na região de adensamento. Pela comprovação da maior capacidade de ajuste aos dados de ASM assimilados durante todo o período do ano e especialmente nos meses de maio e outubro (ver Figura 15), indica-se que este experimento é mais

capaz de manter as feições turbulentas de mesoescala conforme são observadas nos dados.

Porém qual é a relação entre os padrões observados na EKE, a não linearidade na dinâmica e a performance do 4DVAR como observado para cada extensão de janela de assimilação?

No estudo de Chelton *et al.* (2011) o autor realiza uma investigação baseada em dados satelitais de ASM sobre a distribuição da ocorrência de vórtices não lineares. A não linearidade dos vórtices é determinada pela relação entre a velocidade de rotação (U) e a velocidade de translação (c), e caso $U/c > 1$, o vórtice é não linear o que o permite manter uma estrutura coerente ao se propagar. Com isso, os vórtices não lineares são capazes de aprisionar e transportar a parcela d'água em seu interior para outra região. Segundo os autores, cerca de 50-60% da EKE na região de estudo provém de vórtices de mesoescala (Chelton *et al.* 2011, *Figura 7*) e cerca de 95% de todos os vórtices detectados em seu estudo na faixa entre 15° e 25° são não lineares e pode-se observar em Chelton *et al.* (2011 *Figura 17*) que na região da bacia de Santos os valores de U/c variam de 4 a 8 – ou seja, os vórtices na região de análise são em sua maioria não lineares.

Ainda que nem toda a EKE mensurada nesta análise provenha de vórtices, a grande correlação do aumento de EKE com o aumento da distância entre JNL e JTL é um forte indício de que a fonte da não linearidade na dinâmica no período analisado provém de feições turbulentas de mesoescala. Pelos motivos apresentados, conclui-se que a utilização da janela de assimilação de 4 dias possibilita satisfazer a premissa linear em momentos de maior influência de processos não lineares quando comparado com janelas mais extensas.

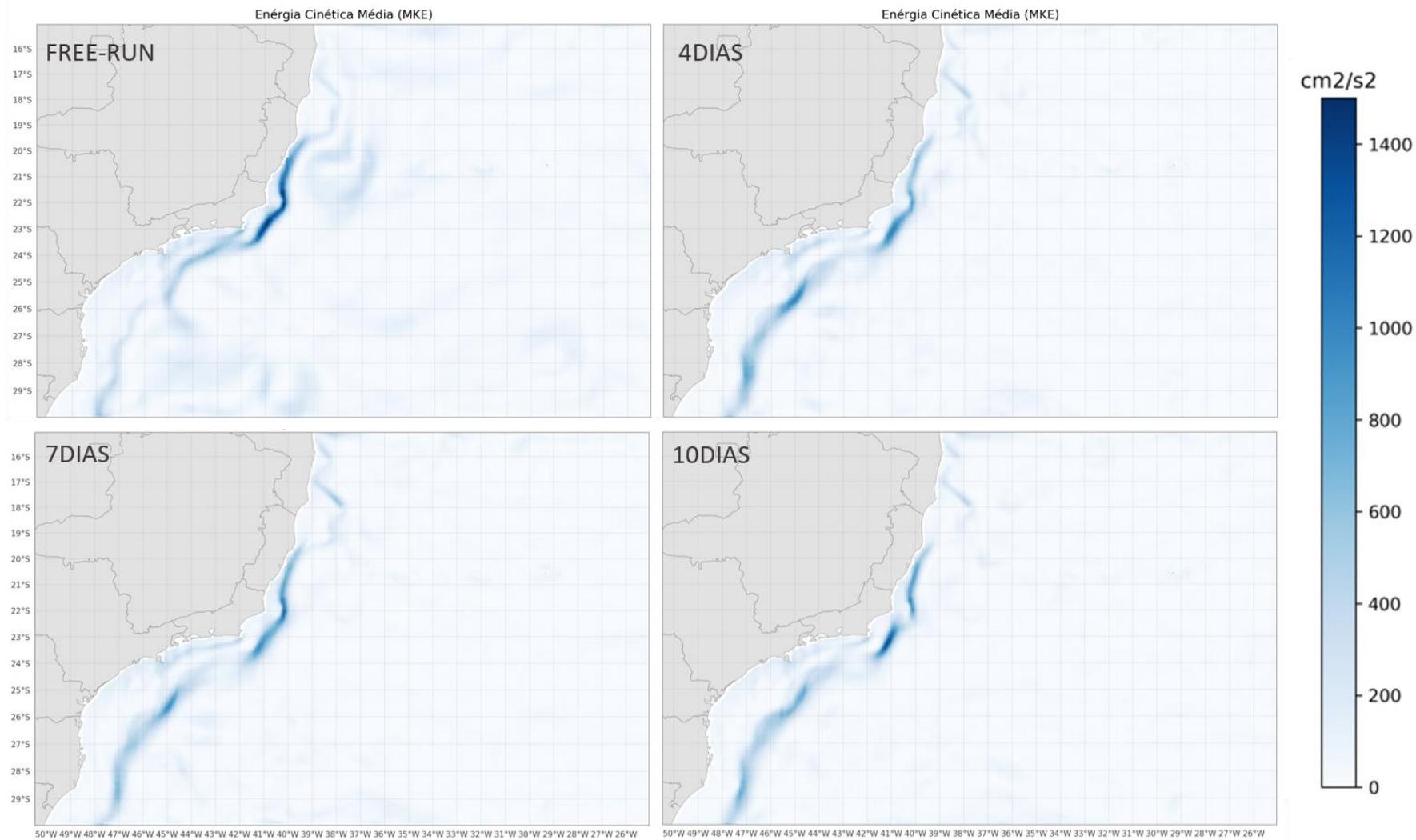


Figura 33: Mapa da energia cinética média (cm^2/s^2) para os três experimentos e para a rodada livre (FREE-RUN).

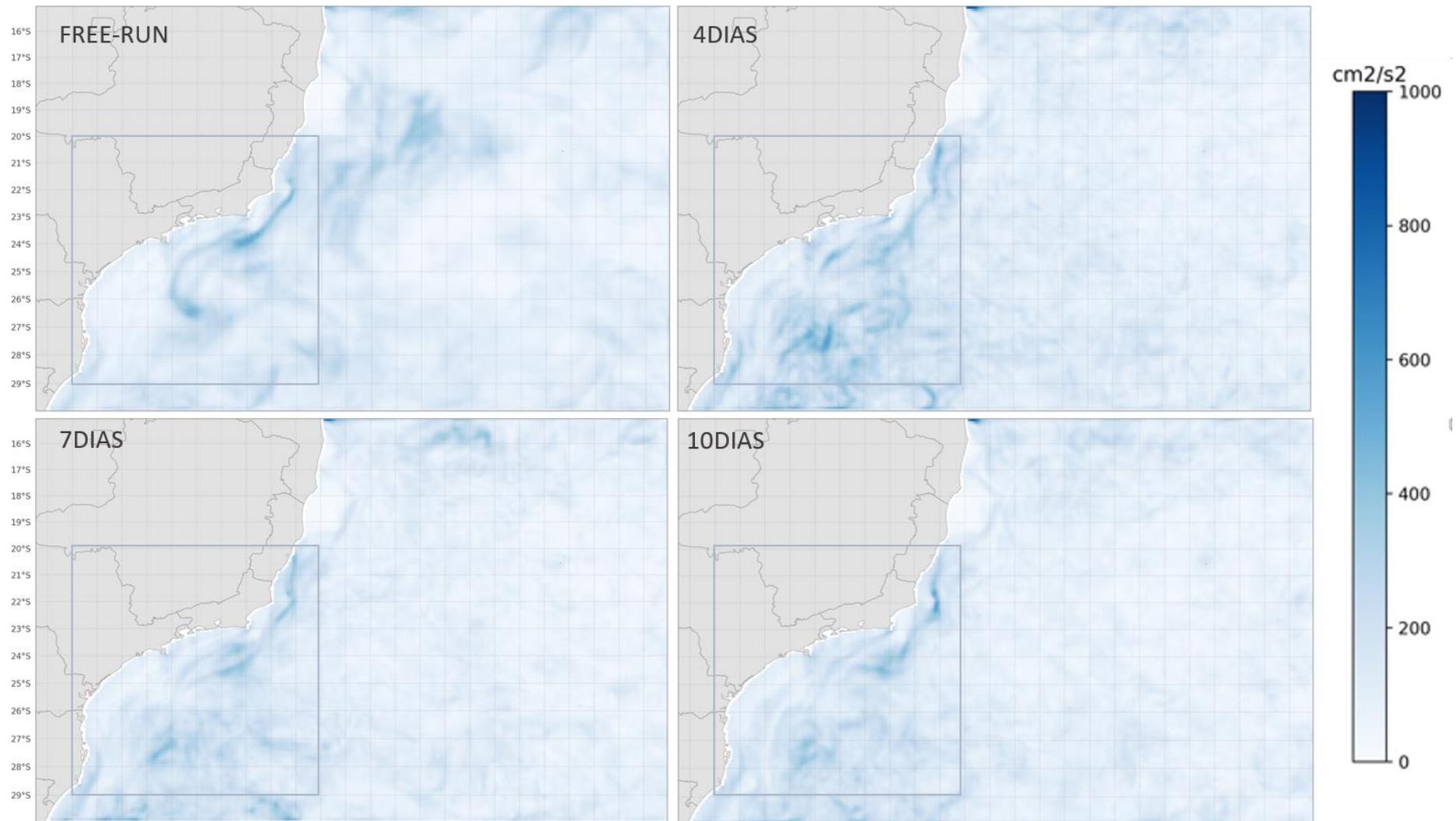


Figura 34: Mapa da energia cinética turbulenta média (cm^2/s^2) para os três experimentos e para a rodada livre (FREE-RUN).

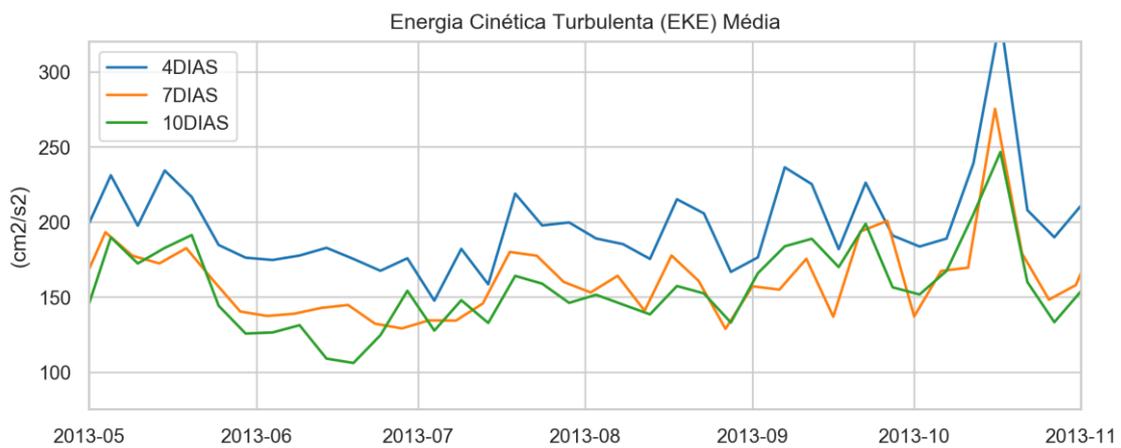
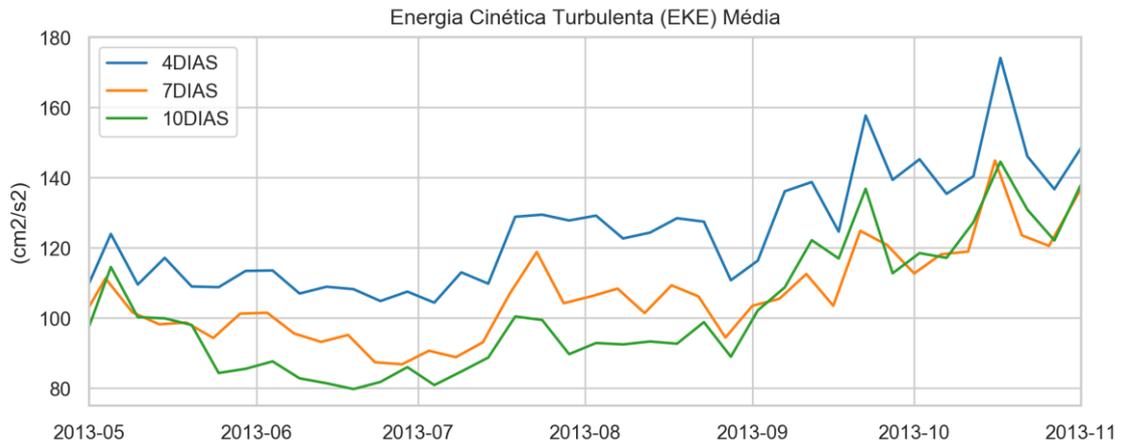


Figura 35: Série temporal da EKE média em todo o domínio do modelo (acima) e na região de maior concentração de EKE definida entre 29°S-19°S e 49°W-39°W (abaixo).

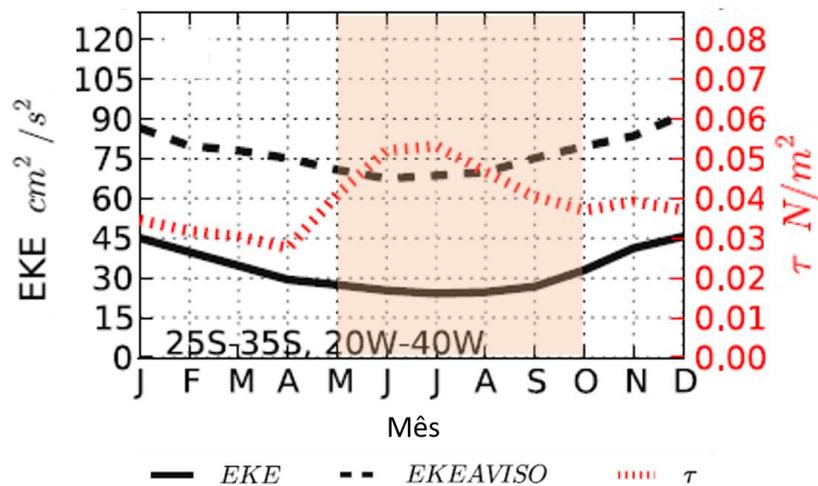


Figura 36: EKE mensal climatológico a partir dos dados ORCA12 (linha sólida preta) e dados de ASM satelital da AVISO (linha pontilhada preta), e tensão de cisalhamento do vento (linha pontilhada vermelha) para a área de 25°S-35°S e 20°W-40°W. O período de maio a outubro é ressaltado pela área em laranja. Adaptado de Rieck, *et al.* (2015).

4.5 Custo Computacional

Neste item é apresentada uma análise do custo computacional envolvido nos três experimentos. Esta análise visa oferecer mais um subsídio para comparação dos experimentos, tendo em vista o objetivo de operacionalização do sistema e fornecimento de análises e previsões diárias. Para tal, o sistema deve ser apto a realizar a análise em tempo hábil para gerar condições iniciais para a previsão. O tempo de análise também é crucial na estimativa de custos do sistema.

Cada ciclo de assimilação dos experimentos realizados – com 2 *outer loops* e 10 *inner loops* - consiste em três integrações do NLROMS, 20 integrações do TLROMS e do ADROMS, além dos algoritmos de minimização empregados em cada *inner loop* e o tempo de leitura dos *inputs* e escrita dos resultados. Desta forma, foi obtido o tempo médio de execução de cada ciclo de assimilação dos três experimentos a partir dos arquivos log e, como cada experimento cobre uma extensão temporal diferente, foi feita a comparação entre os tempos de execução necessários para cobrir 140 dias (MMC entre 4, 7 e 10).

Como esperado, o custo computacional dos três experimentos para realizar a integração de 140 dias foi praticamente igual (~320 horas), e as pequenas diferenças observadas são provavelmente causadas por flutuações no processamento do cluster utilizado.

Tabela 5: Tempo para a execução de 1 ciclo e para a integração de 140 dias, em horas, para cada experimento analisado.

	1 CICLO	140 DIAS
4DIAS	9,1 h	320,8 h
7DIAS	16,23 h	324,8 h
10DIAS	23,24 h	320,43 h

Ainda que o custo computacional seja equivalente entre os experimentos, o tempo de execução de um ciclo de assimilação ainda é muito alto para ser utilizado em uma rotina operacional diária mesmo no experimento de 4 dias. Wilkin *et al.* (2018)

apresenta o sistema operacional utilizando IS4DVAR para a costa nordeste dos Estados Unidos e relata um tempo de ~2 horas de execução da análise que cobre 3 dias com 2 *outer loops* e 8 *inner loops*. Wilkin *et al.* (2018) utiliza uma grade de 140x240x40 com resolução de ~7 km, cobrindo um domínio significativamente menor do que o utilizado neste trabalho, e utiliza um computador de 12 núcleos (3.5 GHz Intel Xeon processor).

Esta análise indica que o sistema ainda não está configurado de forma satisfatória para sua operacionalização e indica alguns caminhos para alcançar esse objetivo, tais como reduzir o domínio do modelo, reduzir o número de *inner loops* e utilizar um número maior de processadores na paralelização das rodadas. Estima-se que um tempo de processamento de 3 horas seja satisfatório para esse objetivo.

5 Considerações finais e conclusões

No presente estudo foram conduzidos experimentos com o objetivo de analisar a influência da extensão da janela de assimilação em um sistema de assimilação de dados 4DVAR para a costa sudeste brasileira. A partir da configuração base apresentada em Fragoso *et al.* (2016) que utilizou uma extensão de 7 dias por ciclo, foram configurados dois experimentos “gêmeos” com janelas de assimilação de 4 e 10 dias. Os três experimentos foram rodados para o período de maio a novembro de 2013 – período com grande quantidade de observações *in-situ* de Temperatura, Salinidade e corrente superficial oriundos da rede de observações do Projeto Azul. Os resultados foram comparados levando-se em conta os dados assimilados e dados independentes (derivadores). Além das comparações diretas, também foram conduzidas análises sobre a minimização da função custo e a representação da energia cinética média e turbulenta, com o objetivo de investigar a relação da extensão da janela de assimilação 4DVAR com a não-linearidade da dinâmica na região de estudo.

Em relação aos dados satelitais assimilados (TSM e ASM), nota-se que:

- O experimento 4DIAS apresenta erros menores que os demais experimentos, seguido pelo experimento 7DIAS e por último o 10DIAS.

- Foi observado um indício de sazonalidade no *Jfit* de ASM e TSM, com maiores valores nos meses de maio e outubro e menores nos meses centrais (junho a agosto).
- Foi observado que o *Jfit* médio por dia dentro do ciclo de assimilação se manteve abaixo de 1 em todos os experimentos, apresentando valores mínimos aproximadamente no dia central dentro da janela de assimilação.
- Recomendam-se trabalhos futuros para testar a utilização de sobreposição entre ciclos, para que os mesmos sejam iniciados a partir da solução com menores erros.
- Para as soluções prévias, notou-se que o *Jfit* se mantém abaixo de 1 por mais tempo (~3 dias) no experimento com janela de 4 dias, indicando uma maior capacidade de previsão com boa acurácia para estas variáveis neste experimento.

Para os perfis hidrográficos assimilados, nota-se uma grande melhoria na representação do desvio padrão de Temperatura e Salinidade em relação ao modelo sem assimilação, principalmente na região da Termoclina. Nesta análise, nota-se que:

- O experimento 7DIAS apresenta vantagem sobre os demais, seguido pelo 4DIAS.
- Os três experimentos apresentam algumas regiões na coluna d'água com erros sistemáticos (*bias*), o que indica a necessidade de realizar uma avaliação da fonte desses erros, possivelmente oriundos do modelo global utilizados como fonte de informação nos contornos laterais.

A comparação com dados independentes de derivadores demonstrou redução do RMSE de velocidade e temperatura superficial nos experimentos em relação ao modelo livre, e grande melhoria nos índices de correlação linear para ambas variáveis. Nota-se que:

- O experimento de 7DIAS apresenta índices melhores que os demais, seguido pelo 4DIAS. O índice de correlação linear, contudo, é sensivelmente melhor nos experimentos de 4 e 7 dias em relação ao de 10 dias.

- A representação da energia cinética média e turbulenta é sensivelmente afetada pelo sistema de assimilação e é melhor no experimento 4DIAS, em relação aos valores calculadas com os dados de derivadores.

A análise da função custo deixa claro que:

- A utilização de 2 *outer loops* é importante para a presente configuração do sistema e que todos os experimentos apresentam valores da função custo não linear bem próximos da função custo linear ao final do ciclo de assimilação.
- Foi observado que nos meses de maio e outubro a diferença entre a função custo não linear e a função custo linear é maior, e que nos meses centrais (junho a agosto) essa diferença é menor. Tal diferença reflete o aumento da importância de feições não-lineares na dinâmica da região.

Foi observada a existência de relação entre esta diferença e o padrão sazonal da energia cinética turbulenta média na região. Nesta análise, observa-se que:

- O experimento 4DIAS apresenta valores de energia cinética turbulenta mais altos do que nos demais experimentos, seguido pelo experimento 7DIAS.
- Como a energia cinética turbulenta na região é composta principalmente por vórtices não lineares de meso escala – feições representadas pelas observações de ASM satelitais assimiladas – e o experimento 4DIAS apresenta menores erros de ASM ao longo de todo o período mas especialmente nos meses de maior EKE, conclui-se que este experimento foi mais capaz de manter as feições turbulentas de mesoescala conforme observadas nos dados.

Indica-se que a utilização da janela de assimilação de 4 dias e de 7 dias são adequadas para este sistema mantendo-se as configurações básicas utilizadas. Pelo comportamento aparentemente sazonal do J_{fit} dos dados satelitais, da energia cinética turbulenta e da diferença entre a função custo não-linear e a função custo linear, sugere-se também como trabalho futuro a ampliação destes experimentos para todo o período anual e a segmentação das análises nas estações.

6 Referências bibliográficas

AMANTE, C. AND B.W. EAKINS. ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. **NOAA Technical Memorandum NESDIS NGDC-24**. National Geophysical Data Center, NOAA. doi:10.7289/V5C8276M. 2009.

BJERKNES, V. Das Problem der Wettervorhersage, betrachtet vom Standpunkte der Mechanik und der Physik (The problem of weather prediction, considered from the viewpoints of mechanics and physics). – *Meteorol. Z.* 21, 1–7, 1904. (translated and edited by VOLKEN E. and S. BRONNIMANN. – *Meteorol. Z.* 18 (2009), 663–667).

BONAVITA, M., LEAN, P., AND HOLM, E.: Nonlinear effects in 4D-Var, **Nonlin. Processes Geophys.**, 25, 713–729, <https://doi.org/10.5194/npg-25-713-2018>, 2018.

BROQUET, G., EDWARDS, C. A., MOORE, A.M., DOYLE, J.D. Application of 4DVARiational data assimilation to the California Current System. *Dynamics of Atmospheres and Oceans*, 48, p. 69-91, 2009a.

CALADO, L. Dinâmica da formação dos meandros e vórtices da Corrente do Brasil ao largo do Sudeste Brasileiro. **Dissertação de Mestrado**. Instituto Oceanográfico Universidade de São Paulo / Brasil, 2001.

CAMPOS, E. J. D. Estudos da circulação oceânica no Atlântico tropical e na região oeste do Atlântico subtropical sul. **Tese (Doutorado)**, Instituto Oceanográfico, Universidade de São Paulo. 114 p, 1995.

CERRONE, B. N. Estatísticas da Circulação do Oceano Atlântico Sudoeste a partir de Dados Lagrangeanos. **Dissertação de Mestrado**. Universidade Federal do Rio de Janeiro. 48 p. 2010.

CHAPMAN, D. C. Numerical treatment of cross-shelf open boundaries in a barotropic ocean model. **Journal of Physical Oceanography** 15, p. 1060-1075, 1985.

CHELTON, D.B. SCHLAX, M.G. SAMELSON, R.M. Global Observations of nonlinear mesoscale eddies. **Progress in Oceanography**. 91(2011). pp. 167-216. 2011.

CHARNEY, J., FJORTOF, J. VON NEUMANN. Numerical Integration of the Barotropic Vorticity Equation. – **Tellus 2**, 237–254, 1950.

COURTIER, P., THÉPAUT, J. N., HOLLINGSWORTH, A. A strategy for operational implementation of 4DVAR, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120, p. 1367-1387, 1994.

DESROZIERS, G., BERRE L., CHAPNIK, B., POLI, P. Diagnosis of observation, background- and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, 131: 3385–3396, 2005.

DOS SANTOS, FA, DE OLIVEIRA ALSC, SOARES, FLM, CARVALHO, GV, SÃO TIAGO PM, SANTOS NG, GARÇÃO HF, PIMENTEL FP, BOECHAT AC, MARIANO LS, BERGMAN C, RINALDI FL, MIRANDA TC, CABRAL MM, DA CRUZ LMM, PELLEGRINI JAC, ASSAD LPF, LANDAU L, FRAGOSO MR. Projeto Azul: Operational Oceanography in an Active Oil and Gas Area Southeastern Brazil. In: Liu Y, Kerkering H, Weisberg RH (ed) *Coastal Ocean Observing Systems*, 1st edn. Elsevier, London, pp 99-116. 2016.

EVENSEN, G. The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53, p. 343-367, 2003.

FALLER, D. G. Modelos Bayesianos Hierárquicos para estimar matrizes de covariância de erro vertical do *background* na assimilação de dados oceanográficos. **Tese (Doutorado)**, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 182 pp, 2016.

FERRARI R. WUNSCH, C. Ocean Circulation Kinetic Energy – Reservoirs, Sources and Sinks. **Annu. Ver. Fluid Mech.** 2008.

FISHER M, COURTIER P. Estimating the covariance matrices of analysis and forecast error in variational data assimilation. **ECMWF Tech Memo** 220:1–26. 1995.

FLATHER, R. A. A tidal model of the northwest European continental shelf. **Memoires, Societe Royale Sciences de Liege**, 6 (10), p. 141-164, 1976.

FRAGOSO, M. R., CARVALHO, G. V., SOARES, F. L. M., FALLER, D. G., ASSAD, L. P. F., TOSTE, R., SANCHO, L. M. B., PASSOS, E. N., BÖCK, C. S., REIS, B.,

LANDAU, L., ARANGO, H., MOORE, A. M. A 4D-variational ocean data assimilation application for Santos Basin, Brazil. *Ocean Dynamics*, 66: 419, 2016.

GOOD, S. A., M. J. MARTIN AND N. A. RAYNER. EN4: quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates, **Journal of Geophysical Research: Oceans**, 118, 6704-6716, doi:10.1002/2013JC009067. 2013.

GODOI, S. S. Dinâmica quase-geostrófica do sistema Corrente do Brasil no embanhamento de São Paulo (23,5°-27 °S). **Tese (Doutorado)**, Universidade de São Paulo, São Paulo, 133 p., 2005.

GLEICK, James. Chaos; making a new science. New York : **Viking**, 1987. 352 p.

HAIKVOGEL, D. B., ARANGO, H., BUDGELL, W. P., CORNUELLE, B. D., CURCHITSER, E., DI LORENZO, E., FENNEL, K., GEYER, W. R., HERMANN, A. J., LANEROLLE, L., LEVIN, J., MCWILLIAMS, J. C., MILLER, A. J., MOORE, A. M., POWELL, T. M., SHCHEPETKIN, A. F., SHERWOOD, C. R., SIGNELL, R. P., WARNER, J. C., WILKIN, J. Ocean forecasting in terrain-following coordinates: formulation and skill assessment of the Regional Ocean Modeling System. *Journal of Computational Physics* 227, 3595–3624, 2008.

IBGE. CENSO DEMOGRÁFICO, 2010. Disponível em <https://censo2010.ibge.gov.br/>. Acessado em 02/2018.

KALNAY, E. Atmospheric Modelling, Data Assimilation and Predictability. *Cambridge University Press*, 2003.

KANAMITSU, M., EBISUKASI, W., WOOLLEN, J., YANG, S., HNILLO, J. J., FIORINO, M., POTTER, G. L, 2002. An Updated NCEP-NCAR Reanalysis, Covering 1979-Present, Features Newer Physics and Observed Soil Moisture Forcing and Also Eliminates Several Previous Errors. *Bulletin of the American Meteorological Society*, v. 82, pp. 1631-1643, 2002.

LI, Y. TOUMI, R. A balanced Kalman filter ocean data assimilation system with application to the South Australian Sea. *Ocean Modelling*, v. 116 , pp.159–172. 2017.

LORENC A.C. Analysis Methods for Numerical Weather Prediction. **Quarterly Journal of the Royal Meteorological Society**, 112, p. 1177-1194, 1986.

LORENC, A. C., PAYNE, T.: 4D-Var and the butterfly effect: Statistical four-dimensional data assimilation for a wide range of scales, **Q. J. Roy. Meteor. Soc.**, 133, 607–614, <https://doi.org/10.1002/qj.36>, 2007.

LORENZ EN. Deterministic nonperiodic flow. **J Atmospher Sci**1963;20:130–41, 1963.

LORENZ EN. Atmospheric predictability as revealed by naturally occurring analogues. **J. Atmos. Sci.** 26, 636-646. 1969.

MANO, M.F., A.M. PAIVA, A.R. TORRES, AND A.L. COUTINHO. Energy Flux to a Cyclonic Eddy off Cabo Frio, Brazil. **J. Phys. Oceanogr.**, 39, 2999–3010, <https://doi.org/10.1175/2009JPO4026.1>. 2009.

MARCHESIELLO, P., MCWILLIAMS, J. C., SHCHEPETKIN, A. Open boundary conditions for long-term integration of regional oceanic models. **Ocean Modelling**, 3, p. 1-20, 2001.

MARIANO, LÍVIA SANT'ANGELO. Avaliação de alterações no desvio padrão do erro de background em um sistema de assimilação de dados oceânicos 4D-variacional / Lívia Sant'Angelo Mariano. – Rio de Janeiro: UFRJ/COPPE, 2019.

MATTERN, J. P; EDWARDS, C. A.; MOORE, A. M. Improving variational data assimilation through background and observation error adjustments. **Monthly Weather Review**. 2018

MOORE, A. M, ARANGO, H. G, BROQUET, G., POWELL, B. S., WEAVER, A. T., ZAVALAGARAY, J. The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems, Part I: Formulation and overview. **Progress in Oceanography**, 91, p. 34-49, 2011a.

MOORE, A. M., ARANGO, H. G., BROQUET, G., EDWARDS, C., VENEZIANI, M., POWELL, B., FOLEY, D., DOYLE, J.D., COSTA, D., ROBINSON, P. The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems,

Part II: Performance and applications to the California Current System. *Progress in Oceanography*, 91, p. 50-73, 2011b.

MOORE, A., ARANGO, H. G., BROQUET, G., EDWARDS, C., VENEZIANI, M., POWELL, B., FOLEY, D., DOYLE, J. D., COSTA, D., ROBINSON, P. The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems, Part III: Observation impact and observation sensitivity in the California Current System. *Progress in Oceanography*, 91, p. 74-94, 2011c.

MOORE, A. M, ARANGO, H. G., DI LORENZO, E., CORNUELLE, B.D., MILLER, A. J., NEILSON, D. J. A comprehensive ocean prediction and analysis system based on the tangent linear and adjoint of a regional ocean model. *Ocean Modelling*, 7:227-258, 2004.

NEVEU, E., MOORE, A. M., EDWARDS, C. A., FIECHTER, J., DRAKE, P., CRAWFORD, W. J., JACOX, M. G., NUSS, E. An historical analysis of the California Current circulation using ROMS 4D-Var: System configuration and diagnostics. *Ocean Modelling*, 99, p. 133-151, 2016.

NCEP-DOE AMIP-II Reanalysis (R-2): M. Kanamitsu, W. Ebisuzaki, J. Woollen, S-K Yang, J.J. Hnilo, M. Fiorino, and G. L. Potter. 1631-1643, Nov 2002, **Bulletin of the American Meteorological Society**. 2002.

NGODOCK H, MUSCARELLA P, CARRIER M, SOUOPGUI I, SMITH S. Assimilation of HF Radar Observations in the Chesapeake-Delaware Bay Region using the Navy Coastal Ocean Model (NCOM) and the Four-Dimensional Variational (4DVAR) Method. In: Liu Y, Kerkering H, Weisberg RH (ed) *Coastal Ocean Observing Systems*, 1st edn. Elsevier, London, pp 373-390, doi: 10.1016/B978-0-12-802022-7.00020-1. 2015

PIRES, C., VAUTARD, R., AND TALAGRAND, O. On extending the limits of variational assimilation in nonlinear chaotic systems, **Tellus A**, 48, 96–121, 1996.

PIRSIG, R. M. Zen e a arte da manutenção de motocicletas. Editora Paz e Terra. ISBN: 9788521903789. 1984.

POND, S. AND PICKARD, G.L. Introductory dynamical oceanography. 2nd Edition, Pergamon Press, Oxford. 1983.

RICHARDSON, L.F. Weather Prediction by Numerical Processes. – **Cambridge University Press**, Cambridge. 1922.

RIECK, J.K., C.W. BÖNING, AND R.J. GREATBATCH. Decadal Variability of Eddy Kinetic Energy in the South Pacific Subtropical Countercurrent in an Ocean General Circulation Model. **J. Phys. Oceanogr.**, 48, 757–771, <https://doi.org/10.1175/JPO-D-17-0173.1>. 2018.

RIECK, J. K., C. W. BÖNING, R. J. GREATBATCH, AND M. SCHEINERT. Seasonal variability of eddy kinetic energy in a global high-resolution ocean model. **Geophys. Res. Lett.**, 42, 9379–9386, <https://doi.org/10.1002/2015GL066152>. 2015.

SHCHEPETKIN, A. F.; MCWILLIAMS, J.C., 1998. Quasi-monotone advection schemes based on explicit locally adaptive dissipation. **Monthly Weather Review**, v. 126, p.1541-1580. 1998.

SHCHEPETKIN, A. F.;MCWILLIAMS, J.C. A method for computing horizontal pressure-gradient force in an oceanic model with a nonaligned vertical coordinate. **Journal of Geophysical Research**, n. 108 (C3), p. 3090-3105. 2003.

SILVEIRA, I. C. A., SCHIMIDT, A. C. K., CAMPOS. E.J.D., GODOI, S.S., IKEDA, Y. A Corrente do Brasil ao largo da Costa Leste Brasileira. **Revista Brasileira de Oceanografia**, 48 (2), p. 171-183, 2000.

SILVEIRA, I. C. A. O Sistema Corrente do Brasil na Bacia de Campos, RJ. **Tese** (Doutorado), Instituto Oceanográfico, Universidade de São Paulo, São Paulo, 160 p., 2007.

SOUZA, R. B.; ROBINSON, I. S. “Lagrangian and satellite observations of the Brazilian Coastal Current”. **Continental Shelf Research** vol. 24 (2) pp. 241-262. 2004.

STEPHEN POND, GEORGE L. PICKARD, 7 - The Role of the Non-linear Terms and the Magnitudes of Terms in the Equations of Motion, Editor(s): STEPHEN POND, GEORGE L. PICKARD, Introductory Dynamical Oceanography (Second Edition),

Butterworth-Heinemann, p. 44-62, ISBN 9780080570549,
<https://doi.org/10.1016/B978-0-08-057054-9.50013-4>. 1983.

STRAMMA, L., ENGLAND, M. On the water masses and mean circulation of the South Atlantic Ocean. *Journal of Geophysical Research*, 104(C9), p. 20863-20883, 1999.

TALAGRAND, O. Assimilation of observations, an introduction. *Journal of Meteorological Society of Japan*, 75(1B), p. 191-209, 1997.

TALAGRAND, O., COURTIER P. Variational assimilation of meteorological observations with the adjoint vorticity equation. i: Theory. *Quarterly Journal of the Royal Meteorological Society*, 113(478), p. 1311-1328, 1987.

WEAVER, A. T., COURTIER, P. Correlation modelling on the sphere using a generalized diffusion equation, *Quarterly Journal of the Royal Meteorological Society*, 127, p. 1815-1846, 2001.

WILKIN, J., ET AL. A coastal ocean forecast system for the U.S. Mid-Atlantic Bight and Gulf of Maine. In "New Frontiers in Operational Oceanography", E. Chassignet, A. Pascual, J. Tintoré, and J. Verron, Eds., GODAE OceanView, 593-624, doi:10.17125/gov2018.ch21. 2018.

ANEXO 1 - Mapas Médios Mensais

Neste anexo, são apresentados mapas de corrente superficial médios mensais para cada experimento. Sobreposto à cada mapa, podem ser observadas as trajetórias dos derivadores do Projeto Azul ao longo do mês. Ainda que as trajetórias dos derivadores sejam reflexo da evolução do campo de corrente ao longo de um período de 1 mês, é possível observar a representação de diversas feições de mesoescala nos resultados médios mensais dos experimentos, condizentes com as trajetórias.

O objetivo deste capítulo é fornecer uma visão mais geral, qualitativa, dos campos de corrente superficial dos experimentos, uma vez que as comparações diretas quantitativas já foram apresentadas em itens anteriores.

De uma forma geral, nota-se uma grande semelhança nos campos médios dos 3 experimentos, com a representação da maior parte das feições de mesoescala com pouca diferença de posicionamento e diferindo um pouco mais nas intensidades.

Destacam-se algumas feições interessantes como a representação do Vórtice de Cabo Frio em maio/2013 em aproximadamente 24°S e 42°W , cujo posicionamento é coerente com as trajetórias dos derivadores observados (Figura 37). No mesmo mês, há um derivador que apresenta trajetória meandrante para leste na latitude de $26,5^{\circ}$, sendo possível observar duas feições com giro ciclônico responsáveis pelo padrão meandrante observado nas trajetórias. Este mesmo derivador retorna ao domínio de modelagem em Junho apresentando um giro ciclônico – feição também representada pelos 3 experimentos no que aparentemente é uma sucessão de vórtices ciclônicos com translação para oeste (Figura 38). Em Julho observa-se a Corrente do Brasil bem definida com pouco meandramento, alta intensidade e nenhum vórtice associado. Esta representação nos resultados dos modelos é corroborada pela trajetória de um derivador que em um mês percorre aproximadamente 5° de latitude transportado pela CB (Figura 39). Em agosto, observam-se duas trajetórias que conduzem os derivadores para a costa e indicam a influência da Corrente Costeira do Brasil, não representada no fluxo médio dos experimentos (Figura 40). Outras feições de meso escala podem ser observadas com grande coerência em relação à trajetória dos derivadores, especialmente na região à leste da CB ao longo dos meses de análise (Figura 41 a Figura 42).

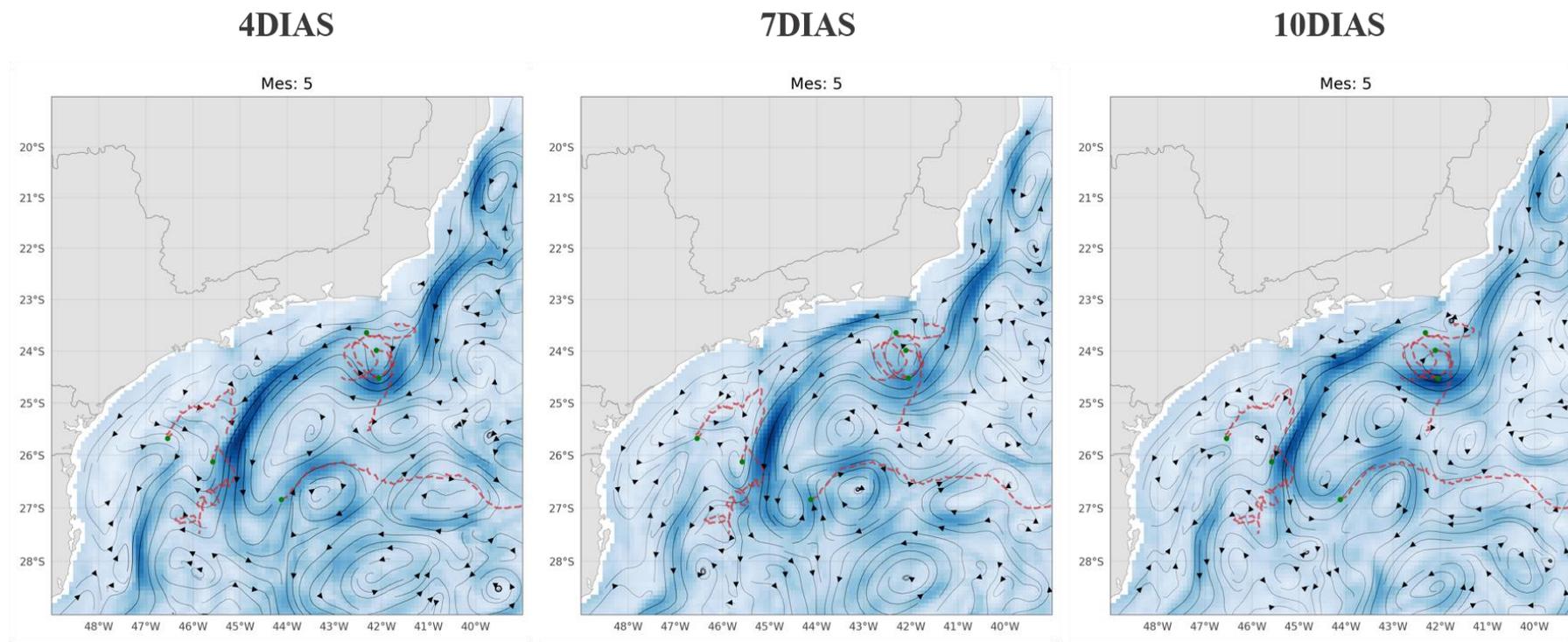


Figura 37: Campos de corrente média extraídas das soluções posteriores (análises) para o mês de Maio . Da esquerda para a direita, resultados do experimento com janela de assimilação de 4, 7 e 10 dias. As linhas tracejadas vermelhas representam as trajetórias dos derivadores do Projeto Azul ao longo do mês de maio.

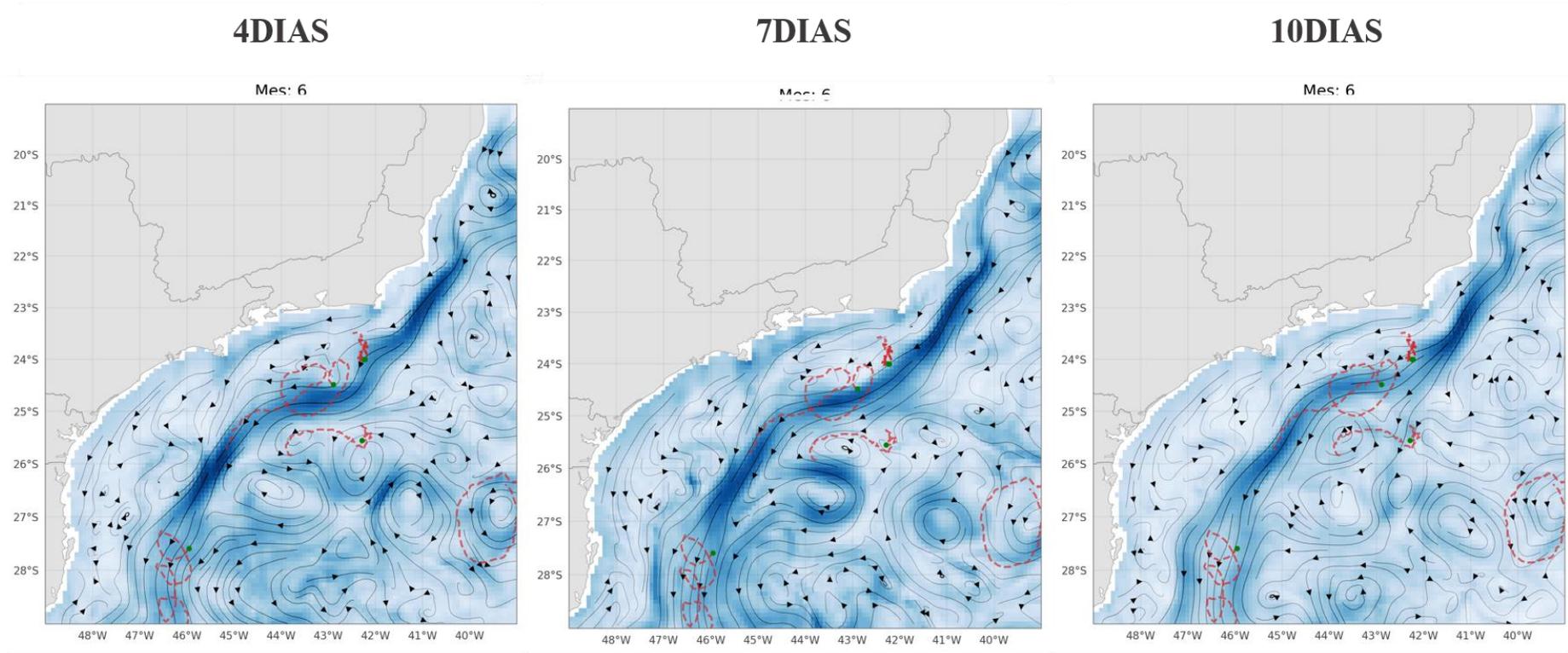


Figura 38: Campos de corrente média extraídas das soluções posteriores (análises) para o mês de Junho . Da esquerda para a direita, resultados do experimento com janela de assimilação de 4, 7 e 10 dias. As linhas tracejadas vermelhas representam as trajetórias dos derivadores do Projeto Azul ao longo do mês de Junho.

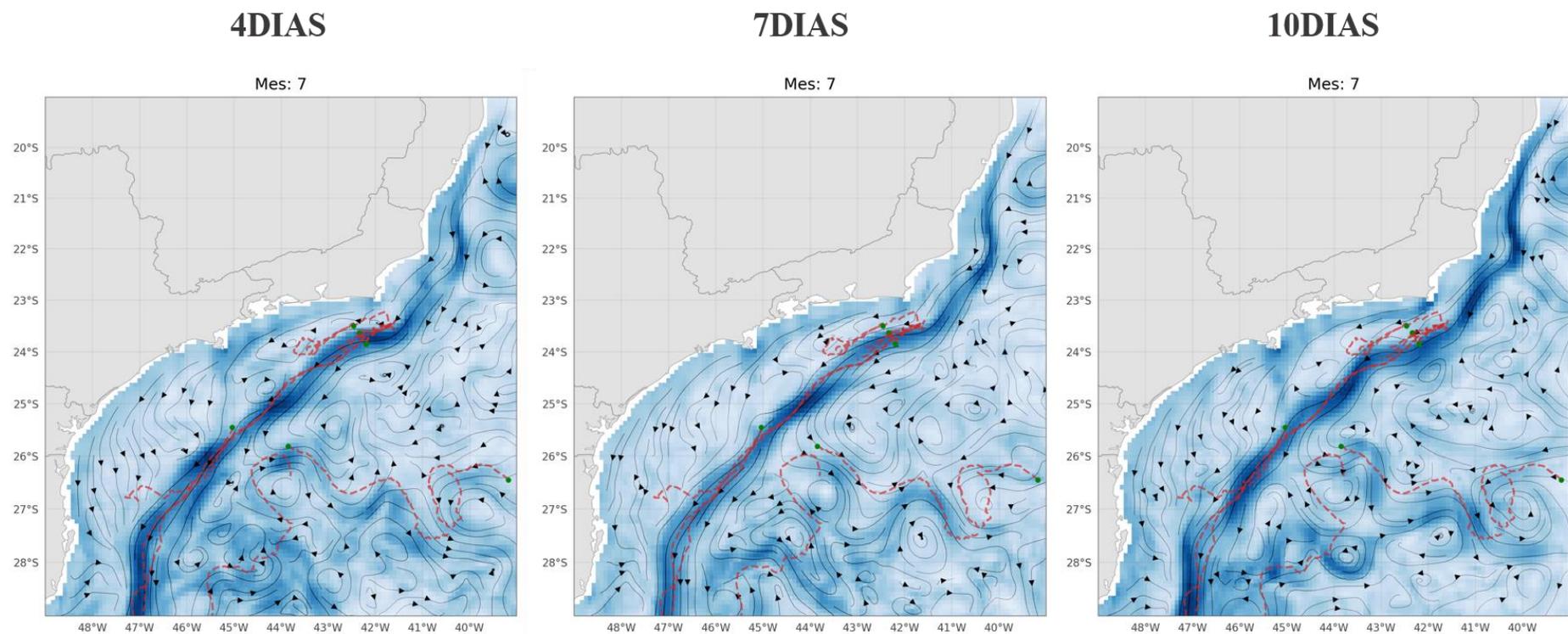


Figura 39: Campos de corrente média extraídas das soluções posteriores (análises) para o mês de Julho . Da esquerda para a direita, resultados do experimento com janela de assimilação de 4, 7 e 10 dias. As linhas tracejadas vermelhas representam as trajetórias dos derivadores do Projeto Azul ao longo do mês de Julho.

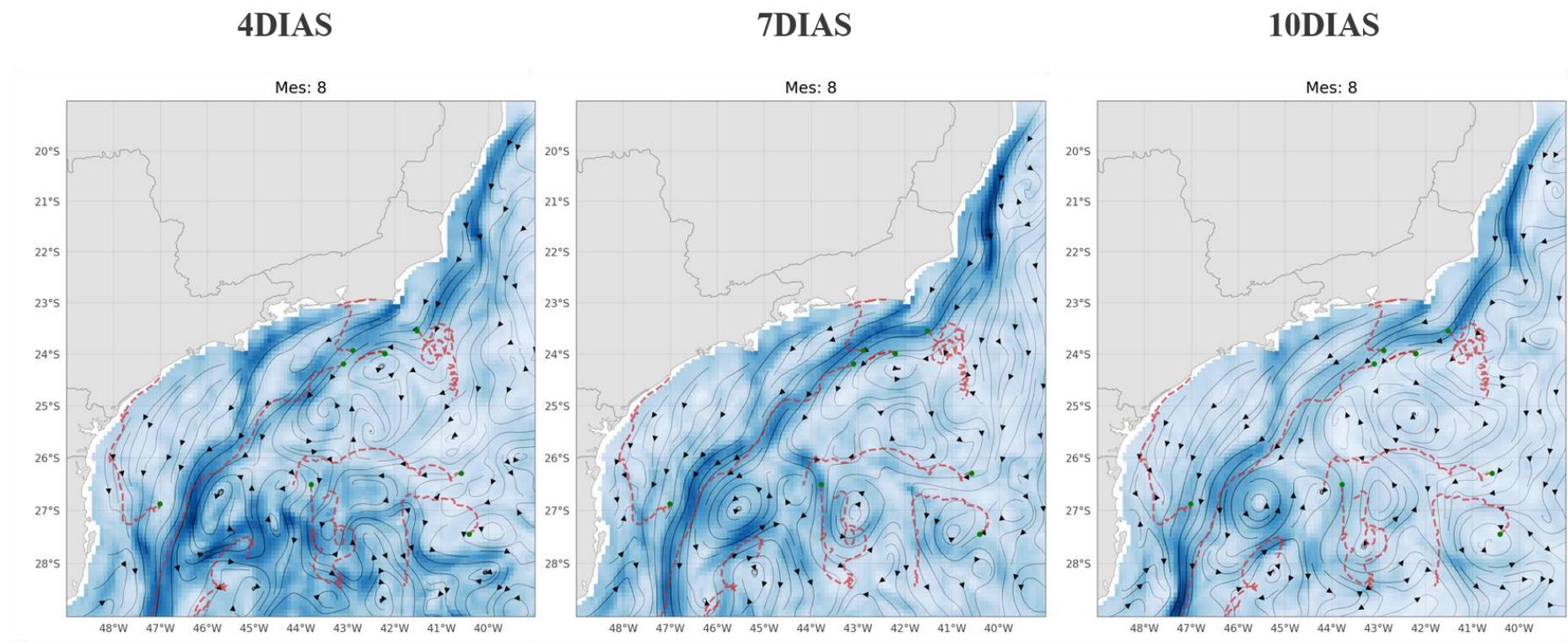


Figura 40: Campos de corrente média extraídas das soluções posteriores (análises) para o mês de Agosto . Da esquerda para a direita, resultados do experimento com janela de assimilação de 4, 7 e 10 dias. As linhas tracejadas vermelhas representam as trajetórias dos derivadores do Projeto Azul ao longo do mês de Agosto.

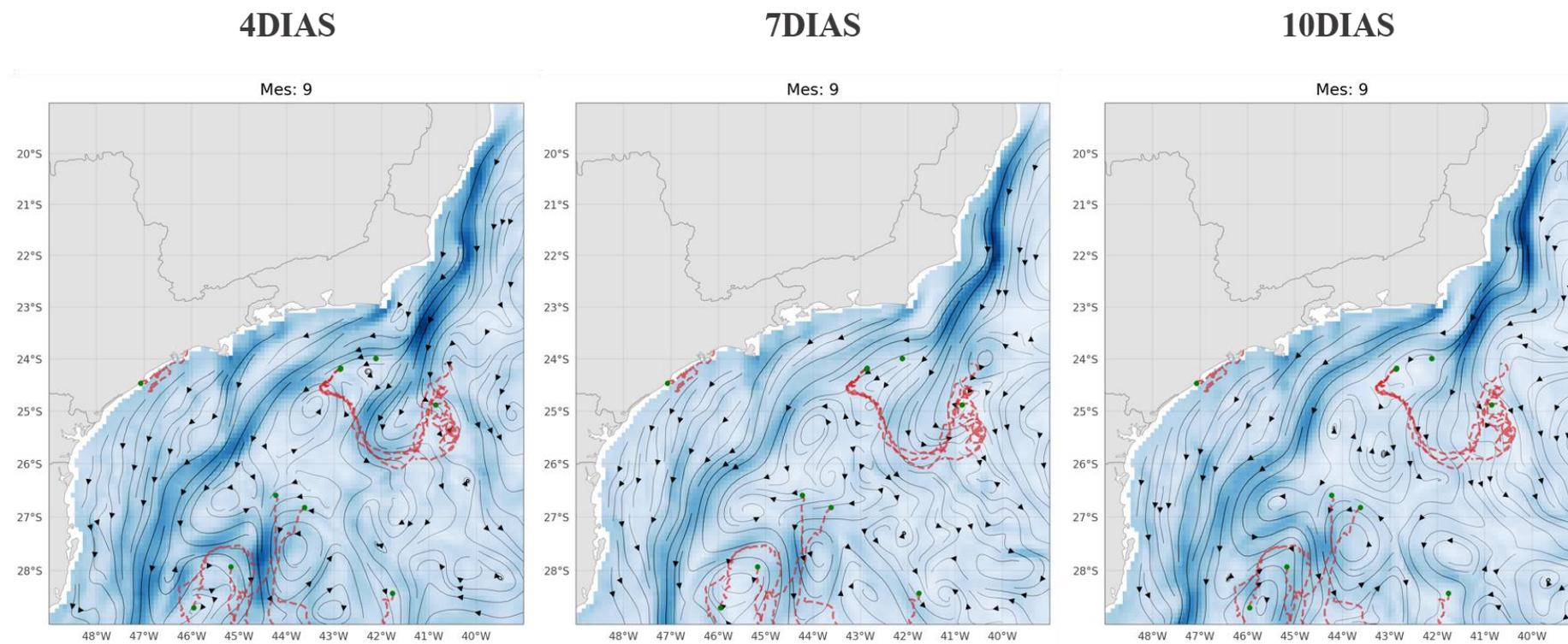


Figura 41: Campos de corrente média extraídas das soluções posteriores (análises) para o mês de Setembro. Da esquerda para a direita, resultados do experimento com janela de assimilação de 4, 7 e 10 dias. As linhas tracejadas vermelhas representam as trajetórias dos derivadores do Projeto Azul ao longo do mês de Setembro.

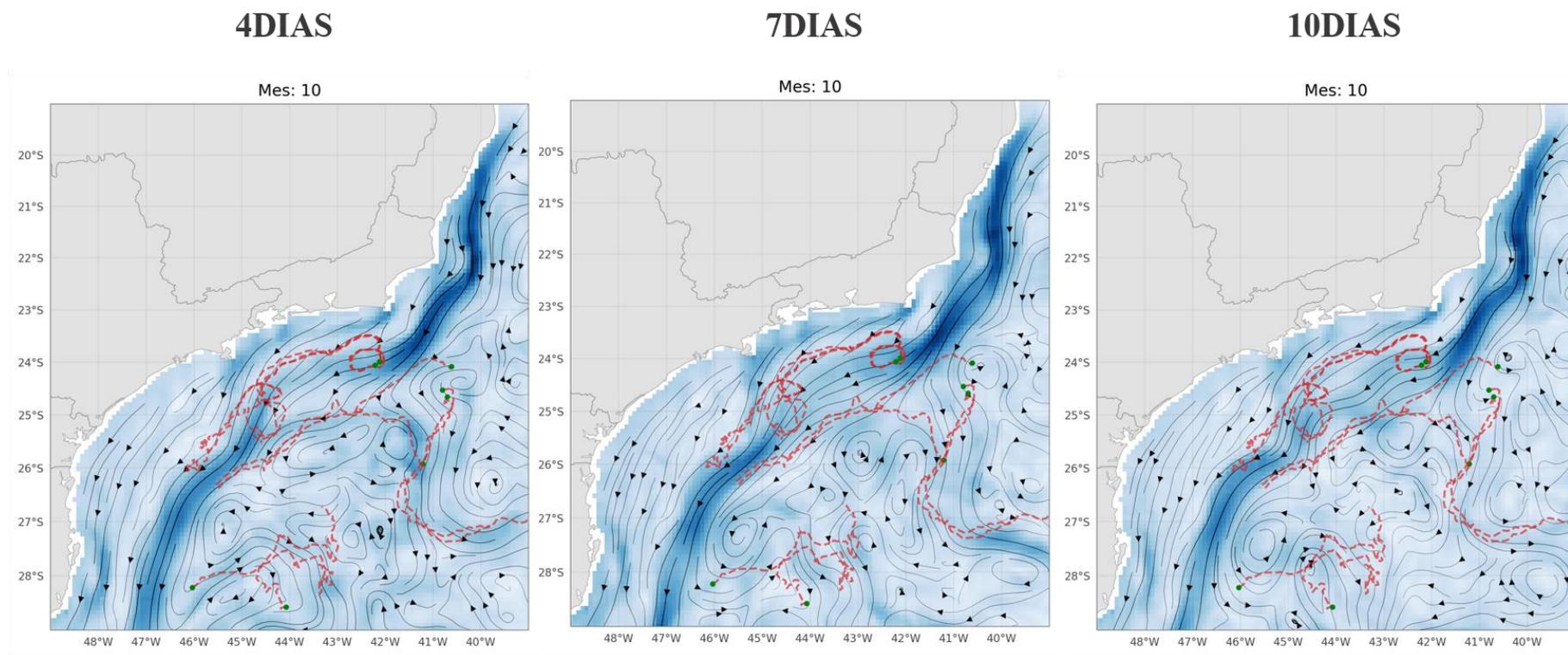


Figura 42: Campos de corrente média extraídas das soluções posteriores (análises) para o mês de Outubro. Da esquerda para a direita, resultados do experimento com janela de assimilação de 4, 7 e 10 dias. As linhas tracejadas vermelhas representam as trajetórias dos derivadores do Projeto Azul ao longo do mês de Outubro

