



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES  
**INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS**

sid.inpe.br/mtc-m21c/2018/10.29.17.51-TDI

## **AJUSTE DINÂMICO PARA ANÁLISE HÍBRIDA ENTRE UM SISTEMA VARIACIONAL E FILTRO DE KALMAN POR CONJUNTO**

Helena Barbieri de Azevedo

Tese de Doutorado do Curso de Pós-Graduação em Meteorologia, orientada pelo Dr. Luis Gustavo Gonçalves de Gonçalves, aprovada em 22 de outubro de 2018.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34R/3S5CNC2>>

INPE  
São José dos Campos  
2018

**PUBLICADO POR:**

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GBDIR)

Serviço de Informação e Documentação (SESID)

CEP 12.227-010

São José dos Campos - SP - Brasil

Tel.:(012) 3208-6923/7348

E-mail: pubtc@inpe.br

**COMISSÃO DO CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO  
DA PRODUÇÃO INTELECTUAL DO INPE (DE/DIR-544):****Presidente:**

Dr. Marley Cavalcante de Lima Moscati - Centro de Previsão de Tempo e Estudos Climáticos (CGCPT)

**Membros:**

Dra. Carina Barros Mello - Coordenação de Laboratórios Associados (COCTE)

Dr. Alisson Dal Lago - Coordenação-Geral de Ciências Espaciais e Atmosféricas (CGCEA)

Dr. Evandro Albiach Branco - Centro de Ciência do Sistema Terrestre (COCST)

Dr. Evandro Marconi Rocco - Coordenação-Geral de Engenharia e Tecnologia Espacial (CGETE)

Dr. Hermann Johann Heinrich Kux - Coordenação-Geral de Observação da Terra (CGOBT)

Dra. Ieda Del Arco Sanches - Conselho de Pós-Graduação - (CPG)

Silvia Castro Marcelino - Serviço de Informação e Documentação (SESID)

**BIBLIOTECA DIGITAL:**

Dr. Gerald Jean Francis Banon

Clayton Martins Pereira - Serviço de Informação e Documentação (SESID)

**REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:**

Simone Angélica Del Ducca Barbedo - Serviço de Informação e Documentação (SESID)

André Luis Dias Fernandes - Serviço de Informação e Documentação (SESID)

**EDITORAÇÃO ELETRÔNICA:**

Ivone Martins - Serviço de Informação e Documentação (SESID)

Murilo Luiz Silva Gino - Serviço de Informação e Documentação (SESID)



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES  
**INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS**

sid.inpe.br/mtc-m21c/2018/10.29.17.51-TDI

## **AJUSTE DINÂMICO PARA ANÁLISE HÍBRIDA ENTRE UM SISTEMA VARIACIONAL E FILTRO DE KALMAN POR CONJUNTO**

Helena Barbieri de Azevedo

Tese de Doutorado do Curso de Pós-Graduação em Meteorologia, orientada pelo Dr. Luis Gustavo Gonçalves de Gonçalves, aprovada em 22 de outubro de 2018.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34R/3S5CNC2>>

INPE  
São José dos Campos  
2018

Dados Internacionais de Catalogação na Publicação (CIP)

---

Azevedo, Helena Barbieri de.

Az25a Ajuste dinâmico para análise híbrida entre um sistema variacional e filtro de Kalman por conjunto / Helena Barbieri de Azevedo. – São José dos Campos : INPE, 2018.

xx + 154 p. ; (sid.inpe.br/mtc-m21c/2018/10.29.17.51-TDI)

Tese (Doutorado em Meteorologia) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2018.

Orientador : Dr. Luis Gustavo Gonçalves de Gonçalves.

1. Assimilação de dados. 2. Sistemas híbridos. 3. Filtro de Kalman por conjunto. 4. Sistema variacional. 5. Previsão numérica de tempo. I.Título.

CDU 551.509.313.3

---



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](https://creativecommons.org/licenses/by-nc/3.0/).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).

Aluno (a): **Helena Barbieri de Azevedo**

Título: "AJUSTE DINÂMICO PARA ANÁLISE HÍBRIDA ENTRE UM SISTEMA VARIACIONAL E FILTRO DE KALMAN POR CONJUNTO"

Aprovado (a) pela Banca Examinadora em cumprimento ao requisito exigido para obtenção do Título de **Doutor(a)** em **Meteorologia**

Dr. José Antonio Aravéquia



Presidente / INPE / Cachoeira Paulista - SP

( ) Participação por Vídeo - Conferência

Aprovado ( ) Reprovado

Dr. Luis Gustavo Gonçalves de Gonçalves



Orientador(a) / INPE / Cachoeira Paulista - SP

( ) Participação por Vídeo - Conferência

Aprovado ( ) Reprovado

Dr. Carlos Frederico Bastarz

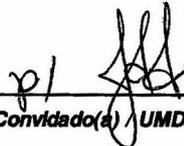


Membro da Banca / INPE / Cachoeira Paulista - SP

( ) Participação por Vídeo - Conferência

Aprovado ( ) Reprovado

Dra. Eugênia Kalnay



Convidado(a) / UMD / Maryland - EUA

Participação por Vídeo - Conferência

Aprovado ( ) Reprovado

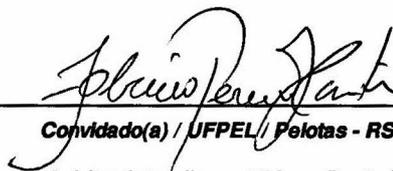
**Este trabalho foi aprovado por:**

( ) maioria simples

unanimidade

Aprovado (a) pela Banca Examinadora  
em cumprimento ao requisito exigido para  
obtenção do Título de **Doutor(a)** em  
**Meteorologia**

Dr. **Fabício Pereira Harter**



Convidado(a) / UFPEL / Pelotas - RS

Participação por Video - Conferência

Aprovado

Reprovado

**Este trabalho foi aprovado por:**

maioria simples

unanimidade

**São José dos Campos, 22 de outubro de 2018**

## AGRADECIMENTOS

Agradeço aos meus pais Ana e Edson pelo apoio em todas as minhas escolhas, pelo carinho e amor dados; ainda que distantes, foram muito importantes.

Aos meus amigos que, com certeza, fizeram a diferença para o bom andamento deste trabalho, seja com o apoio emocional ou com momentos de descontração, em especial à Bianca Maske, Carla Comin, Claudia Medeiros, Giuliana Barbieri, Graziela Luzia, Juliana Arrué, Larissa Antunes, Natalia Sinhori, Marcos Silveira, Mari Firpo, Tassia Sene e Thamiris Santana. Ao meu namorado Isac Satim por todas as conversas, momentos de descontração, paciência e, principalmente, pelo apoio e incentivo sempre.

Ao meu orientador Dr. Luis Gustavo Gonçalves de Gonçalves pela oportunidade, incentivo e confiança, além do apoio, paciência e amizade durante todo esse período.

Aos colegas do Grupo de Desenvolvimento em Assimilação de Dados, em especial ao colega Dr. Carlos Frederico Bastarz por inúmeras discussões, pelo incentivo e pelas muitas ajudas. Ao Dr. José Antonio Aravéquia e ao Dr. Luiz Fernando Sapucci, pelas críticas e sugestões dadas durante a banca de proposta de tese, além de discussões durante o desenvolvimento deste trabalho.

Ao Professor Dr. Fabrício Harter por me receber na Universidade Federal de Pelotas (UFPel) durante o período de estágio docência.

À Professora Dra. Eugênia Kalnay da *University of Maryland* (UMD), pela boa vontade em me receber no DAOS (*Department of Atmospheric and Ocean Science*) durante os 6 meses em que eu estive lá. Agradeço pela oportunidade, orientação, conversas, paciência, por acreditar no trabalho e, principalmente, pelo incentivo. Agradeço também ao colega Matthew Wespetal, também da UMD, por gentilmente ceder o código de trabalho.

Ao Centro de Previsão de Tempo e Estudos Climáticos do Instituto Nacional de Pesquisas Espaciais pelo espaço e suporte computacional fornecido.

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (CAPES), e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão das bolsas de estudos, tanto nacionais quanto internacionais.



## RESUMO

Ao longo dos últimos anos, os sistemas de assimilação de dados Híbridos têm sido alvo de vários estudos. Esses sistemas consistem em combinar dois outros sistemas de assimilação de dados, sendo usualmente um variacional e outro baseado em conjunto. Os sistemas Híbridos tradicionais visam beneficiar um sistema variacional através do uso da covariância do erro do *background* de um sistema por conjunto. Uma abordagem mais recente visa melhorar um sistema por conjunto através da estabilidade vinda do sistema variacional, pela combinação linear entre suas análises, ponderadas por um peso estático arbitrário. Considerando que os sistemas de assimilação de dados variacionais e os sistemas por conjunto apresentam erros diferentes e que uma forma de quantificar a incerteza do sistema por conjunto é através do seu espalhamento, esta tese de doutorado propõe avaliar o ganho ao incluir em um sistema de assimilação de dados Híbrido Médio um coeficiente de ajuste dinâmico, o qual utiliza essas informações de incerteza do sistema por conjunto na configuração desse coeficiente. A metodologia empregada, utiliza como modelo um MCGA simplificado (SPEEDY), como sistema por conjunto o LETKF, e como sistema variacional o 3DVar para comporem o sistema Híbrido Médio. O coeficiente de ajuste dinâmico é calculado com base nos valores normalizados do espalhamento do conjunto. Foram realizados ao todo oito experimentos, sendo dois deles com o LETKF e com o 3DVar separadamente, cinco deles variando os valores de alfa fixo ( $\alpha = 0, 1, \alpha = 0, 3, \alpha = 0, 5, \alpha = 0, 7$  e  $\alpha = 0, 9$ ), e por último, um experimento utilizando o coeficiente de ajuste dinâmico. Nesses experimentos foram avaliadas as análises e previsões de até 120 horas para os períodos de verão e inverno. Os resultados mostraram que o LETKF apresenta erros menores que o 3DVar tanto em suas análises quanto nos primeiros horários de previsão, porém esse padrão se inverte a partir de aproximadamente 48 horas de previsão. Através da avaliação feita com as análises geradas pelos diversos valores de alfa, foi encontrado um valor ótimo de  $\alpha = 0, 1$ , e quando este resultado foi comparado contra o experimento com o coeficiente de ajuste dinâmico, também obteve, na maioria dos casos, menores erros nas análises. No entanto, para o período de verão, o coeficiente de ajuste dinâmico apresentou análises mais balanceadas, enquanto as análises geradas pelo  $\alpha = 0, 3$ , para o período de inverno, foram melhor balanceadas. Ao serem comparadas as previsões geradas a partir das análises dos experimentos com  $\alpha = 0, 1$  e com o coeficiente de ajuste dinâmico, observou-se o seguinte padrão de comportamento: existe uma tendência no aumento do ganho do experimento com o coeficiente de ajuste dinâmico em relação ao experimento com  $\alpha = 0, 1$ . Por fim, experimentos adicionais foram realizados a fim de avaliar o comportamento do sistema Híbrido Médio em relação ao aumento no número de membros e em relação a otimização do intervalo do coeficiente de ajuste dinâmico.

Palavras-chave: Assimilação de Dados. Sistemas Híbridos. Filtro de Kalman por Conjunto. Sistema Variacional. Previsão Numérica de Tempo.



# DYNAMIC FIT FOR HYBRID ANALYSIS BETWEEN A VARIATIONAL SYSTEM AND ENSEMBLE KALMAN FILTER

## ABSTRACT

Over the past few years, Hybrid data assimilation systems have been the subject of several studies. These systems consist of combining two other data assimilation systems, usually one variational and one ensemble based. Traditional Hybrid systems aim to benefit a variational system by using the background error covariance of the ensemble based system. The most recent approach is to improve the ensemble based system, through the stability of the variational system, by the linear combination between both analyses, weighted by an arbitrary static weight. Considering that variational data assimilation systems and ensemble based systems have different errors, one way of quantifying the uncertainty of the system is through the ensemble spread. This PhD thesis proposed to evaluate a Hybrid Gain data assimilation system by including a dynamic fit coefficient, which considers the uncertainty information from the ensemble system. The methodology uses a simplified AGCM (SPEEDY) model, an ensemble based system (LETKF) and a variational system (3DVar) to build up the Hybrid Gain system. The dynamic fit coefficient is calculated based on the normalized values of the ensemble spread. A total of eight experiments were performed, two of them using LETKF and 3DVar separately, five of them varying the fixed alpha values ( $\alpha = 0.1$ ,  $\alpha = 0.3$ ,  $\alpha = 0.5$ ,  $\alpha = 0.7$  and  $\alpha = 0.9$ ), and finally, an experiment using the dynamic fit coefficient. In these experiments, the analyses and forecasts of up to 120 hours for the summer and winter were analyzed. The results showed that LETKF presented smaller errors when compared to the 3DVar in both analyzes and the first forecasting hours, but this pattern reverses starting on approximately 48 hours. By evaluating the various alpha values, an optimal value of  $\alpha = 0.1$  was found, and when this result was compared against the experiment with the dynamic fit coefficient, it also obtained, in most cases, smaller errors in the analyzes. Furthermore, during the summer, the dynamic fit coefficient presented more balanced analyzes, whereas the analyzes generated by the  $\alpha = 0.3$ , for the winter period, were better balanced. When the forecasts generated from the analyzes of the experiments with  $\alpha = 0.1$  and the dynamic fit coefficient were compared, the following was noticed: there is a tendency in the gain of the experiment with the dynamic fit coefficient with respect to the experiment with  $\alpha = 0.1$ . Finally, additional experiments were carried out in order to evaluate the behavior of the Hybrid Gain system with respect to the increase in the number of members and i to the optimization of the range of the dynamic adjustment coefficient.

Keywords: Data Assimilation System. Hybrid System. Ensemble Kalman Filter. Variational System. Numerical Weather Prediction.



## LISTA DE FIGURAS

	<u>Pág.</u>
3.1 Diagrama esquemático do Sistema de Assimilação de Dados Híbrido Médio.	16
3.2 Diagrama esquemático do Sistema de Assimilação de Dados Híbrido Médio utilizando o alfa dinâmico. . . . .	20
3.3 Distribuição das observações sintéticas. . . . .	23
3.4 Desvio padrão do erro das observações para o vento zonal, vento meridional e umidade. . . . .	24
4.1 Alfa médio na área da pressão em superfície ao longo do trimestre DJF. .	26
4.2 Distribuição espacial do alfa médio da pressão em superfície para o período de dezembro, janeiro e fevereiro. . . . .	27
4.3 Valores de alfa médio de temperatura para os sete níveis do modelo e para as quatro regiões de estudo: Globo, Hemisfério Norte, Hemisfério Sul e Trópicos, respectivamente, ao longo do trimestre DJF. . . . .	29
4.4 Idem à Figura 4.2, porém, para a temperatura nos sete níveis do modelo.	30
4.5 Idem à Figura 4.3, porém, para a componente zonal do vento. . . . .	31
4.6 Idem à Figura 4.4, porém, para a componente zonal do vento. . . . .	32
4.7 Idem à Figura 4.3, porém, para a componente meridional do vento. . . .	33
4.8 Idem à Figura 4.4, porém, para a componente meridional do vento. . . .	34
4.9 Idem à Figura 4.3, porém, para a umidade. . . . .	36
4.10 Idem à Figura 4.2, porém, para a umidade nos seis níveis do modelo. . .	37
4.11 Alfa médio na área da pressão em superfície ao longo do trimestre JJA. .	38
4.12 Distribuição espacial do alfa médio da pressão em superfície para o período de junho, julho e agosto. . . . .	39
4.13 Valores de alfa médio de temperatura para os sete níveis do modelo e para as quatro regiões de estudo: Globo, Hemisfério Norte, Hemisfério Sul e Trópicos, respectivamente, ao longo do trimestre JJA. . . . .	41
4.14 Idem à Figura 4.12, porém, para a temperatura nos sete níveis do modelo.	42
4.15 Idem à Figura 4.13, porém, para a componente zonal do vento. . . . .	43
4.16 Idem à Figura 4.14, porém, para a componente zonal do vento. . . . .	44
4.17 Idem à Figura 4.13, porém, para a componente meridional do vento. . . .	45
4.18 Idem à Figura 4.14, porém, para a componente meridional do vento. . . .	46
4.19 Idem à Figura 4.13, porém, para a umidade. . . . .	48
4.20 Idem à Figura 4.12, porém, para a umidade nos seis níveis do modelo. . .	49
4.21 Desvio padrão do alfa da pressão em superfície para o período de dezembro, janeiro e fevereiro. . . . .	50

4.22	Idem à Figura 4.21, porém, para a temperatura nos sete níveis do modelo.	52
4.23	Idem à Figura 4.21, porém, para o vento zonal nos sete níveis do modelo.	53
4.24	Idem à Figura 4.21, porém, para o vento meridional nos sete níveis do modelo. . . . .	54
4.25	Idem à Figura 4.21, porém, para a umidade nos seis níveis do modelo. . .	55
4.26	Desvio padrão do alfa da pressão em superfície para o período de junho, julho e agosto. . . . .	56
4.27	Idem à Figura 4.26, porém, para a temperatura nos sete níveis do modelo.	57
4.28	Idem à Figura 4.26, porém, para o vento zonal nos sete níveis do modelo.	59
4.29	Idem à Figura 4.26, porém, para o vento meridional nos sete níveis do modelo. . . . .	60
4.30	Idem à Figura 4.26, porém, para a umidade nos seis níveis do modelo. . .	61
5.1	Distribuição espacial do RMSE médio da pressão em superfície ao longo do trimestre DJF para as análises geradas pelo LETKF. . . . .	66
5.2	Distribuição espacial do RMSE médio da pressão em superfície ao longo do trimestre DJF para as análises geradas pelo 3DVar. . . . .	66
5.3	Idem à Figura 5.1, porém, para a temperatura nos sete níveis do modelo.	68
5.4	Idem à Figura 5.2, porém, para a temperatura nos sete níveis do modelo.	69
5.5	Idem à Figura 5.1, porém, para o vento zonal nos sete níveis do modelo.	70
5.6	Idem à Figura 5.2, porém, para o vento zonal nos sete níveis do modelo.	71
5.7	Idem à Figura 5.1, porém, para o vento meridional nos sete níveis do modelo. . . . .	72
5.8	Idem à Figura 5.2, porém, para o vento meridional nos sete níveis do modelo. . . . .	73
5.9	Idem à Figura 5.1, porém, para a umidade nos seis níveis do modelo. . .	74
5.10	Idem à Figura 5.2, porém, para a umidade nos seis níveis do modelo. . .	75
5.11	Distribuição espacial do RMSE médio da pressão em superfície ao longo do trimestre JJA para as análises geradas pelo LETKF. . . . .	78
5.12	Distribuição espacial do RMSE médio da pressão em superfície ao longo do trimestre JJA para as análises geradas pelo 3DVar. . . . .	78
5.13	Diferença nos valores de RMSE da pressão em superfície entre o LETKF e 3DVar para as previsões de até 120 horas pertencentes ao período de dezembro, janeiro e fevereiro. . . . .	80
5.14	Idem à Figura 5.13, porém, para o período de junho, julho e agosto. . . .	81
5.15	Perfil vertical de RMSE médio da umidade para quatro regiões de estudo: Globo, Trópicos, Hemisfério Norte e Hemisfério Sul. . . . .	85
5.16	Idem à Figura 5.15, porém, para a temperatura. . . . .	86
5.17	Idem à Figura 5.15, porém, para o vento zonal. . . . .	88

5.18	Idem à Figura 5.15, porém, para o vento meridional. . . . .	89
5.19	Idem à Figura 5.15, porém, para o período de junho, julho e agosto. . . . .	92
5.20	Idem à Figura 5.19, porém, para a temperatura. . . . .	93
5.21	Idem à Figura 5.19, porém para o vento zonal. . . . .	94
5.22	Idem à Figura 5.19, porém, para o vento meridional. . . . .	95
5.23	Média do erro absoluto do vento em 500hPa para os oito experimentos: alfa dinâmico, $\alpha = 0, 1, \alpha = 0, 3, \alpha = 0, 5, \alpha = 0, 7, \alpha = 0, 9$ , LETKF e 3DVar. . . . .	98
5.24	Média do vento ageostrófico em 500hPa para os oito experimentos: alfa dinâmico, $\alpha = 0, 1, \alpha = 0, 3, \alpha = 0, 5, \alpha = 0, 7, \alpha = 0, 9$ , LETKF e 3DVar. . . . .	99
5.25	Idem à Figura 5.23, porém, para o trimestre de JJA. . . . .	100
5.26	Idem à Figura 5.24, porém, para o trimestre JJA. . . . .	101
5.27	O-A da pressão em superfície no Globo, no Hemisfério Norte, no Hemis- fério Sul e na região Tropical, para o período de 1 de dezembro de 1982 a 28 de fevereiro de 1983. . . . .	103
5.28	Idem à Figura 5.27, porém, para os valores de O-F. . . . .	104
5.29	Idem à Figura 5.27, porém, para o período de junho, julho e agosto de 1982. . . . .	105
5.30	Idem à Figura 5.28, porém, para o período de junho, julho e agosto de 1982. . . . .	106
5.31	Ganho percentual nos valores de RMSE da pressão em superfície para as previsões de até 120 horas pertencentes ao período de dezembro, janeiro e fevereiro. . . . .	108
5.32	Distribuição espacial do ganho percentual do RMSE médio da pressão em superfície na análise e nas previsões de 24, 48, 72, 96 e 120 horas para o trimestre DJF. . . . .	110
5.33	Ganho percentual nos valores de RMSE da umidade para as previsões de até 120 horas pertencentes ao período de dezembro, janeiro e fevereiro. . . . .	113
5.34	Idem à Figura 5.33, porém para a temperatura. . . . .	114
5.35	Idem à Figura 5.34, porém, para a componente zonal do vento. . . . .	115
5.36	Idem à Figura 5.34, porém, para a componente meridional do vento. . . . .	116
5.37	Idem à Figura 5.31, porém, para o período de junho, julho e agosto. . . . .	117
5.38	Distribuição espacial do Ganho percentual do RMSE médio da pressão em superfície na análise e nas previsões de 24, 48, 72, 96 e 120 horas, para o trimestre JJA. . . . .	119
5.39	Idem à Figura 5.33, porém para o período de junho, julho e agosto. . . . .	120
5.40	Idem à Figura 5.39, porém para a temperatura. . . . .	122
5.41	Idem à Figura 5.40, porém para a componente zonal do vento. . . . .	123
5.42	Idem à Figura 5.40, porém para a componente meridional do vento. . . . .	124

5.43	Distribuição espacial do alfa médio da pressão em superfície para o período de dezembro, janeiro e fevereiro utilizando 40 membros. . . . .	125
5.44	Perfil vertical da diferença do RMSE médio da umidade para quatro regiões: Globo, Trópicos, Hemisfério Norte e Hemisfério Sul. . . . .	127
5.45	Idem à Figura 5.44, porém, para a temperatura. . . . .	128
5.46	Idem à Figura 5.44, porém, para a componente zonal do vento. . . . .	129
5.47	Idem à Figura 5.44, porém, para a componente meridional do vento. . . . .	130
5.48	Ganho percentual nos valores de RMSE da pressão em superfície para as previsões de até 120 horas pertencentes ao período de dezembro, janeiro e fevereiro. . . . .	131
5.49	Distribuição espacial do Ganho do RMSE médio da pressão em superfície na análise e nas previsões de 24, 48, 72, 96 e 120 horas, para o trimestre DJF, para 40 membros. . . . .	133
5.50	Perfil vertical da diferença do RMSE médio da umidade para quatro regiões: Globo, Trópicos, Hemisfério Norte e Hemisfério Sul. . . . .	136
5.51	Idem à Figura 5.50, porém, para a temperatura. . . . .	136
5.52	Idem à Figura 5.50, porém, para a componente zonal do vento. . . . .	137
5.53	Idem à Figura 5.50, porém, para a componente meridional do vento. . . . .	138
5.54	Ganho percentual nos valores de RMSE da pressão em superfície para as previsões de até 120 horas. . . . .	139

## LISTA DE TABELAS

	<u>Pág.</u>
3.1 Níveis verticais do Modelo SPEEDY. . . . .	14
5.1 RMSE médio da pressão em superfície ( $ps$ ), temperatura ( $T$ ), componente zonal e meridional do vento ( $u$ e $v$ ) e umidade ( $q$ ) para o Globo nos sete níveis do modelo para o trimestre DJF. . . . .	64
5.2 Idem à Tabela 5.1, porém, para o Hemisfério Norte. . . . .	64
5.3 Idem à Tabela 5.1, porém, para o Hemisfério Sul. . . . .	64
5.4 Idem à Tabela 5.1, porém, para a região Tropical. . . . .	65
5.5 RMSE médio da pressão em superfície ( $ps$ ), temperatura ( $T$ ), componente zonal e meridional do vento ( $u$ e $v$ ) e umidade ( $q$ ) para o Globo nos sete níveis do modelo para o trimestre JJA. . . . .	76
5.6 Idem à Tabela 5.5, porém, para o Hemisfério Norte. . . . .	76
5.7 Idem à Tabela 5.5, porém, para o Hemisfério Sul. . . . .	77
5.8 Idem à Tabela 5.5, porém, para a região Tropical. . . . .	77
5.9 RMSE médio da pressão em superfície para as quatro regiões: Globo, Trópicos, Hemisfério Norte e Hemisfério Sul, para os meses de dezembro, janeiro e fevereiro. . . . .	82
5.10 Idem à Tabela 5.9, porém, para os meses de junho, julho e agosto. . . . .	90
5.11 RMSE médio da pressão em superfície para as quatro regiões: Globo, Trópicos, Hemisfério Norte e Hemisfério Sul, para os meses de dezembro, janeiro e fevereiro utilizando 40 membros. . . . .	126
5.12 RMSE médio da pressão em superfície para as quatro regiões: Globo, Trópicos, Hemisfério Norte e Hemisfério Sul. . . . .	134



## LISTA DE ABREVIATURAS E SIGLAS

3DVar	–	<i>3-Dimensional Variational</i>
4DVar	–	<i>4-Dimensional Variational</i>
CPTEC	–	Centro de Previsão de Tempo e Estudos Climáticos
ECMWF	–	<i>European Centre for Medium-Range Weather Forecasts</i>
ENIAC	–	<i>Electronic Numerical Integrator and Computer</i>
EnKF	–	<i>Ensemble Kalman Filter</i>
EnSRF	–	<i>Ensemble Square Root Filter</i>
ETKF	–	<i>Ensemble Transform Kalman Filter</i>
GPSAS	–	<i>Global Physical-space Statistical Analysis System</i>
GSI	–	<i>Gridpoint Statistical Interpolation</i>
INPE	–	Instituto Nacional de Pesquisas Espaciais
IO	–	Interpolação Ótima
KF	–	<i>Kalman Filter</i>
LETKF	–	<i>Local Ensemble Transform Kalman Filter</i>
MCGA	–	Modelo de Circulação Geral da Atmosfera
NCEP	–	<i>National Centers for Environmental Predictions</i>
PNT	–	Previsão Numérica do Tempo
PSAS	–	<i>Physical-space Statistical Analysis System</i>
RMSE	–	<i>Root Mean Square Error</i>
RPSAS	–	<i>Regional Physical-space Statistical Analysis System</i>
SPEEDY	–	<i>Simplified Parameterizations, primitive-Equation DYNamics</i>
VCAN	–	Vórtice Ciclônico de Altos Níveis
ZCIT	–	Zona de Convergência Intertropical



## SUMÁRIO

	<u>Pág.</u>
<b>1 INTRODUÇÃO</b> . . . . .	<b>1</b>
1.1 Motivação . . . . .	3
1.1.1 Objetivo Geral . . . . .	3
1.2 Estrutura do Documento . . . . .	4
<b>2 FUNDAMENTAÇÃO TEÓRICA</b> . . . . .	<b>5</b>
2.1 Assimilação de Dados . . . . .	5
2.2 Sistemas de Assimilação de Dados Variacionais . . . . .	6
2.3 Sistemas de Assimilação de Dados por Conjunto . . . . .	8
2.4 Sistemas de Assimilação de Dados Híbridos . . . . .	10
<b>3 MATERIAIS E METODOLOGIA</b> . . . . .	<b>13</b>
3.1 Modelo SPEEDY . . . . .	13
3.2 Os sistemas de assimilação de dados . . . . .	14
3.2.1 O sistema Híbrido Médio . . . . .	14
3.2.2 O sistema variacional . . . . .	17
3.2.3 O sistema baseado em conjunto . . . . .	17
3.3 Inclusão do coeficiente de ajuste dinâmico . . . . .	18
3.4 Experimentos . . . . .	20
<b>4 O ALFA DINÂMICO</b> . . . . .	<b>25</b>
4.1 Estrutura média do Alfa Dinâmico . . . . .	25
4.1.1 Alfa médio para o trimestre DJF . . . . .	25
4.1.2 Alfa médio para o trimestre JJA . . . . .	37
4.2 Variação do Alfa Dinâmico . . . . .	49
4.2.1 Desvio Padrão do Alfa Dinâmico para o trimestre DJF . . . . .	50
4.2.2 Desvio Padrão do Alfa Dinâmico para o trimestre JJA . . . . .	55
<b>5 RESULTADOS</b> . . . . .	<b>63</b>
5.1 Performance do LETKF e do 3DVar . . . . .	63
5.2 Avaliação das análises do Híbrido Médio . . . . .	81
5.3 Verificação do balanço das análises . . . . .	97
5.4 Comparação entre os coeficientes dinâmico e estático . . . . .	101

5.4.1	Avaliação das análises e estimativa inicial contra as observações . . . .	102
5.4.2	Avaliação das Previsões de até 120 horas . . . . .	107
5.5	Sensibilidade ao número de membros . . . . .	124
5.5.1	Estrutura média do Alfa Dinâmico . . . . .	125
5.5.2	Impacto do número de membros nas análises . . . . .	126
5.5.3	Impacto do número de membros nas previsões . . . . .	131
5.6	Otimização do intervalo do alfa dinâmico . . . . .	134
5.6.1	Erros nas análises . . . . .	134
5.6.2	Erros nas previsões . . . . .	138
<b>6</b>	<b>CONCLUSÕES . . . . .</b>	<b>141</b>
6.1	Trabalhos futuros . . . . .	146
	<b>REFERÊNCIAS BIBLIOGRÁFICAS . . . . .</b>	<b>147</b>
	<b>APÊNDICE A - EQUIVALÊNCIA DA ALTERAÇÃO NA MATRIZ</b>	
	<b>GANHO E O ALGORITMO DO HÍBRIDO MÉDIO . . . . .</b>	<b>153</b>

# 1 INTRODUÇÃO

A assimilação de dados teve início na década de 1950, com a finalidade de gerar condições iniciais balanceadas para os modelos de Previsão Numérica do Tempo (PNT). Ela nada mais é do que uma série de correções que são feitas em previsões de curto prazo, vindas de um modelo de PNT utilizando dados observacionais, levando em consideração o conhecimento dos erros de modelagem e observacionais. Entre as primeiras técnicas que surgiram, cita-se o Método de Correções Sucessivas e o Método de Interpolação Ótima, que apesar de mais simplificadas quando comparadas as atuais representam um avanço na PNT, mais precisamente em assimilação de dados. À medida em que os modelos foram evoluindo e que foram sendo criados novos sistemas de observação, surgiu a necessidade de aprimorar os métodos de assimilação de dados para que eles pudessem acompanhar esse crescimento. Dessa forma, os métodos variacionais e métodos baseados em conjunto foram sendo desenvolvidos. Atualmente, as técnicas mais utilizadas em centros operacionais ao redor do mundo são o 3DVar e o 4DVar, dentro da classe variacional, e o EnKF (*Ensemble Kalman Filter*), no grupo dos métodos por conjunto. Uma terceira categoria seriam os sistemas híbridos, que são uma combinação entre dois sistemas de assimilação de dados, mais comumente um sistema variacional e um sistema baseado em conjunto.

No CPTEC/INPE (Centro de Previsão de Tempo e Estudos Climáticos/Instituto Nacional de Pesquisas Espaciais), os esforços dedicados para a construção de sua própria condição inicial começaram na década de 1990, um ano após a criação do Centro, quando, em meados de 1995 foi implementado o Método de Interpolação Ótima (GANDIN, 1965; JULIAN; THIEBAUX, 1975), onde eram apenas assimilados dados convencionais. Este sistema foi implementado conjuntamente com o Modelo de Circulação Geral da Atmosfera (MCGA) do CPTEC/INPE na resolução T062L28 (com resolução espacial de aproximadamente 200km na horizontal e 28 níveis verticais).

Em 1997, começou a implementação do sistema PSAS (do inglês *Physical-space Statistical Analysis System*) (SILVA et al., 1994) em sua versão regional, sendo chamado então de RPSAS (CINTRA; ARAVÉQUIA, 1998). Esse sistema foi acoplado ao modelo Eta 40km e utilizava como condição de contorno o MCGA do CPTEC/INPE na resolução T062L28. Em 1999, esse sistema já estava operacional no Centro. No ano seguinte começaram os trabalhos no PSAS em sua versão global (GPSAS) no MCGA do CPTEC/INPE na resolução TQ126L28 (HERDIES et al., 2002), o que representa aproximadamente 100km de resolução horizontal. Tal sistema passou a ser

operacional em 2003. Esse foi um passo muito importante para a independência do Centro em relação às condições iniciais, pois, para gerar previsões de tempo, não dependeria mais de análises geradas em outros centros. Nele, além das observações convencionais, passou-se a assimilar dados de *retrievals* do ATOVS (Andreoli et al. (2008)). O PSAS, é um sistema 3DVar que resolve a equação da análise no espaço das observações.

O PSAS se manteve como sistema operacional até o ano de 2013, quando foi substituído pelo GSI (do inglês *Gridpoint Statistical Interpolation*) (WU et al., 2002; KLEIST et al., 2009), o qual foi implementado conjuntamente com o MCGA do CPTEC/INPE na configuração T299L64 (AZEVEDO et al., 2017). Durante esses 14 anos utilizando o PSAS, muitos foram os avanços na assimilação de dados do CPTEC/INPE. Houve a inclusão de dados SATEM e SATOB (CINTRA, 2005), além de uma mudança de resolução do GPSAS, passando para T213L42, sendo aproximadamente 63km na horizontal e 42 níveis verticais, também houve mudança na condição de contorno utilizada pelo RPSAS, além de testes assimilando aerossóis e GPS. Nesse período, ainda começaram os trabalhos com o sistema LETKF (do inglês *Local Ensemble Transform Kalman Filter*) (HUNT et al., 2007) e também com assimilação de dados de superfície, ambos sendo utilizados em modo pesquisa.

Conforme mencionado anteriormente, a comunidade científica tem investido esforços em sistemas de assimilação de dados Híbrido. Comumente, é uma combinação entre um sistema de assimilação de dados variacional e um baseado em conjunto. A abordagem mais antiga desse sistema é de 2000 (HAMILL; SNYDER, 2000). Os autores realizaram uma combinação linear entre as matrizes de covariância dos erros do *background* desses dois sistemas, de modo a incluir em sistemas de variacionais os "erros do dia" vindos da matriz de covariância do sistema por conjunto. Recentemente, no CPTEC/INPE, Bastarz (2017) implementou essa metodologia de sistema Híbrido no Modelo de Circulação Geral da Atmosfera em baixa resolução.

Uma abordagem mais recente, o Híbrido Médio, proposta por Penny (2014), visa melhorar os resultados de um sistema baseado em conjunto através de sistemas variacionais. Nela, não há a necessidade de um novo cálculo para a matriz de covariância dos erros do *background*. Informações climatológicas são inseridas em um sistema por conjunto através da combinação das mesmas com sistemas variacionais. Esse sistema possui uma grande vantagem, que é ser de fácil implementação e a atualização desse sistema é feita à medida que os sistemas envolvidos forem sendo melhorados. Ele permite o desenvolvimento em paralelo dos sistemas, e a medida

que esses desenvolvimentos avancem, o desenvolvimento do Híbrido Médio avança também. Basicamente, essa abordagem realiza uma combinação linear entre duas análises: a análise média vinda do sistema por conjunto e análise vinda de um sistema variacional o qual utilizou como seu *background* a análise média do sistema por conjunto. A contribuição de cada uma dessas análises é feita por meio de um coeficiente de ajuste que varia entre zero e um.

O Híbrido Médio foi utilizado inicialmente no Modelo Lorenz-96, com o LETKF e o 3DVar. Em 2015 foi implementado por Bonavita et al. (2015) no ECMWF (*European Centre for Medium-Range Weather Forecasts*) utilizando o modelo operacional do centro e os sistemas EnKF e o 4DVar. De acordo com os autores, essa nova abordagem é funcionalmente equivalente ao Híbrido tradicional, e eles ainda listaram algumas vantagens desse método, como a fácil implementação, por exemplo. Essa metodologia também trouxe bons resultados quando utilizada em modelos oceânicos (PENNY et al., 2015).

## 1.1 Motivação

Devido as suas diferentes metodologias, os erros do LETKF e do 3DVar são estruturados de forma distinta, conseqüentemente, cada um destes sistemas apresenta uma variabilidade espacial e temporal de erros diferente. Dentro de um sistema baseado em conjunto, o seu grau de incerteza pode ser descrito pelo espalhamento. Essa informação (medida da incerteza pelo espalhamento), não foi completamente explorada dentro do Híbrido Médio proposto por Penny (2014). Em seu estudo, ele realiza apenas uma combinação linear entre as duas análises, não levando em consideração a magnitude dos erros de cada análise para diferentes regiões.

Diante dessa nova abordagem, e a partir do apresentado por Penny (2014), onde a combinação linear entre as análises geradas por diferentes sistemas de assimilação de dados (leia-se um variacional e outro baseado em conjunto) gere uma análise melhor do que as produzidas pelos sistemas de maneira isolada. Dessa forma, a proposta dessa tese visa testar a hipótese de que existiria um ganho em adicionar o espalhamento na forma do coeficiente de ajuste da combinação linear.

### 1.1.1 Objetivo Geral

Incluir em um sistema de assimilação de dados Híbrido Médio um coeficiente de ajuste dinâmico cujo peso leva em consideração o espalhamento do conjunto. Para a obtenção do resultado proposto como objetivo geral dessa tese, algumas etapas

foram listadas e são apresentadas abaixo:

- a) Análise da performance dos sistemas de assimilação de dados envolvidos no Híbrido Médio;
- b) Identificação do valor ótimo para coeficiente de ajuste estático da combinação linear para o Híbrido Médio nas análises;
- c) Inclusão do coeficiente de ajuste dinâmico no sistema Híbrido Médio;
- d) Avaliação da estabilidade do sistema Híbrido Médio utilizando o coeficiente de ajuste dinâmico;
- e) Avaliação do ganho desse novo coeficiente nas análises e previsões de até 120 horas.

## **1.2 Estrutura do Documento**

Na seção seguinte é feita uma breve fundamentação teórica, onde apresenta-se motivações para o uso de assimilação de dados, assim como o uso de sistemas de assimilação de dados variacionais, por conjunto e sistemas híbridos. No capítulo 3, a metodologia empregada é apresentada, bem como uma descrição dos sistemas de assimilação de dados que são envolvidos na construção do sistema híbrido proposto e os experimentos realizados. No capítulo seguinte é apresentada a estrutura do coeficiente de ajuste dinâmico. No capítulo 5 são apresentados os resultados dessa tese e, em seguida, no capítulo 6, as conclusões desse trabalho, bem como as sugestões de trabalhos futuros.

## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1 Assimilação de Dados

A previsão de tempo teve início no começo do século 20, quando Vilhelm Bjerknes (1904), citado por Kalnay (2003), propôs formulações matemáticas para descrever os movimentos atmosféricos, com isso, podendo realizar previsões de tempo. Essas equações eram baseadas em problemas de física e de dinâmica de fluidos. Sua proposta encorajou outros pesquisadores como meteorologistas e físicos a estudarem os distúrbios meteorológicos.

Em 1922, Lewis Fry Richardson (RICHARDSON, 1922) realizou cálculos à mão a fim de realizar uma previsão de tempo baseada em sistemas dinâmicos. Infelizmente, ele não obteve sucesso com seus resultados. Ele encontrou uma variação brusca no campo de pressão em apenas 6 horas de previsão (aproximadamente 146hPa/6h). Um dos problemas encontrados por ele foi o fato das condições iniciais terem sido obtidas a partir de análises subjetivas dos campos meteorológicos. Foi então que, em 1940, com o advento de um computador digital eletrônico, o ENIAC (do inglês *Electronic Numerical Integrator and Computer*), um grupo interdisciplinar de pesquisadores começou a trabalhar na Previsão Numérica de Tempo (PNT) em um contexto computacional e objetivo.

No ano de 1950, Charney, Fjortoft e von Neuman (CHARNEY et al., 1950) reportavam com sucesso a realização da primeira previsão numérica de tempo. Para esse resultado, Panofsky (1949) desenvolveu um algoritmo de análises objetivas baseado em ajustamento de observações aos pontos de grade. Desde então, os esforços foram maiores nessa área, de modo que computadores mais rápidos foram desenvolvidos, bem como a criação de novos sensores e instrumentos, o que aumentou significativamente o número de observações a serem assimiladas. Outras melhorias também aconteceram no âmbito de modelagem e assimilação de dados.

Dessa forma, a assimilação de dados visa corrigir uma previsão de curto prazo (chamado de *background*) originada a partir de uma integração de um modelo de PNT utilizando dados observacionais, com a finalidade de gerar uma condição inicial (análise) apropriada para os modelos. Tal correção é feita levando em consideração os erros do modelo e os erros das observações. A análise é uma representação do estado atual da atmosfera em um determinado instante.

A expressão matemática utilizada na elaboração desta análise é representada pela

Equação 2.1, onde  $\mathbf{x}^a$  é o campo de análise,  $\mathbf{x}^b$  é o campo de *background*,  $\mathbf{K}$  é a matriz peso (chamada de matriz Ganho de Kalman),  $\mathbf{y}$  é o vetor de observações e  $H$  é o operador observação (da maneira como está escrito, ele leva do espaço do modelo para o espaço das observações, para que a diferença entre observado e modelado (inovação) possa ser calculada).

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}[\mathbf{y} - H(\mathbf{x}^b)] \quad (2.1)$$

O cálculo da matriz Ganho de Kalman é dado pela Equação 2.2, onde  $\mathbf{B}$  é a matriz de covariância dos erros do *background*,  $\mathbf{R}$  é a matriz de covariância dos erros das observações e  $\mathbf{H}$  é a matriz com diferenças entre as observações e o *background*. A matriz  $\mathbf{B}$  para alguns sistemas é fixa, enquanto, para outros, ela varia de acordo com os "erros do dia".

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \quad (2.2)$$

De acordo com Talagrand (1997) a assimilação de dados tem por finalidade reunir o maior número possível de informações a fim de obter uma melhor estimativa do estado atual da atmosfera, e assim utilizando-a como condição inicial para modelos de PNT. O propósito final é alcançar a melhor análise e conseqüentemente a melhor previsão.

Atualmente, os métodos mais utilizados pelos centros operacionais são mais sofisticados quando comparados aos primeiros que surgiram como, por exemplo, os métodos variacionais 3DVar e 4DVar, métodos por conjunto como EnKF e ainda os sistemas híbridos. Esses sistemas serão discutidos com maiores detalhes nas seções seguintes.

## 2.2 Sistemas de Assimilação de Dados Variacionais

A abordagem variacional da assimilação de dados é um problema no qual se deseja determinar a menor distância entre dados modelados e dados observados. Esses sistemas surgiram devido à necessidade de considerar os erros de modelagem e também de se utilizar as observações assintóticas. Eles incluem informações dos erros do *background* (matriz de covariância dos erros do *background*) de forma simplificada. Nela, assume-se conhecer como os erros são distribuídos, e ainda sua média e variância.

Os sistemas variacionais mais conhecidos são o 3DVar e o 4DVar. Para o 3DVar o objetivo desse sistema é minimizar uma função custo (Equação 2.3), que nada mais é do que a função que mede a distância entre o observado e o modelado, levando em consideração os erros do modelo e das observações. Nela,  $\mathbf{x}^a$  é o campo de análise,  $\mathbf{x}^b$  é o campo de *background*,  $\mathbf{y}$  é o vetor de observações,  $\mathbf{B}$  é a matriz de covariância

dos erros do *background*,  $\mathbf{R}$  é a matriz de covariância dos erros das observações e  $H$  é o operador observação. Em um sistema 3DVar, tenta-se encontrar um valor para  $\mathbf{x}^a$  que resulte em um menor valor para a função  $J$ . Para isso, é calculado o gradiente da função (Equação 2.4) de modo iterativo.

$$J(\mathbf{x}^a) = (\mathbf{x}^a - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x}^a - \mathbf{x}^b) + (\mathbf{y} - H(\mathbf{x}^a))^T \mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{x}^a)) \quad (2.3)$$

$$\nabla J(\mathbf{x}^a) = 2\mathbf{B}^{-1}(\mathbf{x}^a - \mathbf{x}^b) - 2\mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{x}^a)) \quad (2.4)$$

Esse sistema se torna interessante, pois é de fácil implementação (precisa apenas da função custo e do cálculo do seu gradiente), considerações externas podem ser inseridas (*constraints*), além de assimilar diferentes tipos de observações sinóticas e assinóticas. Contudo, a matriz precisa ser bem definida para todos os pontos de grade do modelo, além dela ser fixa.

A diferença principal entre o 3DVar e o 3DVar-PSAS é basicamente o espaço onde a análise é calculada. Em um 3DVar simples, a análise é calculada no espaço do modelo. Já no PSAS, ela é calculada no espaço das observações. Por esse motivo, se torna difícil a utilização de dados não-convencionais, como por exemplo, dados de radiâncias pelo PSAS.

O 4DVar por sua vez, é uma generalização do 3DVar para as observações que estão distribuídas irregularmente no tempo. Utiliza-se as mesmas equações do 3DVar, porém, o operador observação é generalizado para incluir o tempo, permitindo comparações entre os estados do modelo e as observações no seu tempo, conforme apresentado na Equação 2.5. Nela, o índice  $i$  representa o tempo em que qualquer observação está disponível.

$$J(\mathbf{x}^a) = (\mathbf{x}^a - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x}^a - \mathbf{x}^b) + \sum_{i=0}^n (\mathbf{y}_i - H_i(\mathbf{x}^a_i))^T \mathbf{R}_i^{-1}(\mathbf{y}_i - H_i(\mathbf{x}^a_i)) \quad (2.5)$$

A primeira motivação para a criação de um sistema de assimilação de dados 4DVar foi a necessidade de se utilizar dados assinóticos em seu devido tempo. Ele é um sistema complexo, envolvendo latitude, longitude, altura e tempo, e ainda apresenta um custo computacional alto, pois necessita um modelo adjunto e um tangente linear. O modelo adjunto é utilizado para obter os erros do modelo. De acordo com [Errico \(1997\)](#) modelos adjuntos são poderosas ferramentas para vários estudos de sensibilidade de uma saída de um modelo com relação aos dados utilizados por ele.

Para maiores informações sobre os sistemas variacionais, pode-se encontrar em [Bout-](#)

tier e Courtier (1999). Nesse artigo, os autores fazem uma compilação de diversos materiais sobre os métodos variacionais.

### 2.3 Sistemas de Assimilação de Dados por Conjunto

A abordagem baseada em conjunto surgiu devido a fortes incertezas e aproximações que se tornam difíceis para serem inseridas nos modelos numéricos, fazendo com que os métodos probabilísticos ganhassem espaço. Entre os diversos métodos baseados em conjunto, destaca-se o EnKF e suas variações. A ideia principal é executar o sistema de análise várias vezes para uma determinada data. Cada vez utilizando campos de *background* que diferem por uma característica do erro do *background* e observações que foram perturbadas pela adição de um ruído aleatório. Os campos de *background* são produzidos pela execução de previsões curtas a partir de cada membro do conjunto anterior de análise.

Os métodos por conjunto mais utilizados na previsão numérica do tempo são os métodos baseados em Filtragem de Kalman. O Filtro de Kalman (KF) (KALMAN, 1960) é considerado um estimador recursivo, pois só precisa do estado no tempo passado e a medição atual para se obter a estimativa do estado atual. Esse método é basicamente composto de duas etapas distintas: a predição, que usa a estimativa do estado no tempo anterior para obter uma estimativa no tempo atual; e a fase de atualização, na qual a predição feita no tempo anterior é combinada com as observações atuais para melhorar a estimativa atual, essa estimativa é chamada de estimativa a posteriori. Essa metodologia pode ser muito custosa computacionalmente e é aplicada apenas para sistemas lineares. Em decorrência desses problemas ele não é utilizado operacionalmente na PNT. A partir da criação desse método, variações dele começaram a surgir como por exemplo o *Ensemble Kalman Filter* (EnKF) (EVENSEN, 1994), o *Ensemble Transform Kalman Filter* (ETKF) (BISHOP et al., 2001), o *Ensemble Square Root Filter* (EnSRF) (WHITAKER; HAMILL, 2002) e o *Local Ensemble Transform Kalman Filter* (LETKF).

Os sistemas de filtragem por conjunto como o EnKF, o ETKF, o EnSRF e o LETKF foram desenvolvidos para que pudessem ser utilizados em sistemas não-lineares. Suas formulações são mais eficientes computacionalmente e são mais adequados para serem utilizados em computadores massivamente paralelos (por exemplo, com o uso do MPI - *Message Parcing Interface*), tornando-os mais viáveis para o uso em modo operacional por centros de PNT.

O *Ensemble Kalman Filter* (EnKF) foi inicialmente proposto na década de 1990 por

Evensen (1994) como forma de lidar com os sistemas não-lineares e agora é muito utilizado para aplicações realistas com modelos de equações primitivas para oceano e atmosfera. Evensen (2003) apresenta uma revisão de toda a evolução do EnKF até o ano de 2003 e ainda comenta que esse sistema ganhou popularidade por ter formulações simples e de fácil implementação, pois não precisa de tangente linear ou adjunto e tampouco integrações para trás no tempo. Em contraste com o KF, a ideia principal do EnKF é que um conjunto, o qual é uma amostra estatística dos estados estimados, é usado para calcular o estado da análise para cada membro do conjunto. A partir de um conjunto de previsões, a média da amostra e as covariâncias dos erros são calculados. Durante o estágio de análise, esses valores são utilizados para calcular um único ganho de Kalman, que é usado para assimilar as observações e produzir um conjunto de estados de análise.

Em 2001, surgiu o ETKF. Ele foi proposto por Bishop et al. (2001) e modificado em 2004 por Wang et al. (2004). Uma das vantagens desse sistema (computacionalmente mais rápido que o EnSRF) é que ele é capaz de calcular rapidamente a covariância do erro do *background*. O equacionamento desses sistemas (EnKF, ETKF e EnSRF) é bem apresentado em Petrie (2008). No ano de 2004, o parâmetro de localização foi inserido nesses sistemas. O LEKF (*Local Ensemble Kalman Filter*) foi desenvolvido por Ott et al. (2004). Nele, a análise é calculada localmente no ponto de grade, usando as variáveis do estado do modelo e todas as observações em uma região centrada em cada ponto de grade, tornando esse sistema massivamente paralelizável.

No ano de 2007, Hunt et al. (2007) lançavam o LETKF. Esse sistema foi desenvolvido pelo Grupo *Weather and Chaos* da *University of Maryland*. O LETKF, desde sua criação, foi testado utilizando observações sintéticas em um cenário de modelo perfeito (HARLIM; HUNT, 2005; LIU et al., 2008) e utilizando observações reais (LI et al., 2007; MIYOSHI; SATO, 2007). Szunyogh et al. (2008) e Whitaker et al. (2008) assimilaram dados convencionais em um modelo de baixa resolução do NCEP. Araújo et al. (2011) avaliou estratégias para assimilar dados de radiancias utilizando esse sistema.

A principal vantagem do LETKF em relação aos demais sistemas apresentados até aqui é a combinação da localização com a transformação aplicada ao processo de assimilação de dados. Ele ainda é computacionalmente mais eficiente e fornece uma flexibilidade na localização que é necessária para a assimilação de radiancias de satélite. Ele resolve as equações localmente no espaço do modelo, dividindo o domínio do modelo em domínios locais, dessa forma, a equação da análise é calculada para

cada um desses domínios, sendo essa uma vantagem para poder utilizar computação massivamente em paralelo. Outra vantagem é que o LETKF, diferente de alguns sistemas que usam a matriz de covariância do erro do *background* fixa, calcula a covariância do erro do *background* usando o conjunto de estados do modelo e atualiza a covariância em cada ciclo de assimilação. Por esse motivo, ele pode representar melhor os "erros do dia" dentro do sistema, permitindo uma solução mais adequada. Esses erros representam uma das principais fontes de incertezas do processo de modelagem como um todo.

## 2.4 Sistemas de Assimilação de Dados Híbridos

A primeira abordagem de sistema híbrido se deu com Hamill e Snyder (2000). Os autores propuseram um sistema baseado no EnKF e no 3DVar utilizando um modelo quasi-geostrófico perfeito. Os autores calcularam uma nova matriz de covariância dos erros do *background* ( $\mathbf{B}$ ) utilizando a Equação 2.6. Nela,  $\alpha$  é o peso atribuído aos sistemas ( $0 \leq \alpha \leq 1$ ),  $\mathbf{P}^b$  é a matriz de covariância dos erros do *background* do EnKF e  $\mathbf{B}_{3DVAR}$  é a matriz de covariância dos erros do *background* do 3DVar.

$$\mathbf{B} = (1 - \alpha)\mathbf{P}^b + \alpha\mathbf{B}_{3DVAR} \quad (2.6)$$

Essa abordagem visava incluir os "erros do dia" vindos de sistemas por conjunto em sistemas de assimilação de dados variacionais, através da combinação das matrizes de covariância dos erros do *background*, fazendo com que a matriz  $\mathbf{B}$  do 3DVar se tornasse mais robusta.

A partir dessa primeira abordagem de sistema híbrido, outros pesquisadores passaram a utiliza-la de forma igual, com outros sistemas de assimilação de dados, alterando a Equação 2.6 ou ainda variando os modelos. Em 2004, por exemplo, Etherton e Bishop (2004) realizaram um híbrido utilizando o ETKF e o 3DVar. Wang et al. (2007) utilizaram para a formação do sistema Híbrido o ETKF e IO. Zhang et al. (2009) realizaram o chamado E4DVar através do EnKF e um 4DVar.

Trabalhos mais recentes como os de Wang et al. (2013) e Clayton et al. (2013) apresentam os sistemas híbridos utilizados no NCEP e no MetOffice, respectivamente. No CPTEC, essa metodologia de sistema Híbrido proposta por Hamill e Snyder (2000) foi utilizada por Bastarz (2017). Em seu trabalho, o autor calculou uma nova matriz de covariância do *background* utilizando o EnKF, EnSRF e o 3DVar no Modelo de Circulação Geral da Atmosfera (MCGA) do CPTEC/INPE na resolução TQ0062L028. De acordo com o autor, os resultados obtidos mostram que utilizando

a matriz de covariâncias híbrida, as análises produzidas pelo sistema permitem que o modelo de previsão utilizado desempenhe melhor a previsão de diversas variáveis relacionadas ao aspecto físico e dinâmico da atmosfera.

Esses trabalhos mencionados até então são uma abordagem de sistemas de assimilação de dados híbridos que visam a criação de uma nova matriz de covariância de erros do *background* baseada em outras duas. Porém uma nova classe de híbrido vem sendo trabalhada desde meados de 2014. Penny (2014) sugeriu uma nova metodologia na qual a ideia principal é de utilizar a análise média gerada por um sistema por conjunto como o *background* utilizado por um sistema variacional e por fim realizar uma combinação linear entre ambas análises. Penny (2014) utilizou o LETKF e o 3DVar para a formação do Sistema Híbrido no modelo Lorenz-96. De acordo com o autor, sistemas híbridos tradicionais começam com uma abordagem variacional e incorporam a ele informações de sistemas por conjunto através da matriz de covariância, já sua proposta de sistema híbrido começa com um EnKF e usa a abordagem variacional para realizar correções no espaço do modelo para estabilizar o EnKF. Ou seja, enquanto as outras abordagens visam melhorar um sistema variacional com a informação da covariância de um conjunto, essa abordagem visa melhorar um sistema baseado em conjunto com informações de um 3DVar.

O Híbrido Médio visa resolver a combinação linear apresentada na Equação 2.7, onde  $\mathbf{x}^a_{HM}$  é a análise final,  $\bar{\mathbf{x}}^a_{LETKF}$  é a análise média vinda do sistema de assimilação de dados por conjunto,  $\mathbf{x}^a_{3DVar}$  é a análise vinda do sistema variacional utilizando o  $\bar{\mathbf{x}}^a_{LETKF}$  como seu *background*. Por fim, o  $\alpha$  que é o coeficiente de ajuste linear ( $0 \leq \alpha \leq 1$ ).

$$\mathbf{x}^a_{HM} = (1 - \alpha)\bar{\mathbf{x}}^a_{LETKF} + \alpha\mathbf{x}^a_{3DVar} \quad (2.7)$$

Penny (2014) utilizou  $\alpha = 0,25$ ,  $\alpha = 0,5$  e  $\alpha = 0,75$ , e concluiu que esse valor depende das especificações do problema, como por exemplo, o tamanho do conjunto, a cobertura das observações e ainda o raio de localização.

Essa metodologia foi implementada no início de 2015 no ECMWF por Bonavita et al. (2015). Os autores utilizaram como base para o Híbrido Médio o EnKF e o 4DVar. Os resultados mostraram que o Híbrido Médio é competitivo com o 4DVar-EDA em baixa resolução, que hoje é utilizado operacionalmente no centro. O conjunto de dados assimilados pelos sistemas foi diferente, no EnKF não foram utilizadas apenas radiâncias não afetadas por nuvens. Os autores ainda levantaram a questão desse sistema possuir uma independência com relação ao seu desenvolvimento, pois

ambos os sistemas envolvidos podem ser desenvolvidos em paralelo, e ainda assim beneficiar o Híbrido Médio. Como trabalhos futuros, os autores intencionam tornar o  $\alpha$  um valor dependente da escala, de modo que escalas espaciais pequenas sejam beneficiadas pelo EnKF e escalas grandes pelo 4DVar. A grande vantagem que os autores viram em utilizar esse sistema foi o baixo custo computacional, sendo este determinado basicamente pelo sistema por conjunto.

Penny et al. (2015) e Penny (2017) encontraram ótimos resultados em sistemas oceânicos. Os autores implementaram o Híbrido Médio no Sistema de Assimilação de Dados Global Oceânico do NCEP, o qual chamaram de Hybrid-GODAS. Nele, foram utilizados o 3DVar e o LETKF, e foi implementado no modelo oceânico MOM4 (Modular Ocean Model). Os autores encontraram uma redução do RMSE e do viés na salinidade e temperatura quando comparados aos sistemas 3DVar e LETKF sozinhos.

Mais recentemente, Wespetal et al. (2018) implementaram o Híbrido Médio no modelo SPEEDY. Os autores estavam interessados em encontrar a melhor estatística para a matriz de covariância do erro do 3DVar. Eles compararam o método NMC com um método que utiliza as perturbações de um conjunto de previsões em torno da média do conjunto como os valores para o erro de previsão.<sup>1</sup>

Os sistemas híbridos apresentam-se então como uma ferramenta que combina vantagens de dois sistemas de assimilação de dados. Eles tendem a apresentar uma análise melhor do que cada um dos sistemas de assimilação de dados envolvidos quando utilizados de maneira isolada.

---

<sup>1</sup>Será descrito em maiores detalhes no Capítulo 3.

### 3 MATERIAIS E METODOLOGIA

Neste capítulo são apresentadas as ferramentas utilizadas para o desenvolvimento deste trabalho, assim como a metodologia que foi empregada. Na primeira Seção é feita uma apresentação do modelo SPPEEDY e suas configurações. Na sequência são apresentados os sistemas de assimilação de dados que compõem o trabalho, a começar pelo sistema Híbrido Médio, seguido pelo 3DVar e pelo LETKF, ambos dentro do contexto do Híbrido Médio. Na Seção 3.3, descreve-se como foi feita a inclusão do coeficiente de ajuste dinâmico, o qual é o ponto chave deste trabalho. Por fim, é mostrada a configuração dos experimentos realizados, incluindo nela informações dos dados utilizados.

#### 3.1 Modelo SPEEDY

O modelo de circulação geral da atmosfera SPEEDY (do inglês *Simplified Parameterizations, primitive-Equation Dynamics*) foi criado em 2003 por Molteni (2003) com parametrizações físicas simplificadas que se mostraram computacionalmente eficientes. Ele mantém as mesmas características básicas de um modelo "estado da arte" com uma física complexa. Foi adaptado por Miyoshi (2005) para experimentos de assimilação de dados.

A versão do modelo SPEEDY utilizada possui resolução T30L7 (com truncamento espectral horizontal no número de onda 30 e 7 níveis verticais) a qual corresponde a uma grade de 96x48x7 e coordenada vertical sigma. Este modelo resolve as equações primitivas para cinco variáveis prognósticas: vento zonal ( $u$ ), vento meridional ( $v$ ), temperatura ( $T$ ), umidade específica ( $q$ ) e pressão em superfície ( $ps$ ). As entradas do modelo são todas em níveis sigma, porém as saídas podem ser vistas em níveis sigma ou em níveis de pressão. As alturas específicas de cada nível são apresentadas na Tabela 3.1, onde na primeira coluna consta o índice do nível, na segunda coluna os valores em coordenada sigma e por fim, na última coluna, seu equivalente em coordenadas de pressão.

As condições de contorno do modelo SPEEDY incluem a altura topográfica e máscara de terra-mar, as quais são constantes, Temperatura da Superfície do Mar (TSM), fração de mar e gelo, temperatura da superfície na primeira camada do solo, umidade na primeira camada do solo e na camada de zona de raiz, e profundidade de neve. Todas estas informações apresentam-se através de médias mensais, exceto albedo da superfície e fração de superfície coberta por vegetação, que são apresentados como campos de média anual. As condições de contorno inferior, como

a TSM, são obtidas à partir da reanálise do ECMWF no período 1981-1990, O fluxo de entrada de radiação solar e as condições de contorno restantes, exceto o albedo de superfície e a fração de vegetação, são atualizados diariamente.

Tabela 3.1 - Níveis verticais do Modelo SPEEDY.

Índice do Nível	Altura Sigma ( $\sigma$ )	Nível de Pressão (hPa)
1	0,950	925
2	0,835	850
3	0,685	700
4	0,510	500
5	0,340	300
6	0,200	200
7	0,080	100

Fonte: Adaptado de Miyoshi (2005).

O SPEEDY inclui componentes básicos de parametrizações físicas que também são utilizadas em modelos globais mais complexos, como convecção, condensação de larga escala, nuvens, radiação de onda curta, radiação de onda longa, fluxos de superfície e difusão vertical. Maiores informações sobre o modelo podem ser obtidas na página do *ICTP Atmospheric General Circulation Model SPEEDY-Net* (<http://users.ictp.it/~kucharsk/speedy-net.html>), para detalhes sobre os esquemas de parametrizações físicas simplificadas pode-se recorrer ao Apêndice desse mesmo link.

## 3.2 Os sistemas de assimilação de dados

### 3.2.1 O sistema Híbrido Médio

O sistema de assimilação de dados que foi utilizado no presente trabalho foi o sistema Híbrido Médio proposto por Penny (2014) e que foi implementado por Wespetal et al. (2018) no modelo SPEEDY. Esse sistema visa utilizar como *background* de um sistema variacional a análise média de um sistema por conjunto, e por fim combinar ambas análises: a gerada pelo sistema por conjunto e a gerada pelo sistema variacional que utilizou como *background* o membro médio. Para esta tese, foi utilizado o LETKF como sistema por conjunto e o 3DVar como sistema variacional, ambos serão brevemente descritos na sequência. A combinação deles se da conforme

a Equação 3.1:

$$\mathbf{x}_{HM}^a = (1 - \alpha)\bar{\mathbf{x}}_{LETKF}^a + \alpha\mathbf{x}_{3DVar}^a \quad (3.1)$$

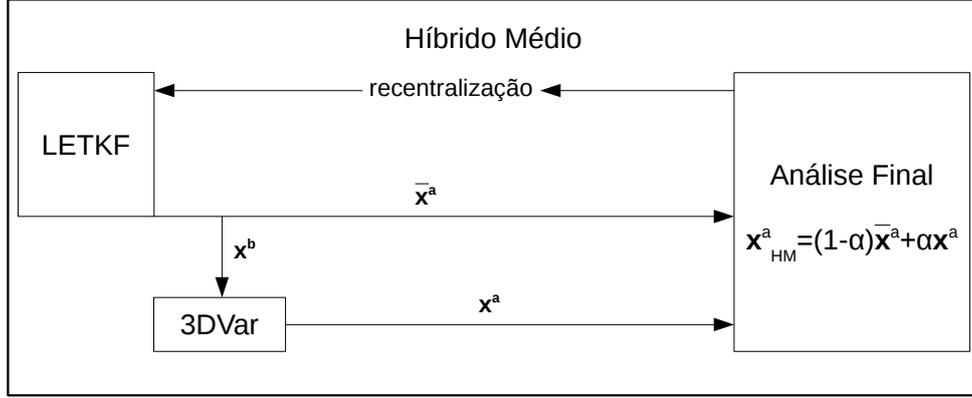
onde:

- $\mathbf{x}_{HM}^a$ : é a análise final do sistema Híbrido Médio;
- $\bar{\mathbf{x}}_{LETKF}^a$ : é a análise média do LETKF;
- $\mathbf{x}_{3DVar}^a$ : é a análise do 3DVar utilizando o membro médio da análise do LETKF como *background*;
- $\alpha$ : é o coeficiente de ajuste que determina o peso que cada sistema terá na análise final variando entre 0 e 1.

De acordo com a Equação 3.1, quando os valores de alfa forem próximos de zero, significa que houve uma contribuição maior do LETKF na geração da análise do Híbrido Médio, quando  $\alpha = 0$ , tem-se o resultado da análise gerada apenas pelo LETKF. Quando os valores de alfa forem próximos de 1, significa que a análise do 3DVar contribuiu mais na geração da análise final e, diferentemente de quando utiliza-se  $\alpha = 0$ , quando se utiliza  $\alpha = 1$ , a análise gerada pelo Híbrido Médio utilizou aproximadamente 100% de 3DVar. Como o 3DVar dentro do Híbrido Médio utiliza como seu *background* a análise média do LETKF, ele é diferente de quando se utiliza apenas o 3DVar sem passar pelo Híbrido.

Na Figura 3.1 é mostrado um diagrama esquemático do sistema Híbrido Médio para elucidar a Equação 3.1. Nele mostra como a análise final do Híbrido Médio ( $\mathbf{x}_{HM}^a$ ) é gerada. Para isso, inicialmente é calculada a análise média gerada pelo LETKF ( $\bar{\mathbf{x}}^a$ ). Uma vez tendo essa análise, ela é reservada para a composição da análise final, e também servirá de *background* para uma realização do 3DVar. Dessa forma, o 3DVar gera uma segunda análise (a partir da análise média do LETKF) que é também encaminhada para compor o resultado (análise) final. O valor de  $\alpha$ , que representa o peso de cada análise na combinação linear, é determinado de forma subjetiva dentro dos *namelists* do sistema e mantendo-se constante para todos os ciclos subsequentes. A análise do Híbrido Médio é gerada então com a combinação linear de ambas análises: LETKF e 3DVar. Para inicializar o LETKF no passo seguinte de assimilação, essa análise é somada às perturbações geradas anteriormente pelo LETKF, repetindo-se os passos anteriores sucessivamente.

Figura 3.1 - Diagrama esquemático do Sistema de Assimilação de Dados Híbrido Médio.



Fonte: Produção do autor.

Outra forma de entender esse sistema é do ponto de vista da matriz de ganho. A matriz de ganho, também chamada de Ganho de Kalman, visa pesar a inovação de observações ( $\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}^b$ ), ou seja, determina o quanto que o *background* será corrigido pelas observações. A matriz de ganho do Híbrido Médio ( $\widehat{\mathbf{K}}$ ) é uma combinação entre as matrizes de ganho do 3DVar ( $\mathbf{K}^B$ ) e do LETKF ( $\mathbf{K}$ ), conforme a Equação 3.2.

$$\widehat{\mathbf{K}} = \beta_1 \mathbf{K} + \beta_2 \mathbf{K}^B + \beta_3 \mathbf{K}^B \mathbf{H} \mathbf{K} \quad (3.2)$$

De acordo com Penny (2014), quando utilizado  $\beta_1 = 1$ ,  $\beta_2 = \alpha$  e  $\beta_3 = -\alpha$ , é utilizada a análise média do sistema por conjunto como *background* do sistema variacional, já quando se escolhe  $\beta_1 = (1 - \alpha)$ ,  $\beta_2 = \alpha$  e  $\beta_3 = 0$  o sistema variacional é alimentado pelo membro médio das previsões do conjunto<sup>1</sup>.

Atribuindo os valores apresentados de  $\beta$ , pode-se escrever a matriz de ganho do sistema Híbrido Médio da seguinte forma:

$$\widehat{\mathbf{K}} = \mathbf{K} + \alpha \mathbf{K}^B - \alpha \mathbf{K}^B \mathbf{H} \mathbf{K} \quad (3.3)$$

$$\widehat{\mathbf{K}} = \mathbf{K} + \alpha \mathbf{K}^B (\mathbf{I} - \mathbf{H} \mathbf{K}) \quad (3.4)$$

<sup>1</sup>Aqui será tratada apenas a primeira combinação apresentada. Para maiores informações sobre a segunda combinação pode-se procurar no Apêndice B de Penny (2014).

Dessa forma, temos a equação da análise para o Híbrido Médio apresentada em 3.5.<sup>2</sup>

$$\mathbf{x}_{HM}^a = \bar{\mathbf{x}}^b + \widehat{\mathbf{K}}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b) \quad (3.5)$$

### 3.2.2 O sistema variacional

Como sistema de assimilação de dados variacional foi utilizado o 3DVar de Barker et al. (2004), com uma climatologia de erro de previsão ajustado para SPEEDY desenvolvida por Wespetal et al. (2018). Esse novo cálculo utiliza um conjunto de previsões de uma reanálise do SPEEDY-LETKF como amostras de erros.

O 3DVar dentro do Híbrido Médio visa minimizar a função custo apresentada na Equação 3.6,

$$J(\mathbf{x}^a) = (\mathbf{x}^a - \bar{\mathbf{x}}^a)^T \mathbf{B}^{-1} (\mathbf{x}^a - \bar{\mathbf{x}}^a) + (\mathbf{y} - H(\mathbf{x}^a))^T \mathbf{R}^{-1} (\mathbf{y} - H(\mathbf{x}^a)) \quad (3.6)$$

onde:

- $\mathbf{x}^a$ : representa os campos de análise do 3DVar;
- $\bar{\mathbf{x}}^a$ : que é a análise média do conjunto vinda do LETKF;
- $\mathbf{B}$ : é a matriz de covariância do erro de *background*;
- $\mathbf{y}$ : é o vetor de observações;
- $H$ : é o operador observação não linear, que é responsável por transformar as variáveis do espaço do modelo para o espaço das observações;
- $\mathbf{R}$ : é a matriz de covariância do erro das observações.

Diferentemente da função custo de um sistema de assimilação de dados variacional tradicional apresentada na Equação 2.3, no Híbrido Médio o 3DVar é alimentado pela análise média do LETKF, ou seja, onde era  $\mathbf{x}^b$  na equação tradicional, no Híbrido Médio é  $\bar{\mathbf{x}}^a$ .

### 3.2.3 O sistema baseado em conjunto

O sistema baseado em conjunto utilizado foi o LETKF. Esse sistema foi desenvolvido pelo grupo *Weather and Chaos* da Universidade de Maryland (UMD) por Hunt et

---

<sup>2</sup>As demonstrações para a obtenção da equação 3.5 são apresentadas no Apêndice A.

al. (2007). O cálculo do LETKF é feito aqui como proposto por Hunt et al. (2007). Primeiro é calculada a covariância do erro da análise no espaço de perturbação do conjunto (Equação 3.7) onde  $\mathbf{Y}^b = \mathbf{H}(\mathbf{X}^b)$  e  $\mathbf{X}^b$  são as perturbações do conjunto,  $\rho$  é o parâmetro de inflação local,  $k$  é o tamanho do conjunto e  $\mathbf{R}$  é a matriz de covariância do erro das observações.

$$\tilde{\mathbf{P}}^a = \left[ \frac{(k-1)}{\rho} \mathbf{I} + (\mathbf{Y}^b)^T \mathbf{R}^{-1} (\mathbf{Y}^b) \right]^{-1} \quad (3.7)$$

Os pesos atribuídos para cada membro da análise são dados pela Equação 3.8.

$$\mathbf{W}^a = [(k-1)\tilde{\mathbf{P}}^a]^{1/2} \quad (3.8)$$

Para transformar do espaço do conjunto para o espaço do modelo, cada membro do conjunto é multiplicado pelos pesos de perturbações do *background*.

$$\mathbf{X}^a = \mathbf{X}^b \mathbf{W}^a \quad (3.9)$$

As perturbações de análise calculadas na Equação 3.9 são adicionadas a diferentes estimativas de média do conjunto, e a análise final é obtida pela Equação 3.10, onde  $\bar{\mathbf{x}}^b$  é o campo de *background* médio.

$$\bar{\mathbf{x}}^a = \mathbf{X}^b [\tilde{\mathbf{P}}^a (\mathbf{Y}^b)^T \mathbf{R}^{-1} (\mathbf{y}^o - \bar{\mathbf{y}}^b)] + \bar{\mathbf{x}}^b \quad (3.10)$$

Foram utilizados 16 e 40 membros, uma inflação da covariância do erro Multiplicativa e, inflação do espalhamento Aditiva. Essas configurações foram utilizadas conforme recomendações já vindas no sistema.

### 3.3 Inclusão do coeficiente de ajuste dinâmico

O coeficiente de ajuste dinâmico proposto, tem a finalidade de ponderar a combinação entre as análises originadas pelo LETKF e pelo 3DVar, no contexto do Híbrido Médio, utilizando um critério objetivo. Verifica-se, a partir da distribuição espacial (e temporal) dos erros de ambos sistemas, que a qualidade das condições iniciais também varia espacialmente e temporalmente. Consequentemente, um peso constante e arbitrário não necessariamente representa a combinação ótima entre as duas análises. Este trabalho de tese parte da hipótese de que, o espalhamento do LETKF é um bom estimador das incertezas do modelo numérico em um dado ciclo de análise tornando-se também uma potencial base para ponderação entre as duas análises.

Esse novo coeficiente, que é uma matriz com as mesmas dimensões e variáveis do modelo, e é ajustada a cada novo horário de análise, substituiu o alfa fixo, que na metodologia proposta por Penny (2014) é um escalar escolhido de forma subjetiva. Por ter se tornado uma matriz de alfa que varia a cada novo horário, a partir de agora será chamado de alfa dinâmico. Esse novo coeficiente consiste em normalizar os campos de espalhamento vindos do LETKF para cada variável ( $u$ ,  $v$ ,  $T$ ,  $q$  e  $ps$ ), ponto de grade e nível do modelo, com valores entre 0 e 1, e em seguida aplicar esses novos campos na Equação 3.1.

De acordo com Houtekamer e Zhang (2016), quanto menor é o tamanho do conjunto usado e quanto maior é o seu espalhamento, mais baixa será a qualidade da análise. Por esse motivo, quanto menor for o espalhamento, menor será o valor de alfa. Dessa forma, ao utilizarmos o alfa dinâmico, quando o espalhamento for baixo, irá favorecer o LETKF, onde o espalhamento for alto, favorecerá a análise variacional.

Esse processo pode ser entendido através das Equações 3.11 e 3.12, ou então associado com o diagrama esquemático apresentado na Figura 3.2. De acordo com Miyoshi (2005), o espalhamento do conjunto representa o desvio padrão dos erros das análises e das previsões. Além disso, Lima (2018) coloca que o espalhamento pode ser representado pelo desvio padrão das perturbações em relação a média do conjunto. Dessa forma, a Equação 3.11 calcula o desvio padrão ( $\sigma_{\mathbf{x}}$ ), utilizando as informações vindas das Equações 3.9 e 3.10. Em seguida, esse valor é passado na forma de espalhamento para a Equação 3.12, onde é minimizado para se transformar em alfa.

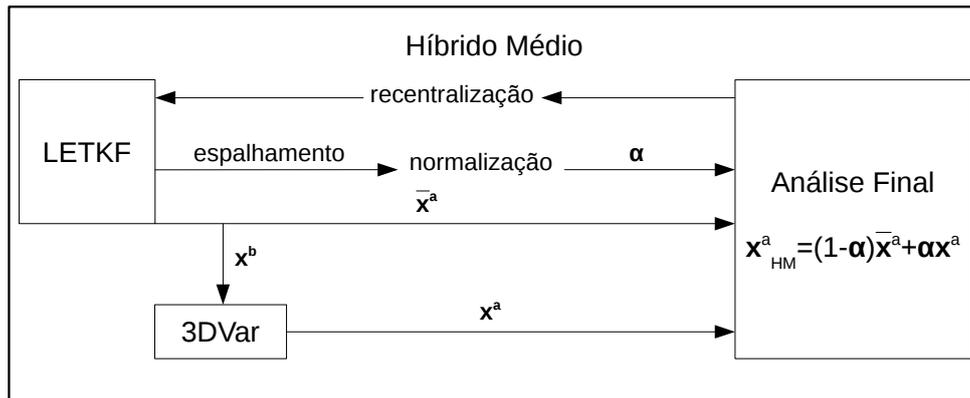
$$\sigma_{\mathbf{x}} = \sqrt{\frac{1}{k} \sum_{i=1}^k (\mathbf{X}^a - \bar{\mathbf{x}}^a)^2} \quad (3.11)$$

$$\alpha = \frac{\sigma_{\mathbf{x}} - \min(\sigma_{\mathbf{x}})}{\max(\sigma_{\mathbf{x}}) - \min(\sigma_{\mathbf{x}})} \quad (3.12)$$

O diagrama esquemático (Figura 3.2) segue o mesmo padrão do diagrama apresentado na Figura 3.1. Nele mostra o ciclo do Híbrido Médio, agora, com a inclusão do novo coeficiente. A análise média que é gerada pelo LETKF, assim com na Figura 3.1, é usada para compor a análise final do Híbrido Médio, bem como para alimentar o 3DVar, que por sua vez, gera sua análise variacional e essa também é encaminhada para compor a análise final. Além disso, como produto do LETKF, existe o espalhamento que no Híbrido Médio tradicional não foi utilizado. Aqui, ele é normalizado e entregue na forma do alfa dinâmico para contribuir de forma objetiva para a composição da análise final. A partir dessas três informações, a análise do Híbrido Médio

é gerada pela Equação 3.1. Para que o passo seguinte comece, a análise do Híbrido Médio é somada às perturbações para ser re-centralizada e alimentar o LETKF no próximo horário.

Figura 3.2 - Diagrama esquemático do Sistema de Assimilação de Dados Híbrido Médio utilizando o alfa dinâmico.



Fonte: Produção do autor.

### 3.4 Experimentos

O experimentos dessa tese fora divididos em dois grupos, sendo o primeiro com os experimentos principais e o segundo com os experimentos adicionais. Os experimentos do primeiro grupo foram divididos em três etapas para um melhor desenvolvimento do trabalho. A primeira etapa foi de comparação entre os sistemas de assimilação de dados envolvidos no Híbrido Médio: LETKF<sup>3</sup> e 3DVar<sup>4</sup>, a segunda visando encontrar o valor ótimo de alfa nas análises para que o resultado deste seja comparado com o passo seguinte, e por fim o experimento utilizando o alfa dinâmico. Estes oito experimentos foram realizados utilizando um conjunto com 16 membros, sendo sumarizados aqui:

- experimento utilizando apenas o LETKF;
- experimento utilizando o 3DVar;
- experimento com o Híbrido Médio, utilizando o  $\alpha = 0, 1$ ;

<sup>3</sup>Sendo o mesmo que  $\alpha = 0, 0$

<sup>4</sup>Execução do 3DVar sem passar pelo Híbrido Médio, ou seja, diferente de  $\alpha = 1, 0$

- experimento com o Híbrido Médio, utilizando o  $\alpha = 0,3$ ;
- experimento com o Híbrido Médio, utilizando o  $\alpha = 0,5$ ;
- experimento com o Híbrido Médio, utilizando o  $\alpha = 0,7$ ;
- experimento com o Híbrido Médio, utilizando o  $\alpha = 0,9$ ;
- experimento com o Híbrido Médio, utilizando o alfa dinâmico.

Para cada um dos experimentos apresentados, foram geradas análises para dois períodos: inverno austral (compreendendo de 1 de maio de 1982 a 31 de agosto de 1982) e verão austral (compreendendo de 1 de novembro de 1982 até 28 de fevereiro de 1983). Para ambos os casos o primeiro mês foi excluído para fins de *spin up*. Foram então avaliados os meses de junho, julho e agosto de 1982 (para o caso de inverno) e dezembro de 1982, janeiro e fevereiro de 1983 (para o caso de verão). Além das análises, foram geradas também previsões de até 120 horas a partir do experimento utilizando o alfa ótimo e utilizando o alfa dinâmico.

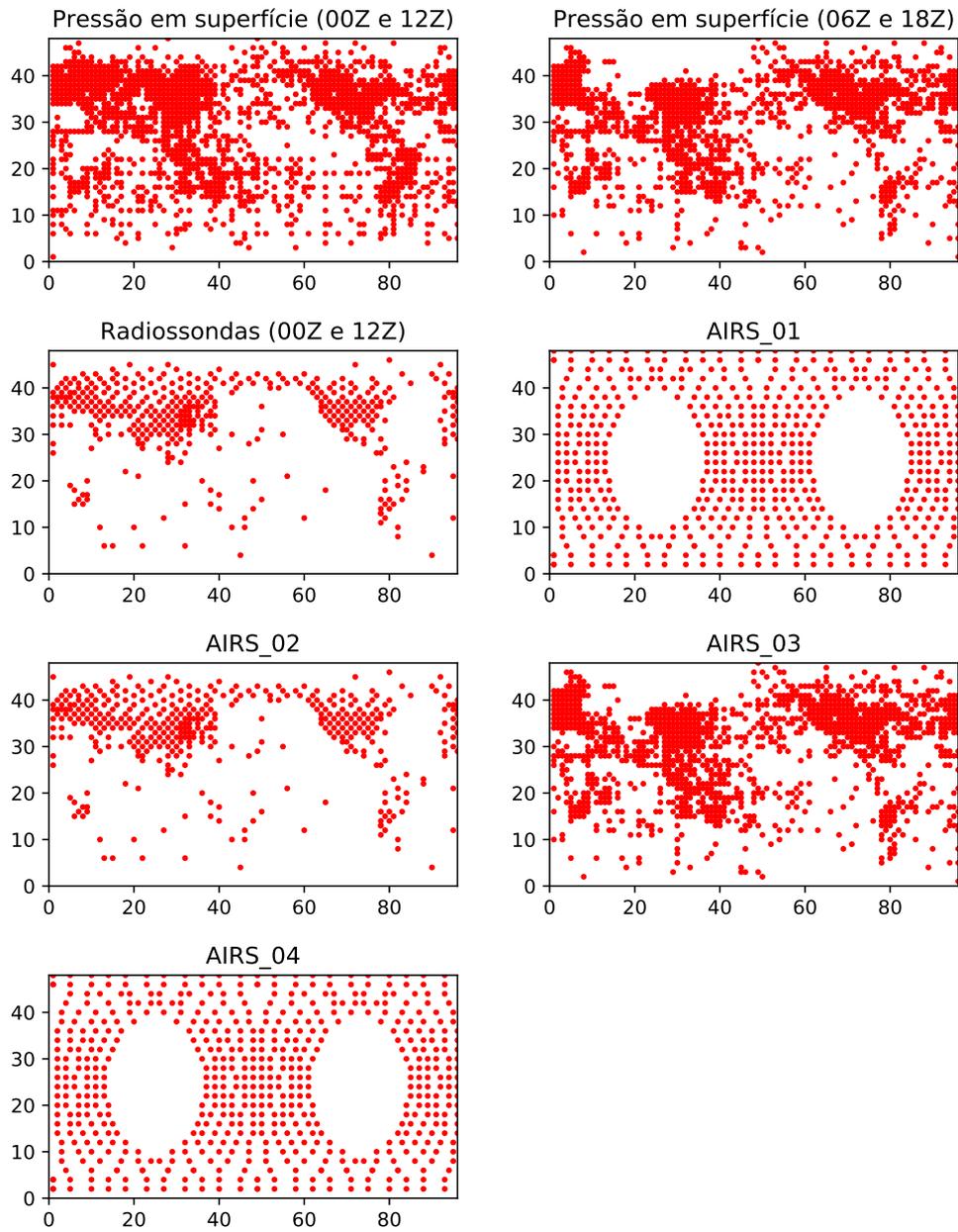
Três experimentos adicionais foram realizados, sendo ambos para o período de verão austral. Esses experimentos tem a finalidade de testar a sensibilidade do sistema Híbrido Médio ao aumento no número de membros e também a sensibilidade ao restringir o intervalo de alfa a valores menores. Para ambos experimentos, foram geradas análises e suas respectivas previsões para até 120 horas. Estes experimentos estão sumarizados aqui:

- experimento com o Híbrido Médio, utilizando o  $\alpha = 0,1$ , com 40 membros;
- experimento com o Híbrido Médio, utilizando o alfa dinâmico, com 40 membros;
- experimento com o Híbrido Médio, utilizando o alfa dinâmico com seu intervalo de valores alterados, utilizando 16 membros.

Os experimentos mencionados aqui utilizaram neles observações sintéticas retiradas a partir de uma rodada livre do modelo (*Nature Run*). Essas observações visam simular observações reais como observações de radiossondas, satélite (AIRS) e estações de superfície. A distribuição espacial delas é mostrada na Figura 3.3. Os eixos são mostrados em pontos de grade, sendo o eixo  $x$  até 96 pontos e o eixo  $y$  até 48 pontos. Nela, é possível notar que as observações não estão distribuídas uniformemente no

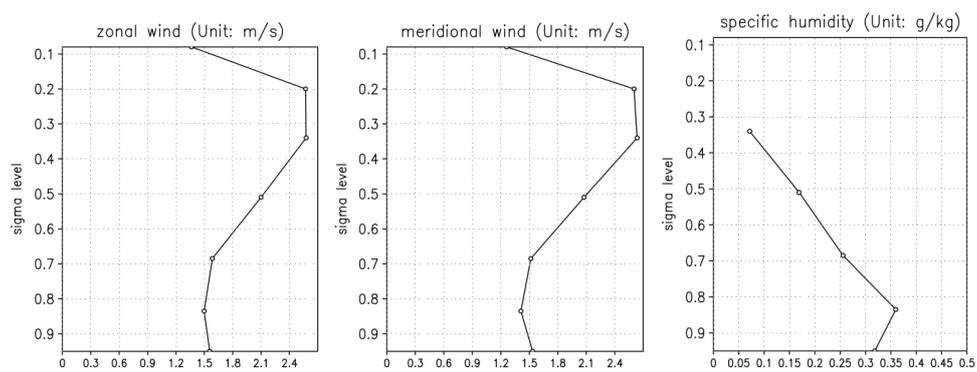
globo, sendo compatíveis com a distribuição real destas observações. Elas contém dados de pressão em superfície (para os 4 horários sinóticos), perfis de temperatura, umidade e vento (zonal e meridional). As observações ditas de radiossondas são apenas das 00Z e das 12Z. Tais observações seguem conforme feito em [Liu et al. \(2009\)](#). Segundo os autores, as observações foram criadas através da adição de um ruído aleatório gaussiano na *Nature Run*. Os desvios padrões que foram assumidos para os campos de vento e umidade específica foram cerca de 30% da variabilidade normal para essas variáveis, e são mostradas na Figura 3.4. Foi assumido um desvio padrão de 0,8K em todos os níveis verticais de temperatura. O desvio padrão assumido ao erro da pressão em superfície foi de 1 hPa.

Figura 3.3 - Distribuição das observações sintéticas.



Fonte: Produção do autor.

Figura 3.4 - Desvio padrão do erro das observações para o vento zonal, vento meridional e umidade.



Fonte: Liu et al. (2009)

## 4 O ALFA DINÂMICO

O alfa dinâmico é uma matriz de mesmas dimensões do modelo, possuindo 7 níveis verticais, exceto para a pressão em superfície que possui apenas um nível e para umidade, que possui seis níveis, e 5 variáveis, sendo elas  $ps$ ,  $T$ ,  $u$ ,  $v$  e  $q$ . Para cada nova análise, é gerada uma nova matriz de alfa. Diante disso, esse Capítulo é destinado ao alfa dinâmico e suas características. Aqui são apresentadas as estruturas do alfa dinâmico para os experimentos que utilizaram 16 membros e um intervalo entre 0 e 1.

### 4.1 Estrutura média do Alfa Dinâmico

Nessa seção, são apresentados valores médios para os períodos de estudo (verão e inverno) bem como seus valores médios na área. Os valores médios foram calculados conforme a Equação 4.1, onde  $n$  é o número total de valores e  $x_i$  é cada valor onde  $i$  varia de 1 até  $n$ .

$$m\acute{e}dia = \frac{1}{n} \sum_{i=1}^n x_i \quad (4.1)$$

Para a série temporal da média do alfa dinâmico são avaliados o Globo, Hemisfério Norte (compreendendo de 20°N a 80°N), Hemisfério Sul (compreendendo de 80°S a 20°S) e região Tropical (compreendendo de 20°S a 20°N).

#### 4.1.1 Alfa médio para o trimestre DJF

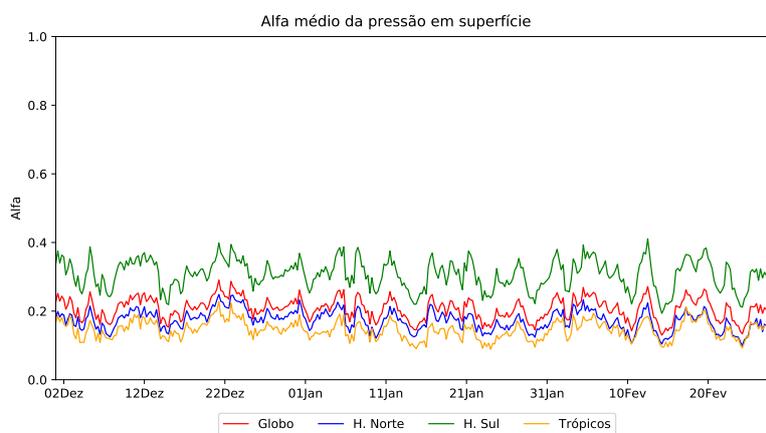
Nas Figuras 4.1, 4.3, 4.5, 4.7 e 4.9, são apresentados os valores médios na área que foram atribuídos ao alfa dinâmico para as 4 regiões de estudo: Globo, Hemisfério Norte, Hemisfério Sul e Trópicos, ao longo do período de estudo: dezembro, janeiro e fevereiro. Já nas Figuras 4.2, 4.4, 4.6, 4.8 e 4.10, são apresentadas as distribuições espaciais dos alfas médios no período, para a pressão em superfície, temperatura, componente zonal e meridional do vento e umidade, respectivamente. Exceto a figura da pressão em superfície, as demais apresentam a distribuição espacial para cada nível do modelo. Nessas figuras, quanto mais azul, menor o valor do alfa e por consequência, maior é o peso dado a análise gerada pelo LETKF na geração da análise final do Híbrido Médio. Por outro lado, quanto mais vermelho, maior é o valor do alfa e consequentemente maior é a contribuição do 3DVar na composição da análise final do Híbrido Médio.

A Figura 4.1 mostra os valores médios de alfa para de pressão em superfície para as 4 regiões: em vermelho para o Globo, em azul para o Hemisfério Norte, em verde para o Hemisfério Sul e em laranja para os Trópicos. O eixo  $x$  apresenta o período

de estudo e o eixo  $y$  os valores médios de alfa. Pode-se observar que para todo o período os valores que prevaleceram de alfa foram valores abaixo de 0,5, ou seja, de acordo com esses valores de alfa, infere-se que para a formação da análise final do sistema Híbrido Médio, as análises geradas pelo LETKF contribuíram mais do que as geradas pelo 3DVar. Além disso, pode-se observar que para o Hemisfério Sul esses valores foram maiores (não ultrapassando 0,4). Na Figura 4.2, pode-se ver essa região onde o alfa médio é maior. Na faixa entre 60°S e 35°S, tem-se valores médios para o período acima de 0,5. Para essa faixa de latitude, o 3DVar teve um peso maior na contribuição da análise final do Híbrido Médio. A curva verde, referente ao Hemisfério Sul, na Figura 4.1, só não foi acima do 0,5 pois no gráficos de médias ao longo do tempo a faixa que compreende o Hemisfério Sul é entre 80°S e 20°S.

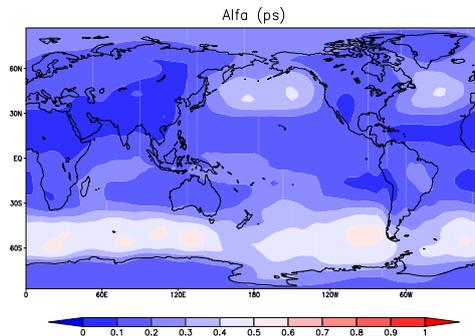
Um ponto importante a ser lembrado aqui é que o espalhamento reflete o grau de incerteza do sistema baseado em conjunto, ou seja, possíveis regiões onde a confiabilidade é maior ou menor. O mesmo é visto para os valores de alfa já que esse é um reflexo do espalhamento. Em regiões com valores de alfa baixos (próximo do zero) utilizam mais o LETKF e em regiões com alfa alto (próximos de 1) utilizam mais o 3DVar. Na Figura 4.2, pode-se associar os valores de alfa mais elevados às regiões onde se observam os ciclones. No caso do Hemisfério Sul, é a região dos ciclones transientes que atuam durante todo o ano, enquanto no Hemisfério Norte, sendo inverno nessa região, os maiores valores de alfa ficaram localizados nas regiões de atuação de frentes. Isso estar relacionado com uma baixa destreza dos sistemas e/ou do modelo em simular determinados sistemas meteorológicos.

Figura 4.1 - Alfa médio na área da pressão em superfície ao longo do trimestre DJF.



Fonte: Produção do autor.

Figura 4.2 - Distribuição espacial do alfa médio da pressão em superfície para o período de dezembro, janeiro e fevereiro.



Fonte: Produção do autor.

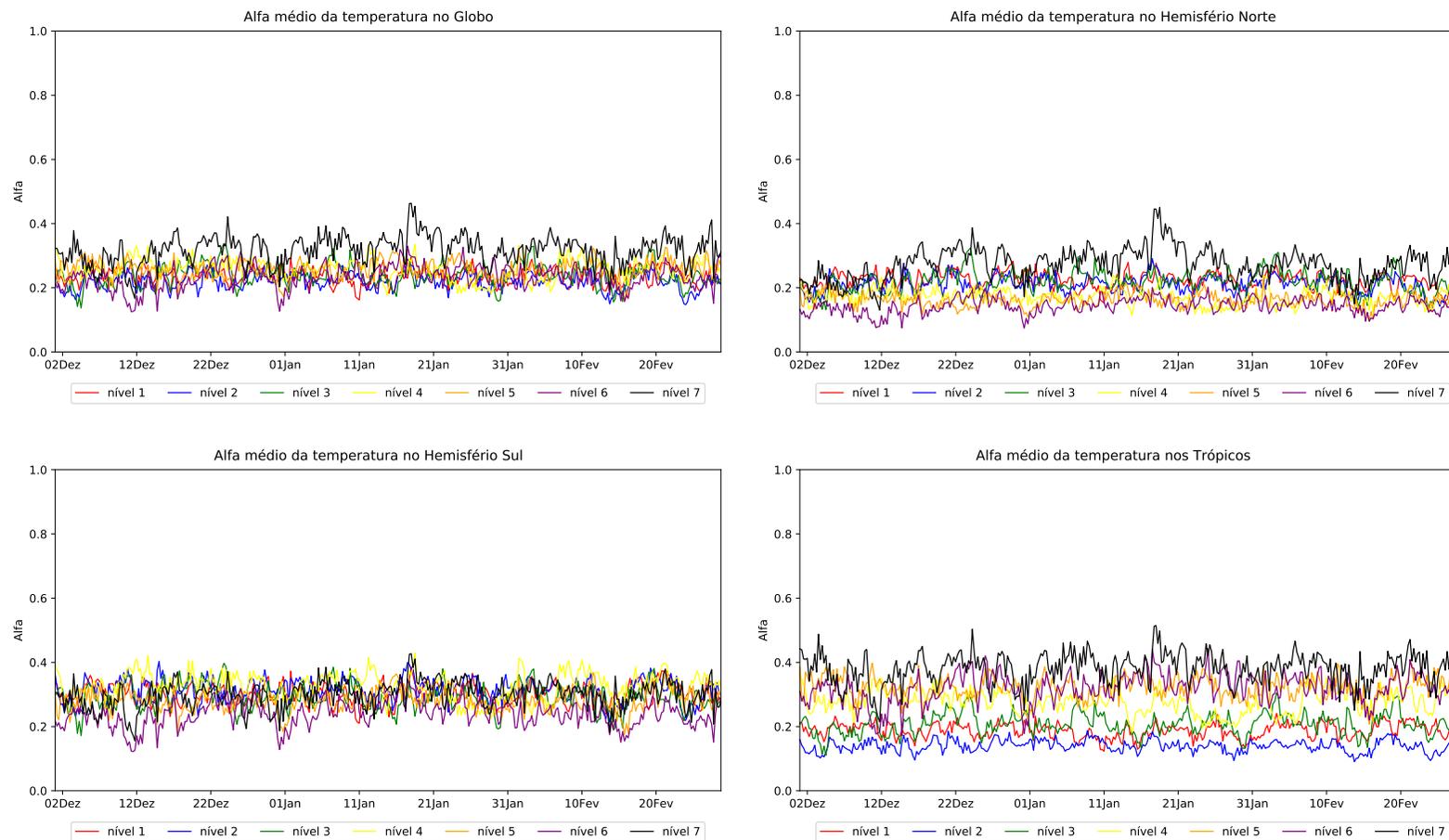
As Figuras 4.3, 4.5, 4.7 e 4.9, diferentemente da anterior, apresentam cada região de forma separada e contendo os valores médios na área para cada um dos sete níveis do modelo, sendo: em vermelho o nível 1, em azul o nível 2, em verde o nível 3, em amarelo o nível 4, em laranja o nível 5, em roxo o nível 6 e em preto o nível 7. Na Figura 4.3, são apresentados os valores referentes à temperatura. Nelas, nota-se que os valores são um pouco maiores do que os apresentados para o alfa da pressão em superfície, principalmente para valores nos Trópicos. Nessa região, consegue-se observar uma distinção nos valores, os quais giram em torno de 0,2 para os níveis 1, 2 e 3, e à medida que o nível aumenta, os valores de alfa também aumentam, alcançando 0,5 em alguns casos.

Observando a figura da distribuição espacial (Figura 4.4), nota-se a partir do quarto nível do modelo, que na região Tropical já começam a aparecer valores acima de 0,5 (manchas avermelhadas) e à medida que vai subindo na atmosfera, os valores vão crescendo, chegando a valores entre 0,7 e 0,8 sobre o Oceano Pacífico Leste no sétimo nível. No Hemisfério Sul são mais homogêneos com relação aos níveis, apresentando valores de alfas para os 7 níveis mais próximos, mantendo-se, na maioria dos horários, entre a faixa de 0,2 e 0,4. No caso do Hemisfério Norte, não se observou regiões onde o alfa foi acima de 0,5, mostrando a grande influência do LETKF para essa região. As regiões onde nota-se um valor de alfa maior, ou seja, a confiabilidade do LETKF é menor, são regiões onde existem um aquecimento elevado devido as convecções que ocorrem na Zona de Convergência Intertropical (ZCIT) (5°, 6° e 7° nível).

As médias para o vento zonal são mostradas nas Figuras 4.5 e 4.6. Observando os gráficos da média ao longo do tempo, nota-se que tanto o Hemisfério Sul quanto o Globo os valores de alfa variaram em torno de 0,2 e 0,3. No caso do Hemisfério Norte, ele atribui valores ao alfa, na maioria das vezes abaixo de 0,2, e para o caso da região Tropical, mantém valores que vão desde 0,1 até um pouco acima de 0,5. Ao observarmos a distribuição espacial do alfa, pode-se encontrar valores de até 0,7 no sétimo nível do modelo na faixa entre 30°S e a Linha do Equador. Valores entre 0,5 e 0,6 são vistos desde o segundo nível em regiões como a faixa Equatorial do continente Africano e do Oceano Pacífico. No primeiro nível, pode-se relacionar os maiores valores de alfa com os ciclones que atuam naquela região no Hemisfério Sul. Da mesma forma, pode-se relacionar os valores do Hemisfério Norte vistos no Oceano Atlântico Norte e no Oceano Pacífico Norte. E assim como a confiabilidade na temperatura nos altos níveis pode ter influência sobre o alfa, o mesmo ocorre para os ventos, podendo estar relacionados à convecção que ocorre na região devido a ZCIT.

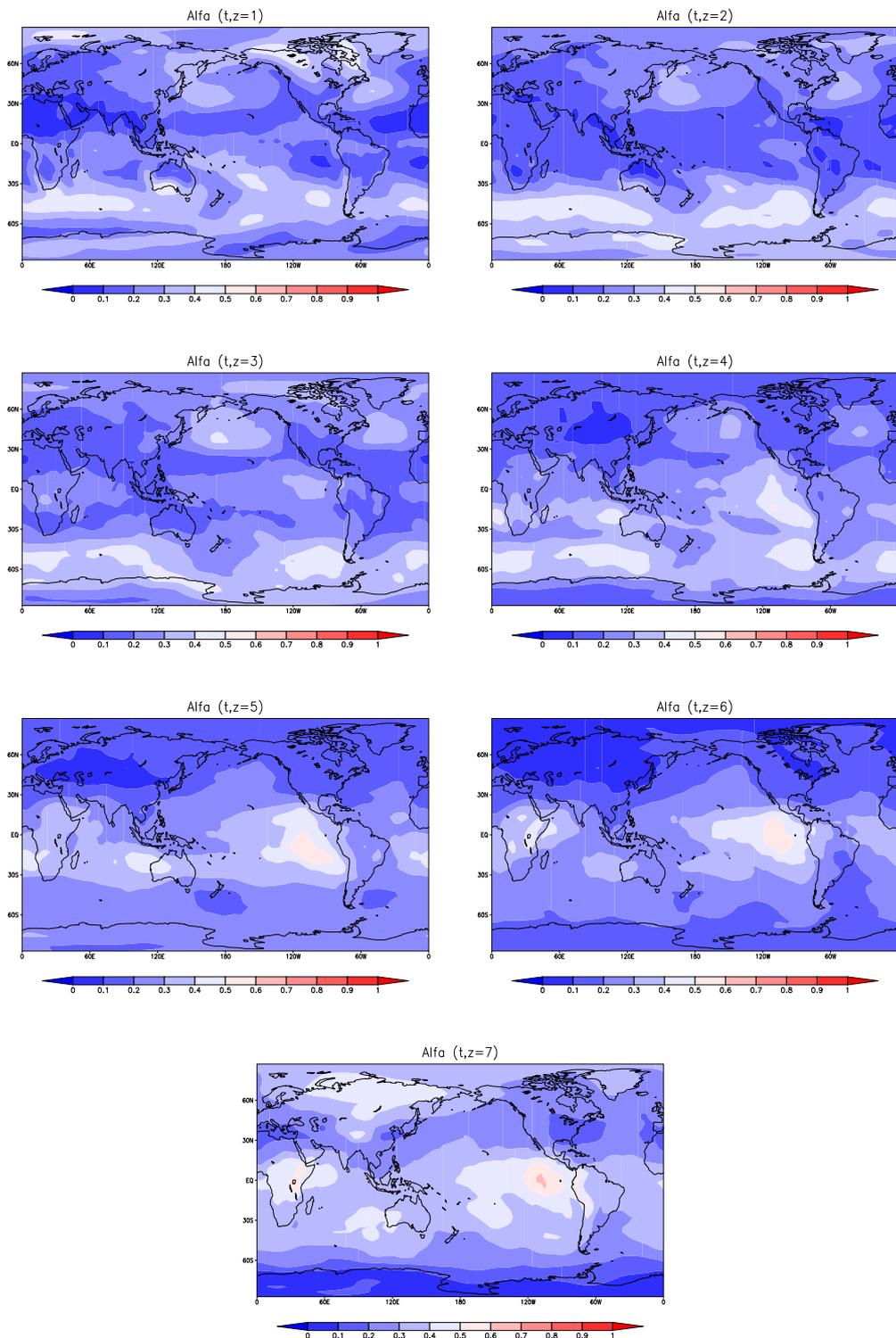
Na Figura 4.7, são apresentados os valores médios para a componente meridional do vento nos sete níveis. Observamos nela o mesmo padrão visto no vento zonal, com valores de alfa maiores nos Trópicos, menores no Hemisfério Norte e em torno de 0,2 e 0,3 no Globo e Hemisfério Sul. A localização dos alfas também foi bem similar ao visto anteriormente no vento zonal, porém, no último nível, observamos valores de alfa chegando a 0,8 sobre a América do Sul. Esses valores altos no último nível podem estar relacionados à baixa confiabilidade do LETKF com relação à Alta da Bolívia e com o Vórtice Ciclônico de Altos Níveis (VCAN) do Nordeste, assim como pode haver alguma relação do valor elevado de alfa sobre o continente Africano no segundo nível, já que tal região é afetada por um Jato de Baixos Níveis que pode causar uma diminuição na confiabilidade. Do mesmo modo como falado para a componente zonal, em baixos níveis também podem ter relação com os ciclones atuantes, tanto no Hemisfério Norte quanto no Hemisfério Sul.

Figura 4.3 - Valores de alfa médio de temperatura para os sete níveis do modelo e para as quatro regiões de estudo: Globo, Hemisfério Norte, Hemisfério Sul e Trópicos, respectivamente, ao longo do trimestre DJF.



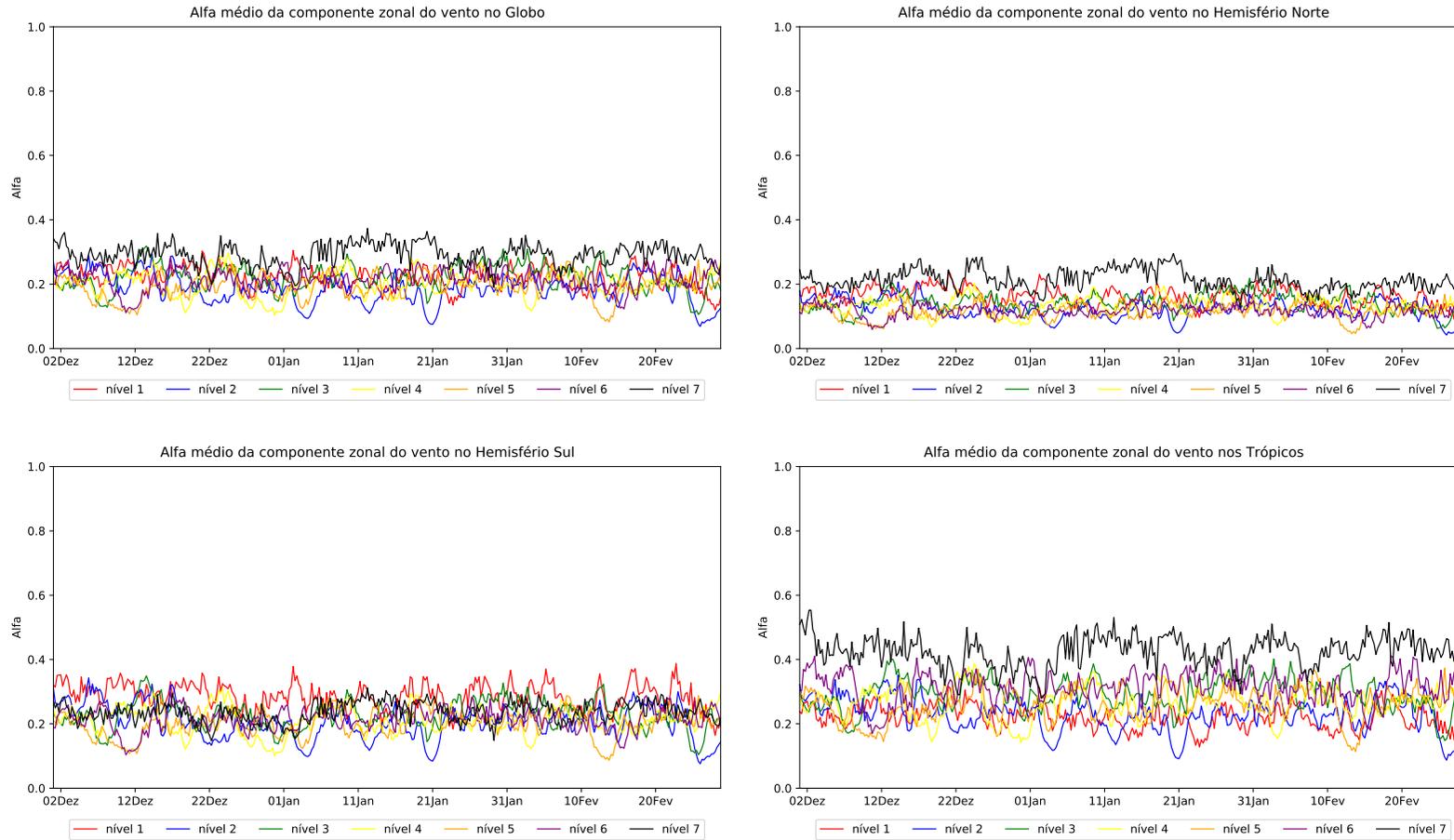
Fonte: Produção do autor.

Figura 4.4 - Idem à Figura 4.2, porém, para a temperatura nos sete níveis do modelo.



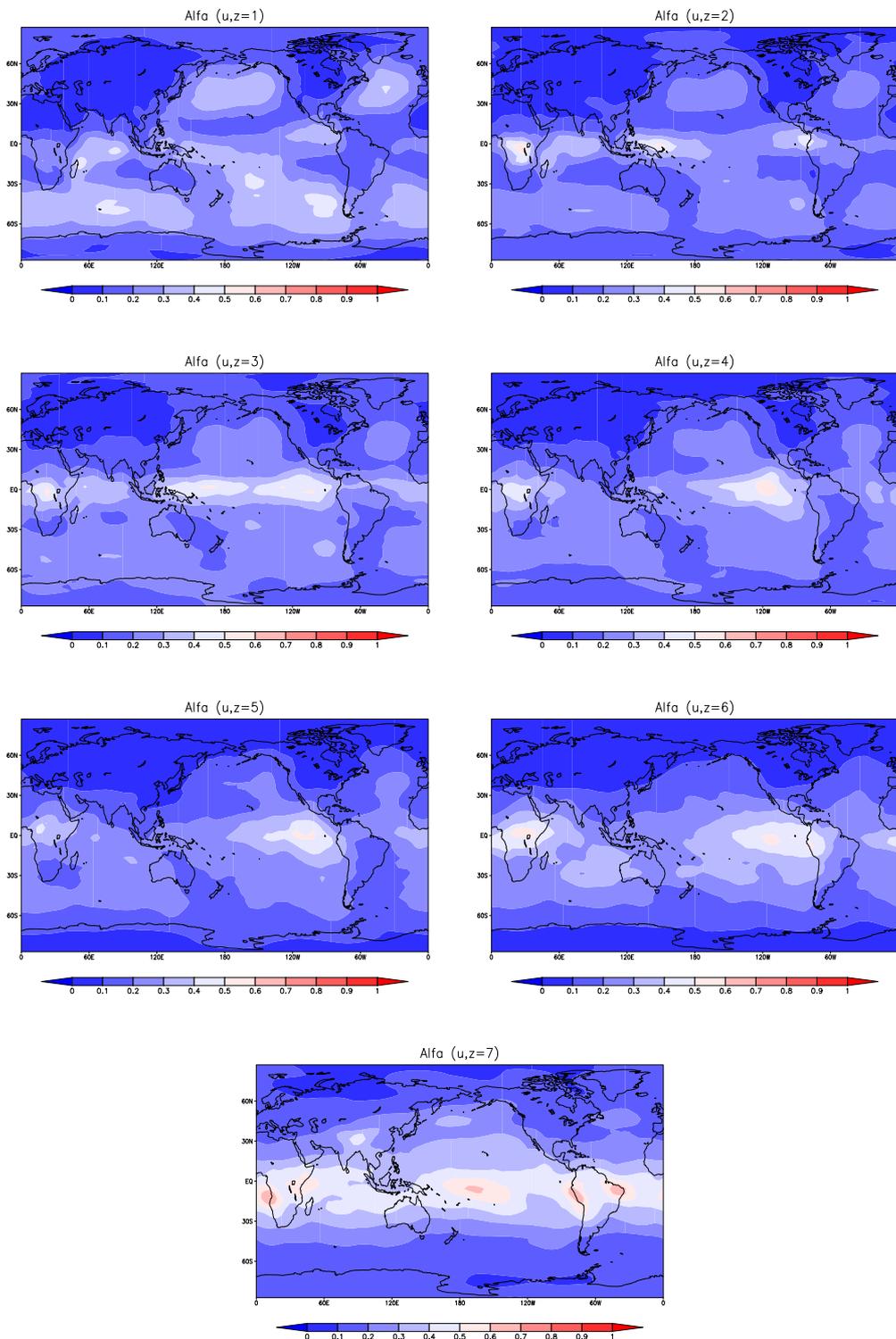
Fonte: Produção do autor.

Figura 4.5 - Idem à Figura 4.3, porém, para a componente zonal do vento.



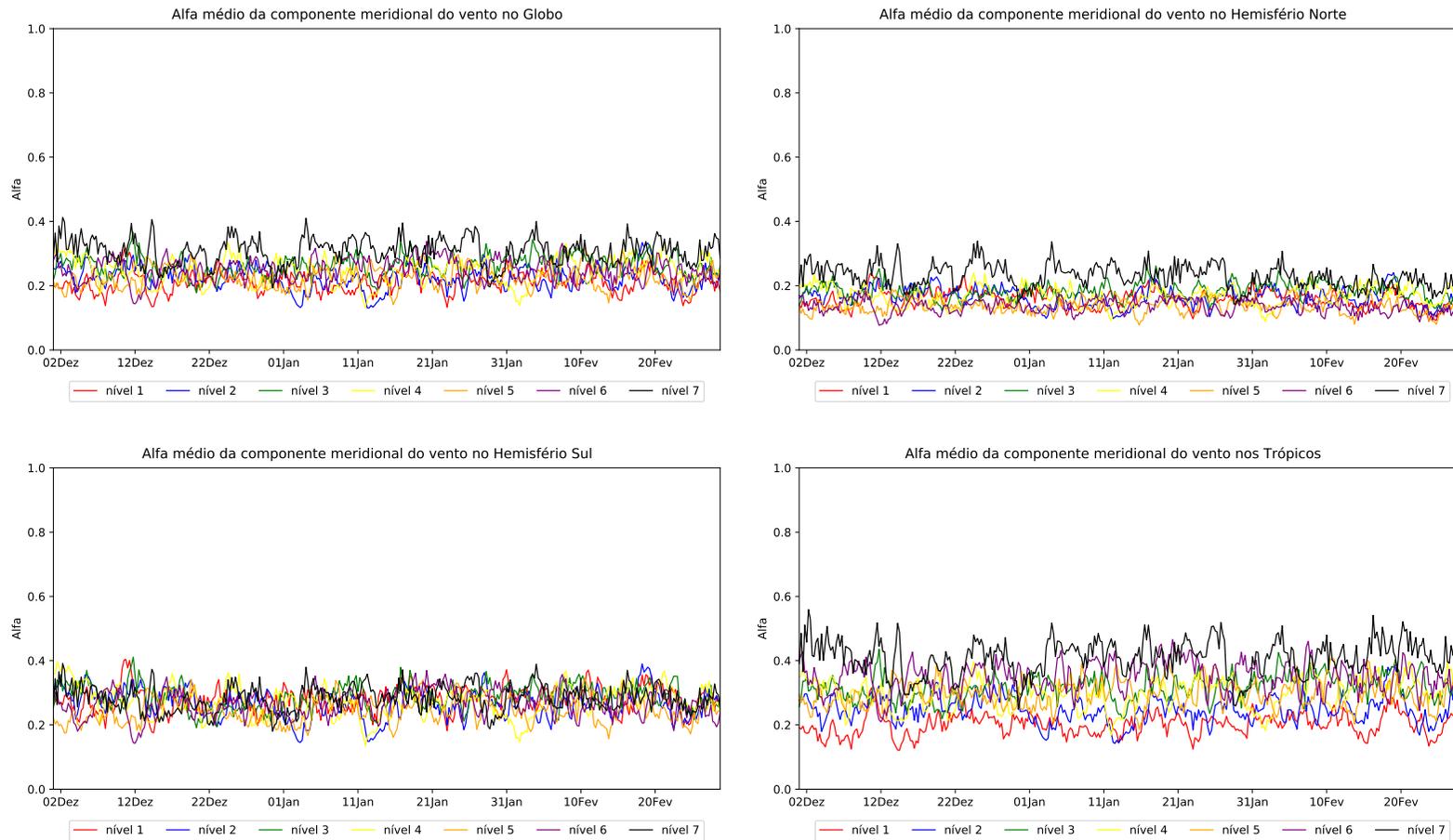
Fonte: Produção do autor.

Figura 4.6 - Idem à Figura 4.4, porém, para a componente zonal do vento.



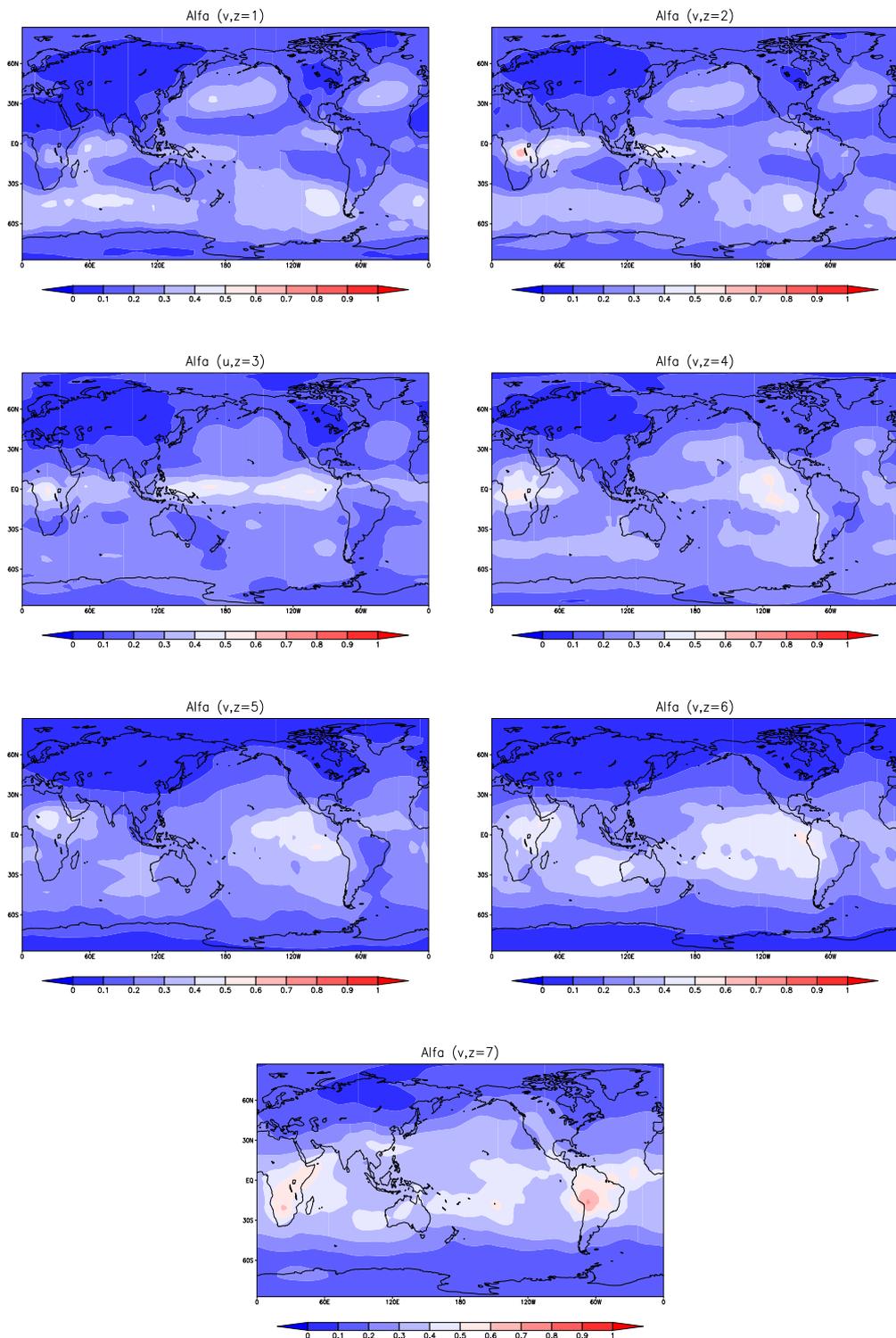
Fonte: Produção do autor.

Figura 4.7 - Idem à Figura 4.3, porém, para a componente meridional do vento.



Fonte: Produção do autor.

Figura 4.8 - Idem à Figura 4.4, porém, para a componente meridional do vento.

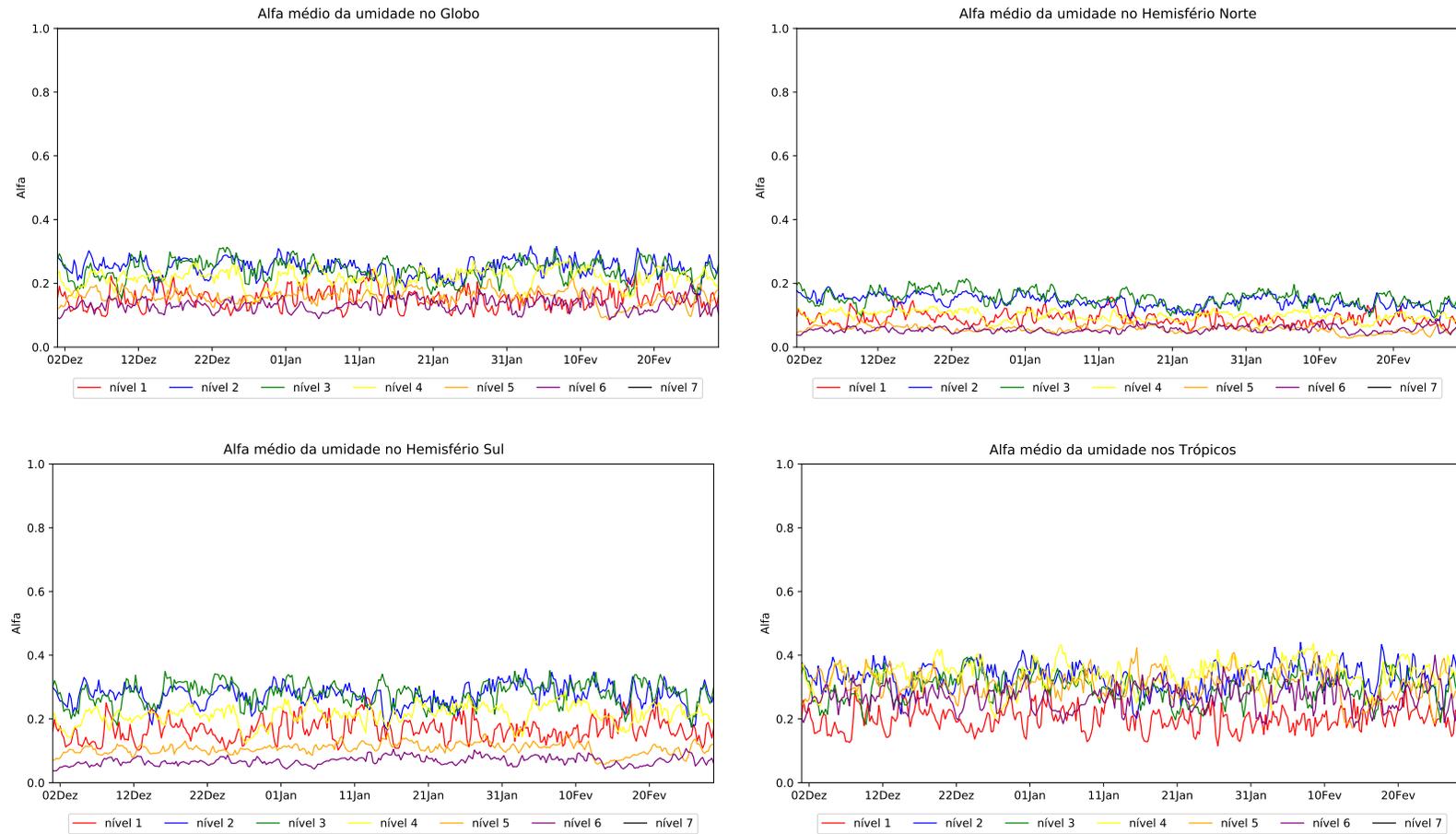


Fonte: Produção do autor.

Os alfas utilizados nos campos de umidade são apresentados nas Figuras 4.9 e 4.10. É possível observar que os valores de alfa ao longo do período de estudo oscilaram entre 0,1 e 0,3 na média Global, porém, observando as regiões separadamente, nota-se que para o Hemisfério Norte foram utilizados valores abaixo de 0,2, os Trópicos utilizaram valores acima de 0,2 e o Hemisfério Sul alcançou os valores utilizados tanto no Hemisfério Norte quando nos Trópicos. Ao olharmos para a distribuição espacial dessas médias, notamos que até o terceiro nível do modelo, em média, não foi utilizado o 3DVar. No entanto, a partir do quarto nível notamos valores de até 0,6 sobre o Oceano Pacífico Leste, mostrando que nessa região tivemos a influência do sistema variacional. Nessa região, acredita-se que o valor elevado de alfa tenha relação também com a convecção apresentada na região. Caso haja uma baixa confiabilidade na temperatura naquela região, isso pode afetar os ventos e conseqüentemente o transporte de umidade dos baixos níveis para os níveis mais altos da atmosfera. Nota-se que para essa variável não há avaliações para o sétimo nível do modelo. Isso ocorre devido ao fato de não serem assimilados dados de umidade nesse nível, e por esse motivo não temos valores de espalhamento. Foi então atribuído para todo o nível valores de  $\alpha = 1$ . Além disso, observam-se valores de alfa próximos do zero em altas latitudes. Se relacionarmos o alfa com o espalhamento, e o espalhamento com incerteza no sistema por conjunto, espera-se que nessas regiões apresentem erros pequenos, uma vez que a umidade nessas regiões é baixa em comparação as demais áreas do globo.

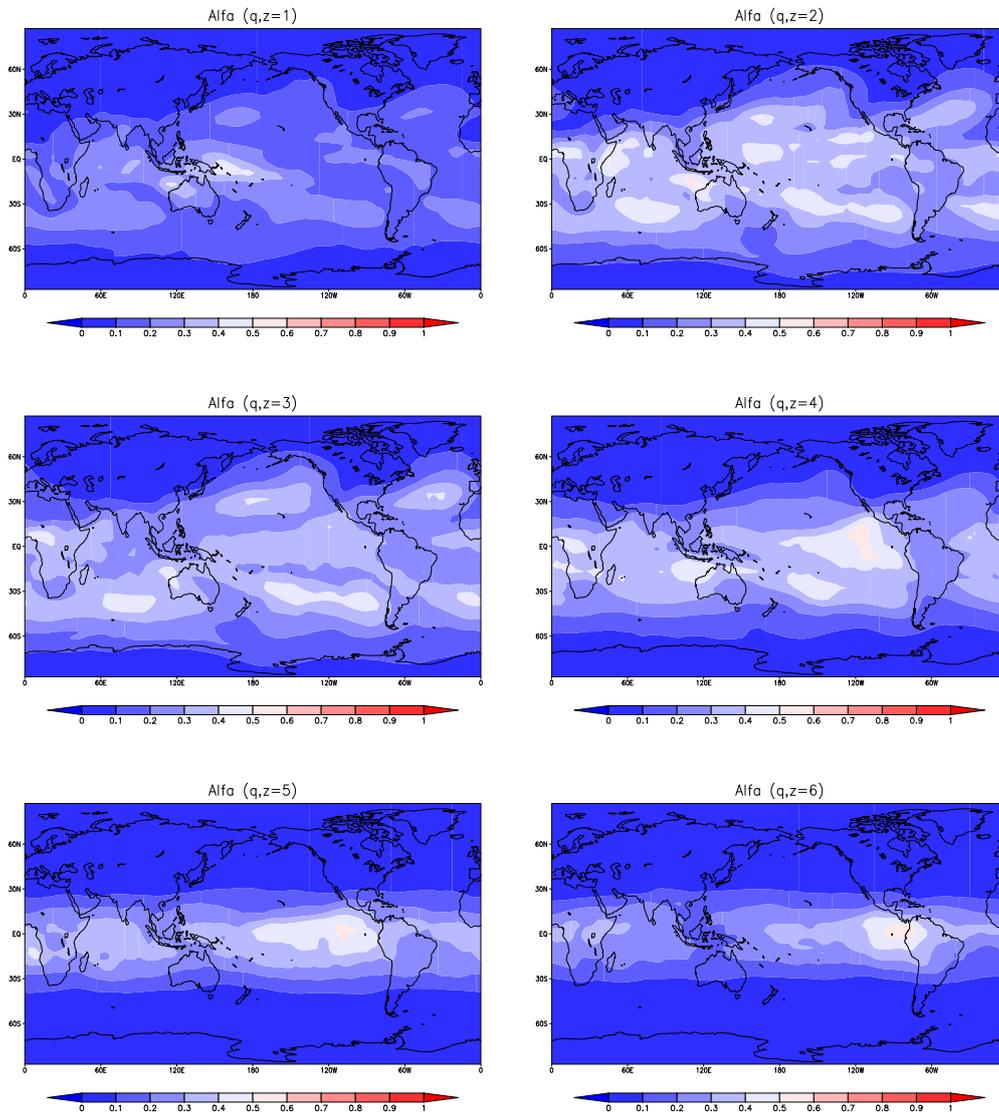
Com os resultados apresentados aqui, para o período de dezembro, janeiro e fevereiro, pode-se notar que o alfa dinâmico, na maioria dos casos, apresentou valores abaixo de 0,5, o que significa um uso maior do LETKF do que do 3DVar na composição da análise final do Híbrido Médio. É possível também relacionar um uso maior do 3DVar em regiões onde o LETKF, através do espalhamento, demonstra ter uma menor confiabilidade, coincidindo com regiões em que atuam determinados sistemas meteorológicos, como por exemplo ciclones, Jatos de Baixos Níveis, Alta da Bolívia e ZCIT. Além disso, pode-se relacionar os valores de alfa com a distribuição das observações ao longo do globo. Apesar de ser um experimento onde se utilizam observações sintéticas, tais observações (conforme mostrado na Figura 3.3) possuem distribuição realística. Conforme visto na Figura 3.3, o Hemisfério Sul possui uma cobertura menor de observações do que o Hemisfério Norte, isso pode ser um fator para que a confiabilidade do LETKF seja menor e acabe beneficiando o 3DVar em regiões oceânicas e em regiões no Hemisfério Sul.

Figura 4.9 - Idem à Figura 4.3, porém, para a umidade.



Fonte: Produção do autor.

Figura 4.10 - Idem à Figura 4.2, porém, para a umidade nos seis níveis do modelo.



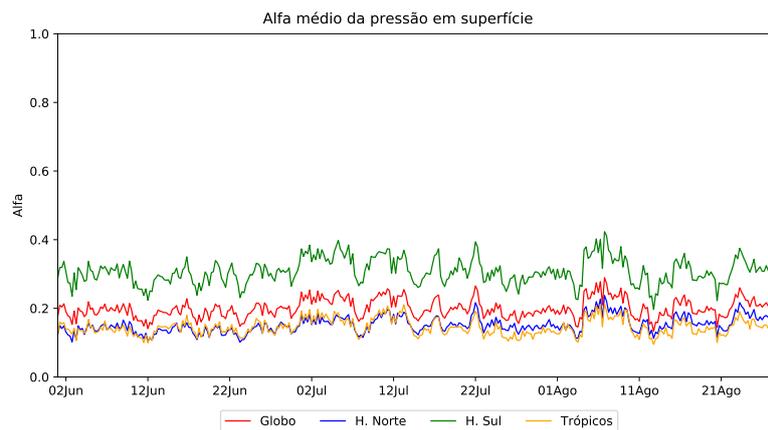
Fonte: Produção do autor.

#### 4.1.2 Alfa médio para o trimestre JJA

Assim como foram apresentados resultados com relação à média do alfa para o período de verão austral, aqui serão apresentados os mesmos resultados, porém para o trimestre de junho, julho e agosto. Nas Figuras 4.11 e 4.12, são apresentados esses resultados para a pressão em superfície. Na primeira figura, pode-se notar valores de alfa similares aos visto para o período de verão austral, mantendo valores médios abaixo de 0,4. Os valores maiores foram atribuídos no Hemisfério Sul e os

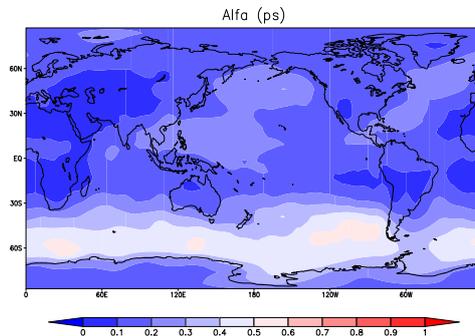
menores valores oscilaram entre o Hemisfério Norte e a região Tropical, mostrando aqui uma diferença com relação ao período anterior, em que os menores valores ficaram localizados na região Tropical. Pode ser observado na figura da distribuição espacial (Figura 4.12) que sobre o Oceano Atlântico Norte e Oceano Pacífico Norte, onde haviam regiões com valores de alfa mais elevados, agora, para o período de inverno austral, esses valores diminuíram aproximadamente 0,2. Possivelmente, isso ocorreu devido a não ocorrência de ciclones nessa região durante esse período, já que esses sistemas são comuns durante o inverno no Hemisfério Norte, porém é possível de se notar em azul mais claro, as regiões de ocorrência dos ciclones explosivos. É possível também observar valores de alfa acima de 0,5 na faixa entre 70°S e 40°S. Essa região é vinculada à ocorrência de ciclones, que por sua vez podem afetar a confiabilidade e por consequência os valores de alfa.

Figura 4.11 - Alfa médio na área da pressão em superfície ao longo do trimestre JJA.



Fonte: Produção do autor.

Figura 4.12 - Distribuição espacial do alfa médio da pressão em superfície para o período de junho, julho e agosto.



Fonte: Produção do autor.

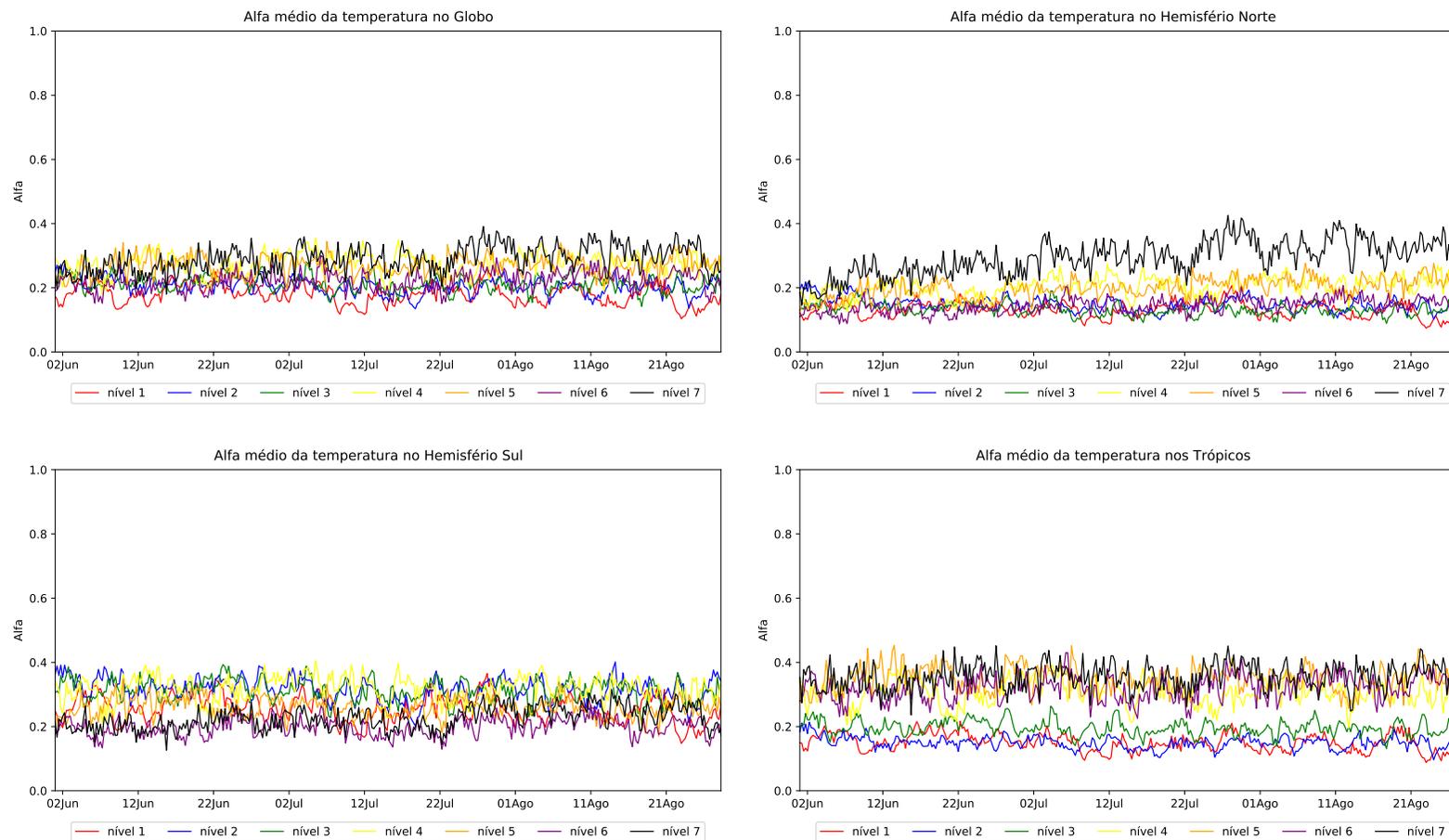
Os valores médios do alfa da temperatura ao longo do trimestre JJA são mostrados na Figura 4.13. Esses gráficos seguem o mesmo padrão dos gráficos apresentados na seção anterior. Já as distribuições espaciais dos valores de alfa, para cada nível, são mostradas na Figura 4.14. Nota-se nos gráficos de média, em geral, para o Globo valores próximos aos vistos no trimestre DJF. Para o Hemisfério Norte, notam-se valores mais elevados de alfa no último nível do modelo. No Hemisfério Sul vemos valores médios entre 0,2 e 0,4, porém, ao observar a distribuição espacial, notamos valores de até 0,7 em regiões do Polo Sul no primeiro nível, 0,6 no segundo e no quarto nível. Esses valores não foram observados no período de verão, o que pode ter relação com a confiabilidade do LETKF na representação dos sistemas transientes do período de inverno, que nessa época são mais intensos, sendo essa a região onde ocorrem as trocas de massas de ar, e por consequência dessa confiabilidade, nessas regiões foi utilizado mais o 3DVar. Na região Tropical, em altos níveis, o mesmo sinal visto como no trimestre DJF é visto para esse período, onde possivelmente os valores mais elevados de alfa estejam relacionados à baixa confiabilidade com relação à convecção na região da ZCIT.

Nas Figuras 4.15 e 4.16, são mostrados os resultados do alfa médio para a componente zonal do vento. Comparando com o trimestre DJF, nota-se um aumento nos valores médios de alfa no sétimo nível do modelo para o Globo, o que ao observarmos o gráfico do Hemisfério Norte fica claro que esse aumento no Globo veio desse Hemisfério. Quando se olha a distribuição espacial nesse nível, observa-se a presença de uma região, localizada sobre a Ásia, com valores de até 0,7, que possivelmente

teve influência para fazer com que os valores aumentassem na média. Esses valores elevados de alfa podem estar relacionados com o sistema de alta pressão que ocorre sobre o Planalto do Tibete durante o período de verão no Hemisfério Norte, mas por algum motivo o sistema por conjunto tem uma baixa confiabilidade e por consequência o 3DVar acaba sendo mais utilizado. Nesse nível, ainda observamos duas outras regiões com esse mesmo valor, uma sobre o leste do continente africano e a outra a oeste da América do Sul. Esses valores ajudam a elevar os valores médios de alfa ao longo do período na região Tropical. O aumento destes valores pode estar relacionado com o relevo da região ou ainda com convecção apresentada nelas. Com relação aos demais níveis, pode-se observar valores acima de 0,5, exceto para o primeiro nível que em média apresentou valores abaixo de 0,5. Nos níveis 2 e 3, especula-se que os ventos Alísios possam interferir na confiabilidade, e por isso apresentarem maiores valores de alfa.

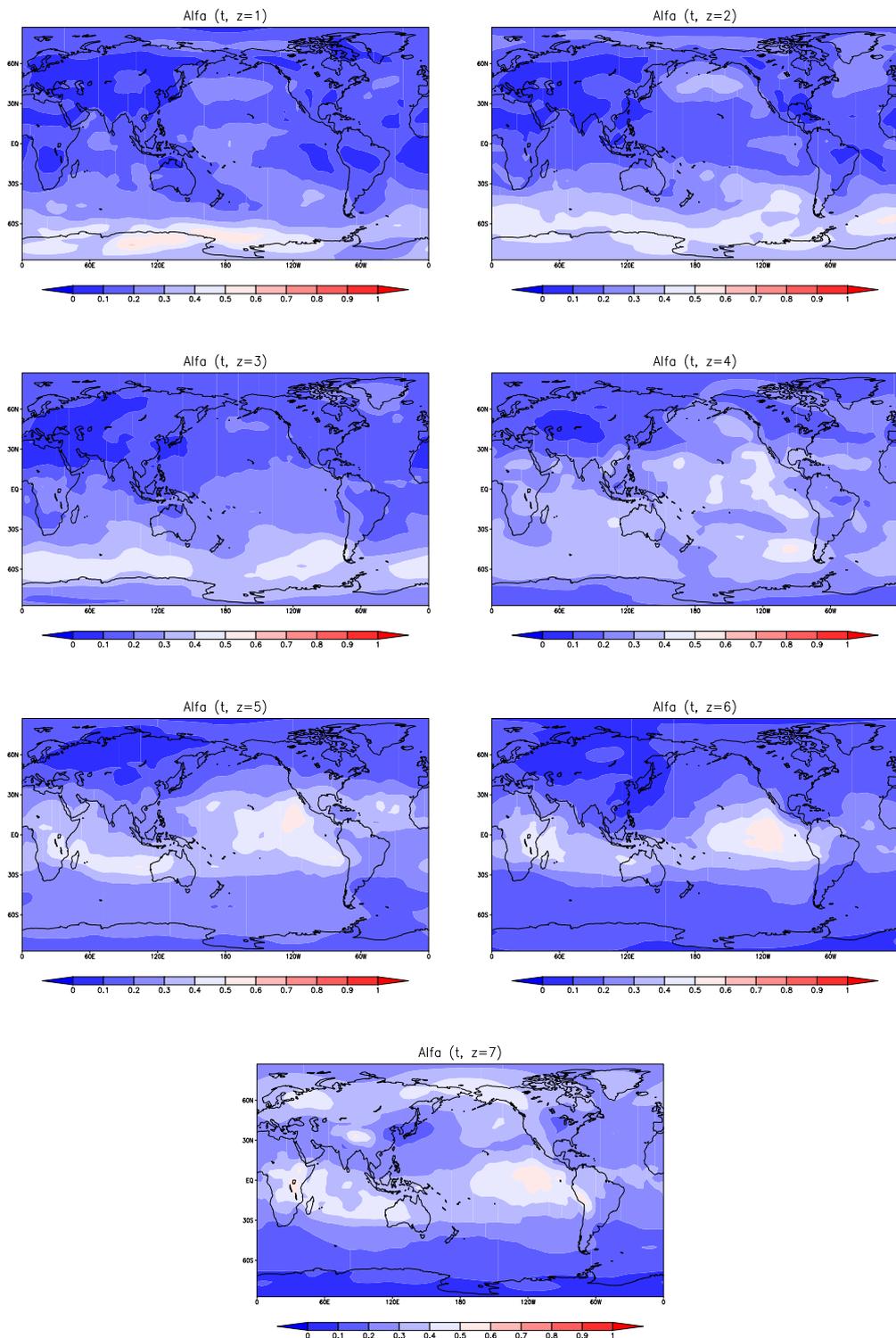
Os resultados para a componente meridional do vento são mostrados nas Figuras 4.17 e 4.18. Nota-se novamente um aumento nos valores Globais no último nível do modelo, e esse aumento continua sendo reflexo de um aumento nos valores de alfa no Hemisfério Norte. Avaliando a distribuição espacial, pode-se observar valores de alfa de até 0,5 sobre a Ásia e Oceano Pacífico Norte, além de valores que chegam a 0,7 sobre o oeste dos Estados Unidos. Isso justifica o aumento no gráfico. No Hemisfério Sul, nota-se um leve aumento nos valores, porém, ao olharmos a distribuição espacial, é possível notar apenas um aumento nas regiões que apresentam um alfa maior, ainda assim ultrapassando 0,5 apenas no último nível, o que também foi observado para o caso de verão austral. Esses valores, tanto no Hemisfério Norte quanto no Hemisfério Sul, assim como mencionado para a componente zonal do vento, podem ter relação com a confiabilidade em relação ao relevo dessas regiões, com o Planalto do Tibete na Ásia, as Montanhas Rochosas na América do Norte e a Cordilheira dos Andes na América do Sul. Na região Tropical, há uma diminuição na média dos valores do último nível com relação ao trimestre DJF. Isso se dá devido a uma diminuição nos valores de alfa sobre o continente africano, para o qual durante o período de verão alcançou valores de até 0,7, enquanto que no trimestre JJA esses valores ficaram abaixo de 0,5.

Figura 4.13 - Valores de alfa médio de temperatura para os sete níveis do modelo e para as quatro regiões de estudo: Globo, Hemisfério Norte, Hemisfério Sul e Trópicos, respectivamente, ao longo do trimestre JJA.



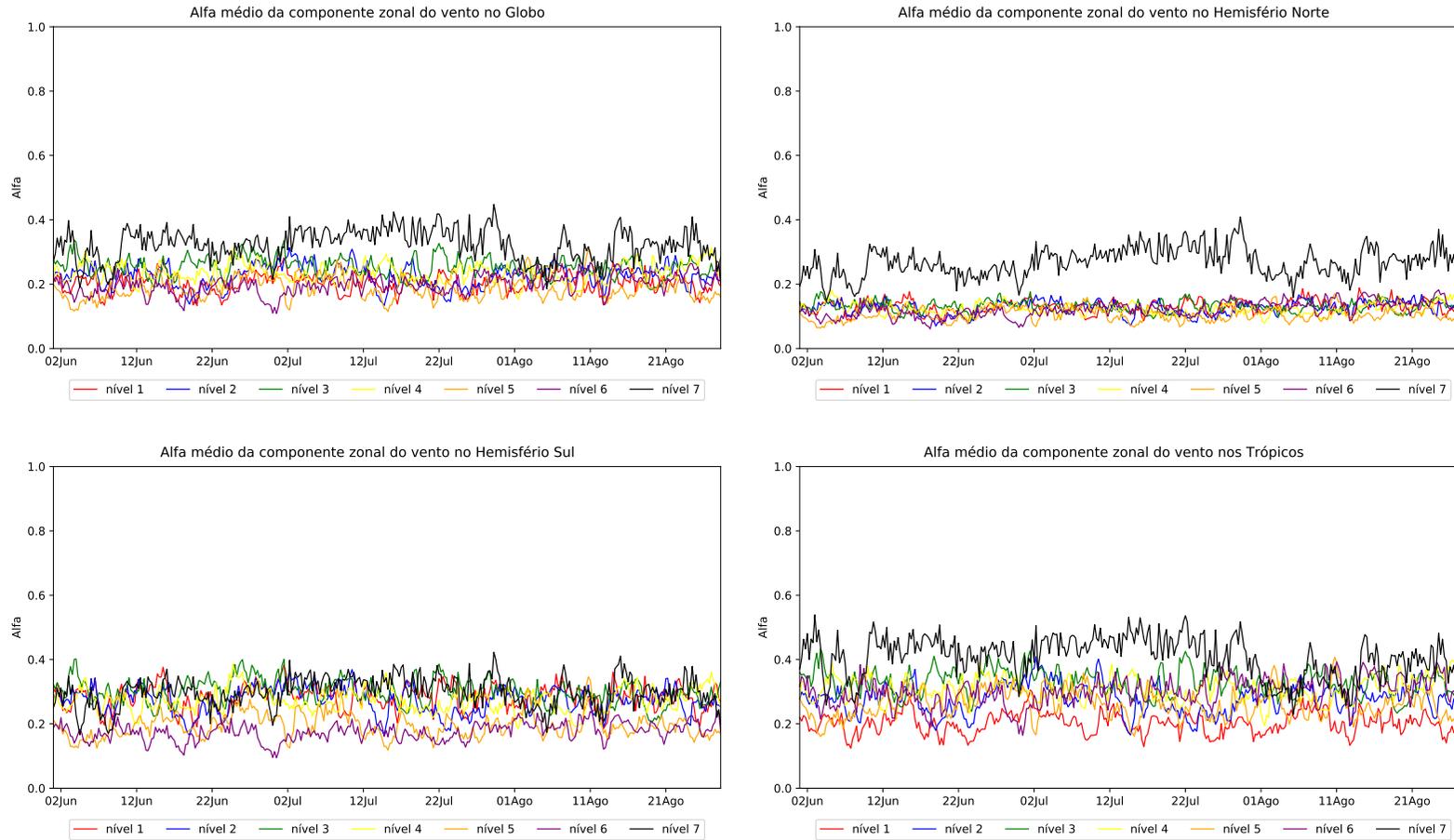
Fonte: Produção do autor.

Figura 4.14 - Idem à Figura 4.12, porém, para a temperatura nos sete níveis do modelo.



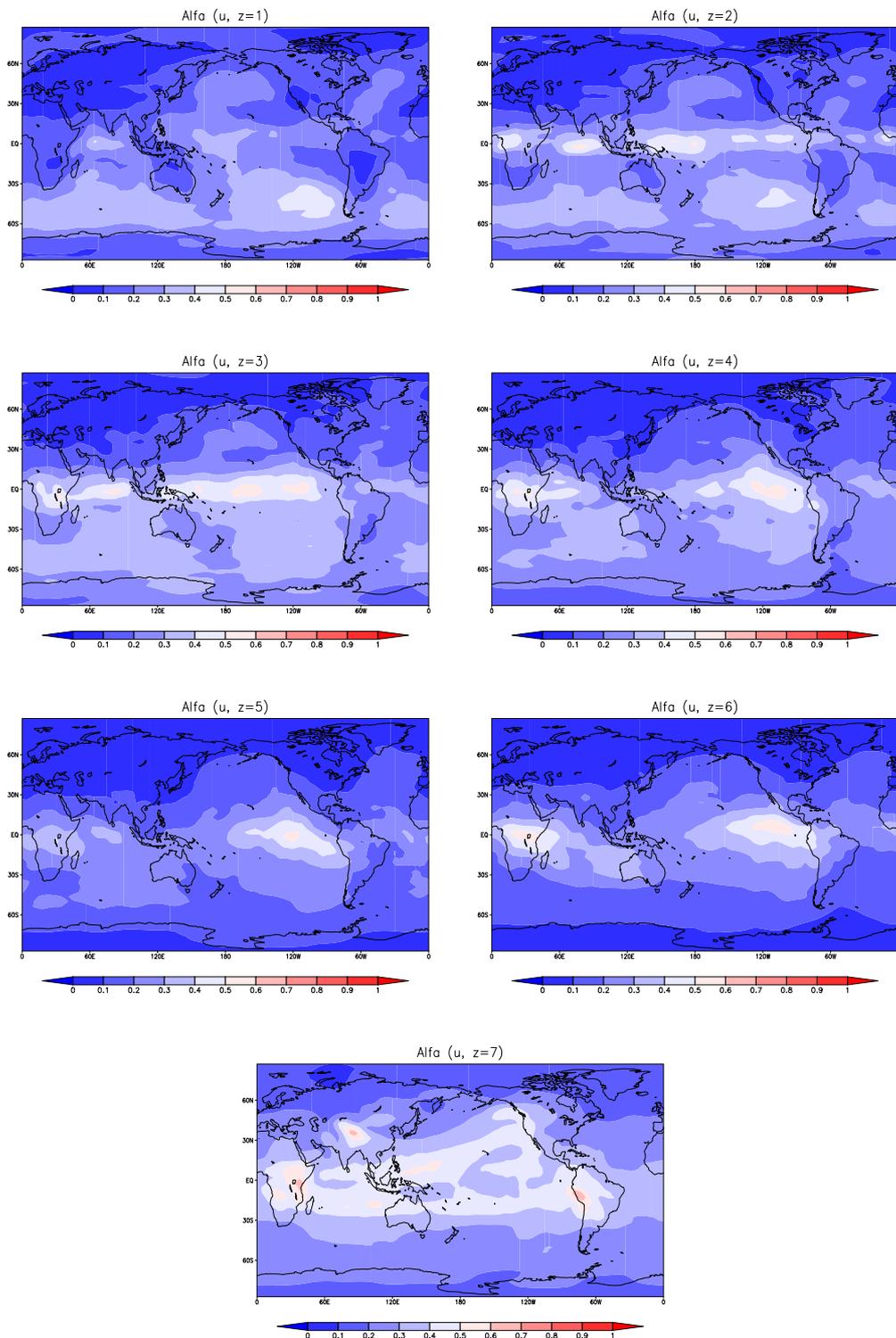
Fonte: Produção do autor.

Figura 4.15 - Idem à Figura 4.13, porém, para a componente zonal do vento.



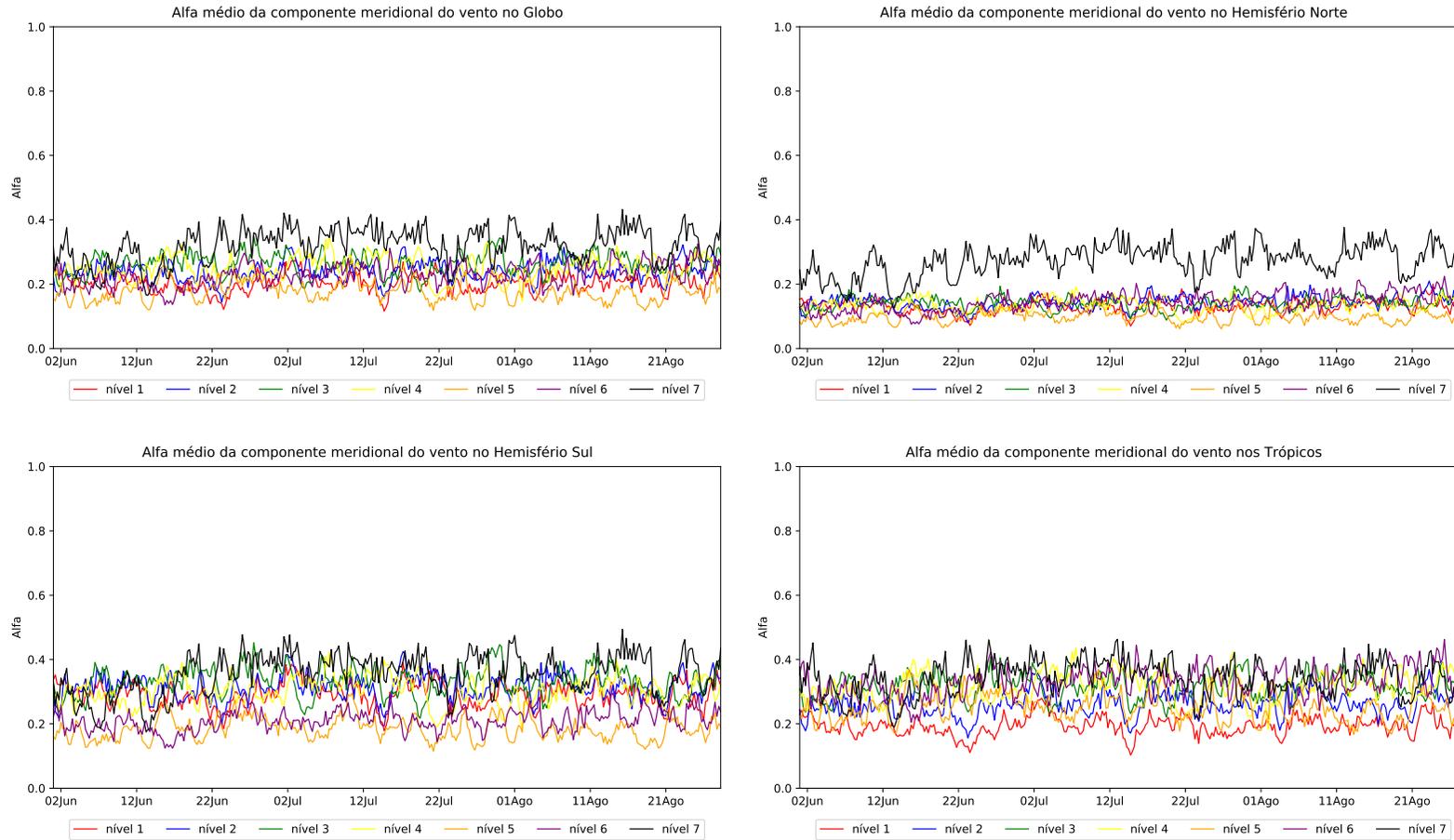
Fonte: Produção do autor.

Figura 4.16 - Idem à Figura 4.14, porém, para a componente zonal do vento.



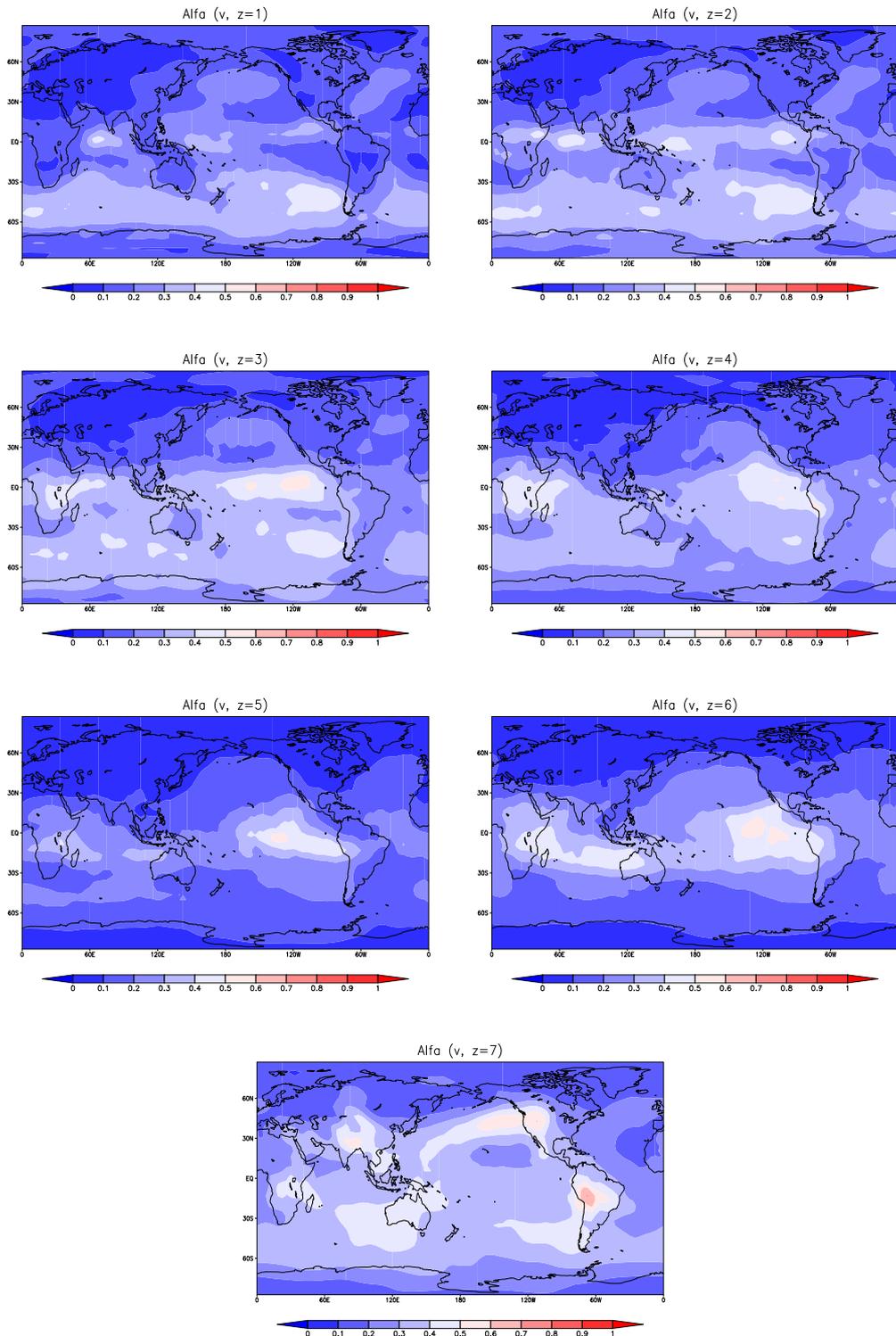
Fonte: Produção do autor.

Figura 4.17 - Idem à Figura 4.13, porém, para a componente meridional do vento.



Fonte: Produção do autor.

Figura 4.18 - Idem à Figura 4.14, porém, para a componente meridional do vento.

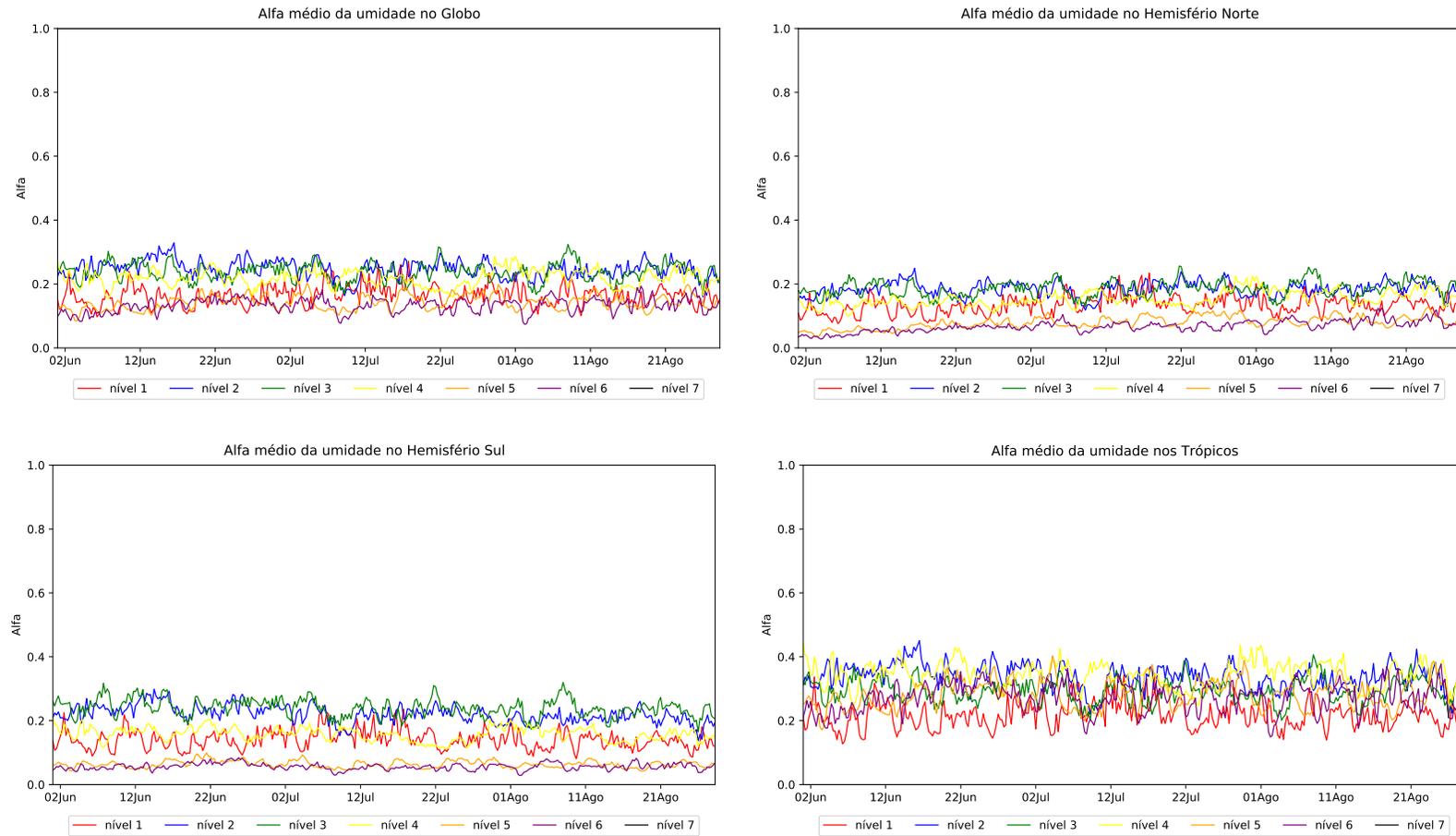


Fonte: Produção do autor.

Nas Figuras 4.19 e 4.20 são apresentados valores médios de alfa para a umidade nos seis níveis. Essa variável foi a que apresentou menores diferenças em relação ao período anterior. Nota-se uma pequena diminuição nos valores do Hemisfério Sul para os níveis 2 e 3. Em contrapartida, ao observarmos a distribuição espacial do alfa, notamos valores maiores de alfa na região Tropical. Assim como no trimestre DJF, as regiões próximas aos polos apresentam valores baixos de alfa uma vez que são regiões mais áridas e, por consequência, a confiabilidade do LETKF é maior.

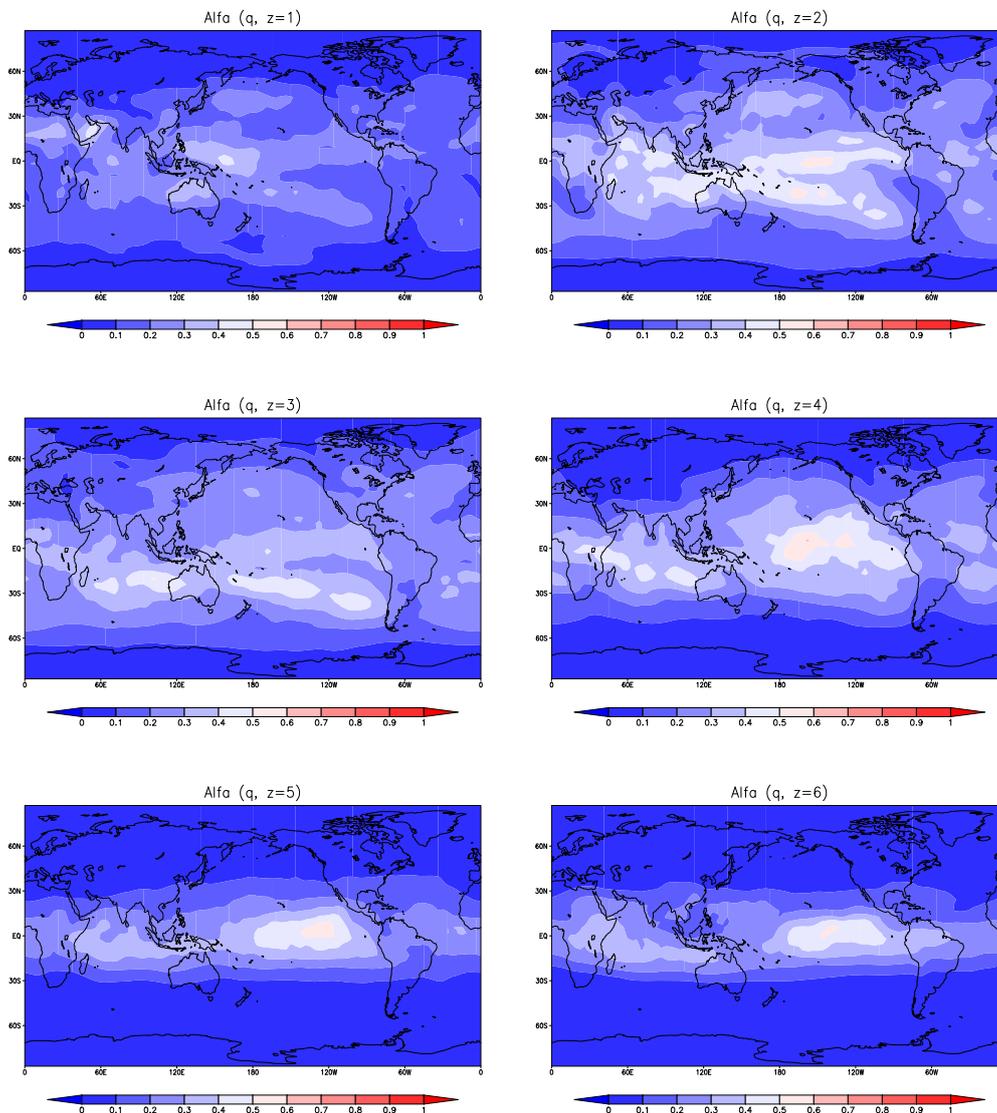
Os resultados mostrados aqui são referentes ao período de junho, julho e agosto. Neles, observamos em média um aumento nos valores de alfa no Hemisfério Norte quando comparados com o período anterior, principalmente com relação às componentes do vento e com relação à temperatura no Hemisfério Sul. Também foi observada uma diminuição do alfa com relação à pressão em superfície no Hemisfério Norte. Esses valores de alfa podem ser relacionados aos sistemas sinóticos atuantes em determinadas regiões, como, por exemplo, a ZCIT, a Alta do Planalto do Tibete, ciclones e ainda com o relevo de determinadas regiões. Essa suposição pode ser pensada se levarmos em conta a relação de confiabilidade do LETKF (espalhamento) com esses sistemas, ainda que com essas diferenças apresentadas aqui, entre o período de inverno e o período de verão em média tem-se maior contribuição do LETKF do que do 3DVar na composição da análise final do Híbrido Médio.

Figura 4.19 - Idem à Figura 4.13, porém, para a umidade.



Fonte: Produção do autor.

Figura 4.20 - Idem à Figura 4.12, porém, para a umidade nos seis níveis do modelo.



Fonte: Produção do autor.

## 4.2 Variação do Alfa Dinâmico

Nessa Seção são mostradas as variações do alfa dinâmico em cada um dos períodos de estudo através do desvio padrão, que é dado pela Equação 4.2, onde  $n$  é o número total de valores,  $x_i$  é cada valor onde  $i$  varia de 1 até  $n$  e o  $\mu$  é a média calculada na Equação 4.1. O objetivo é mostrar o quanto os valores de alfa variam em torno dos seus valores médios, ou seja, pode-se inferir as regiões onde o alfa se manteve mais estático e regiões onde ele se manteve mais dinâmico, e estimar qual foi a sua

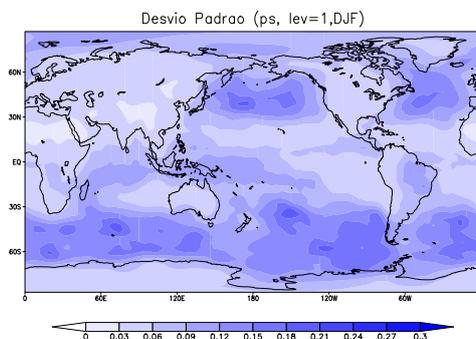
variação dentro de cada período de estudo.

$$Desvio\ Padr\tilde{a}o = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2} \quad (4.2)$$

#### 4.2.1 Desvio Padrão do Alfa Dinâmico para o trimestre DJF

Nas Figuras 4.21, 4.22, 4.23, 4.24 e 4.25 são apresentados os desvios padrões do alfa para o período de dezembro, janeiro e fevereiro para as cinco variáveis e sete níveis do modelo. Esses resultados são apresentados em tons de azul; quanto mais claro, menores os valores, e quanto mais escuro, maiores os valores de desvio padrão. Na primeira delas, pode-se observar o desvio padrão do alfa da pressão em superfície. Nota-se que nas regiões onde estão situados os sistemas transientes no Hemisfério Sul e os ciclones no Hemisfério Norte os valores de desvio padrão são maiores do que no restante do globo, chegando a 0,27. Essas regiões também apresentaram valores elevados de alfa, como mostra a Figura 4.2. Em regiões onde o alfa médio foi baixo, observam-se também valores baixos de desvio padrão.

Figura 4.21 - Desvio padrão do alfa da pressão em superfície para o período de dezembro, janeiro e fevereiro.



Fonte: Produção do autor.

Os valores de desvio padrão da temperatura nos sete níveis são apresentados na Figura 4.22. Nela pode-se notar valores de desvio padrão maiores no primeiro nível do modelo e até o terceiro nível, seguindo o mesmo padrão que o visto no desvio padrão da pressão em superfície, no qual os maiores valores ficaram concentrados nas regiões dos sistemas de baixas pressões que atuam no Hemisfério Sul e no He-

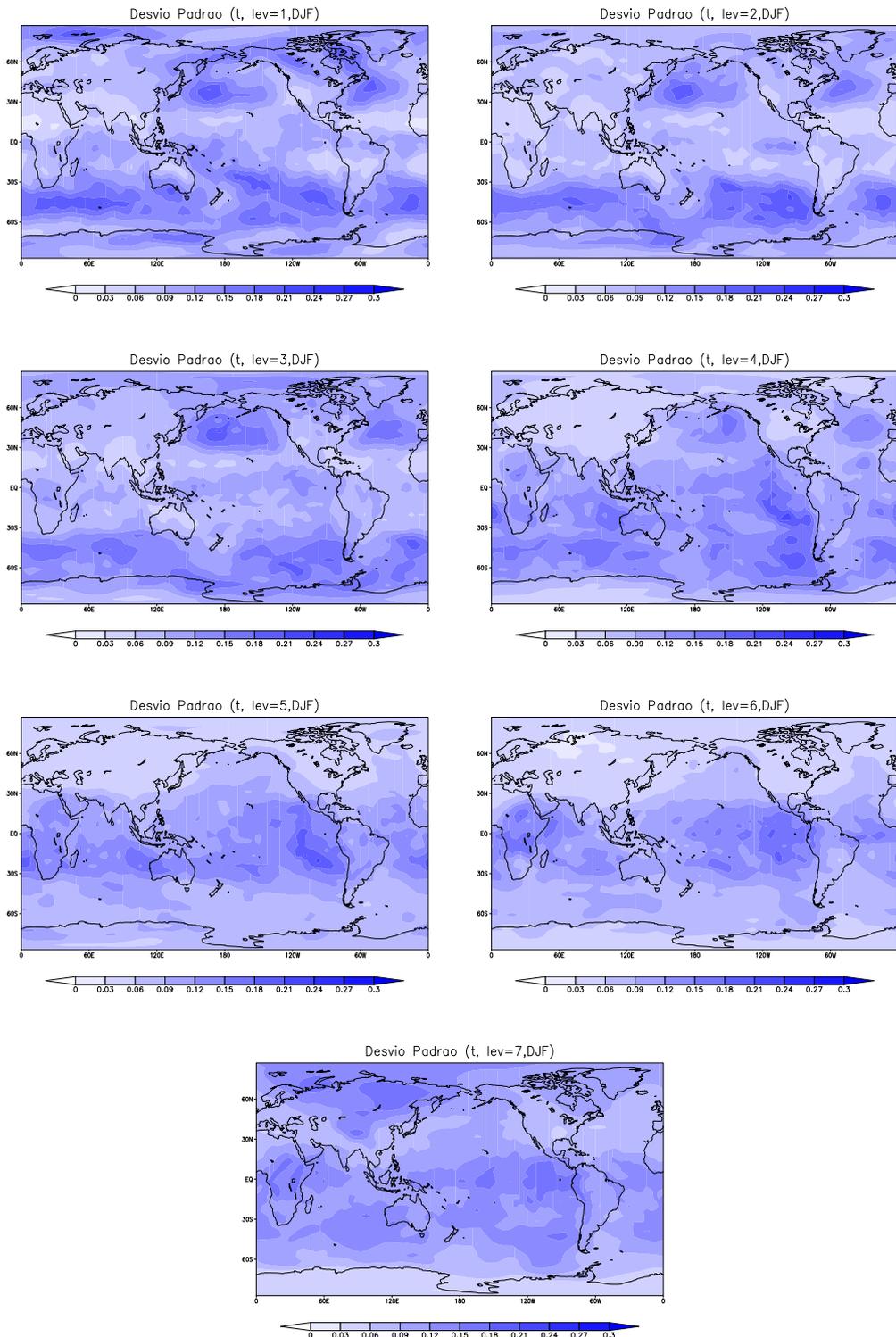
misfério Norte. Esses valores estão relacionados aos altos valores de alfa que também foram observados nessas regiões. Igualmente, é possível observar nos demais níveis, porém, com valores de desvio padrão mais baixos, mostrando, com isso, que para a temperatura o alfa varia mais em superfície do que em níveis mais altos durante o período de dezembro, janeiro e fevereiro.

Nas Figuras 4.23 e 4.24 são mostrados os desvios padrões para as componentes zonal e meridional do vento, respectivamente. Nota-se que para ambas as variáveis os maiores valores de desvio padrão estão associados aos maiores valores de alfa, como visto na seção anterior. Esses valores mais elevados acompanham regiões de sistemas transientes, ciclones e região dos ventos alísios no primeiro e segundo nível do modelo, e valores mais elevados também foram vistos nos níveis seguintes na região da ZCIT e em altos níveis na região da Alta da Bolívia e sobre a região do VCAN do Nordeste.

A Figura 4.25 mostra os campos de desvio padrão para a umidade nos seis níveis do modelo. Para essa variável, nota-se que em alguns casos nos quais se observou um desvio padrão alto com relação às demais regiões, nem sempre foram observados valores altos de alfa como visto até então, como é o caso do primeiro nível do modelo sobre o Oceano Índico.

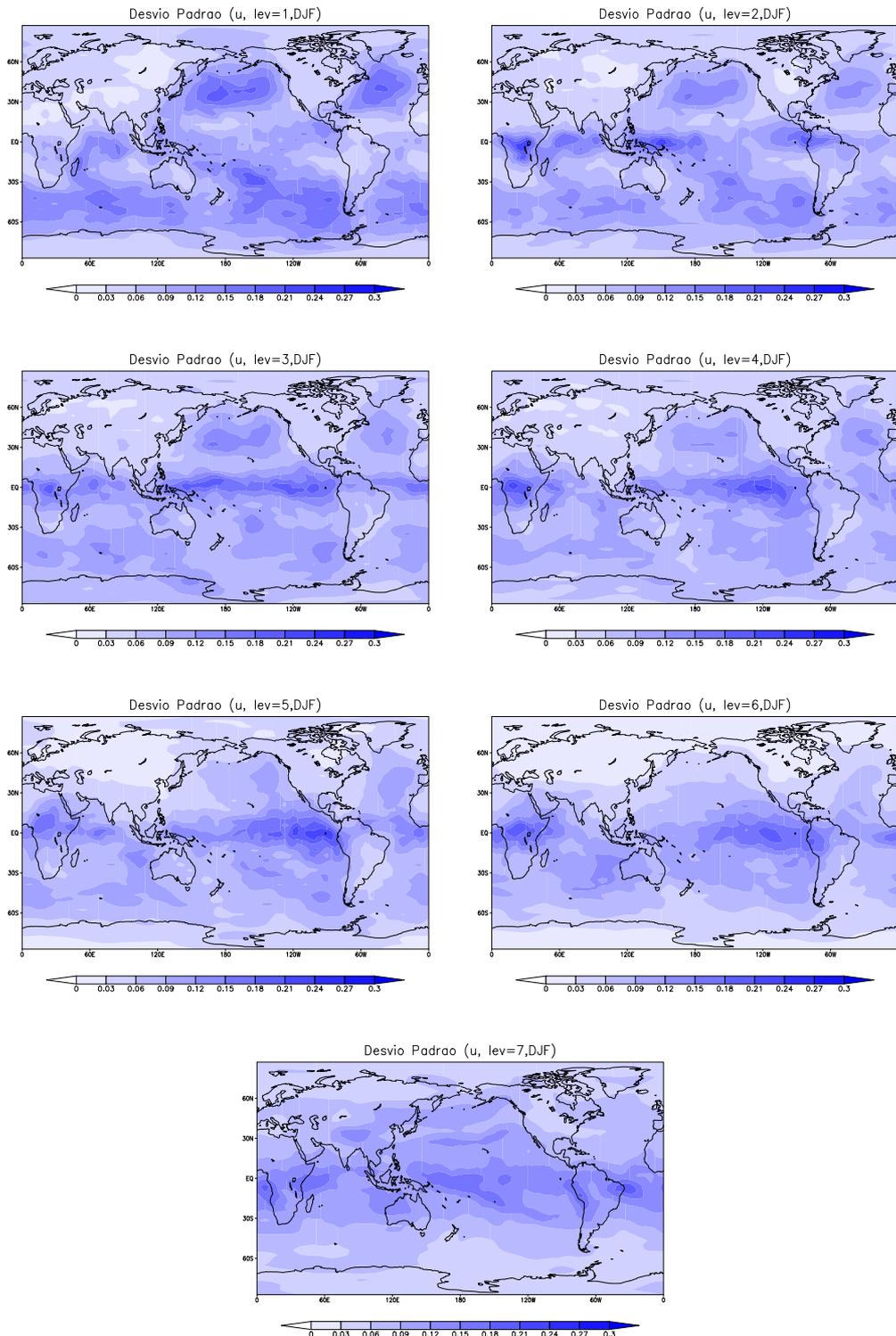
Foram mostrados aqui os valores de desvio padrão para os experimentos realizados no período de dezembro, janeiro e fevereiro. Pode-se notar uma relação entre o alfa médio e o desvio padrão. Na maioria das variáveis, observou-se que, quanto maior o valor médio do alfa dinâmico, maiores foram os valores de desvio padrão encontrados. O contrário também foi visto: regiões com baixos valores de alfa apresentaram um valor baixo de desvio padrão. As regiões com valores mais altos de desvio padrão estão relacionadas às regiões onde ocorrem determinados sistemas meteorológicos, como, por exemplo, ciclones e ZCIT. Isso mostra a tentativa do alfa em acompanhar esses sistemas, em alguns momentos utilizando mais o 3DVar nessas regiões e, em outros, fazendo um uso maior do LETKF. Já nas regiões onde o desvio padrão foi baixo, há estabilidade maior do alfa, ou seja, regiões onde ele mantém uma constância maior nos seus valores, atribuindo, na maioria das vezes, mais peso às análises do LETKF durante a geração da análise final do Híbrido Médio.

Figura 4.22 - Idem à Figura 4.21, porém, para a temperatura nos sete níveis do modelo.



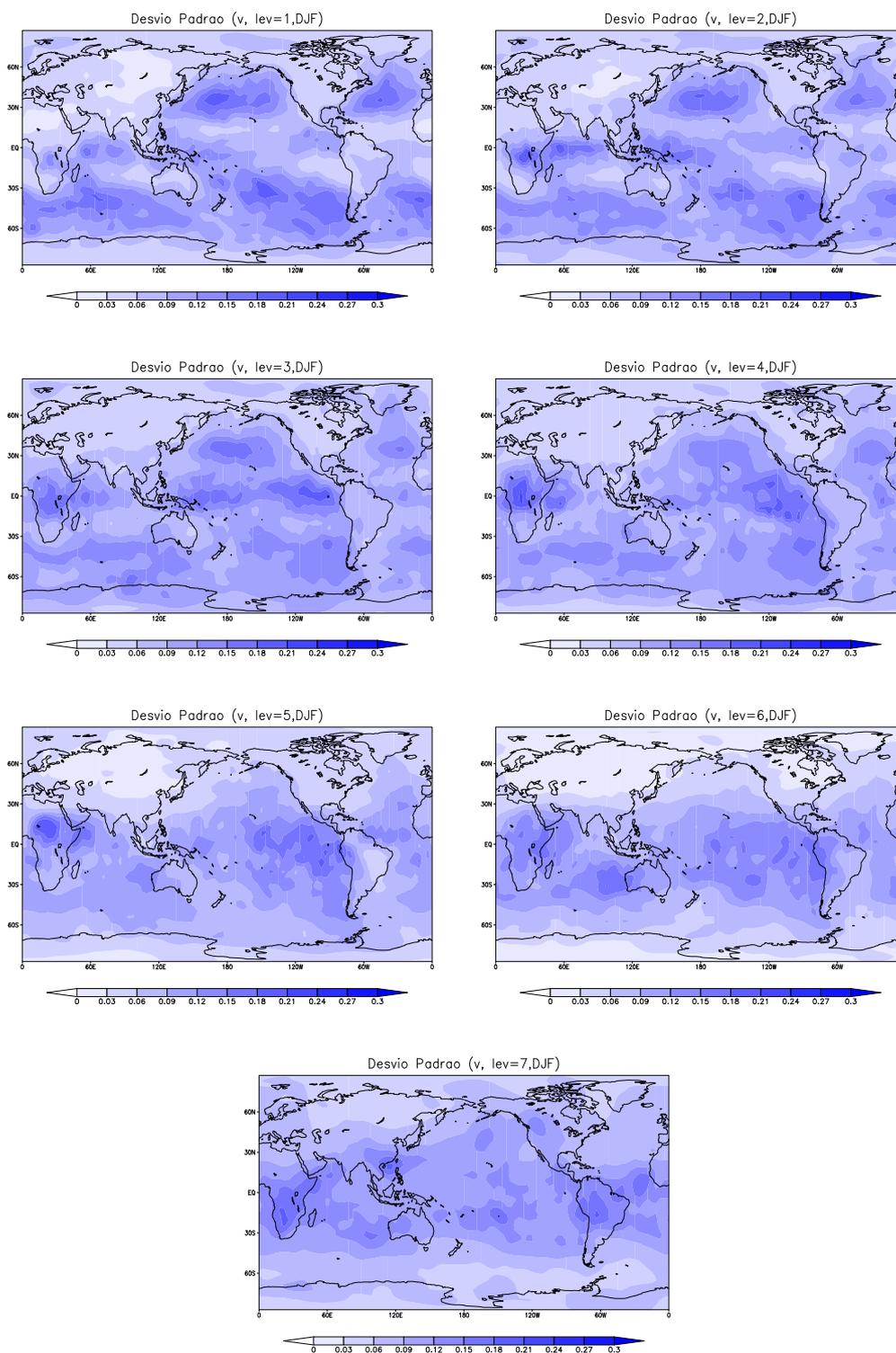
Fonte: Produção do autor.

Figura 4.23 - Idem à Figura 4.21, porém, para o vento zonal nos sete níveis do modelo.



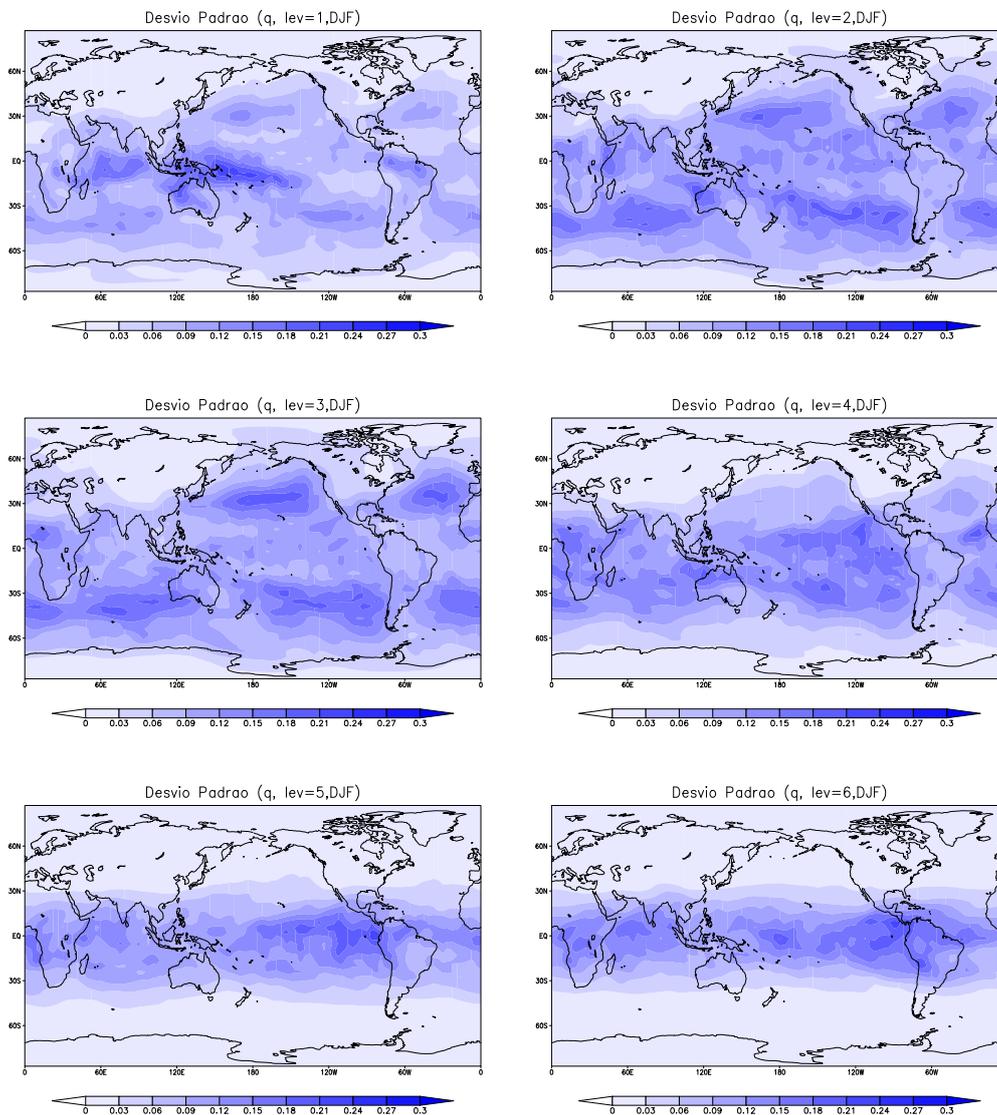
Fonte: Produção do autor.

Figura 4.24 - Idem à Figura 4.21, porém, para o vento meridional nos sete níveis do modelo.



Fonte: Produção do autor.

Figura 4.25 - Idem à Figura 4.21, porém, para a umidade nos seis níveis do modelo.



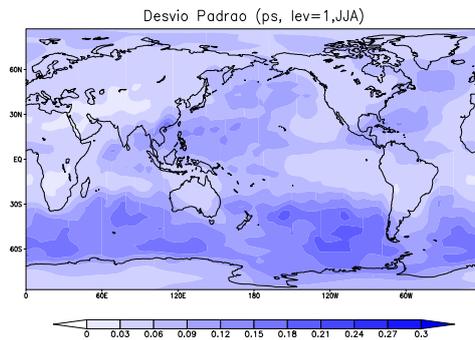
Fonte: Produção do autor.

#### 4.2.2 Desvio Padrão do Alfa Dinâmico para o trimestre JJA

Nas Figuras a seguir, serão apresentados os valores de desvio padrão do alfa para o período de inverno austral. Na Figura 4.26 é mostrada a distribuição horizontal do desvio padrão do alfa da pressão em superfície para o período de estudo. Notam-se os maiores valores de desvio padrão sobre os oceanos, e principalmente na faixa entre aproximadamente 60°S e 30°S, região essa onde tem a ocorrência de sistemas transientes, e que, ao observarmos a Figura 4.11, notam-se os maiores valores de alfa.

Os valores de desvio padrão para essa variável chegaram até 0,25 sobre o Oceano Atlântico Sul. Ao compararmos com a mesma figura, porém, para o trimestre DJF, nota-se que o atual período, ainda que atingindo uma pequena região o valor de 0,25, apresentou uma estabilidade maior com relação aos valores de alfa, mostrando valores menores ao longo do globo.

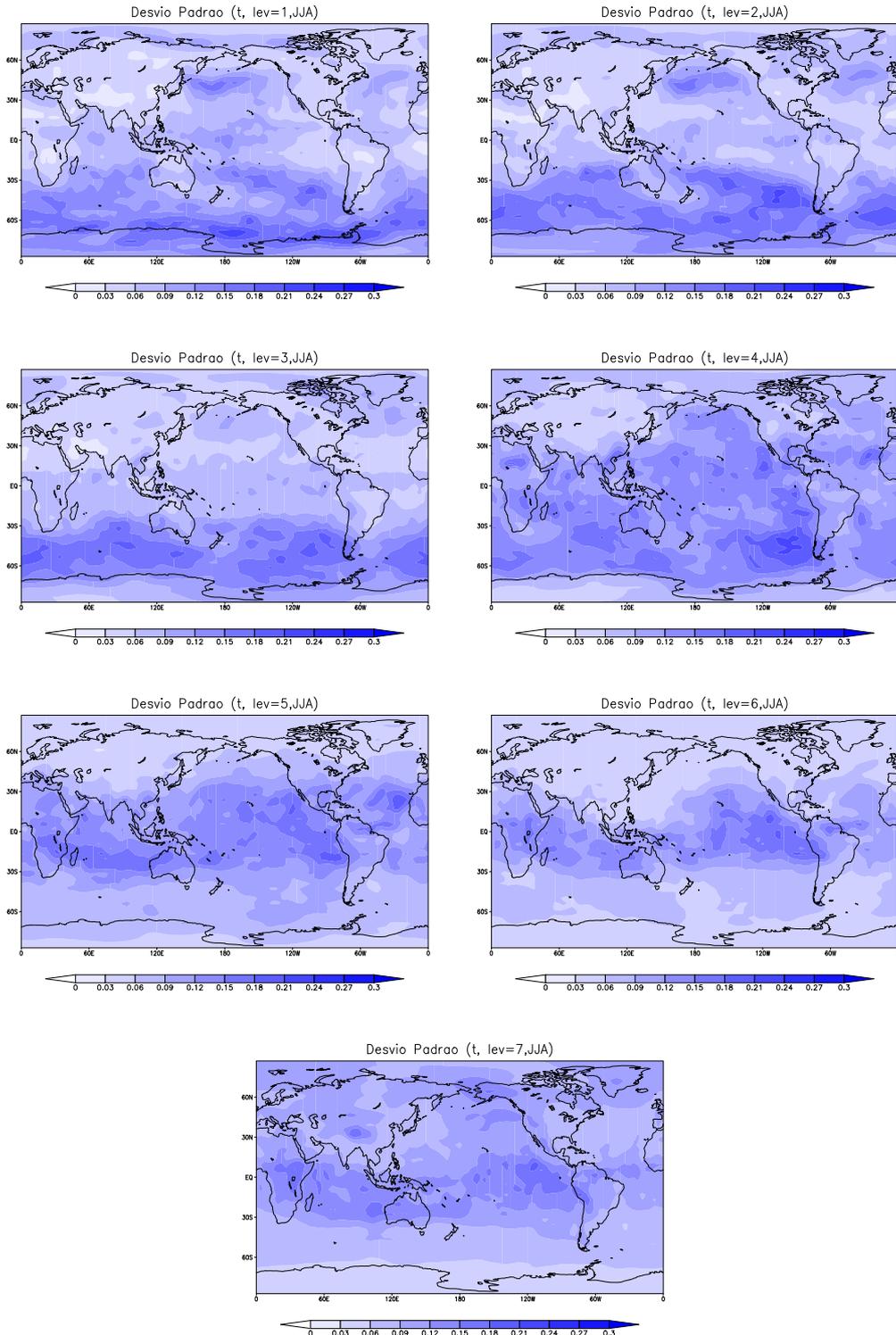
Figura 4.26 - Desvio padrão do alfa da pressão em superfície para o período de junho, julho e agosto.



Fonte: Produção do autor.

Na Figura 4.27 são mostrados os valores de desvio padrão para os sete níveis da temperatura. Assim como visto para a pressão em superfície, aqui também nota-se uma relação entre os valores elevados de alfa e os valores mais altos de desvio padrão. Pode-se observar até o terceiro nível do modelo que os valores mais elevados são vistos no Hemisfério Sul, chegando a 0,25 nesse região. À medida que sobe na atmosfera, os maiores valores são vistos na região Tropical e Hemisfério Norte, porém, alcançando apenas 0,2. Pode-se observar que há uma grande variação nos valores de alfa em regiões onde a atmosfera também varia bastante: foram observados os maiores valores em baixos níveis que é uma região onde há a ocorrência de sistemas transientes, e a região apontada em médios e altos níveis sofre a influência da ZCIT.

Figura 4.27 - Idem à Figura 4.26, porém, para a temperatura nos sete níveis do modelo.



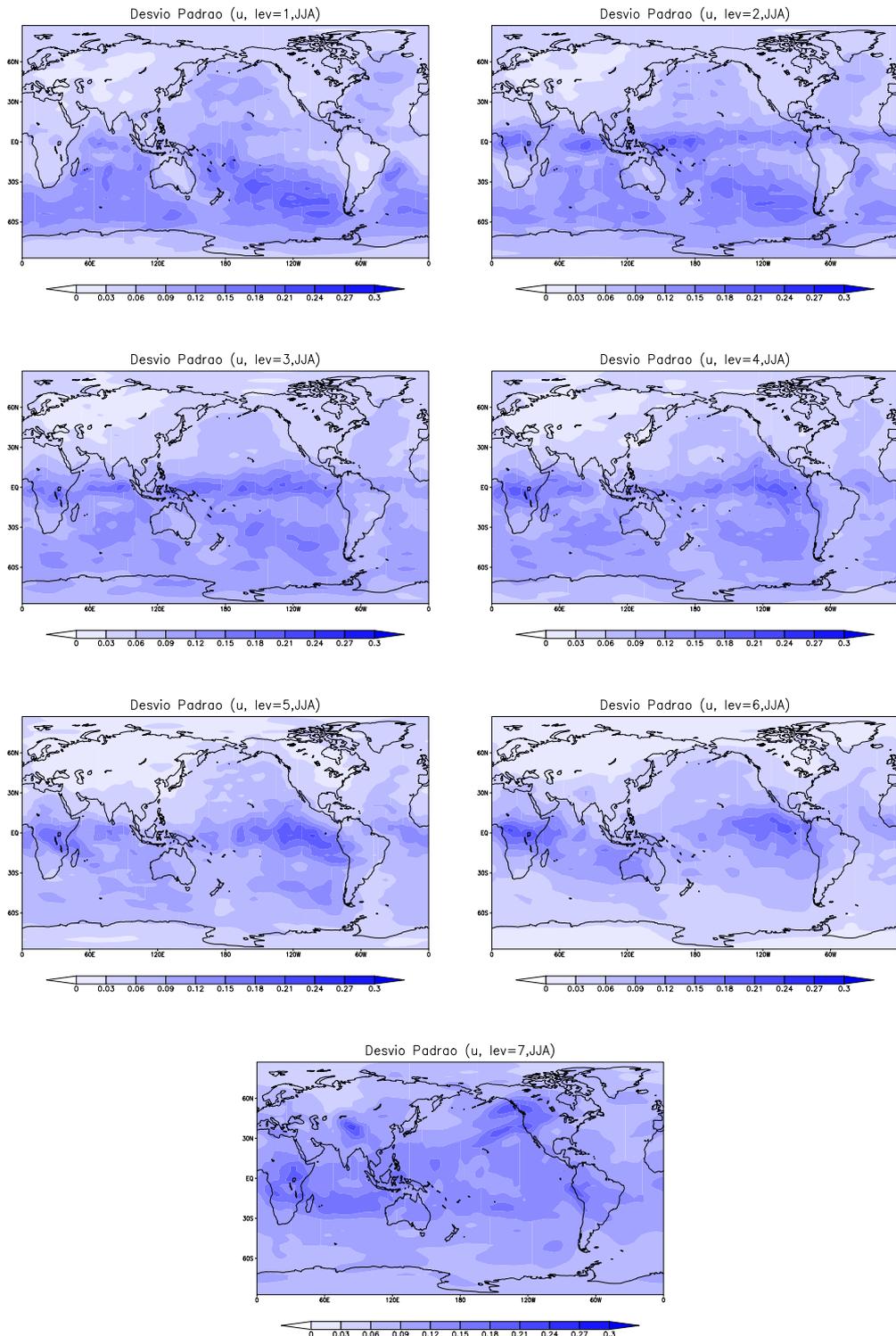
Fonte: Produção do autor.

Os valores de desvio padrão para a componente zonal do vento nos sete níveis são mostrados na Figura 4.28. Nela, pode-se observar valores menores do que os vistos para o período de dezembro, janeiro e fevereiro, porém, a relação já vista entre altos valores de alfa com valores mais elevados de desvio padrão continua. Pode-se observar, por exemplo, valores elevados de desvio padrão sobre a região Tropical nos níveis 2 e 3, e em altos níveis sobre o continente africano, Ásia, oeste da América do Sul e oeste dos Estados Unidos. Tais regiões também apresentaram valores elevados de alfa (Figura 4.15). Já para a componente meridional do vento (apresentada na Figura 4.29), os maiores valores de desvio padrão ficaram localizados no Hemisfério Sul, exceto no nível 7, no qual os valores mais elevados ficaram localizados nas mesmas regiões que o vento zonal: Africa, Ásia, oeste da América do Sul e oeste dos Estados Unidos. Assim como as demais variáveis, o desvio padrão da componente meridional também teve forte relação com os valores de alfa.

Na Figura 4.30 destaca-se a distribuição horizontal do desvio padrão da umidade para os seis níveis do modelo. Nela, é possível observar uma coerência com os valores vistos de alfa (Figura 4.19), apresentando valores elevados de desvio padrão onde foram vistos valores elevados de alfa. Os valores de desvio padrão nos seis níveis não ultrapassaram 0, 2.

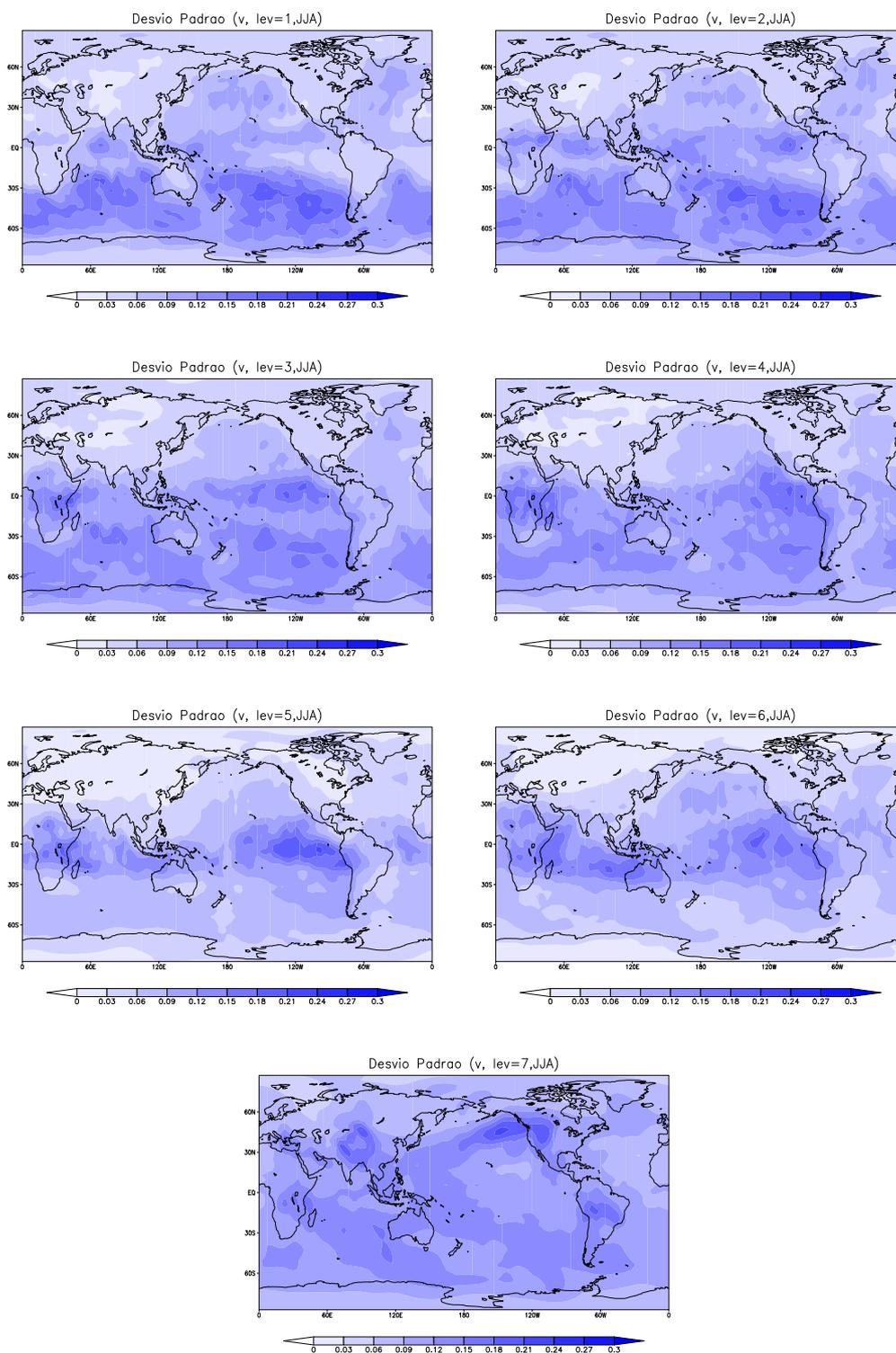
Aqui foram mostradas as distribuições horizontais do desvio padrão do alfa para o período de junho, julho e agosto. Observou-se uma forte relação entre os valores médios de alfa e os valores do desvio padrão. Também notou-se que, em regiões onde o alfa apresenta valores altos, o desvio padrão também apresenta valores elevados em relação as demais regiões. Por consequência, se observou uma maior variação do alfa em regiões onde são vistos alguns sistemas sinóticos, como, por exemplo, ciclones, ZCIT e sistemas de alta pressão, além de regiões com orografia acentuada. Por outro lado, em regiões onde a atmosfera é mais estável, os valores de alfa variaram pouco, mostrando uma estabilidade maior.

Figura 4.28 - Idem à Figura 4.26, porém, para o vento zonal nos sete níveis do modelo.



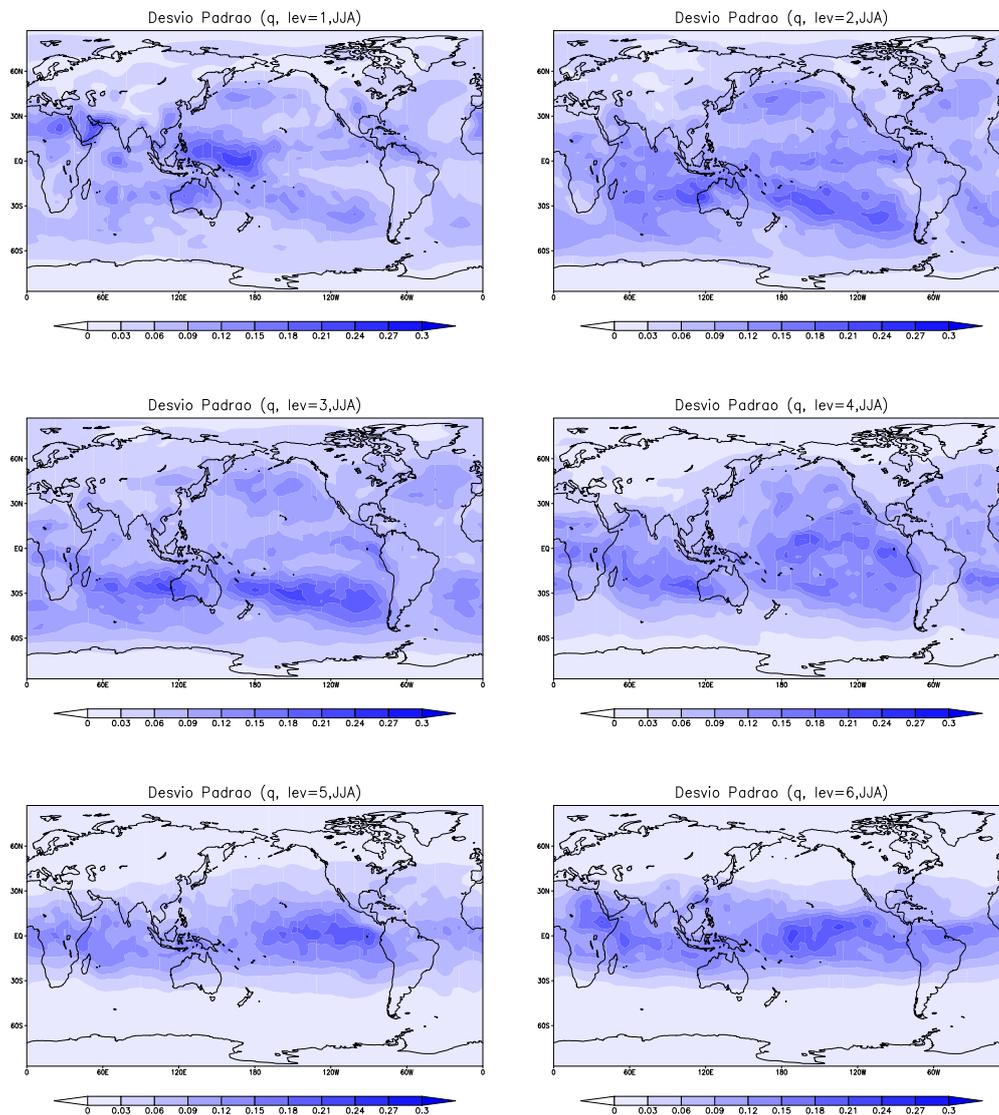
Fonte: Produção do autor.

Figura 4.29 - Idem à Figura 4.26, porém, para o vento meridional nos sete níveis do modelo.



Fonte: Produção do autor.

Figura 4.30 - Idem à Figura 4.26, porém, para a umidade nos seis níveis do modelo.



Fonte: Produção do autor.



## 5 RESULTADOS

Neste capítulo são apresentados os resultados obtidos a partir dos experimentos descritos no Capítulo 3 utilizando o LETKF, o 3DVar e o Híbrido Médio. A Seção 5.1 apresenta a performance do sistema por conjunto e do sistema variacional. A Seção 5.2 traz uma comparação entre as análises geradas por todos os experimentos de alfa fixo e alfa dinâmico. Em seguida, na Seção 5.3, o balanço das análises é avaliado, e, na Subseção 5.4.2 faz-se a comparação entre as previsões geradas pelo experimento com  $\alpha = 0, 1$  e com alfa dinâmico. Além disso, na Seção 5.5 são mostrados resultados em relação a sensibilidade ao número de membros, e, por fim, na Seção 5.6 são mostrados resultados em relação ao intervalo do alfa dinâmico.

### 5.1 Performance do LETKF e do 3DVar

Em Yang et al. (2009), a performance do LETKF foi avaliada em comparação com 3DVar utilizando um modelo quasigeostrófico, enquanto no trabalho de Miyoshi (2005) as performances do EnSRF e do LEKF foram avaliadas contra a performance do 3DVar no modelo SPEEDY. Em ambos os trabalhos concluiu-se que o sistema baseado em conjunto apresenta um melhor desempenho que o sistema variacional em três dimensões. Com base nesses resultados, realizou-se uma avaliação para atestar que os sistemas de assimilação de dados que fizeram parte do Híbrido Médio apresentam resultados esperados conforme a literatura revela.

Os experimentos foram avaliados utilizando como métrica a raiz do erro médio quadrático (do inglês *Root Mean Square Error* - RMSE) do incremento de análise (O-A), para as 5 variáveis ( $ps$ ,  $T$ ,  $u$ ,  $v$  e  $q$ ) e 7 níveis do modelo (exceto para a pressão em superfície, a qual tem valores apenas no primeiro nível do modelo, e a umidade, que não possui informações no último nível do modelo). Através do incremento de análise, pode-se observar o quão próximo da "verdade" a análise gerada conseguiu aproximar, sendo que aqui, a *Nature Run* é considerada como a "verdade". Foram avaliadas quatro regiões: Globo, Hemisfério Norte (compreendendo a faixa entre 20°N e 80°N), Hemisfério Sul (compreendendo a faixa entre 80°S e 20°S) e Trópicos (compreendendo a faixa entre 20°S e 20°N).

Os resultados para o trimestre DJF são apresentados nas Tabelas 5.1, 5.2, 5.3 e 5.4 para o Globo, Hemisfério Norte, Hemisfério Sul e Trópicos, respectivamente. É possível observar para as 4 regiões de estudo que em todas as variáveis e em todos os níveis o LETKF (representado nas tabelas como EKF) apresentou valores de RMSE menores do que os apresentados pelo 3DVar (nas tabelas, representado como VAR),

mostrando, assim, uma coerência com o esperado pela literatura.

Tabela 5.1 - RMSE médio da pressão em superfície ( $ps$ ), temperatura ( $T$ ), componente zonal e meridional do vento ( $u$  e  $v$ ) e umidade ( $q$ ) para o Globo nos sete níveis do modelo para o trimestre DJF.

	$ps$		$T$		$u$		$v$		$q (\times 10^{-4})$	
	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR
1	28,5	45,8	0,13	0,20	0,40	0,67	0,40	0,71	1,14	1,72
2			0,15	0,24	0,42	0,73	0,40	0,74	1,21	1,20
3			0,17	0,28	0,46	0,80	0,44	0,82	0,86	1,49
4			0,23	0,36	0,61	1,08	0,60	1,13	0,85	1,42
5			0,24	0,38	0,80	1,46	0,83	1,54	0,62	1,07
6			0,26	0,42	0,81	1,66	0,83	1,66	0,22	0,38
7			0,27	0,47	0,36	0,78	0,36	0,83		

Tabela 5.2 - Idem à Tabela 5.1, porém, para o Hemisfério Norte.

	$ps$		$T$		$u$		$v$		$q (\times 10^{-4})$	
	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR
1	22.83	38.02	0,11	0,16	0,28	0,53	0,28	0,60	0,60	0,95
2			0,13	0,21	0,27	0,54	0,28	0,61	0,67	1,15
3			0,14	0,23	0,28	0,57	0,30	0,65	0,57	0,94
4			0,16	0,26	0,38	0,72	0,40	0,82	0,41	0,69
5			0,17	0,27	0,46	0,88	0,50	1,00	0,22	0,39
6			0,17	0,30	0,45	0,90	0,47	1,00	0,09	0,16
7			0,19	0,36	0,27	0,58	0,28	0,66		

Tabela 5.3 - Idem à Tabela 5.1, porém, para o Hemisfério Sul.

	$ps$		$T$		$u$		$v$		$q (\times 10^{-4})$	
	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR
1	32,71	53,03	0,14	0,23	0,42	0,75	0,43	0,83	1,00	1,57
2			0,17	0,30	0,38	0,72	0,38	0,82	1,06	1,90
3			0,17	0,30	0,39	0,74	0,40	0,86	0,82	1,43
4			0,23	0,38	0,52	0,94	0,55	1,08	0,69	1,17
5			0,23	0,37	0,67	1,23	0,73	1,42	0,37	0,65
6			0,23	0,37	0,68	1,40	0,73	1,60	0,11	0,18
7			0,23	0,43	0,31	0,84	0,33	0,98		

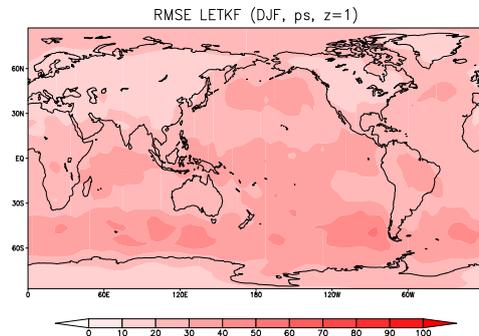
Tabela 5.4 - Idem à Tabela 5.1, porém, para a região Tropical.

	$ps$		$T$		$u$		$v$		$q (\times 10^{-4})$	
	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR
1	29.19	45.63	0,13	0,20	0,47	0,71	0,46	0,69	1,58	2.34
2			0,13	0,21	0,56	0,90	0,50	0,79	1,67	2.64
3			0,19	0,30	0,62	1,01	0,57	0,92	1,11	1,91
4			0,27	0,42	0,83	1,43	0,79	1,39	1,22	2.04
5			0,30	0,47	1,12	2.00	1,13	2.01	0,96	1,69
6			0,34	0,55	1,14	2.32	1,13	2.15	0,34	0,60
7			0,36	0,58	0,47	0,88	0,45	0,82		

As distribuições espaciais dos erros são mostradas nas Figuras 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9 e 5.10. São erros de cada um dos sistemas de assimilação (LETKF e 3DVar) para cada nível e variável do modelo. Nessas figuras, regiões com valores próximos do zero estão em tons mais claros, e regiões que apresentam um vermelho mais intenso estão com valores de RMSE maiores.

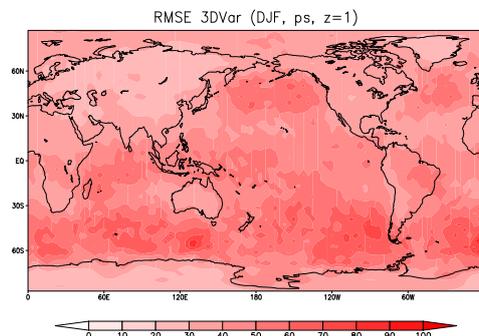
Os erros da pressão em superfície para ambos os sistemas de assimilação de dados são apresentados nas Figuras 5.1 e 5.2. Nelas, é possível observar não somente sua localização, mas também a grande diferença entre os erros de cada sistema. Para o caso do LETKF (Figura 5.1), os valores de RMSE chegam a 50 Pa, enquanto os erros vistos no 3DVar chegam a 100 Pa. A localização dos erros é aparentemente similar: ambos os sistemas apresentam seus maiores erros sobre os oceanos, ficando esses erros mais restritos ao Hemisfério Sul. Esse resultado ajuda a reafirmar o que foi mostrado nas tabelas anteriores, nas quais os maiores valores de erro na pressão em superfície foram observados no Hemisfério Sul. Entretanto, como já mostrado nesse estudo, o 3DVar apresenta erros maiores. Esses erros possivelmente estão relacionados aos sistemas sinóticos que atuam nessas regiões, como, por exemplo, os ciclones nos Oceanos Pacífico e Atlântico Norte, bem como a região dos ciclones transientes no Hemisfério Sul.

Figura 5.1 - Distribuição espacial do RMSE médio da pressão em superfície ao longo do trimestre DJF para as análises geradas pelo LETKF.



Fonte: Produção do autor.

Figura 5.2 - Distribuição espacial do RMSE médio da pressão em superfície ao longo do trimestre DJF para as análises geradas pelo 3DVar.



Fonte: Produção do autor.

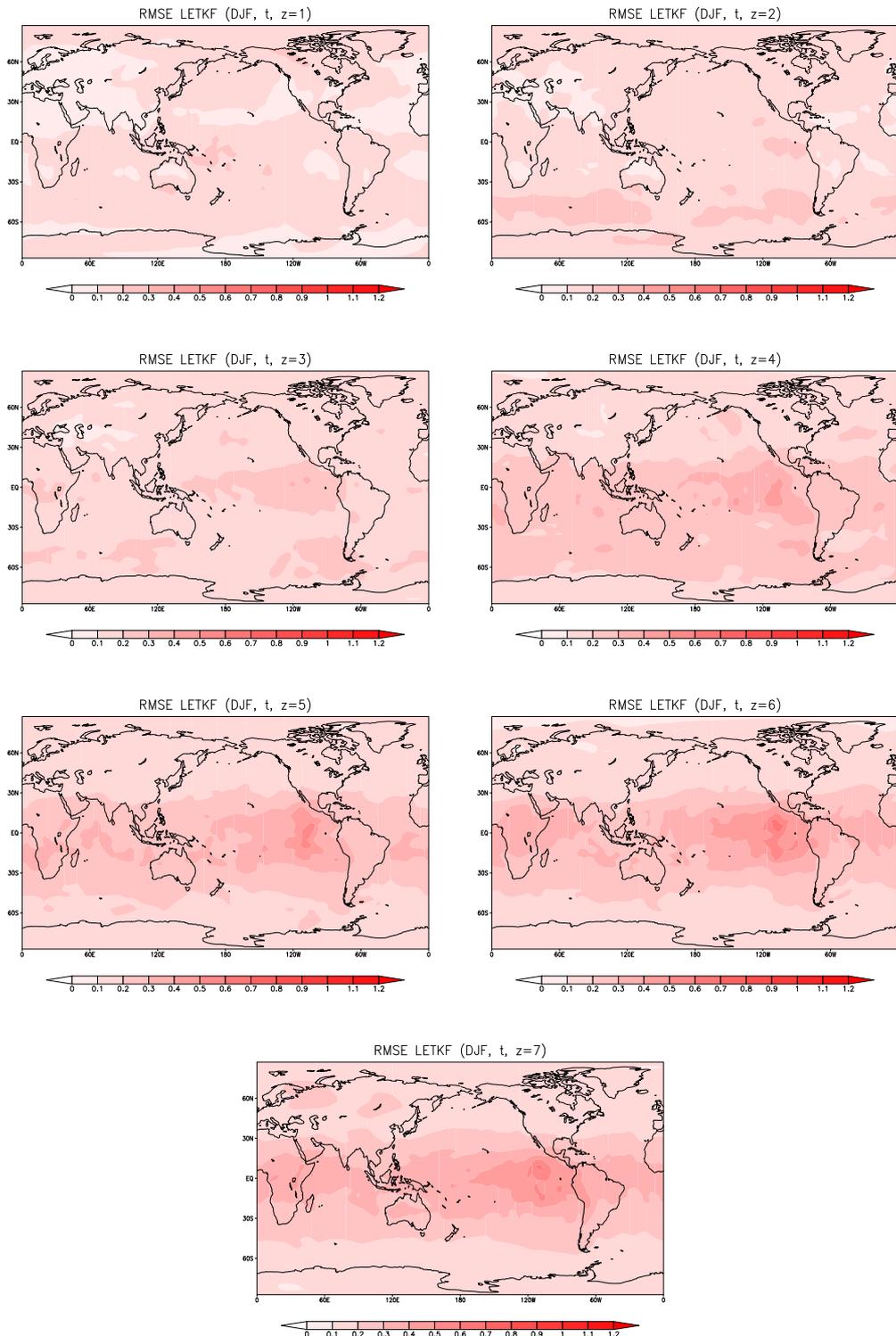
As Figuras 5.3 e 5.4 mostram os erros para a temperatura nos sete níveis do modelo. Para ambos os casos nota-se um aumento no erro com aumento da altura na atmosfera. Esses erros, assim como mostrado em relação à pressão em superfície, são maiores para o 3DVar do que para o LETKF. Os erros do LETKF até o quarto nível do modelo ficaram concentrados mais na região Tropical e Hemisfério Sul; já a partir do quinto nível os erros passam a se concentrar na faixa entre 30°S e 30°N. Para o caso do 3DVar, além de serem maiores, os erros se apresentam de forma mais

espalhada no mapa, e as regiões com erros maiores, a partir também do quinto nível do modelo, ficaram na mesma faixa que os erros do LETKF, entre 30°S e 30°N. Esses erros mencionados em baixos níveis, assim como para a pressão em superfície, podem ter relação com a estrutura dos ciclones transientes no Hemisfério Sul e, no caso do 3DVar, com erros nas regiões dos ciclones do Hemisfério Norte. Já os erros apresentados por ambos na região tropical em médios e altos níveis podem estar relacionados à ZCIT, na qual esses sistemas podem não representar bem os processos relacionados à convecção nessa região.

Os erros para a componente zonal do vento são apresentados nas Figuras 5.5 e 5.6. Eles continuam apresentando valores maiores de erro para o 3DVar do que para o LETKF. Para os dois sistemas nota-se que o erro ficou concentrado na região Tropical e que o RMSE foi aumentando à medida que aumenta o nível. Contudo, para ambos o último nível apresentou erros comparáveis ao primeiro nível. Notou-se também duas regiões nos dois sistemas em que o erro aumenta significativamente sobre o continente africano e sobre o Oceano Pacífico Leste. Esses erros possivelmente estão relacionados à ZCIT. Como a temperatura não está bem representada nessa mesma região, é esperado que o campo de vento também não esteja representado de forma correta. Assim como ocorreu na componente zonal do vento, o mesmo é visto na componente meridional (Figuras 5.7 e 5.8) com relação aos erros no primeiro e último nível, assim como os menores erros foram vistos utilizando o LETKF e os maiores utilizando o 3DVar. Para o LETKF, os maiores erros ficaram localizados na faixa entre 30°S e 30°N; já para o 3DVar, os erros são vistos nessa mesma faixa e cobrindo parte dos oceanos no Hemisfério Sul.

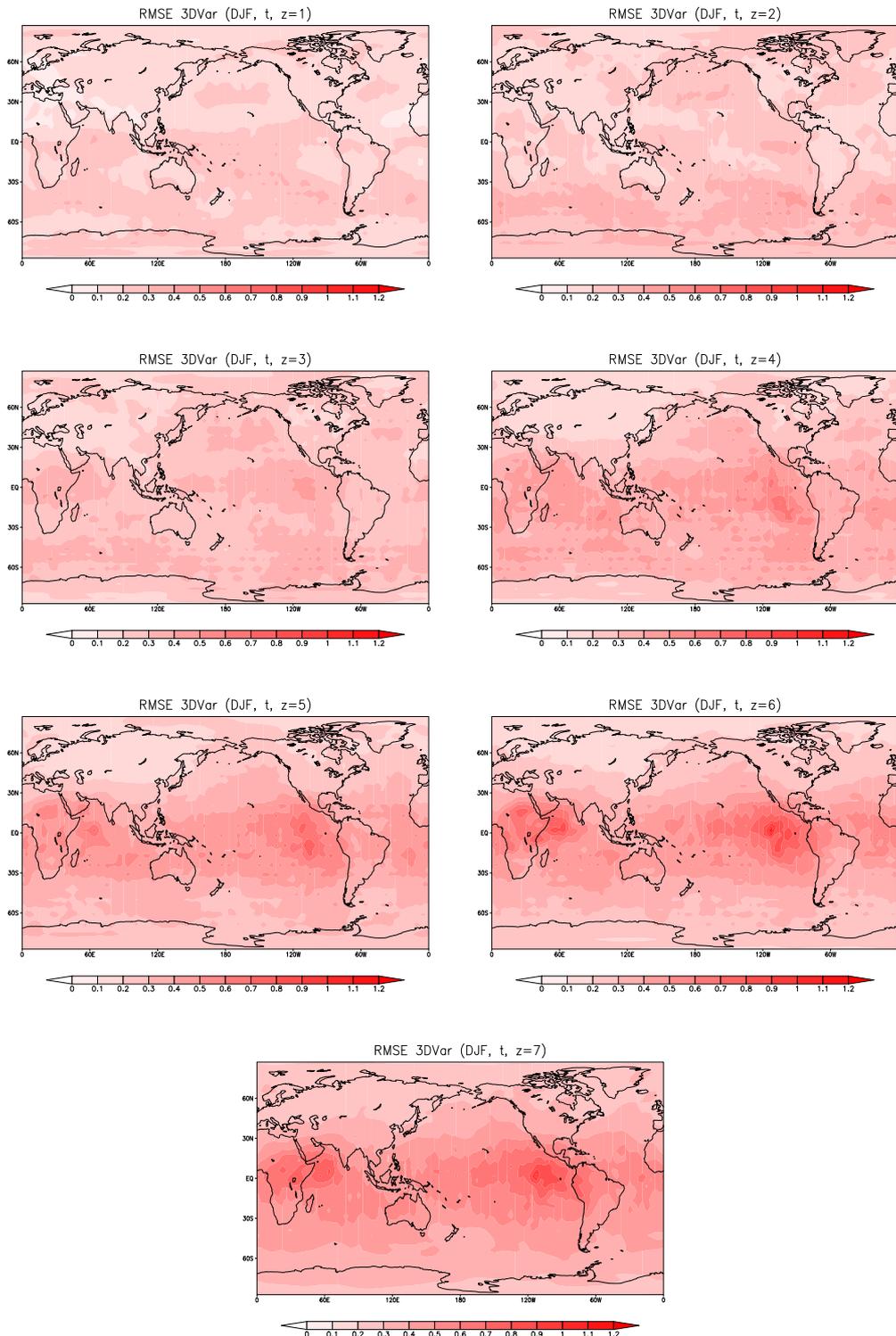
Os erros dos campos de umidade são apresentados nas Figuras 5.9 e 5.10. Para os dois casos, observam-se os maiores erros no primeiro nível do modelo, chegando o LETKF a 0,0005 kg/kg e o 3DVar a valores superiores a isso. Nota-se ainda que os erros do 3DVar abrangem uma área maior do que os erros do LETKF, alcançando até 60°S. Isso já é esperado, uma que vez os maiores valores de umidade estão localizados na região Tropical.

Figura 5.3 - Idem à Figura 5.1, porém, para a temperatura nos sete níveis do modelo.



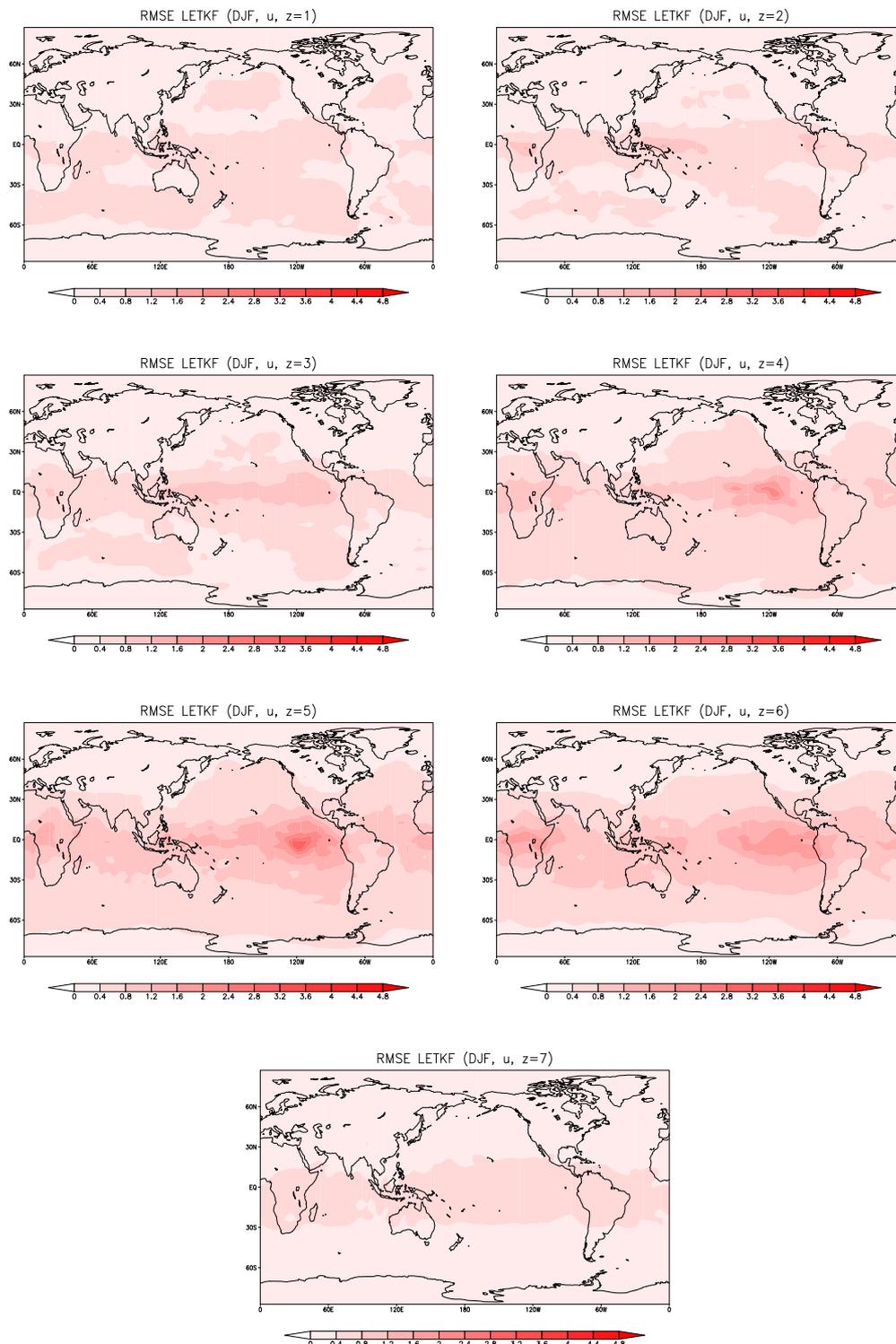
Fonte: Produção do autor.

Figura 5.4 - Idem à Figura 5.2, porém, para a temperatura nos sete níveis do modelo.



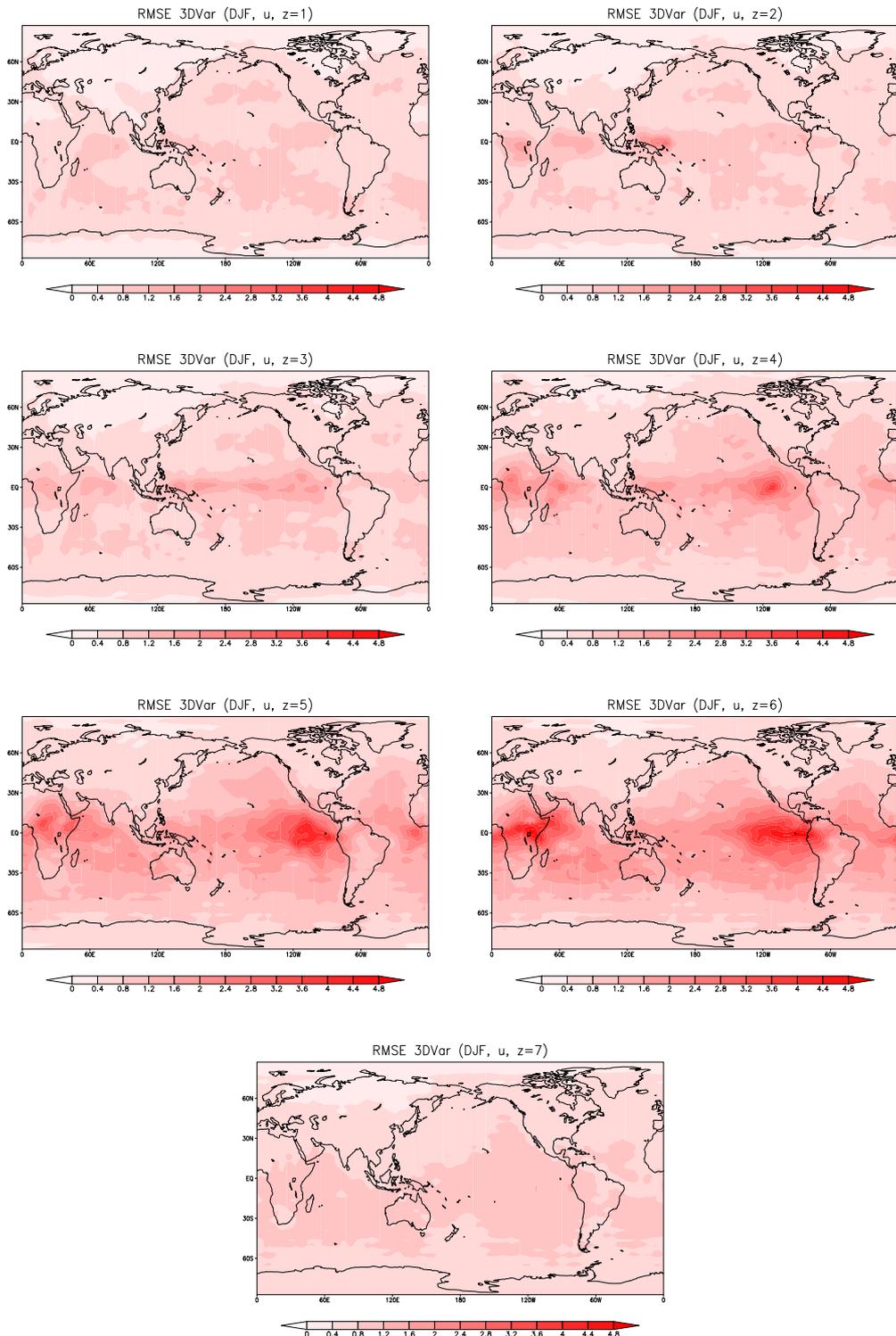
Fonte: Produção do autor.

Figura 5.5 - Idem à Figura 5.1, porém, para o vento zonal nos sete níveis do modelo.



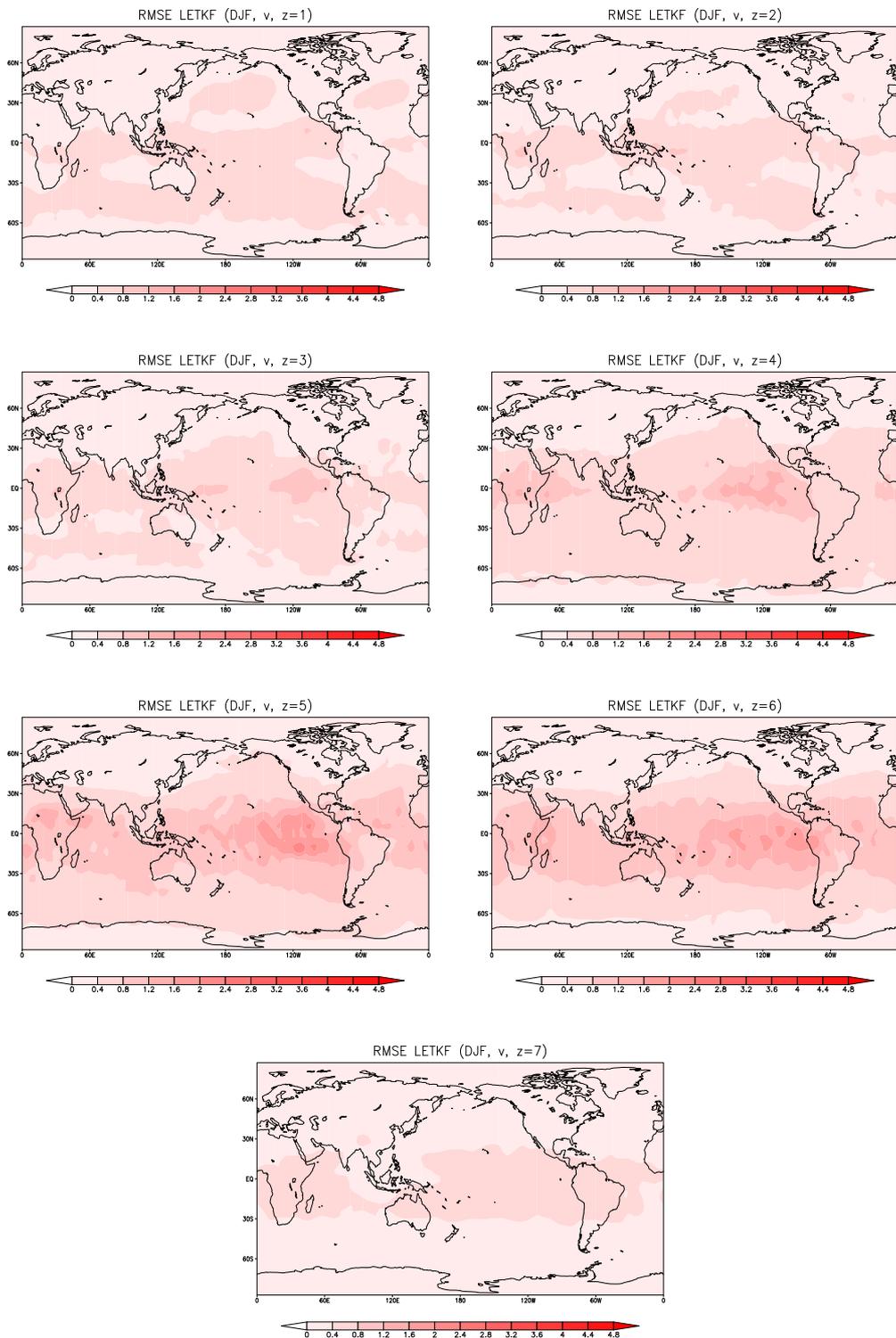
Fonte: Produção do autor.

Figura 5.6 - Idem à Figura 5.2, porém, para o vento zonal nos sete níveis do modelo.



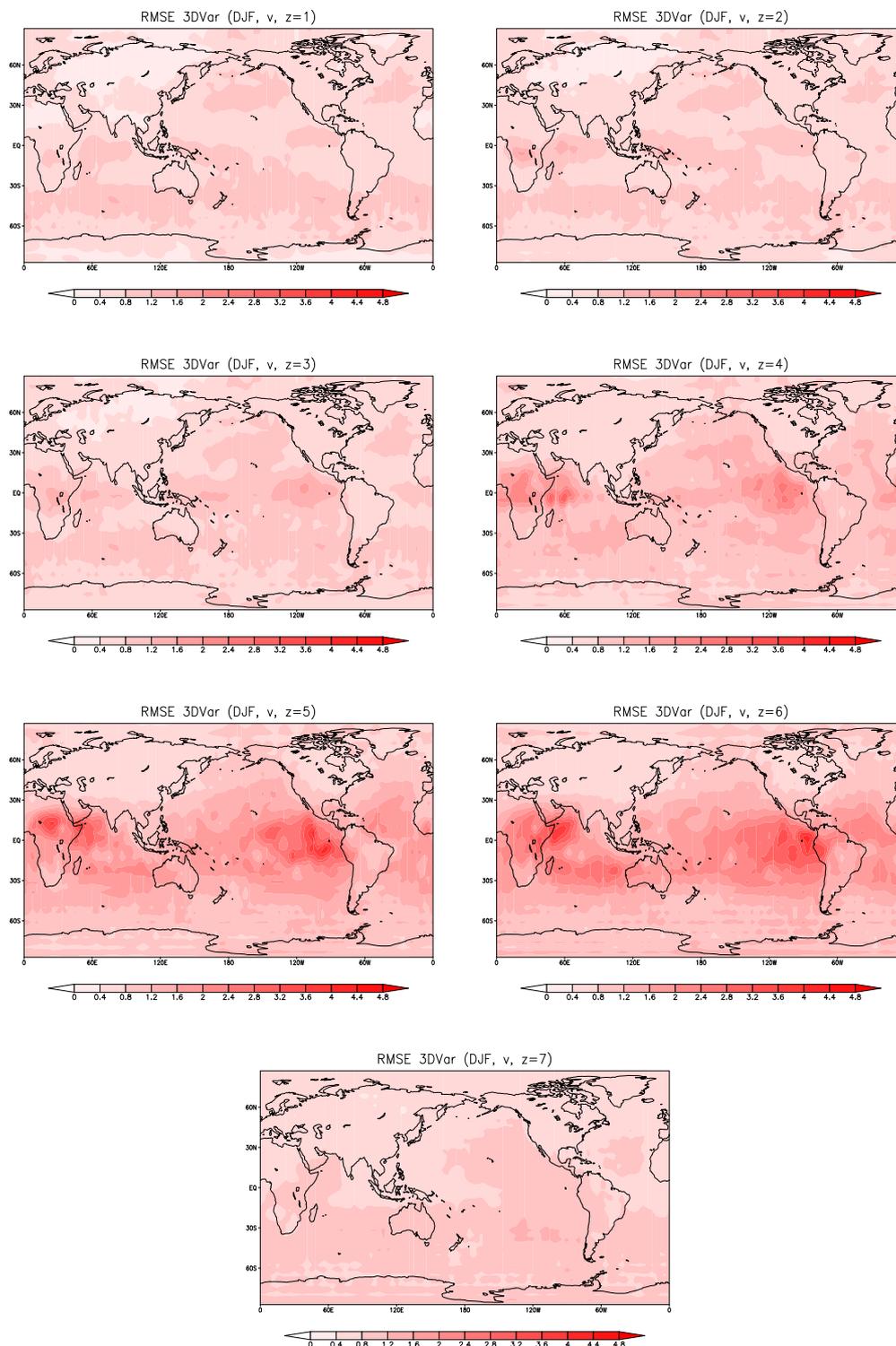
Fonte: Produção do autor.

Figura 5.7 - Idem à Figura 5.1, porém, para o vento meridional nos sete níveis do modelo.



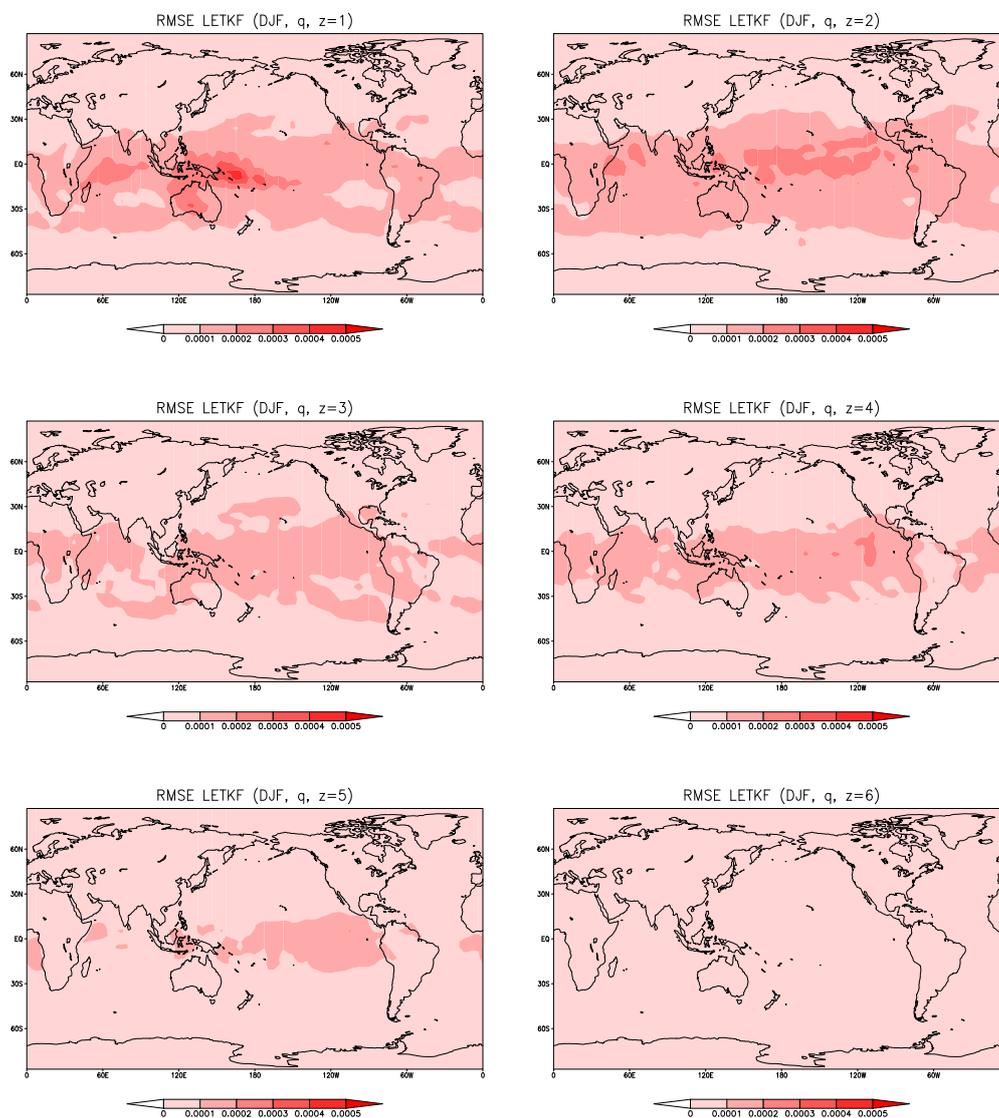
Fonte: Produção do autor.

Figura 5.8 - Idem à Figura 5.2, porém, para o vento meridional nos sete níveis do modelo.



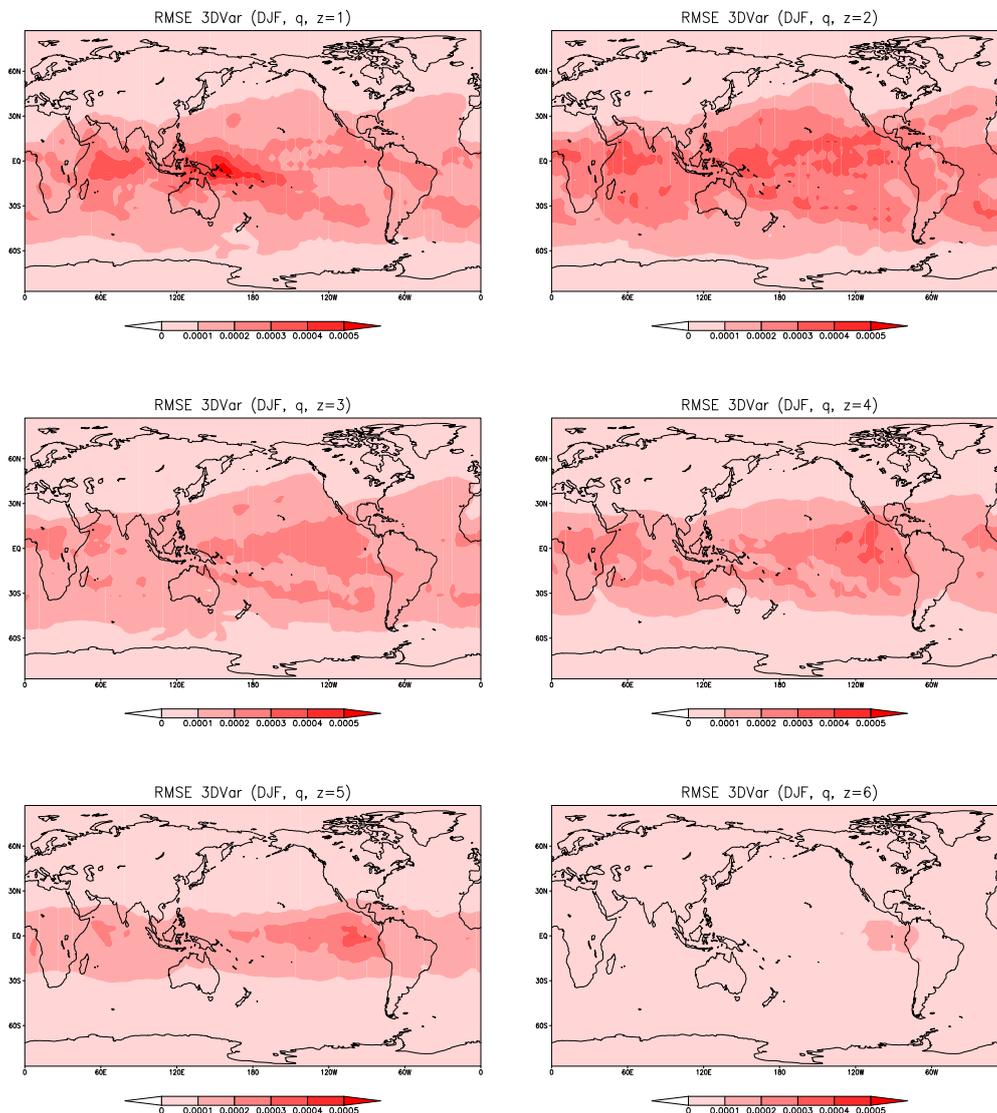
Fonte: Produção do autor.

Figura 5.9 - Idem à Figura 5.1, porém, para a umidade nos seis níveis do modelo.



Fonte: Produção do autor.

Figura 5.10 - Idem à Figura 5.2, porém, para a umidade nos seis níveis do modelo.



Fonte: Produção do autor.

Com esses resultados, foi possível observar a localização dos erros médios de cada um dos sistemas de assimilação de dados envolvidos no Híbrido Médio para o período de dezembro, janeiro e fevereiro. Foram observados valores mais elevados de RMSE para o 3DVar, corroborando com os resultados nas tabelas mostradas anteriormente. Esses erros se concentraram mais sobre regiões oceânicas, sobre baixas latitudes e sobre o Hemisfério Sul. Por se tratarem de regiões com uma vasta área oceânica, elas estão relacionadas às regiões onde se tem um menor número de observações. Além disso, são regiões onde são vistos determinados sistemas sinóticos atuando (como,

por exemplo, sistemas de baixa pressão) e, com essa precariedade de observações, o sistema de assimilação de dados acaba por não ter uma "referência" para corrigir a trajetória do modelo.

Os resultados apresentados até então estão relacionados aos experimentos realizados no período do verão austral. A seguir, serão apresentados alguns resultados referentes as performances do LETKF e do 3DVar para o período de junho, julho e agosto. As Tabelas 5.5, 5.6, 5.7 e 5.8 apresentam valores de RMSE para todas as variáveis e níveis do modelo para as quatro regiões de estudo: Globo, Hemisfério Norte, Hemisfério Sul e região Tropical, respectivamente. Nelas pode-se observar que o padrão visto para o período de verão austral se mantém, ou seja, o LETKF apresenta erros menores que o 3DVar para as quatro regiões. Esses erros, em sua grande maioria, são maiores para o período de inverno do que para o período de verão em ambos os sistemas.

Tabela 5.5 - RMSE médio da pressão em superfície ( $ps$ ), temperatura ( $T$ ), componente zonal e meridional do vento ( $u$  e  $v$ ) e umidade ( $q$ ) para o Globo nos sete níveis do modelo para o trimestre JJA.

	$ps$		$T$		$u$		$v$		$q (\times 10^{-4})$	
	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR
1	30,07	40,08	0,13	0,21	0,42	0,71	0,42	0,75	1,14	1,78
2			0,15	0,25	0,45	0,78	0,42	0,78	1,24	1,21
3			0,18	0,30	0,47	0,83	0,46	0,86	0,90	1,56
4			0,23	0,38	0,64	1,19	0,66	1,25	0,87	1,49
5			0,27	0,42	0,88	1,75	0,95	1,96	0,67	1,22
6			0,28	0,48	0,89	1,90	0,90	1,94	0,26	0,48
7			0,30	0,50	0,38	0,77	0,37	0,81		

Tabela 5.6 - Idem à Tabela 5.5, porém, para o Hemisfério Norte.

	$ps$		$T$		$u$		$v$		$q (\times 10^{-4})$	
	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR
1	23,78	38,76	0,09	0,15	0,28	0,54	0,28	0,60	0,86	1,26
2			0,10	0,19	0,26	0,54	0,26	0,61	0,80	1,30
3			0,11	0,21	0,26	0,56	0,26	0,63	0,61	1,00
4			0,17	0,27	0,34	0,69	0,34	0,77	0,54	0,83
5			0,19	0,29	0,48	0,89	0,49	0,97	0,33	0,55
6			0,19	0,30	0,51	0,94	0,52	1,02	0,12	0,20
7			0,21	0,36	0,30	0,58	0,31	0,66		

Tabela 5.7 - Idem à Tabela 5.5, porém, para o Hemisfério Sul.

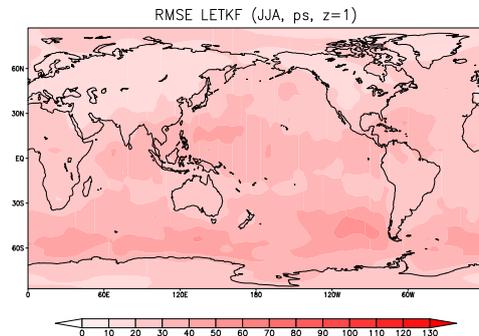
	$ps$		$T$		$u$		$v$		$q (\times 10^{-4})$	
	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR
1	34,72	56,49	0,15	0,24	0,45	0,79	0,45	0,87	0,84	1,41
2			0,18	0,31	0,42	0,76	0,42	0,85	0,94	1,79
3			0,21	0,34	0,44	0,80	0,45	0,91	0,80	1,45
4			0,23	0,40	0,58	1,06	0,61	1,19	0,60	1,10
5			0,23	0,40	0,73	1,43	0,76	1,56	0,29	0,56
6			0,23	0,42	0,68	1,54	0,71	1,71	0,12	0,23
7			0,27	0,46	0,34	0,83	0,35	0,96		

Tabela 5.8 - Idem à Tabela 5.5, porém, para a região Tropical.

	$ps$		$T$		$u$		$v$		$q (\times 10^{-4})$	
	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR	EKF	VAR
1	30,82	47,73	0,14	0,22	0,49	0,78	0,49	0,76	1,57	2,42
2			0,14	0,23	0,59	0,96	0,53	0,85	1,76	2,85
3			0,20	0,33	0,63	1,05	0,60	1,00	1,19	2,04
4			0,28	0,45	0,88	1,62	0,89	1,63	1,26	2,17
5			0,33	0,53	1,25	2,50	1,37	2,84	1,06	1,93
6			0,38	0,65	1,28	2,73	1,28	2,67	0,41	0,77
7			0,39	0,63	0,47	0,87	0,44	0,79		

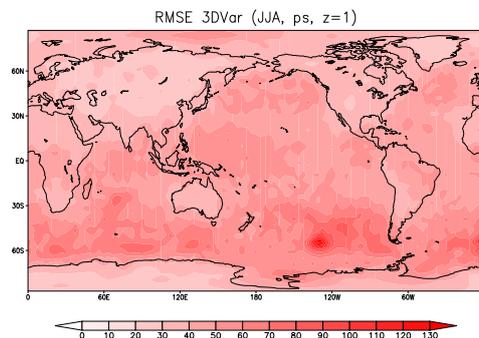
Nas Figuras 5.11 e 5.12 são apresentadas as distribuições espaciais dos erros de pressão em superfície para o LETKF e para o 3DVar, respectivamente. Pode-se observar erros maiores para o 3DVar, alcançando até 130 Pa sobre o Oceano Pacífico. Além disso, nota-se que os erros para o período de inverno foram maiores do que para o período de verão. Enquanto que para o trimestre DJF os erros de pressão em superfície foram até 100 Pa, no trimestre de JJA esses erros aumentaram em 30 Pa, ainda que bem localizado. Os maiores valores de RMSE foram vistos sobre os oceanos e especialmente os erros do 3DVar apresentaram uma área sobre o Oceano Pacífico Sul com valores elevados de erro. As demais variáveis não são mostradas aqui, pois a distribuição espacial de seus erros mantém o mesmo padrão do período de verão, apresentando, como visto já nas Tabelas, erros maiores para esse período e erros maiores para o 3DVar.

Figura 5.11 - Distribuição espacial do RMSE médio da pressão em superfície ao longo do trimestre JJA para as análises geradas pelo LETKF.



Fonte: Produção do autor.

Figura 5.12 - Distribuição espacial do RMSE médio da pressão em superfície ao longo do trimestre JJA para as análises geradas pelo 3DVar.



Fonte: Produção do autor.

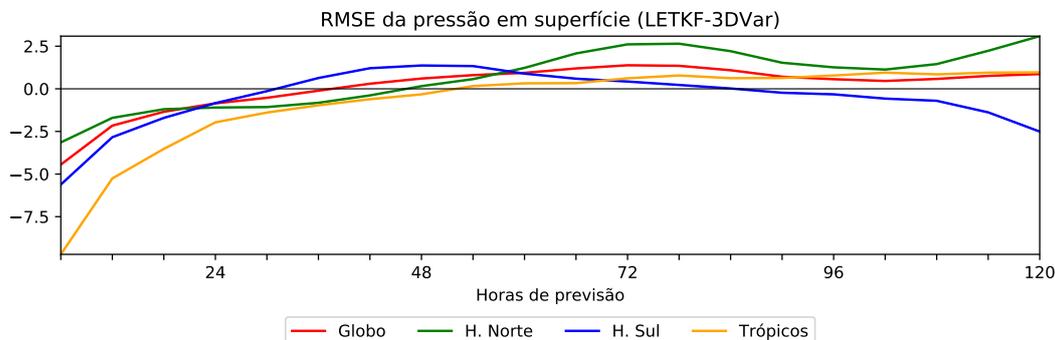
Pode-se observar, em ambos os períodos, erros maiores para o 3DVar. Os erros, tanto para o 3DVar quanto para o LETKF, ficaram localizados mais sobre o Hemisfério Sul e região Tropical, acompanhando sistemas sinóticos atuantes na região, como ciclones e ZCIT. Os erros foram maiores durante o período de inverno austral. Além disso, observou-se uma relação com as regiões carentes de observações. Um exemplo disso é a falta de observações de pressão em superfície sobre os oceanos, fazendo com que, para essa variável, o modelo não consiga ser corrigido nessas regiões. De acordo

com Miyoshi (2005), o desempenho do 3DVar está relacionado mais diretamente à densidade de observações do que os erros do LETKF. Vale ressaltar a diferença principal desses sistemas: o LETKF possui uma matriz de covariância do erro do *background* a qual é atualizada a cada novo ciclo de assimilação de dados, incluindo o que chamamos de "erros do dia" do modelo, enquanto o sistema variacional em três dimensões possui sua matriz estática, o que torna o sistema 3DVar mais estável. De acordo com Yang et al. (2009), além desse motivo, existe também o fato de que o LETKF utiliza a média de um conjunto de possibilidades, melhorando assim o seu desempenho sobre o sistema variacional em situações de carência de dados.

Com a finalidade de avaliar o comportamento das previsões geradas pelas análises do sistema variacional e do sistema por conjunto, previsões de até 120 horas foram geradas, com saídas a cada 6 horas para ambos os experimentos. Essas previsões foram avaliadas através das diferenças dos valores de RMSE previamente calculados (LETKF - 3DVar). Aqui serão mostrados os resultados apenas para a pressão em superfície, uma vez que o comportamento nos resultados das demais variáveis se assemelha aos dessa. Na Figura 5.13 são mostrados os resultados para o trimestre DJF. Já na Figura 5.14 são mostrados os resultados para o trimestre JJA. Nelas as linhas em vermelho representam o Globo, em verde o Hemisfério Norte, em azul o Hemisfério Sul e, em laranja os Trópicos. Nessas figuras, quando os valores forem negativos, significa que o LETKF apresentou erros menores que o 3DVar e, quando os valores forem positivos, significa que os erros do 3DVar foram menores que os do LETKF.

Na Figura 5.13 nota-se nas primeiras horas de previsão o LETKF com erros menores que o 3DVar e, com o avanço das horas, o 3DVar passa a ter um desempenho melhor, mantendo o seu desempenho até 120 horas, exceto para o Hemisfério Sul, onde ocorre uma queda a partir de 84 horas de previsão. No entanto, essa região é a primeira a apresentar erros menores quando comparado ao LETKF já nas 30 horas de previsão, seguido do Globo em aproximadamente 36 horas, do Hemisfério Norte em 48 horas e a região Tropical em 54 horas, aproximadamente. Embora o Hemisfério Sul tenha sido a região com maiores erros na análise do 3DVar, nas previsões foi a primeira a melhorar seu desempenho em comparação com o LETKF.

Figura 5.13 - Diferença nos valores de RMSE da pressão em superfície entre o LETKF e 3DVar para as previsões de até 120 horas pertencentes ao período de dezembro, janeiro e fevereiro.

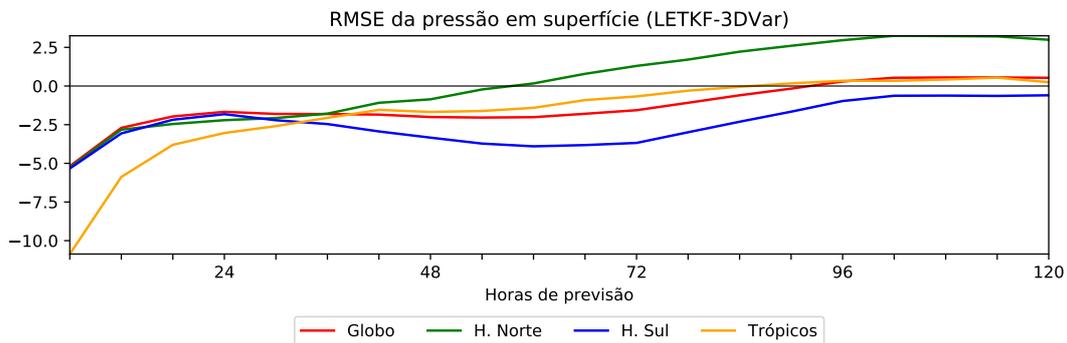


No eixo  $x$  estão as horas de previsão e no eixo  $y$  os valores da diferença. Cada região está representada por uma cor: em vermelho o Globo, em verde o Hemisfério Norte, em azul o Hemisfério Sul e em amarelo os Trópicos. Quando os valores forem positivos significa que o RMSE do 3DVar foi menor que o do LETKF, quando os valores forem negativos, significa que o LETKF apresentou um desempenho melhor que o 3DVar.

Fonte: Produção do autor.

Na Figura 5.14 o padrão visto no período de verão se mantém para o período de inverno, mostrando o LETKF com erros menores que o sistema variacional e, à medida que as horas avançam, o 3DVar mostra uma gradativa melhora, inclusive revelando a queda no desempenho visto no Hemisfério Sul no período anterior. Para esse período o 3DVar apresenta um desempenho melhor que o LETKF a partir de 60 horas de previsão para o Hemisfério Norte, a partir de 90 horas para os Trópicos e em 96 horas para o Globo; já para o Hemisfério Sul o LETKF teve um desempenho melhor em todos os horários.

Figura 5.14 - Idem à Figura 5.13, porém, para o período de junho, julho e agosto.



Fonte: Produção do autor.

Pode-se concluir aqui que, para os experimentos realizados com o LETKF e com o 3DVar, o sistema por conjunto apresenta um melhor desempenho nas primeiras horas de previsão em relação ao sistema variacional. Já para as previsões mais longas, o sistema variacional apresenta erros menores que o sistema por conjunto. Os primeiros horários de previsão são, normalmente, dominados pelas informações trazidas das observações pelo sistema de assimilação de dados, já os horários mais distantes são dominados pelos erros do próprio modelo. Nesse sentido, infere-se que os erros menores vistos no LETKF sejam em decorrência de sua matriz de covariância dos erros do *background*, que é constantemente atualizada, acompanhando assim as oscilações do dia como sugerido pela literatura.

## 5.2 Avaliação das análises do Híbrido Médio

Cinco experimentos foram realizados com valores de alfa fixo ( $\alpha = 0, 1$ ,  $\alpha = 0, 3$ ,  $\alpha = 0, 5$ ,  $\alpha = 0, 7$  e  $\alpha = 0, 9$ ) a fim de encontrar o valor de alfa que apresentaria um menor erro nas análises, e mais um sexto experimento, que utilizou o alfa dinâmico. Essas análises foram avaliadas a partir do RMSE do incremento de análise (O-A) para as cinco variáveis do modelo ( $ps$ ,  $q$ ,  $T$ ,  $u$  e  $v$ ) e para 4 regiões: Globo, Trópicos (compreendendo a faixa de 20°S a 20°N), Hemisfério Norte (compreendendo a faixa de 20°N a 80°N) e Hemisfério Sul (compreendendo a faixa de 80°S a 20°S). Esses resultados para o período de dezembro, janeiro e fevereiro são apresentados na Tabela 5.9 (para a pressão em superfície) e, nas Figuras 5.15, 5.16, 5.17 e 5.18, revelam-se os perfis verticais do erro para a umidade, temperatura, vento zonal e vento meridional, respectivamente. Os resultados do período de junho, julho e agosto são apresentados na Tabela 5.10 e nas Figuras 5.19, 5.20, 5.21 e 5.22, seguindo o mesmo

padrão das anteriores. Para facilitar as comparações entre os resultados do Híbrido Médio em suas várias configurações com os resultados expostos na Seção anterior, foram incluídos nas figuras de perfis verticais, os perfis do LETKF e do 3DVar, que anteriormente foram apresentados na forma de tabela.

A Tabela 5.9 apresenta os valores de RMSE para a pressão em superfície nas quatro regiões de estudo. Na primeira coluna encontram-se os seis experimentos e, nas demais colunas, os erros para cada uma das regiões: Globo, Trópicos, Hemisfério Norte e Hemisfério Sul, respectivamente. É possível observar nessa Tabela que o experimento utilizando o  $\alpha = 0,1$  obteve menores erros para a variável apresentada em todas as regiões de estudo. Um padrão pode ser observado: quanto maior os valores de alfa, maiores foram os erros dessa variável, ou seja, quanto mais se utilizou o 3DVar na elaboração da análise final, maiores foram os erros observados. Já com o experimento utilizando o alfa dinâmico, notou-se que os erros ficaram bem próximos àqueles do experimento com  $\alpha = 0,3$ , apresentando uma diferença entre os erros inferiores a 1 Pa, enquanto a diferença contra o  $\alpha = 0,9$  foi de até 15 Pa. É possível observar também que os maiores erros para os seis experimentos ficaram localizados no Hemisfério Sul, da mesma forma que os erros de cada um dos sistemas que compõem o Híbrido Médio na Seção 5.1. Esses maiores erros podem estar relacionados aos sistemas transientes que ocorrem na região e, pela falta de dados, torna-se difícil a correção do *background* por parte dos sistemas de assimilação. Nota-se ainda que os experimentos que utilizam  $\alpha = 0,1$  e  $\alpha = 0,3$  apresentaram valores de erro menores do que quando se utiliza apenas o LETKF e apenas o 3DVar, mostrando com isso, a eficácia de se realizar a combinação dessas análises.

Tabela 5.9 - RMSE médio da pressão em superfície para as quatro regiões: Globo, Trópicos, Hemisfério Norte e Hemisfério Sul, para os meses de dezembro, janeiro e fevereiro.

	Globo	Trópicos	H. Norte	H. Sul
$\alpha = 0,1$	27,08	27,81	21,87	31,03
$\alpha = 0,3$	28,89	29,05	23,69	33,41
$\alpha = 0,5$	33,07	32,91	27,44	38,35
$\alpha = 0,7$	37,85	37,31	31,67	44,05
$\alpha = 0,9$	43,13	42,33	36,36	50,15
alfa dinâmico	29,75	29,41	23,84	35,27

Nas Figuras dos perfis verticais são apresentados resultados para cada uma das

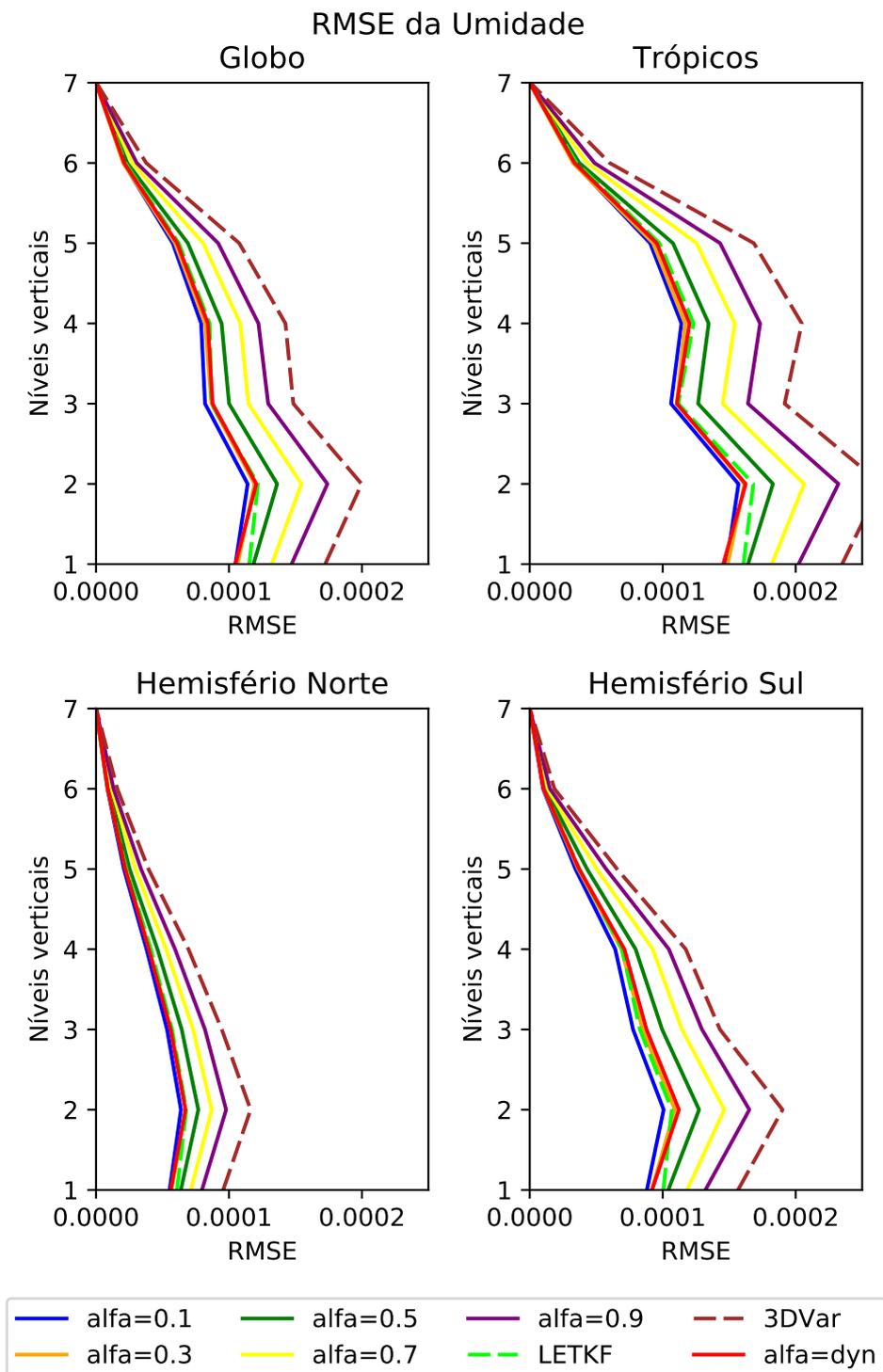
quatro regiões: Globo (esquerda) e Trópicos (direita) na parte superior do painel e Hemisfério Norte (esquerda) e Hemisfério Sul (direita) na parte inferior do painel. As curvas em azul são referentes ao experimento que utiliza  $\alpha = 0,1$ , em laranja ao experimento que utiliza  $\alpha = 0,3$ , em verde ao experimento que utiliza  $\alpha = 0,5$ , em amarelo ao experimento que utiliza  $\alpha = 0,7$ , em roxo ao experimento que utiliza  $\alpha = 0,9$ , em vermelho ao experimento com alfa dinâmico, a linha tracejada verde clara é referente ao LETKF e a linha tracejada em marrom ao 3DVar. Aqui, foram incluídos os valores do experimento com o LETKF e com o 3DVar a fim de facilitar comparações com os valores apresentados anteriormente. Os valores do RMSE são apresentados nos eixos  $x$ , e nos eixos  $y$  estão os 7 níveis verticais.

O perfil vertical da umidade é mostrado na Figura 5.15. Nela observa-se em sua grande maioria valores de erro menores para o experimento que utiliza  $\alpha = 0,1$  e valores maiores para o experimento utilizando  $\alpha = 0,9$ . É possível observar também que no último nível do modelo os valores de erro convergem para o zero. Isso ocorre devido ao fato de não existirem informações de umidade nesse nível. Os valores dos erros na região Tropical são maiores porque nessa região vê-se valores mais elevados de umidade, enquanto nos Hemisférios esses valores normalmente são menores, gerando, então, valores de erros menores. Essa relação também pode ser observada com relação à altura. Os valores de erro decrescem conforme aumenta a altura na atmosfera e isso ocorre devido à diminuição de umidade em níveis mais elevados. Com relação ao alfa dinâmico, nota-se um melhor resultado apenas no primeiro nível na região dos Trópicos, mas uma diferença muito pequena em relação aos experimentos com  $\alpha = 0,1$  e  $\alpha = 0,3$  em diversas situações, como, por exemplo, no perfil do Hemisfério Norte. É possível observar também o comportamento do 3DVar e do LETKF em relação às diversas composições de Híbrido Médio. O 3DVar, como visto na seção anterior, revela erros maiores que o LETKF sempre, porém, nota-se que, para que se tenham erros menores, a combinação entre eles é necessária, como é possível observar para as quatro regiões. O experimento que utiliza  $\alpha = 0,1$ , apresenta os menores erros para essa variável nas quatro regiões. Em algumas situações o experimento com  $\alpha = 0,3$  também se mostra melhor, como é o caso no Globo, Trópicos e Hemisfério Norte. O mesmo também é visto para o experimento com o alfa dinâmico.

Na Figura 5.16 são apresentados os valores de RMSE para o perfil vertical da temperatura. Analisando apenas os experimentos com o sistema Híbrido Médio, observa-se que o experimento com  $\alpha = 0,1$  apresentou erros menores, enquanto o  $\alpha = 0,9$  em todas as situações apresentou os erros maiores. O experimento com o alfa dinâmico

apenas apresentou erros menores que o experimento com  $\alpha = 0,1$  nos dois primeiros níveis do modelo na região Tropical. Quando se compara as curvas dos sistemas Híbridos com aquelas do LETKF e do 3DVar, nota-se então que a utilização do 3DVar sozinho apresenta os maiores erros, ou seja, a utilização de qualquer valor de alfa já mostra benefícios para o 3DVar. No entanto, para que apresente uma melhora no LETKF, é necessário que o alfa seja igual a 0,1. Utilizar o alfa dinâmico ou  $\alpha = 0,3$ , mostra melhorias quando comparado ao LETKF nos Trópicos e Hemisfério Norte.

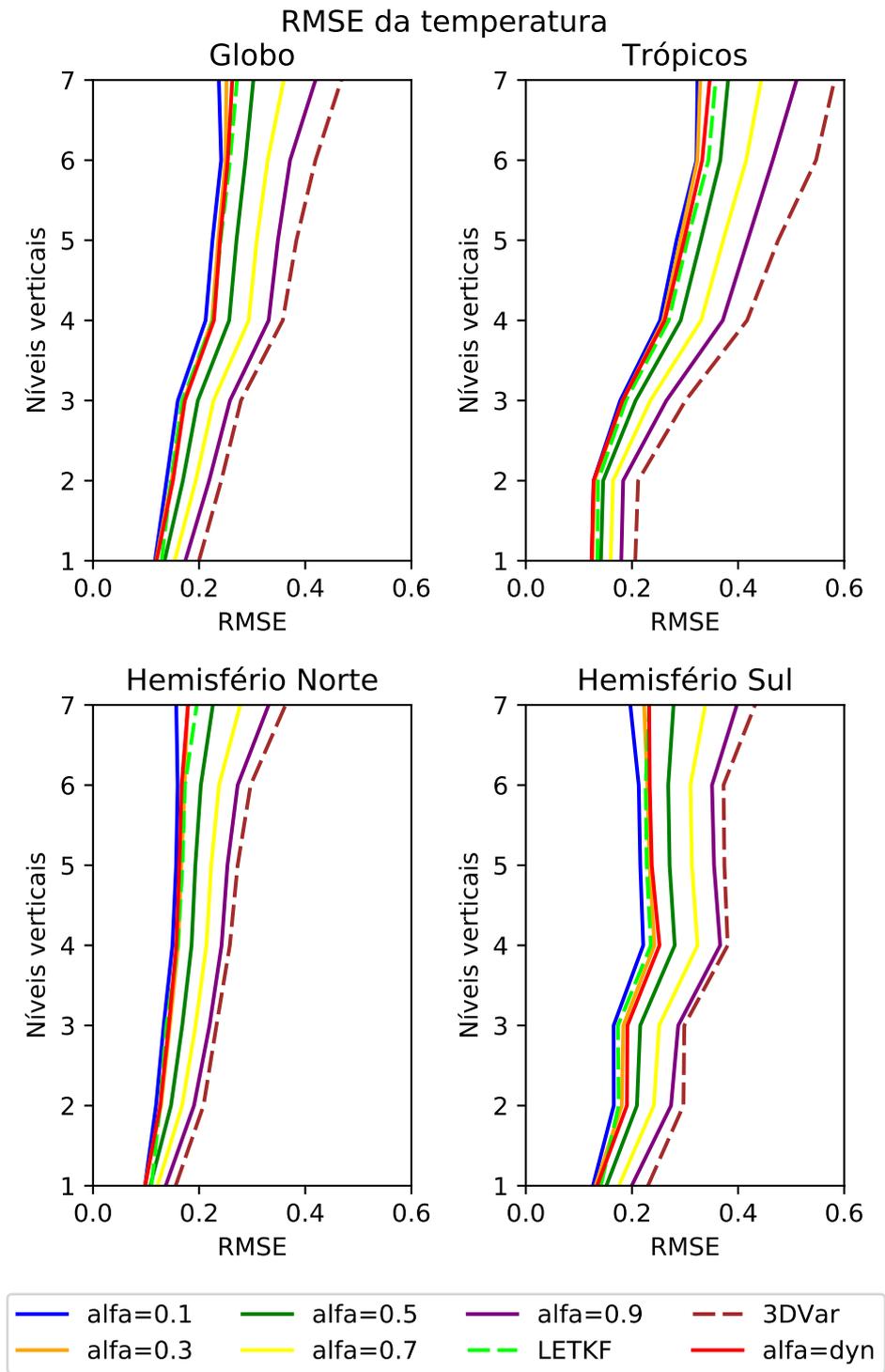
Figura 5.15 - Perfil vertical de RMSE médio da umidade para quatro regiões de estudo: Globo, Trópicos, Hemisfério Norte e Hemisfério Sul.



Em azul o experimento utilizando  $\alpha = 0,1$ , em laranja o experimento utilizando  $\alpha = 0,3$ , em verde experimento utilizando  $\alpha = 0,5$ , em amarelo experimento utilizando  $\alpha = 0,7$ , em roxo experimento utilizando  $\alpha = 0,9$ , em vermelho o experimento com alfa dinâmico, em verde claro o LETKF e em marrom o 3DVar. Os eixos  $x$  são os valores do erro e nos eixos  $y$  os sete níveis do modelo.

Fonte: Produção do autor.

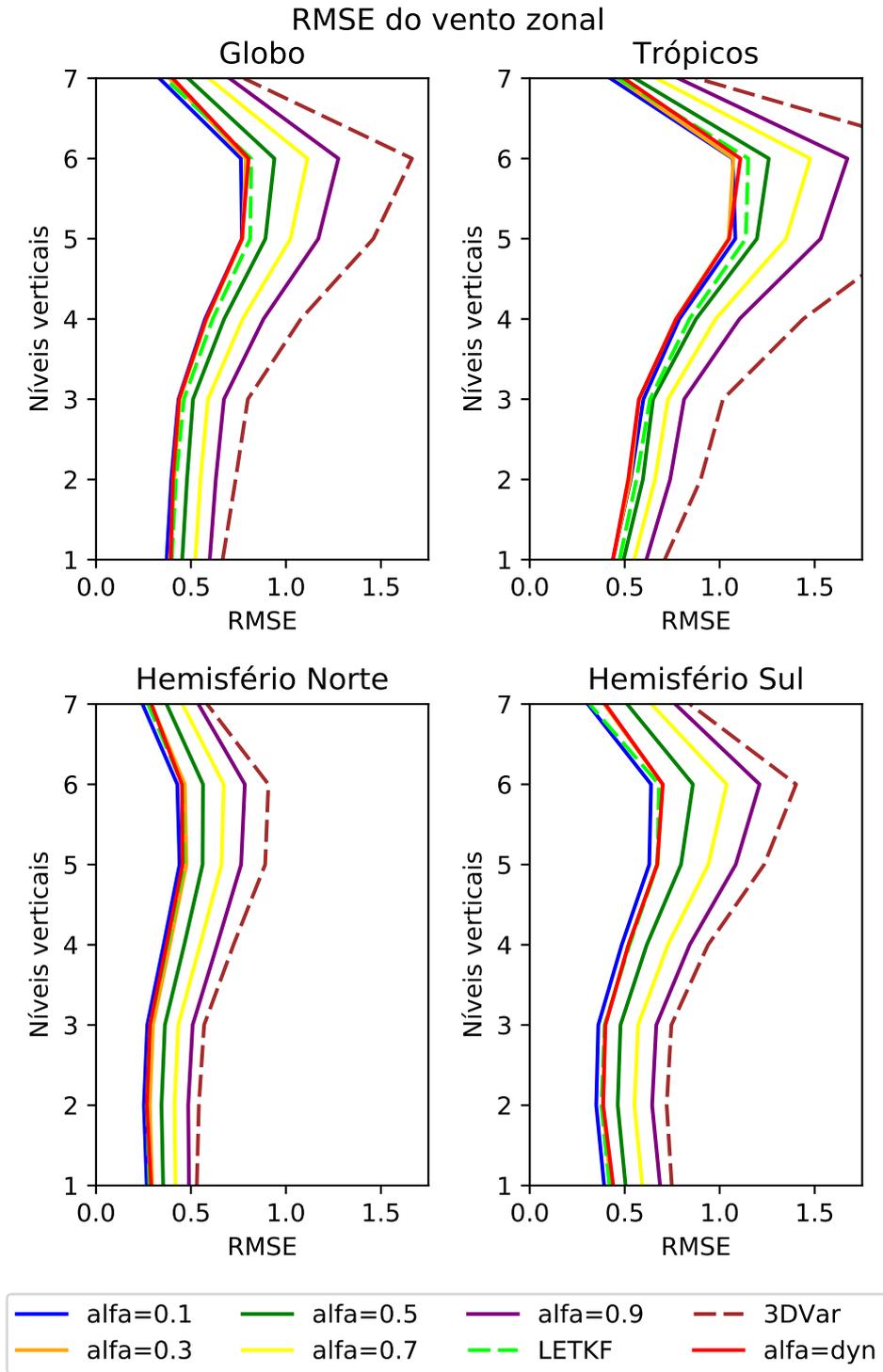
Figura 5.16 - Idem à Figura 5.15, porém, para a temperatura.



Os perfis de vento zonal e meridional são apresentados nas Figuras 5.17 e 5.18, respectivamente. Para ambos os casos nota-se que o experimento utilizando  $\alpha = 0,9$  apresentou os maiores erros dentre aqueles que utilizaram o Híbrido Médio. Com relação aos menores erros, ainda o experimento com  $\alpha = 0,1$  foi melhor, porém, ao observar o perfil do vento zonal na região Tropical, nota-se que o experimento com alfa dinâmico só não obteve erros menores que os demais nos dois últimos níveis. No perfil Global, o alfa dinâmico também se manteve bem próximo do experimento com  $\alpha = 0,1$ . Com relação ao vento meridional, o perfil é um pouco diferente: para o Trópicos, por exemplo, as curvas dos experimentos com  $\alpha = 0,1$ ,  $\alpha = 0,3$  e com alfa dinâmico ficam sobrepostas até o terceiro nível do modelo, posto que o alfa dinâmico acompanha o experimento com  $\alpha = 0,3$  até o quarto nível. Para o vento, o mesmo padrão visto anteriormente em relação aos sistemas LETKF e 3DVar foi visto aqui: o 3DVar quando utilizado sozinho apresenta erros maiores. O LETKF não se mostra melhor que o experimento com  $\alpha = 0,1$  nunca, porém, ele apresenta erros menores que os experimentos que utilizaram  $\alpha = 0,3$  e o alfa dinâmico em alguns níveis dos Hemisférios para ambas as componentes.

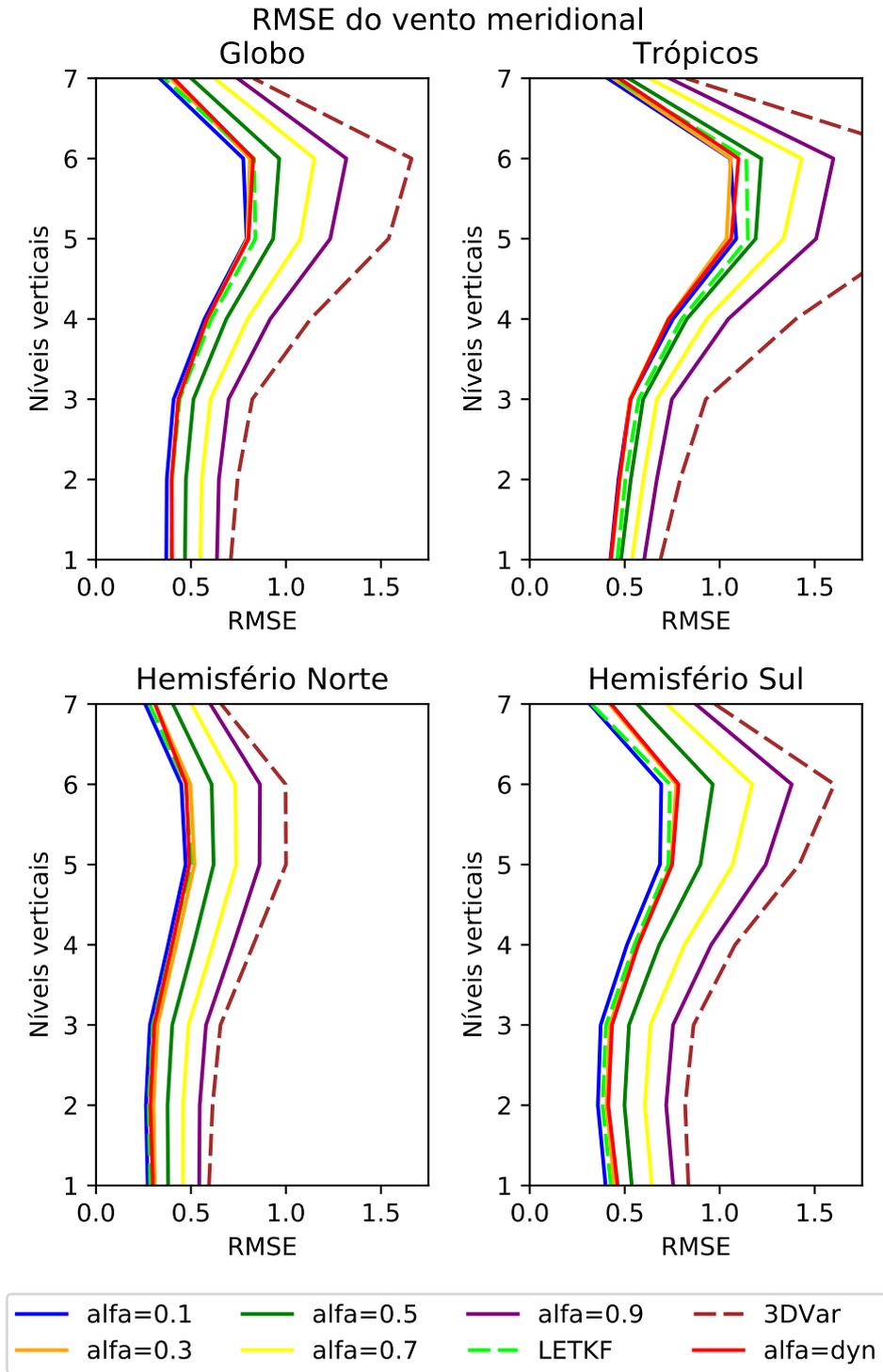
Nessas figuras foi possível observar o mesmo padrão apresentado nos erros de pressão em superfície em que o erro cresce quanto mais 3DVar é utilizado na geração das análises do sistema Híbrido Médio. Com relação apenas aos experimentos com alfa fixo, esse padrão só não foi observado no caso do vento nos trópicos (Figuras 5.17 e 5.18), o qual em médios níveis o experimento utilizando  $\alpha = 0,3$  obteve um desempenho melhor do que com  $\alpha = 0,1$ , mais precisamente nos níveis 2, 3, 4 e 5 para o vento zonal e níveis 3, 4 e 5 para o vento meridional. Já para o alfa dinâmico, observou-se que ele apresentou os menores erros apenas em alguns pontos: primeiro nível da umidade nos trópicos, primeiro e segundo nível da temperatura nos trópicos, vento zonal no quinto nível do Globo e nos níveis 1, 2, 3, 4 e 5 da região Tropical. Em relação aos sistemas de assimilação de dados utilizados no Híbrido Médio, foi possível concluir que a combinação de ambos, independente do valor utilizado para o alfa, gera resultados positivos quando comparado com o 3DVar puro. No entanto, o LETKF se mostrou com erros menores nas análises do que o Híbrido Médio quando utilizou  $\alpha = 0,9$ ,  $\alpha = 0,7$  e  $\alpha = 0,5$ . Ou seja, para que o resultado da combinação linear de ambos sistemas traga resultados melhores do que o LETKF sozinho, é preciso que seja utilizado o  $\alpha = 0,1$ , o  $\alpha = 0,3$  ou o alfa dinâmico.

Figura 5.17 - Idem à Figura 5.15, porém, para o vento zonal.



Fonte: Produção do autor.

Figura 5.18 - Idem à Figura 5.15, porém, para o vento meridional.



Fonte: Produção do autor.

A Tabela 5.10 mostra os valores de RMSE para a pressão em superfície para o trimestre JJA. Nela são mostrados os seis experimentos para as 4 regiões de estudo. Pode-se observar que o experimento utilizando  $\alpha = 0,1$  é o que apresenta os menores valores de erro e, diferentemente do visto para o trimestre DJF, o experimento com alfa dinâmico foi o terceiro a mostrar os menores erros, ficando o experimento com  $\alpha = 0,3$  em segundo lugar. O experimento com alfa dinâmico apenas ganha do que experimento que utiliza  $\alpha = 0,3$  no Hemisfério Norte. A diferença entre os erros dos experimentos com alfa dinâmico e com  $\alpha = 0,1$  é de aproximadamente 1 Pa no Hemisfério Norte, 2 Pa no Globo e da região Tropical, chegando a aproximadamente 4Pa no Hemisfério Sul. Esses valores são baixos se comparados aos do experimento com  $\alpha = 0,1$  e do experimento com  $\alpha = 0,9$ , que novamente apresentou os maiores erros. Essa diferença chega a aproximadamente 16 Pa no Globo, 14 Pa nos Trópicos e Hemisfério Norte e a 19 Pa no Hemisfério Sul. Além disso, foi possível observar que os valores de RMSE para esse período (JJA) foram maiores do que para o trimestre DJF. Nota-se também, quando comparada com as tabelas da Seção 5.1, que a combinação de análises gera erros menores. Isso foi visto em todas as regiões para o experimento que utiliza  $\alpha = 0,1$ . Já o 3DVar apresenta os maiores erros, exceto para o Globo, região em que ele apresentou melhores resultados do que o experimento que utiliza  $\alpha = 0,9$ . Esse resultado deve ter influência dos valores de erros nos Polos, já que estes não foram avaliados, e os Hemisférios e Trópicos apresentam valores menores de erro para o experimento com  $\alpha = 0,9$ .

Tabela 5.10 - Idem à Tabela 5.9, porém, para os meses de junho, julho e agosto.

	Globo	Trópicos	H. Norte	H. Sul
$\alpha = 0,1$	28,50	29,42	22,93	32,55
$\alpha = 0,3$	30,22	30,62	24,54	34,90
$\alpha = 0,5$	34,16	34,19	27,91	39,76
$\alpha = 0,7$	38,86	38,59	31,92	45,40
$\alpha = 0,9$	44,01	43,53	36,30	51,47
alfa dinâmico	30,79	31,03	23,96	36,46

Na Figura 5.19 são apresentados os perfis verticais para a umidade nas quatro regiões. Pode-se notar primeiramente que o experimento utilizando o  $\alpha = 0,1$  apresentou os menores erros e, o experimento utilizando  $\alpha = 0,9$  apresentou os maiores erros, quando avaliados apenas os experimentos envolvendo o Híbrido Médio. O padrão dos erros segue conforme apresentado para o trimestre DJF, apenas com va-

lores maiores nos Hemisférios. Para esse período, o experimento com o alfa dinâmico apenas revelou o menor erro no sexto nível no Hemisfério Norte. Em relação aos experimentos do LETKF e do 3DVar, notam-se os maiores erros para o 3DVar em todas as regiões de estudo. Já o LETKF apresenta erros maiores que o experimento com  $\alpha = 0,1$  sempre e mostra erros menores que o experimento com alfa dinâmico e  $\alpha = 0,3$  apenas no Hemisfério Sul. No entanto, ele apresenta erros menores que os experimentos quando utilizam  $\alpha = 0,5$ ,  $\alpha = 0,7$  e  $\alpha = 0,9$ .

Os perfis de temperatura são apresentados na Figura 5.20. Através desses gráficos, podemos notar que, dentre os experimentos utilizando o sistema Híbrido Médio, para todos os casos o experimento com  $\alpha = 0,1$  apresentou melhores resultados e, novamente, o experimento com  $\alpha = 0,9$  apresentou os piores resultados. Além disso, notam-se os maiores erros quando o 3DVar é utilizado sozinho. Quando se compara o experimento com LETKF em relação aos demais, nota-se que em todos os casos a combinação linear das análises é benéfica ao LETKF quando é utilizado  $\alpha = 0,1$ . Em relação aos experimentos com  $\alpha = 0,3$  e com o alfa dinâmico, nota-se o LETKF melhor apenas no Hemisfério Sul.

Os campos de vento zonal e meridional são apresentados nas Figuras 5.21 e 5.22, respectivamente. Nelas permanece o perfil sobre o experimento com  $\alpha = 0,1$  e com  $\alpha = 0,9$ , quando analisados apenas os experimentos utilizando o Híbrido Médio. Observa-se no perfil de vento zonal (Figura 5.21) que, em todos os casos, o experimento utilizando  $\alpha = 0,1$  apresentou os menores erros e, comparando com o período DJF, a curva do experimento com o alfa dinâmico se aproximou da curva com  $\alpha = 0,3$ . Para o caso do vento meridional (Figura 5.22), nota-se que o experimento com  $\alpha = 0,3$  apresenta menores valores de RMSE no quarto nível na região Tropical e que em situação alguma o alfa dinâmico apresentou melhores resultados. O mesmo padrão descrito até aqui sobre o experimento utilizando apenas o 3DVar é observado: ele apresenta os maiores erros. Já o padrão visto para o LETKF para as componentes do vento está diferente: para a componente zonal do vento, observa-se na região Tropical situações em que o experimento com  $\alpha = 0,5$  alcança os erros do LETKF, além de o experimento com apenas o LETKF apresentar menores erros do que quando utilizado  $\alpha = 0,3$  no Hemisfério Norte; para a componente meridional do vento, o experimento com  $\alpha = 0,5$  aproxima-se novamente dos erros do LETKF e, ainda, nos primeiros níveis do Hemisfério Norte o LETKF se aproxima dos erros do experimento com  $\alpha = 0,1$ .

Figura 5.19 - Idem à Figura 5.15, porém, para o período de junho, julho e agosto.

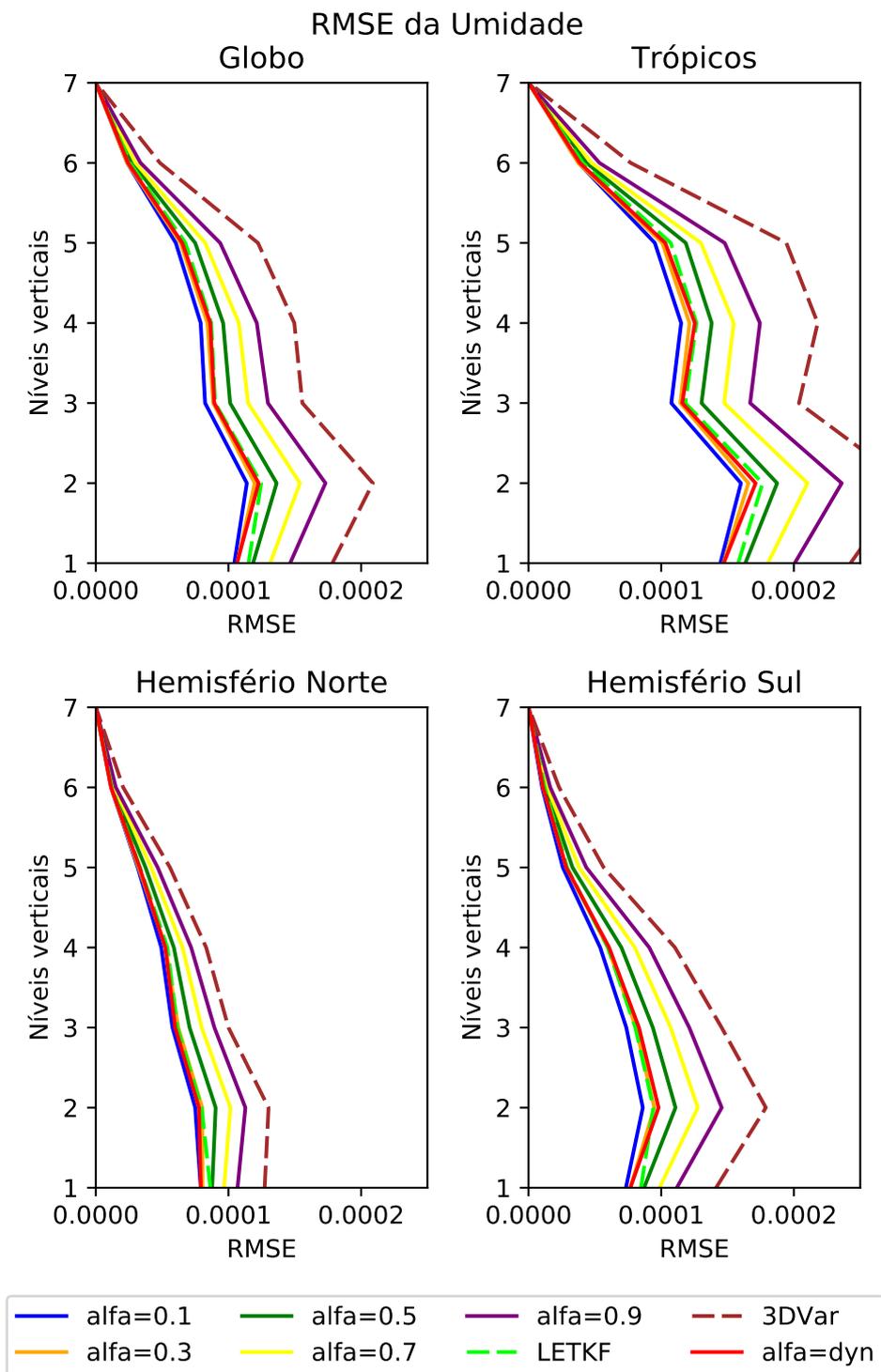


Figura 5.20 - Idem à Figura 5.19, porém, para a temperatura.

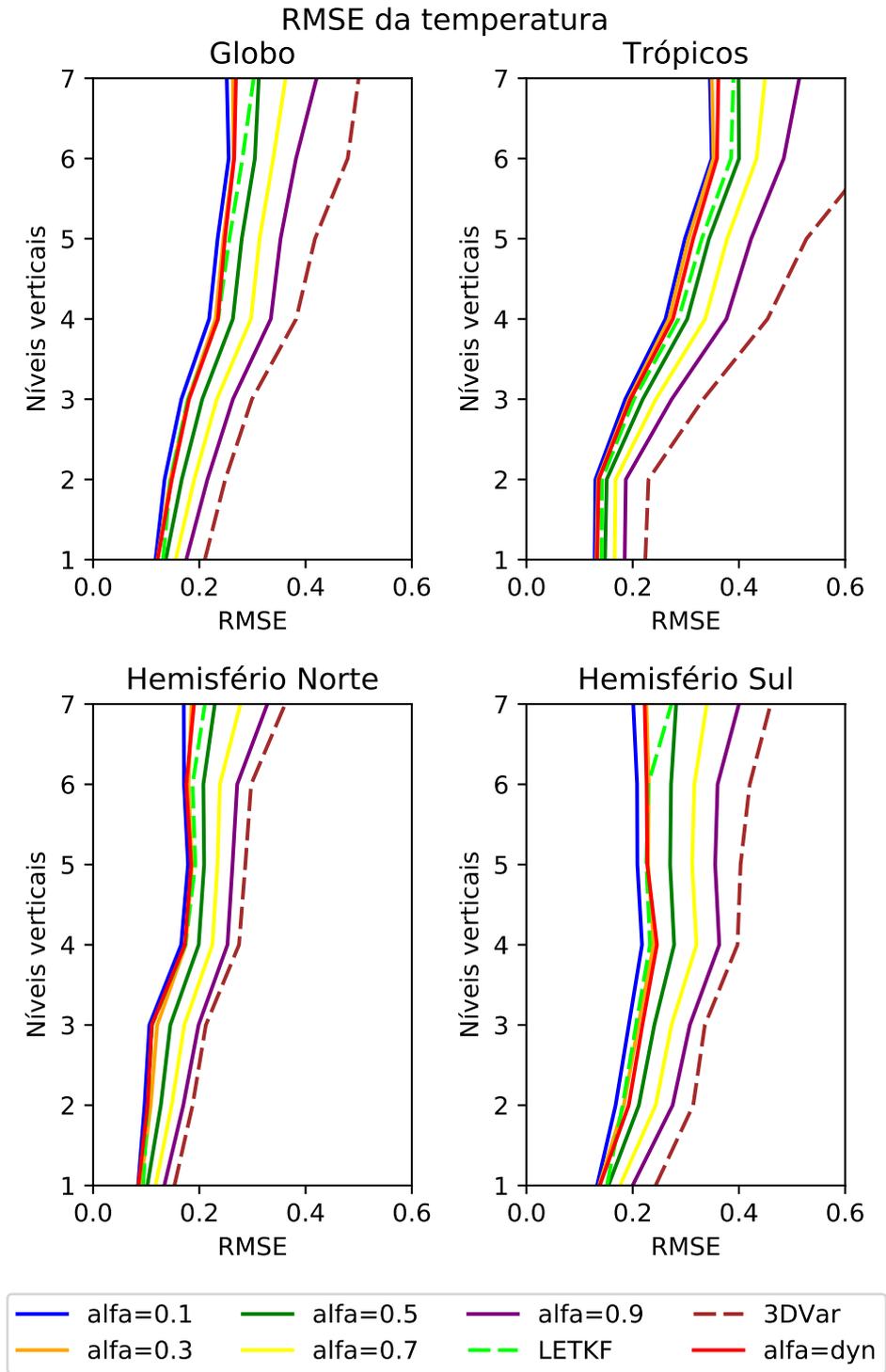


Figura 5.21 - Idem à Figura 5.19, porém para o vento zonal.

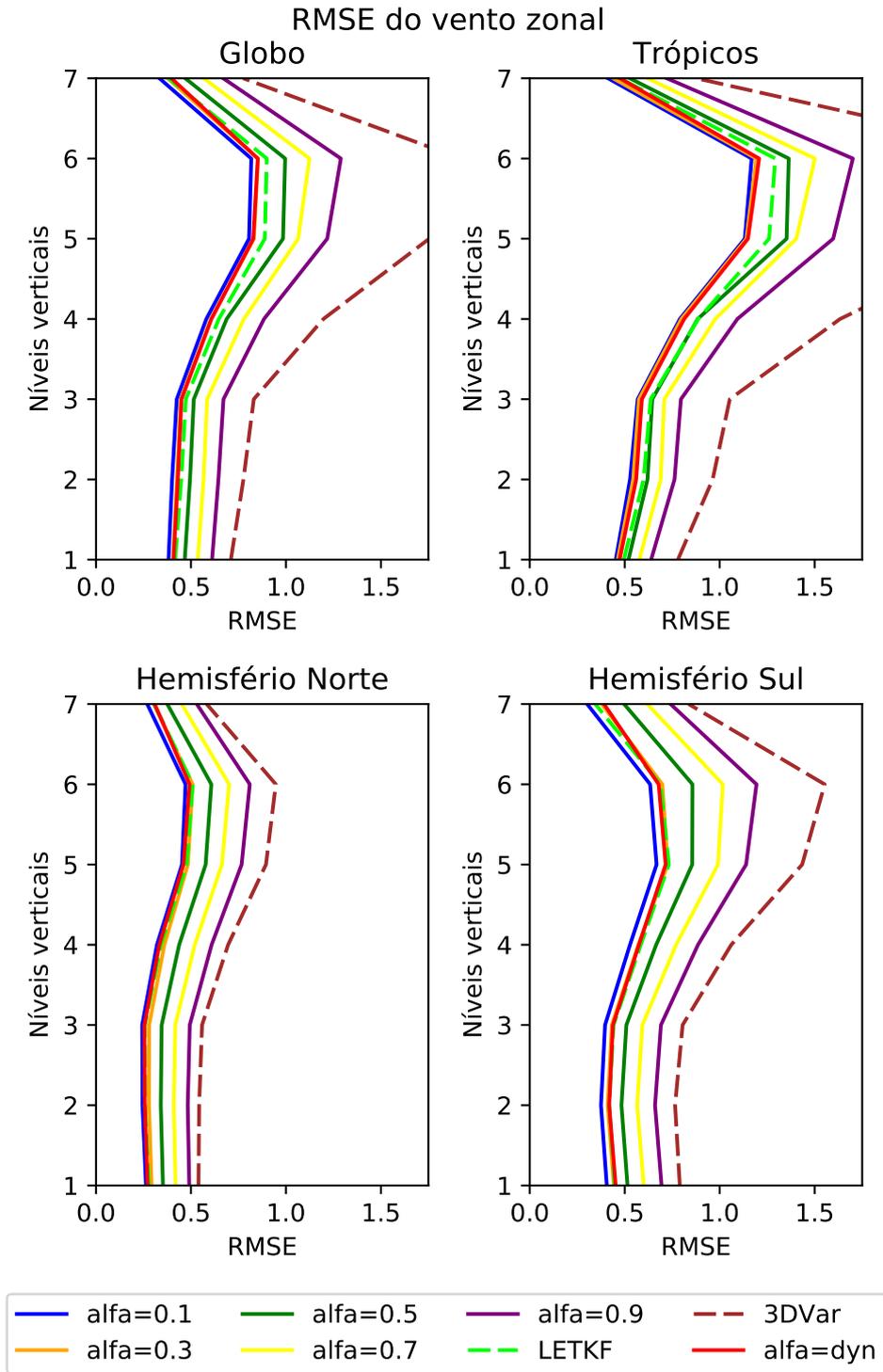
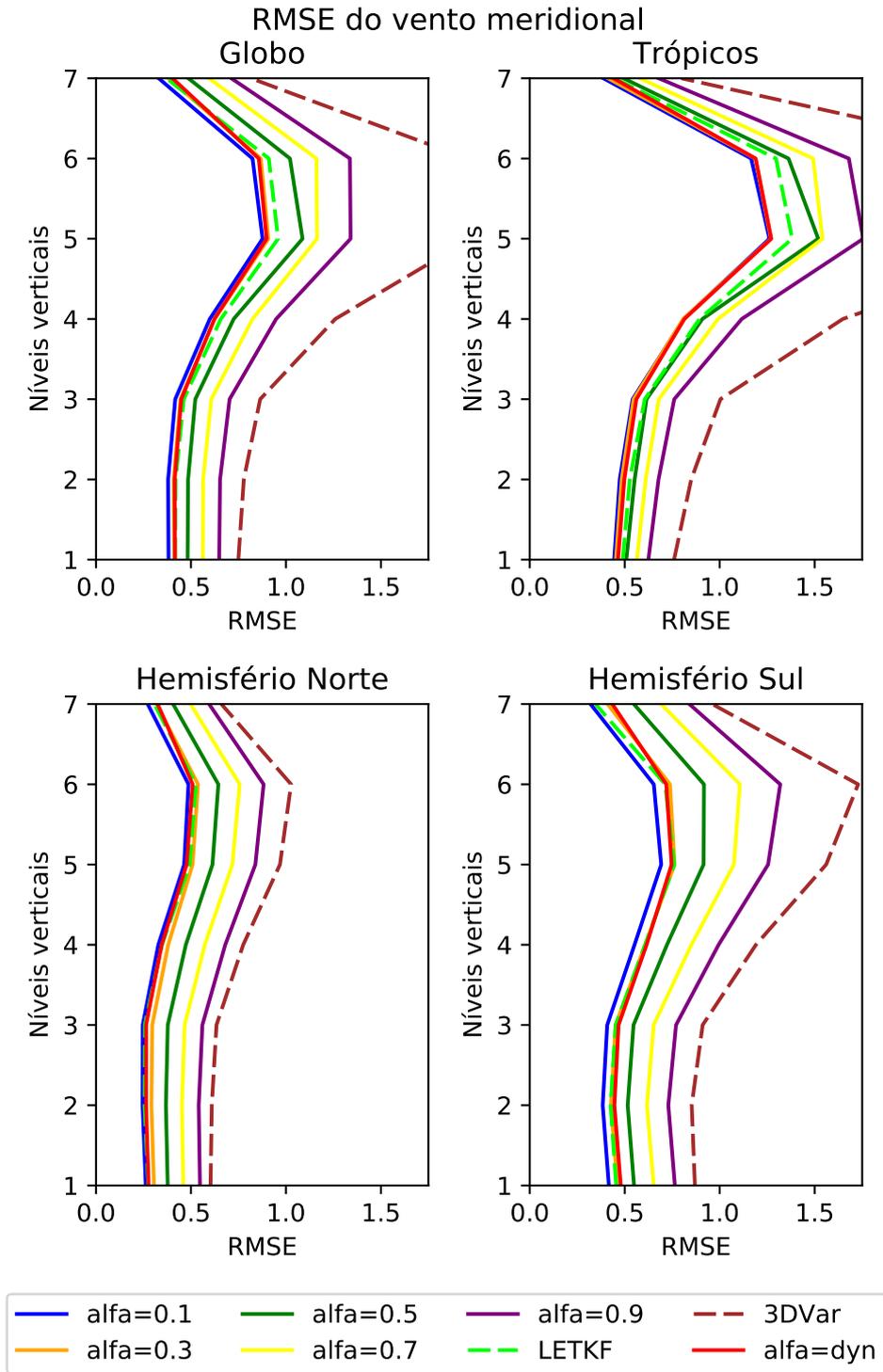


Figura 5.22 - Idem à Figura 5.19, porém, para o vento meridional.



Fonte: Produção do autor.

Pode-se concluir que, independentemente do período de estudo, seja inverno ou verão, o valor de alfa fixo que apresentou os melhores resultados foi o experimento que utilizou  $\alpha = 0,1$ . Também nota-se que para o trimestre DJF houve mais situações em que o experimento com alfa dinâmico se sobressaiu aos demais. Ainda assim, majoritariamente o alfa dinâmico apresentou resultados inferiores às análises dos experimentos com  $\alpha = 0,1$ . Outro ponto observado foi o fato de que quanto maior é o peso atribuído à análise do 3DVar na geração da análise final do Híbrido Médio, maior é o erro e quanto maior o peso para o LETKF, melhores resultados são vistos. Porém, ao analisar os resultados da Seção anterior, nota-se que, quando se utiliza  $\alpha = 0,0$  (que, de acordo com a equação 2.7 resulta no mesmo LETKF puro), os erros são maiores do que quando se utiliza  $\alpha = 0,1$  sempre, e na maioria dos casos os experimentos utilizando o  $\alpha = 0,3$  e o alfa dinâmico também apresentam erros menores que o LETKF puro. O contrário também é visto: em sua grande maioria, os erros vistos no 3DVar<sup>1</sup> são maiores do que os vistos quando se utiliza  $\alpha = 0,9$ . O sistema variacional apenas é melhor do que o experimento com o sistema Híbrido Médio utilizando  $\alpha = 0,9$  para a pressão em superfície no Globo durante o período de inverno. Conclui-se, então, que realizar a combinação linear de ambos os sistemas beneficia a análise final do sistema de assimilação de dados Híbrido Médio, porém, dando mais peso para o LETKF ou utilizando o alfa dinâmico. Vale ressaltar aqui, que, de acordo com Penny (2014), o custo computacional do Híbrido Médio, é o mesmo do que o sistema baseado em conjunto sozinho, o que reforça, então, o uso do sistema Híbrido.

Os trabalhos existentes com o Híbrido Médio até então utilizaram apenas os seguintes valores de alfa:  $\alpha = 0,25$ ,  $\alpha = 0,5$  e  $\alpha = 0,75$ , o que resulta em usar apenas 25%, 50% ou 75% de um sistema ou de outro. No entanto, Hamill e Snyder (2000) concluíram que para o sistema Híbrido tradicional (leia-se o sistema Híbrido que recalcula a matriz  $\mathbf{B}$ ), os melhores resultados foram obtidos quando alfa variou entre 0,1 e 0,4. Ou seja, os autores estabelecem que, para o caso do cálculo da matriz  $\mathbf{B}$ , utilizar mais o sistema por conjunto na combinação de matrizes trouxe melhores resultados. Ainda que os sistemas Híbridos não sejam iguais, ambos realizam uma combinação entre dois outros sistemas: um sistema baseado em conjunto e outro variacional. Dessa forma, esta seção teve a finalidade de avaliar o comportamento das análises geradas pelo sistema Híbrido Médio quando submetido a diversos valores de alfa fixo a fim de encontrar um valor ótimo e, avaliar as análises geradas pela meto-

---

<sup>1</sup>Aqui vale ressaltar que  $\alpha = 1$  é diferente do 3DVar puro apresentado na Seção anterior.  $\alpha = 1$  representa o 3DVar que utilizou como seu *background* a análise média vinda do LETKF, ou seja, sofreu influências do sistema por conjunto.

dogia de alfa dinâmico proposta. Os resultados encontrados aqui, vão ao encontro ao visto por Hamill e Snyder (2000): quando se utiliza mais o sistema baseado por conjunto, melhores são os resultados. Aqui, tanto para o período de verão quanto para o período de inverno, utilizar  $\alpha = 0,1$  apresentou análises com menores erros e, à medida que o alfa se aproxima do 3DVar puro, os erros aumentaram.

### 5.3 Verificação do balanço das análises

Sabe-se que o processo de assimilação de dados introduz um certo desbalanço à previsão numérica de tempo e que o sucesso das condições iniciais depende do quão balanceadas elas estão. Condições iniciais bem balanceadas evitam a geração de ondas de gravidade espúrias, como as observadas nas previsões geradas por Richardson em 1922. Para evitar esse problema, surgiram melhorias na inicialização dos sistemas, porém, ainda com essas melhorias, não é possível garantir o sucesso das previsões. Além dessas melhorias, alguns sistemas de assimilação de dados ainda utilizam técnicas de localização e inflação do espalhamento para evitar essas ondas espúrias.

Greybush et al. (2011) utilizaram algumas métricas para avaliar o balanço de seus experimentos com o modelo SPEEDY como, por exemplo, o erro absoluto do vento em 500hPa e o vento ageostrófico também em 500hPa para ambos os Hemisférios, sendo o Hemisfério Norte a faixa entre 30°N e 60°N, e o Hemisfério Sul a faixa entre 60°S e 30°S. Nesta seção são apresentados os resultados sobre o balanço dos experimentos através das métricas utilizadas por Greybush et al. (2011) para os dois períodos de estudo: trimestre DJF e trimestre JJA.

O vento ageostrófico é, por definição, a diferença entre o vento real ( $\mathbf{V}$ ) e o vento geostrófico ( $\mathbf{V}_g$ ) (conforme a Equação 5.1). Para o cálculo do vento geostrófico a altura geopotencial se faz necessária (conforme a Equação 5.2). Como mostrado na Seção 3.1 do Capítulo 3, essa variável não é dada pelo modelo: foi estimada a partir da Equação Hipsométrica.

$$\mathbf{V}_a = \mathbf{V} - \mathbf{V}_g \quad (5.1)$$

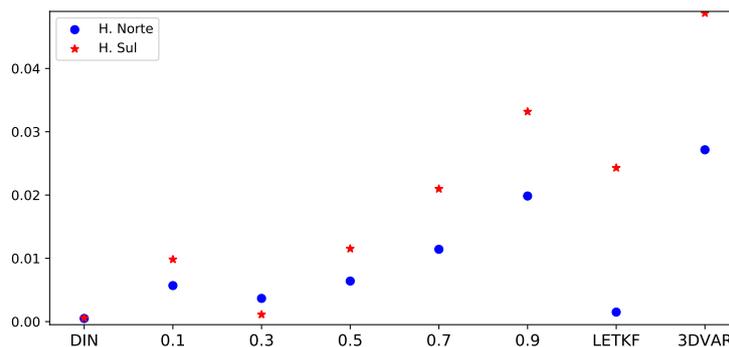
$$\mathbf{V}_g = \frac{\mathbf{k}}{f} \times \nabla_p \phi \quad (5.2)$$

Os resultados da avaliação do balanço nas análises são mostrados a seguir para o período de DJF e, na sequência, para o período de JJA.

A Figura 5.23 apresenta a média do erro absoluto do vento em 500hPa para os

oito experimentos do trimestre DJF. Os pontos em azul são referentes ao Hemisfério Norte e as estrelas em vermelho são referentes ao Hemisfério Sul. No eixo  $x$  estão os experimentos e no eixo  $y$  os valores do erro. Por se tratar de erro absoluto, quanto mais próximo do zero, melhor. É possível observar que o experimento utilizando alfa dinâmico apresentou erro menor para ambos os hemisférios, seguido pelo experimento que utilizou o  $\alpha = 0,3$ . Nota-se, ainda, que o experimento menos balanceado nas análises, de acordo com essa métrica, foi o que utiliza apenas o 3DVar, mostrando os maiores valores de erro para ambos os hemisférios. Já quando se utiliza  $\alpha = 0,5$ , que é o valor típico mais utilizado na literatura (Penny et al. (2015), Bonavita et al. (2015), Wespetal et al. (2018)), ele apresenta erros em ambos os hemisférios, porém erros menores que quando se utiliza  $\alpha = 0,7$ ,  $\alpha = 0,9$ , só o LETKF e só o 3DVar.

Figura 5.23 - Média do erro absoluto do vento em 500hPa para os oito experimentos: alfa dinâmico,  $\alpha = 0,1$ ,  $\alpha = 0,3$ ,  $\alpha = 0,5$ ,  $\alpha = 0,7$ ,  $\alpha = 0,9$ , LETKF e 3DVar.



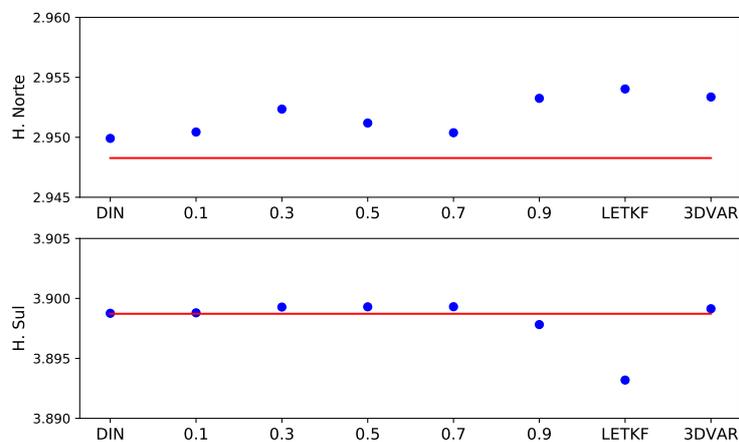
Os pontos em azul são referentes ao hemisfério norte e as estrelas em vermelho são referentes ao hemisfério sul. No eixo  $x$  estão os experimentos e no eixo  $y$  os valores do erro.

Fonte: Produção do autor.

Na Figura 5.24 é mostrada a média do vento ageostrófico em 500hPa também para os oito experimentos. No topo são valores relacionados ao Hemisfério Norte e, na parte de baixo, os valores relacionados ao Hemisfério Sul. Os pontos são referentes a cada experimento e a linha vermelha ao valor médio da *Nature Run*. No eixo  $x$  estão os experimentos e no eixo  $y$  os valores do vento em m/s. Como o vento ageostrófico é um vento residual, espera-se que ele seja o mais próximo do zero possível. Porém, de acordo com Greybush et al. (2011), quando o desbalanço das

análises é maior do que aquele da "verdade", a assimilação está de fato introduzindo um desbalanço ao sistema. Portanto, quanto mais próximo da linha da *Nature Run*<sup>2</sup>, menos desbalanço aquela condição inicial adiciona ao sistema. Pode-se observar que tanto para o Hemisfério Norte quanto para o Hemisfério Sul o experimento com alfa dinâmico ficou mais próximo da linha vermelha, seguido pelo experimento que utiliza  $\alpha = 0,1$ . Embora na métrica anterior o 3DVar tenha apresentado os maiores erros, aqui ele se mostra melhor que o LETKF sozinho e que o experimento quando se utiliza o  $\alpha = 0,9$ .

Figura 5.24 - Média do vento ageostrófico em 500hPa para os oito experimentos: alfa dinâmico,  $\alpha = 0,1$ ,  $\alpha = 0,3$ ,  $\alpha = 0,5$ ,  $\alpha = 0,7$ ,  $\alpha = 0,9$ , LETKF e 3DVar.



No topo estão valores relacionados ao hemisfério norte e, na parte de baixo, os valores relacionados ao hemisfério sul. Os pontos são referentes a cada experimento e a linha vermelha o valor médio do vento ageostrófico em 500 hPa da *nature run*. No eixo  $x$  estão os experimentos e no eixo  $y$  os valores do vento em m/s.

Fonte: Produção do autor.

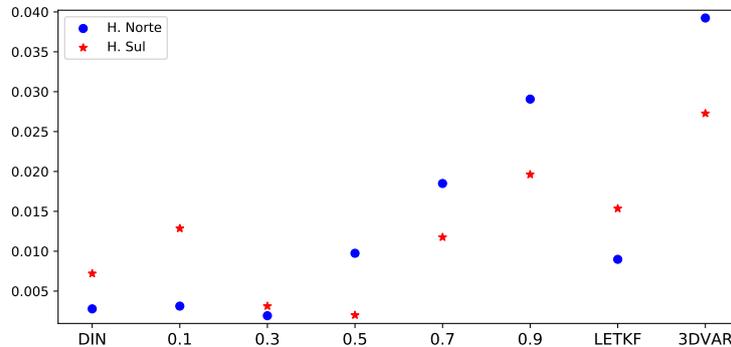
Conforme o que foi visto aqui, para o trimestre DJF o experimento que se mostrou melhor balanceado foi o que utilizou o Híbrido Médio com o alfa dinâmico, seguido pelos experimentos com valores baixos de alfa. Pode-se inferir, então, que o processo de multiplicar diferentes valores em diferentes pontos de grade, para o período de verão, não aumenta o desbalanço existente no processo de assimilação de dados. Além disso, pode-se dizer que não necessariamente a análise que apresenta os menores erros é a melhor balanceada, ou seja, os erros das análises não estão relacionadas

<sup>2</sup>Lembrando que a *Nature Run* é utilizada aqui como a verdade.

ao balanço encontrado nelas.

As Figuras 5.25 e 5.26 apresentam os valores médios do erro absoluto do vento em 500hPa e do vento ageostrófico também em 500 hPa para o período de junho, julho e agosto, conforme apresentado anteriormente para os meses de verão austral. Observando a Figura 5.25, encontra-se o menor valor de erro para o Hemisfério Norte quando se utiliza  $\alpha = 0,3$  e para o Hemisfério Sul quando se utiliza  $\alpha = 0,5$ . Os maiores valores de erro foram encontrados no experimento que utiliza apenas o 3DVar, da mesma forma como no trimestre DJF, porém com o Hemisfério Norte apresentando erros maiores. Além dele, os experimentos com o Híbrido Médio com  $\alpha = 0,9$ ,  $\alpha = 0,7$ ,  $\alpha = 0,5$  e  $\alpha = 0,3$  também apresentaram uma inversão; para o  $\alpha = 0,3$ , agora o Hemisfério Sul é que apresenta os maiores erros, enquanto o restante deles apresenta os maiores erros no Hemisfério Norte. O experimento que utiliza alfa dinâmico apresentou o segundo menor erro para o Hemisfério Norte e o terceiro menor erro para o Hemisfério Sul; ainda assim, revelou erros pequenos quando comparado ao restante.

Figura 5.25 - Idem à Figura 5.23, porém, para o trimestre de JJA.

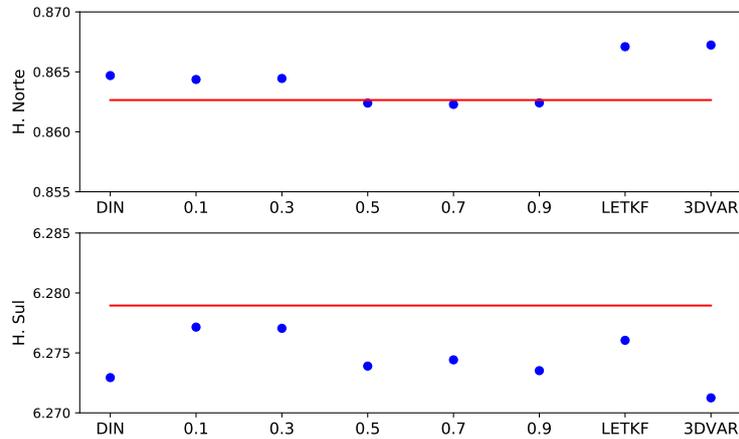


Fonte: Produção do autor.

Ao observar a Figura 5.26, nota-se que o experimento que utiliza alfa dinâmico ficou distante da curva da *Nature Run*, proporcionando desbalanço ao sistema. Para o caso do Hemisfério Norte, os experimentos que mostraram um melhor balanço foram aqueles que utilizam  $\alpha = 0,5$ ,  $\alpha = 0,7$  e  $\alpha = 0,9$ ; já para o Hemisfério Sul, os experimentos com melhor balanço foram aqueles que utilizam  $\alpha = 0,1$  e  $\alpha = 0,3$ , uma vez que ambos os experimentos apresentaram uma coerência maior em relação a apresentar a mesma distância da curva em vermelho para ambos os Hemisférios.

O experimento que apresentou os maiores valores, tanto no Hemisfério Sul quanto no Hemisfério Norte foi o experimento que utilizou apenas o 3DVar.

Figura 5.26 - Idem à Figura 5.24, porém, para o trimestre JJA.



Fonte: Produção do autor.

De uma forma geral, comparando os dois períodos, pode-se concluir que os experimentos utilizando o Híbrido Médio, com valores baixos de alfa ou o alfa dinâmico, foram melhor balanceados que os demais. Por outro lado, para o período de DJF o experimento com o alfa dinâmico foi o melhor balanceado. Já para o período de JJA não existe uma coerência entre as métricas utilizadas. Foi possível observar que o balanço das análises não está diretamente ligado aos erros encontrados nelas (Seção 5.3), uma vez que as análises do experimento utilizando  $\alpha = 0, 1$  apresentaram valores de RMSE menores em ambos os períodos.

#### 5.4 Comparação entre os coeficientes dinâmico e estático

Conforme visto até aqui, o experimento com alfa fixo que apresentou uma melhor performance foi o que utiliza  $\alpha = 0, 1$ . Sendo assim, essa seção destina-se a comparações entre os experimentos que utilizam o  $\alpha = 0, 1$  e o alfa dinâmico. Serão apresentadas aqui comparações entre observação menos análise (O-A) e observação menos estimativa inicial (O-F), além de resultados do Ganho percentual do RMSE nas previsões para até 120 horas. Esses resultados serão apresentados para os dois períodos de estudo (i.e., inverno e verão).

### 5.4.1 Avaliação das análises e estimativa inicial contra as observações

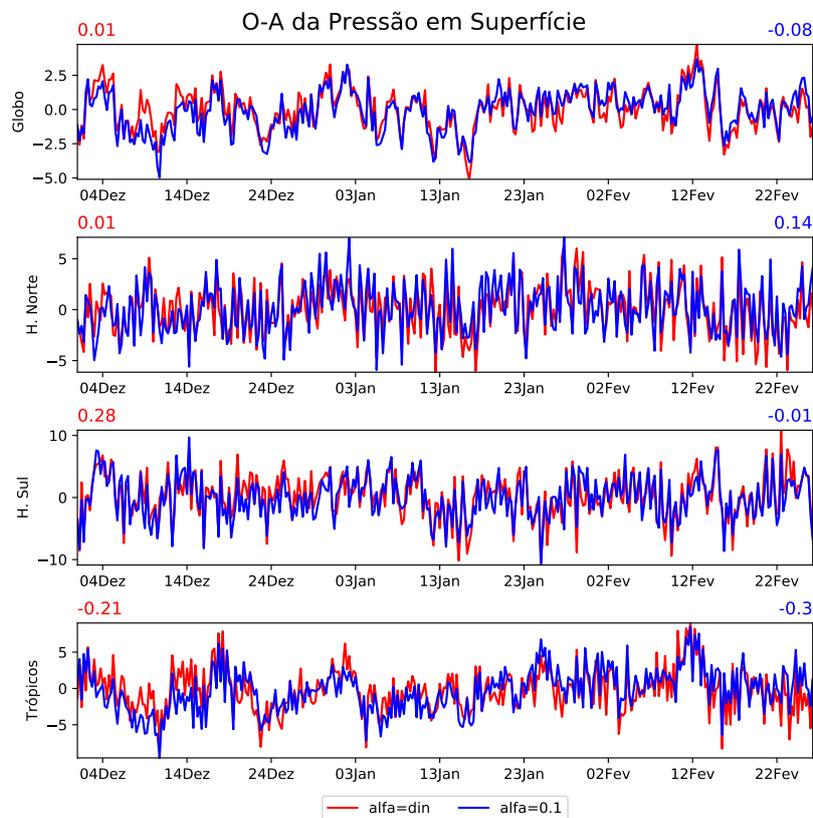
Uma análise mais detalhada foi feita com os experimentos que utilizam o  $\alpha = 0, 1$  e o alfa dinâmico. Foi analisado o comportamento médio do O-A e do O-F durante todo o período de estudos. Essas são estatísticas úteis para se avaliar o esforço do sistema de assimilação de dados em corrigir a previsão de curto prazo. Valores estimados de O-F mostram o quanto que a estimativa inicial precisaria para alcançar a verdade. Já o O-A mostra o quão distante a análise ainda está da "verdade". Aqui, são esperados valores próximos de zero. Essa avaliação foi feita para todos os níveis e variáveis do modelo, para as quatro regiões de estudo e para os dois trimestres (DJF e JJA).

Nas Figuras 5.27 e 5.28, são apresentadas as estatísticas com relação a O-A e O-F da pressão em superfície, respectivamente, para o período de dezembro de 1982 e, janeiro e fevereiro de 1983. Nelas as linhas em azul representam o valores para o experimento que utiliza  $\alpha = 0, 1$  e, as linhas em vermelho, representam o experimento que utiliza o alfa dinâmico. Cada gráfico representa uma região de estudo seguindo a seguinte ordem: Globo, Hemisfério Norte, Hemisfério Sul e região Tropical. Os valores mostrados no topo de cada gráfico representam a média do período. No topo à esquerda, em vermelho, estão os valores médios do experimento com alfa dinâmico; já no topo, à direita, em azul, está o experimento que utiliza  $\alpha = 0, 1$ .

Através desses resultados podemos inferir que o experimento com alfa dinâmico apresenta, na maioria dos casos, valores absolutos mais baixos que o experimento com  $\alpha = 0, 1$ . Apenas no Hemisfério Sul são vistos os menores valores absolutos para o experimento com alfa fixo. Esse resultado mostra que, ainda que os resultados vistos na Seção 5.2 tenham apresentado erros maiores para o experimento com alfa dinâmico, infere-se que pelo fato desse experimento estar melhor condicionado (como visto na Seção 5.3) ele apresenta esses resultados melhores do que o experimento com  $\alpha = 0, 1$ . Além disso, é possível observar que para o experimento com alfa dinâmico, a região em que ele apresenta os maiores valores médios de O-A e O-F é, para ambos os casos, o Hemisfério Sul, apresentando valores iguais a 0,28 e 0,43, respectivamente. Essa é uma região onde normalmente são encontrados erros maiores devido aos sistemas que ocorrem nela, como, por exemplo, os ciclones transientes. Também vale lembrar que essa região é, em sua grande maioria, oceânica; sendo assim, existem menos dados a serem assimilados, o que dificulta a correção da estimativa inicial por parte dos dados. Além da quantidade, a forma como os sistemas assimilam as observações são diferentes. O 3DVar costuma ser melhor utilizado quando há uma densidade grande de observações, e como neste estudo o alfa dinâmico utiliza

mais 3DVar do que o experimento com  $\alpha = 0,1$ , o experimento com alfa fixo pode estar utilizando esses dados de forma melhor. Observando o experimento que utiliza  $\alpha = 0,1$ , nota-se o maior valor de O-A e O-F na região Tropical (-0,3 e -0,27, respectivamente). Essa região, quando comparada com o Hemisfério Sul, apresenta uma densidade de dados maior e, ainda, os erros para a pressão em superfície nessa região costumam ser menores, uma vez que nela a pressão em superfície não varia tanto quanto em médias e altas latitudes.

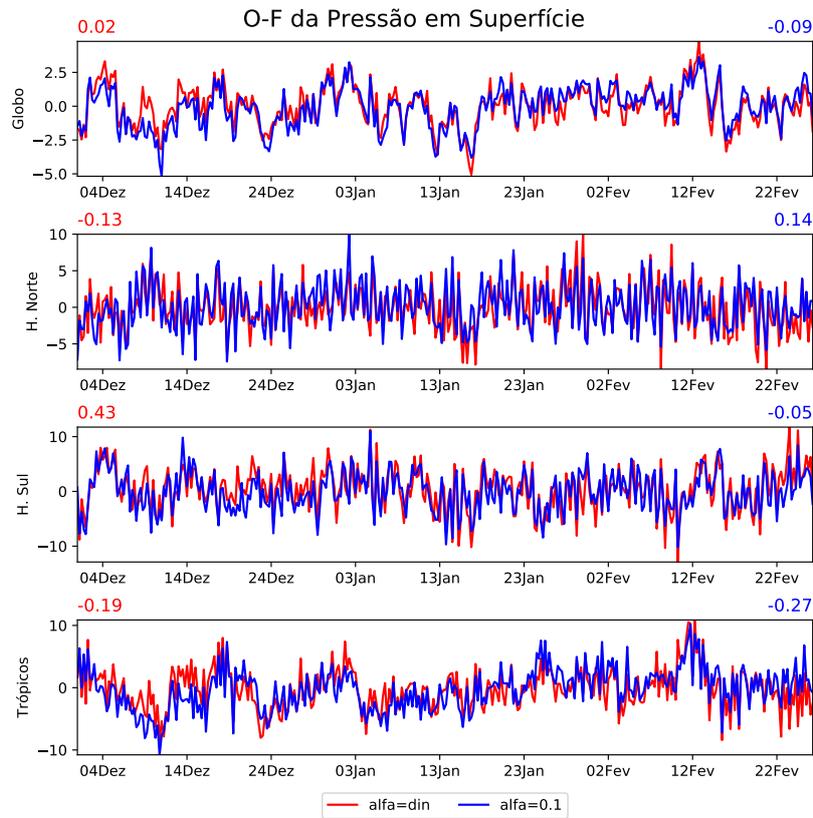
Figura 5.27 - O-A da pressão em superfície no Globo, no Hemisfério Norte, no Hemisfério Sul e na região Tropical, para o período de 1 de dezembro de 1982 a 28 de fevereiro de 1983.



Em azul estão os valores do experimento utilizando  $\alpha = 0,1$  e em vermelho os valores do experimento utilizando o  $\alpha$  dinâmico. No canto superior à direita de cada gráfico (em azul) o valor médio da série para o experimento que utiliza  $\alpha = 0,1$  e à esquerda o valor médio da série para o experimento que utiliza o alfa dinâmico. No eixo  $x$  encontra-se o período e no eixo  $y$  os valores da diferença.

Fonte: Produção do autor.

Figura 5.28 - Idem à Figura 5.27, porém, para os valores de O-F.

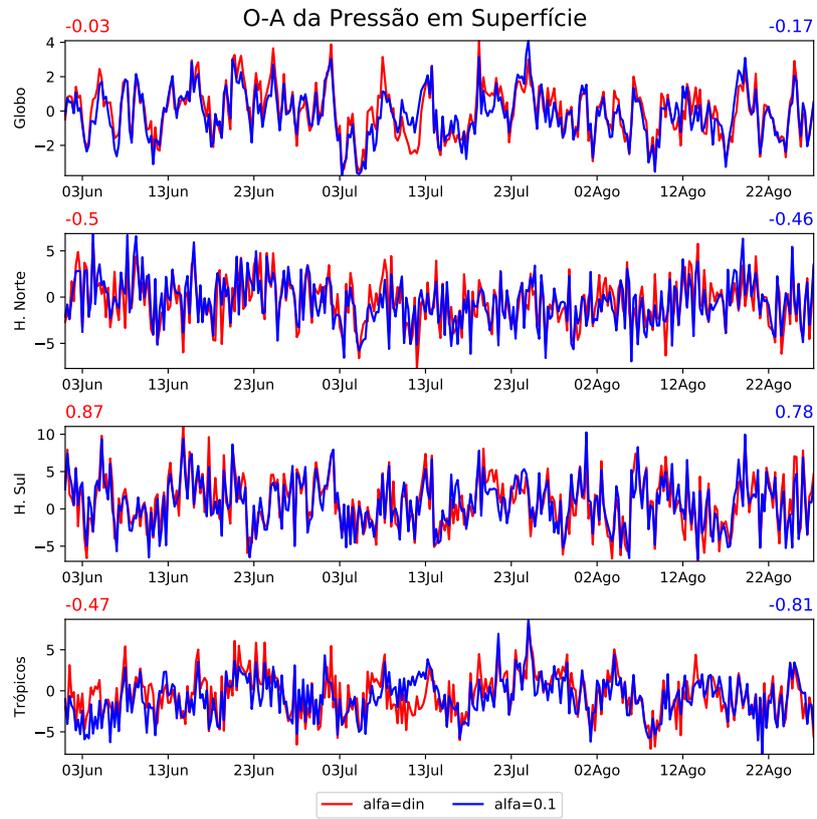


Fonte: Produção do autor.

As Figuras 5.29 e 5.30 mostram os resultados de O-A e O-F para o período de junho, julho e agosto de 1982, seguindo o mesmo padrão das Figuras 5.27 e 5.28 apresentadas anteriormente. Nelas pode-se observar que os valores estão mais elevados quando comparado ao período anterior. Nota-se ainda na Figura 5.29 que no Globo e Trópicos o experimento que utiliza o alfa dinâmico mostra valores médios absolutos menores que o experimento que utiliza o  $\alpha = 0, 1$ . O contrário é visto para o caso dos Hemisférios, onde os menores valores foram obtidos pelo experimento com  $\alpha = 0, 1$  nessa região. Esse padrão é visto também na Figura 5.30. Nessa época do ano espera-se que os maiores erros com relação a pressão em superfície estejam localizados no Hemisfério Sul devido à frequente passagem de sistemas transientes, o que explica os maiores valores para o experimento com alfa dinâmico. No período de dezembro, janeiro e fevereiro, esses sistemas também atuam, porém com intensidade menor e localizados mais ao sul. No período de inverno austral, além deles serem mais intensos, eles avançam mais para norte, entrando mais na região que é

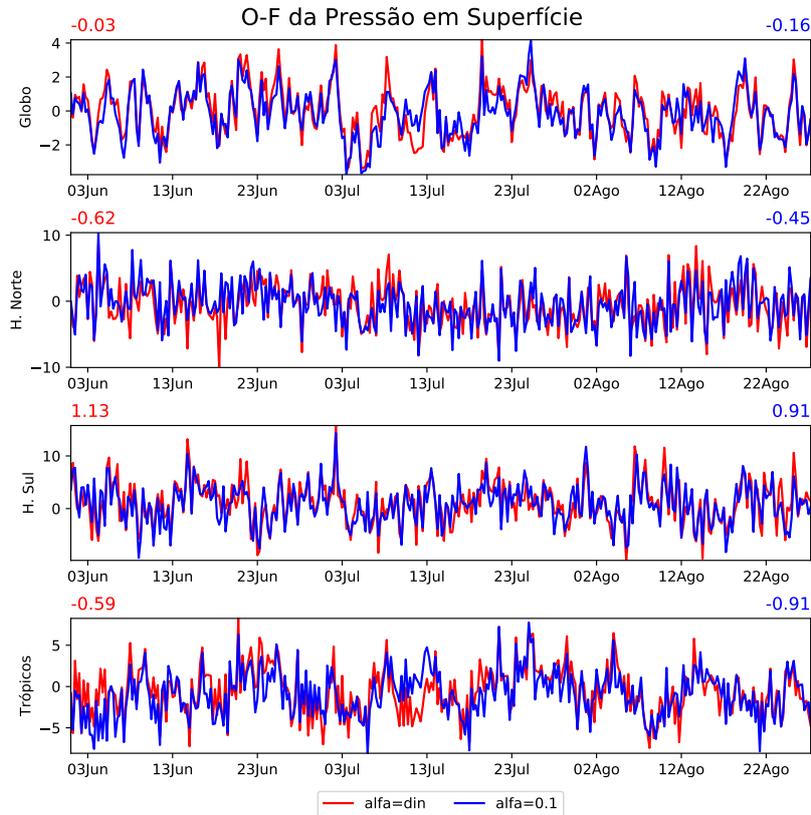
chamada de Hemisfério Sul aqui (entre 80°S e 20°S).

Figura 5.29 - Idem à Figura 5.27, porém, para o período de junho, julho e agosto de 1982.



Fonte: Produção do autor.

Figura 5.30 - Idem à Figura 5.28, porém, para o período de junho, julho e agosto de 1982.



Fonte: Produção do autor.

Com os resultados apresentados aqui, conclui-se que, independentemente da época do ano, o experimento utilizando o alfa dinâmico, apresenta estatísticas melhores do que o experimento com alfa fixo. No entanto, a maior deficiência em relação ao experimento que utiliza o alfa dinâmico encontra-se no Hemisfério Sul para o caso de verão austral e em ambos os Hemisférios para o caso do inverno austral. Uma hipótese para isso é o fato de que no Hemisfério Sul existe um número menor de observações para serem assimiladas e, conseqüentemente, não consegue melhorar o desempenho de ambos os sistemas Híbridos, já que em regiões com uma densidade grande de dados o sistema que utiliza o alfa dinâmico apresenta um bom desempenho.

Os gráficos referentes ao O-A e O-F das demais variáveis não foram mostrados aqui, pois, na grande maioria, apresentaram os mesmos valores para o experimento utilizando o alfa dinâmico e para o experimento utilizando alfa fixo, sendo eles

majoritariamente igual a zero.

#### 5.4.2 Avaliação das Previsões de até 120 horas

Esta seção destina-se a avaliar o ganho percentual no RMSE que o sistema Híbrido Médio, utilizando o alfa dinâmico, tem nas previsões de até 120 horas em relação ao RMSE do Híbrido Médio que utiliza o  $\alpha = 0, 1$ , da mesma forma utilizada por [Baños \(2017\)](#) e [Silveira \(2017\)](#). Esses valores foram estimados a partir de valores de RMSE previamente calculados. O Ganho é calculado conforme é mostrado na Equação 5.3:

$$GANHO_a^{RMSE} = \frac{RMSE_a^{exp} - RMSE_a^{ctr}}{RMSE_a^{perfeito} - RMSE_a^{exp}} \times 100\% \quad (5.3)$$

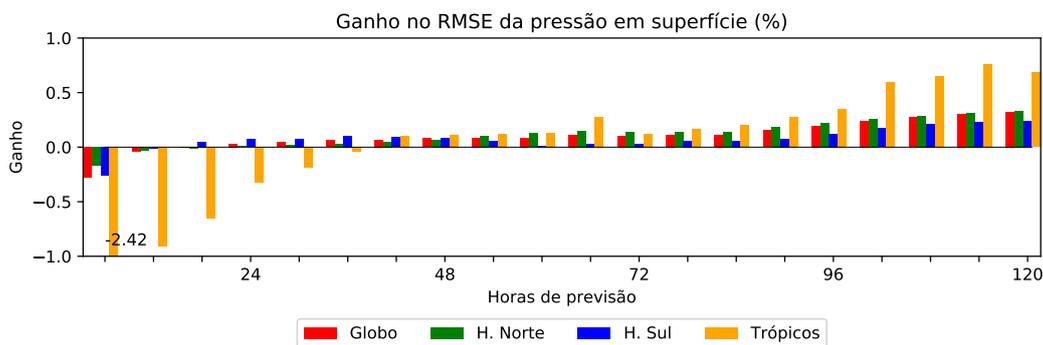
onde o índice  $a$  representa uma determinada variável; o índice  $exp$  representa o experimento que está sendo avaliado; para o caso desta tese, foram utilizados os valores de RMSE para o experimento com alfa dinâmico; o índice  $ctr$  representa o experimento controle, que, aqui, são os experimentos com alfa fixo; e o índice  $perfeito$  representa o RMSE esperado para que se obtenha o menor erro - neste caso, espera-se um RMSE igual a zero. O resultados serão apresentados para todas as variáveis e níveis do modelo, para as quatro regiões de estudo. Quando o ganho for positivo, o experimento que utiliza o alfa dinâmico apresenta um melhor desempenho em relação ao experimento que utiliza  $\alpha = 0, 1$ ; quando o ganho é negativo, é visto o contrário.

A Figura 5.31 mostra os valores do Ganho percentual do RMSE para a pressão em superfície para todos os horários de previsão (06, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66, 72, 78, 84, 90, 96, 102, 108, 114 e 120 horas) nas quatro regiões de estudo, sendo em vermelho o Globo, em verde o Hemisfério Norte, em azul o Hemisfério Sul e em amarelo a região Tropical. Na figura nota-se, na grande maioria dos horários de previsão, ganhos positivos. Observa-se ganhos negativos até 36 horas para a região Tropical até as primeiras 18 horas no Hemisfério Norte, nas primeiras 12 horas para o Globo e nas primeiras 6 horas para o Hemisfério Sul. Quando comparada com a Tabela 5.9, nota-se que a partir da análise até as primeiras 6 horas de previsão houve uma melhora grande por parte do sistema que utiliza o alfa dinâmico, e que existe uma tendência de melhora no ganho do RMSE ao longo dos horários de previsão.

Avaliando do ponto de vista de porcentagem, o Globo apresentou 85% de ganhos positivos, 10% de ganhos negativos e 5% dos horários não apresentaram ganho. O Hemisfério Norte apresentou 80% de ganhos positivos, 15% de ganhos negativos e 5% dos horários não apresentaram ganhos nem positivos e nem negativos. Já para o

Hemisfério Sul houve 85% de ganhos positivos, contra apenas 5% de ganhos negativos e apenas 10% com zero de ganho. Para o região Tropical 70% dos ganhos foram positivos, 30% negativos e não houve horários sem ganho. Esses valores mostram que o Hemisfério Sul foi a região onde a utilização do sistema Híbrido com o alfa dinâmico teve uma melhor performance, apresentando um ganho positivo em quase todas as situações. Por outro lado, a região Tropical foi a menos afetada pelo uso do alfa dinâmico. Ainda assim, essa região apresentou ganhos positivos na maioria dos horários. Embora visto nas análises (apresentado na Tabela 5.9) que o experimento que utiliza  $\alpha = 0,1$  tenha se mostrado melhor que o experimento utilizando alfa dinâmico, infere-se que essa melhora no ganho ao longo das horas de previsão se dê pela estabilidade do sistema híbrido quando utilizado o alfa dinâmico. Como mostrado anteriormente, esse sistema é mais balanceado em relação ao experimento que utiliza  $\alpha = 0,1$ .

Figura 5.31 - Ganho percentual nos valores de RMSE da pressão em superfície para as previsões de até 120 horas pertencentes ao período de dezembro, janeiro e fevereiro.



No eixo  $x$  estão as horas de previsão e no eixo  $y$  os valores de ganho relativos (%). Cada região está representada por uma cor: em vermelho o Globo, em verde o Hemisfério Norte, em azul o Hemisfério Sul e em amarelo os Trópicos.

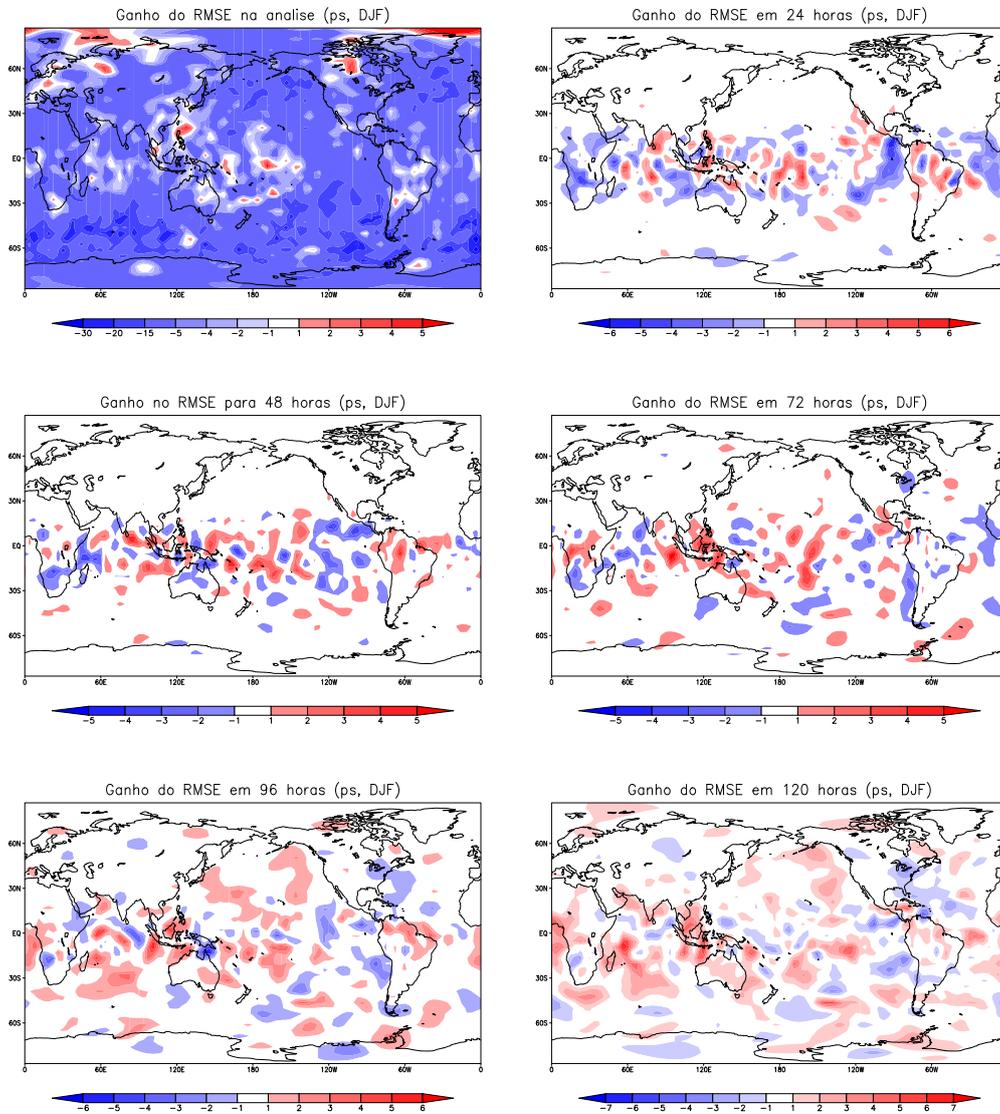
Fonte: Produção do autor.

Ressalta-se que, a intenção do sistema Híbrido Médio é utilizar o sistema variacional para tornar o sistema por conjunto mais estável. Esse padrão observado nas previsões em que as primeiras horas de previsão são melhor representadas pelo experimento com alfa fixo e, em seguida, as previsões do experimento com alfa dinâmico dominam os resultados pode estar relacionado a essa combinação de ambos os sistemas. Nas primeiras horas de previsão as observações tem uma importância muito grande e,

o LETKF, por utilizar os erros do dia do modelo, consegue representar melhor os sistemas de alta frequência. Já no decorrer das previsões a solução é dominada mais pelo modelo do que pelas observações e, com isso, o sistema que utiliza erros climatológicos do modelo acaba por ter um peso maior. Quando utilizado  $\alpha = 0,1$ , aproximadamente 90% do experimento é de LETKF; por esse motivo, infere-se que o ganho negativo visto nos primeiros horários seja decorrente do melhor uso dos erros do dia.

Na Figura 5.32 são mostradas as distribuições espaciais do ganho do RMSE médio mostrado anteriormente. Os tons em azul significam um ganho negativo, ou seja, o experimento que utiliza o alfa dinâmico não beneficia o sistema, e, os tons em vermelho significam ganhos positivos, ou seja, regiões onde houve uma melhora no RMSE por parte do alfa dinâmico. É possível observar no campo de análise que em sua grande maioria o ganho visto é negativo, corroborando o resultado visto na Tabela 5.9 da seção de análise. Nela encontram-se valores maiores de erros para o Hemisfério Sul. No escopo das avaliações realizadas, os Hemisférios variam entre 80°S e 20°S para o Hemisfério Sul e entre 20°N e 80°N para o Hemisfério Norte, além da região Tropical abranger a faixa entre 20°S e 20°N. Dessa forma, vê-se que as poucas regiões com ganhos positivos ou estão localizadas, na maioria das vezes, dentro da região Tropical, como, por exemplo, no Oceano Pacífico Oeste, ou ainda no Polo Norte, região essa que não entra na faixa que abrange o Hemisfério Norte. Também são vistos pontos alcançando 5% de ganho no Hemisfério Norte, sobre o Canadá, e pontos sobre a Europa. No Hemisfério Sul, as regiões de ganhos positivos ficaram localizadas principalmente ao sul da Austrália e América do Sul. Nessa região destaca-se um ponto sobre o sul da Austrália, onde apresenta ganhos positivos. Essa é uma região onde ambos os sistemas de assimilação de dados (LETKF e 3DVar) apresentam erros, uma vez que o 3DVar apresenta um erro maior. Ao observar o campo de alfa médio (apresentado na Figura 4.2), nota-se que nessa região o alfa é maior que 0,5 no oceano, ou seja, para essa região o sistema utiliza mais o 3DVar do que o LETKF. Com isso, pode-se inferir que, por se tratar de sistemas não lineares, embora haja um erro maior no 3DVar, isso não se reflete na combinação de ambos.

Figura 5.32 - Distribuição espacial do ganho percentual do RMSE médio da pressão em superfície na análise e nas previsões de 24, 48, 72, 96 e 120 horas para o trimestre DJF.



Fonte: Produção do autor.

Ao comparar os resultados das figuras de previsões com a Figura 5.31, nota-se coerência. Em 24 horas de previsão, por exemplo, o ganho maior (em módulo) é visto na região Tropical, região essa que contém a maior quantidade de áreas com ganhos. Já em 48 horas percebe-se ainda um maior ganho na região Tropical, porém, as regiões com ganhos estendem-se ao Hemisfério Sul. Para esse horário, a maioria desses ganhos, em média, foram positivos. Para as 72 horas de previsão, observam-se os

pontos de ganhos mais intensos ainda sobre a região Tropical, principalmente sobre a região da Indonésia, no meio do Oceano Pacífico e sobre a África. O Hemisfério Sul foi a região que apresentou menor ganho, ainda que positivo. Observando a figura, pode-se notar vários pontos positivos e negativos, sendo a maioria sobre a região oceânica. Em 96 horas de previsão nota-se mais pontos vermelhos do que azuis no globo, representando assim uma predominância de valores de ganhos positivos. A região Tropical continua sendo a região que apresentou um maior ganho, apresentando áreas de maior valor ainda sobre a região da Indonésia, Oceano Índico Norte, Oceano Atlântico Central e norte da América do Sul. Para o último horário de previsão, é possível observar a grande maioria de regiões com ganhos positivos, sendo mais intensos na região Tropical, e os ganhos negativos mais intensos ficaram localizados sobre o Oceano Pacífico Leste e Oceano Atlântico Central, leste dos Estados Unidos e Canadá.

Ao observar ambas as figuras, pode-se concluir em que à medida que as horas de previsão avançam, há uma tendência de aumentar o ganho positivo. Na análise há um ganho negativo grande. Em contrapartida, já em 24 horas há uma melhora nos Hemisférios e, em seguida, o ganho positivo passa a dominar inclusive na região Tropical.

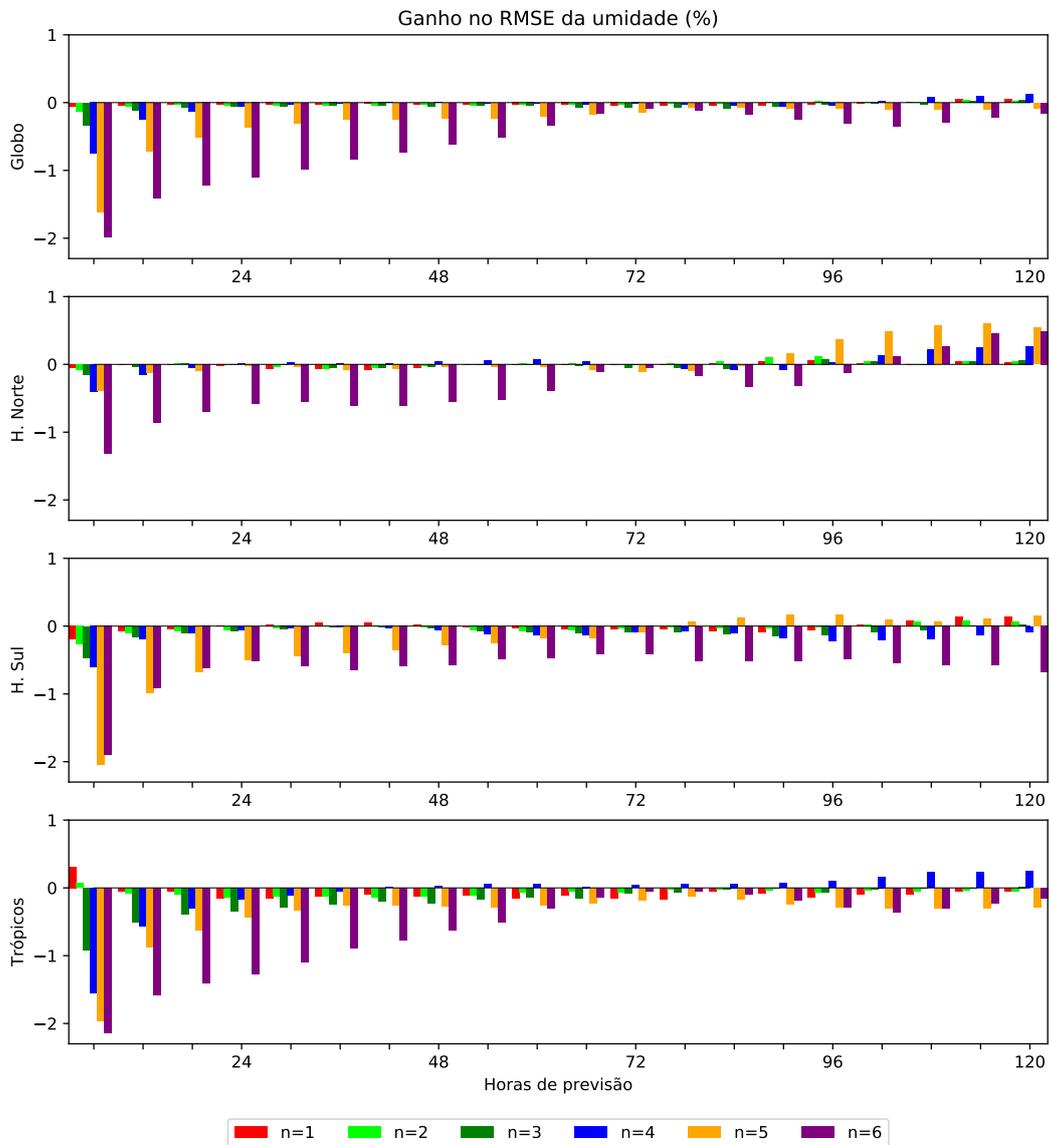
Na Figura 5.33 é apresentado o Ganho do RMSE nas previsões de até 120 horas para a umidade nos seis níveis do modelo para as quatro regiões de estudo. Nela e nas figuras seguintes, cada cor representa um nível, sendo em vermelho o nível 1, em verde claro o nível 2, em verde escuro o nível 3, em azul o nível 4, em amarelo o nível 5 e em roxo o nível 6. É possível notar que na grande maioria dos casos o ganho foi negativo, mostrando que o experimento com o alfa dinâmico não foi superior ao experimento com  $\alpha = 0,1$  na maioria dos casos. O Ganho positivo apareceu no Globo nos últimos horários de previsão, sendo a partir de 108 horas para o nível 4 e a partir de 114 horas para os níveis 1, 2 e 3. No Hemisfério Norte os valores positivos aparecem já em 18 horas nos níveis 2 e 3. O nível 4 apresenta valores positivos na maioria dos horários de previsão, sendo eles 24, 30, 36, 42, 48, 54, 60, 66, 96, 102, 108, 114 e 120 horas. O restante dos valores positivos foram vistos a partir de 78 horas para os demais níveis. Com relação ao Hemisfério Sul os valores positivos foram vistos no nível 1, de 30 a 48 horas de previsão, e a partir de 102 horas. O nível 5 apresentou valores de ganhos positivos a partir de 78 horas e o nível 2 a partir de 102 horas. O nível 3 foi melhor apenas no último horário de previsão. Já na região Tropical esses valores positivos ficaram restritos ao primeiro horário (6 horas) para os níveis 1 e 2, e a partir de 42 horas o nível 4. Exceto pelos resultados vistos no Hemisfério Sul, o

padrão de diminuição do ganho negativo visto no campo da pressão em superfície continua sendo observado aqui, mostrando uma melhora no desempenho do Híbrido Médio quando utiliza o alfa dinâmico em relação ao experimento com alfa fixo. Porém, ainda assim os resultados mostram que o experimento com o alfa dinâmico não apresentou um desempenho melhor do que quando utilizado  $\alpha = 0, 1$ .

O ganho no RMSE da temperatura nos sete níveis é mostrado na Figura 5.34. Nela, é possível notar valores altos (em módulo) nos primeiros horários de previsão, e, à medida que as horas passam, esses valores diminuem; em alguns casos, como o Globo e o Hemisfério Norte, eles decrescem rapidamente. Porém, no Hemisfério Sul são vistos ganhos positivos já a partir de 18 horas de previsão em alguns níveis do modelo. Na região Tropical os valores de ganho também decrescem com o passar dos horários de previsão, porém não tão acentuados quanto no Globo e Hemisfério Norte.

A Figura 5.35 mostra o ganho do RMSE para a componente zonal do vento. Nela a grande maioria dos valores de ganho foram positivos, exceto para o Hemisfério Sul, onde todos os níveis apresentaram maioria de valores negativos. Essa foi a variável que mais apresentou ganhos positivos. Nota-se valores elevados de ganho negativo nos primeiros horários de previsão e esses valores diminuem à medida que avançam os horários, tornando-se positivos para todos os níveis, exceto no Hemisfério Sul, que apresenta valores negativos de ganho sempre para o nível 4. Aqui, como visto até então, os resultados seguem o mesmo padrão, apresentando uma melhora do experimento que utiliza o alfa dinâmico conforme avançam os horários de previsão. Esse padrão continua sendo observado na Figura 5.36 para o Globo e a região Tropical, onde são apresentados os resultados para a componente meridional do vento. Nela é possível observar os maiores valores para a região Tropical, tanto em ganhos negativos quanto em ganhos positivos. Observa-se também ganhos positivos em todos os horários de previsão para o nível 2, e na maioria dos níveis 1 e 4.

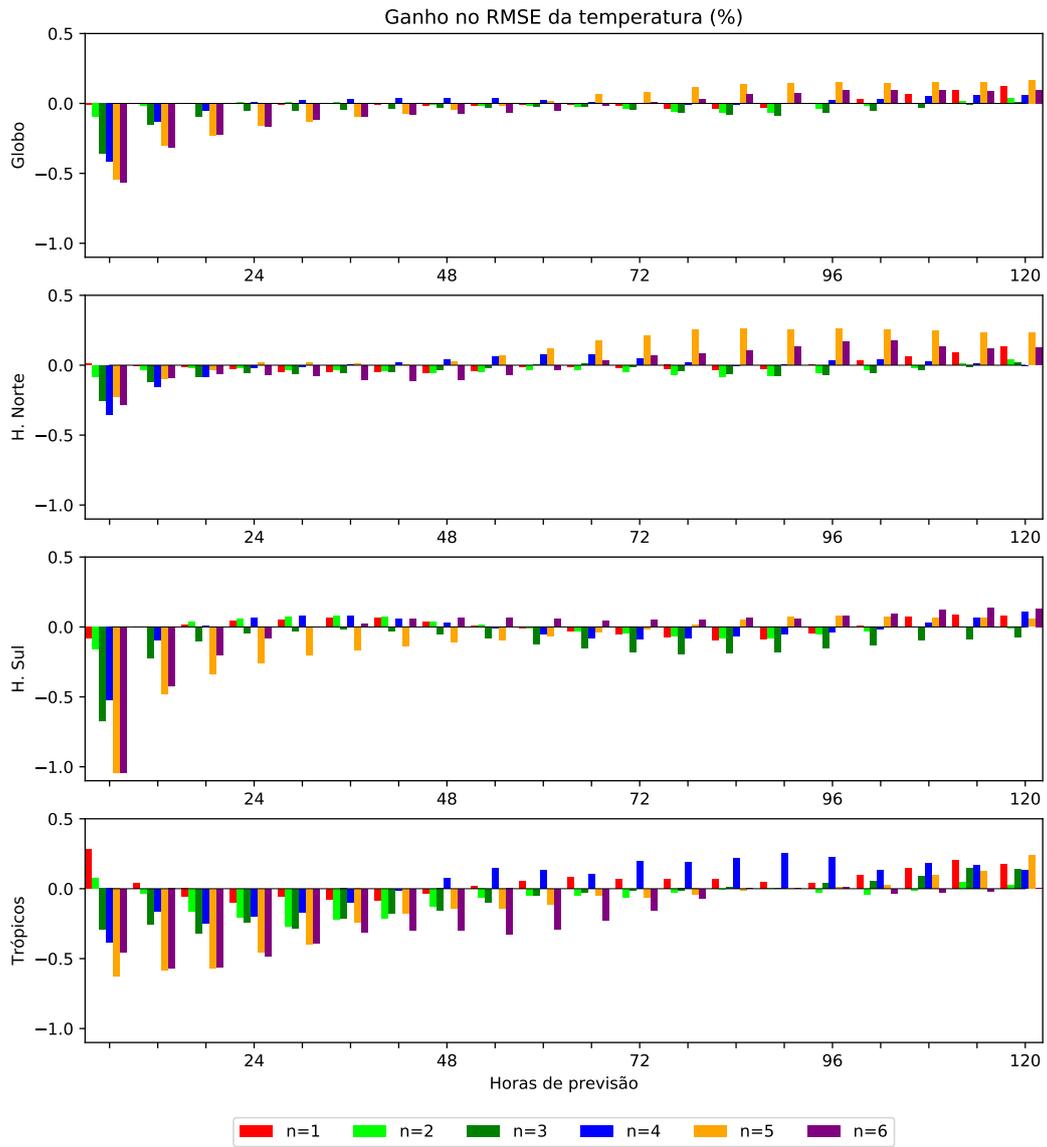
Figura 5.33 - Ganho percentual nos valores de RMSE da umidade para as previsões de até 120 horas pertencentes ao período de dezembro, janeiro e fevereiro.



No eixo  $x$  estão as horas de previsão e no eixo  $y$  os valores de ganho relativos (%). O primeiro gráfico é referente ao Globo, o segundo é referente ao Hemisfério Norte, o terceiro é referente ao Hemisfério Sul e o último a região Tropical. Cada nível está representado por uma cor: nível 1 em vermelho, nível 2 em verde claro, nível 3 em verde escuro, nível 4 em azul, nível 5 em amarelo e nível 6 em roxo.

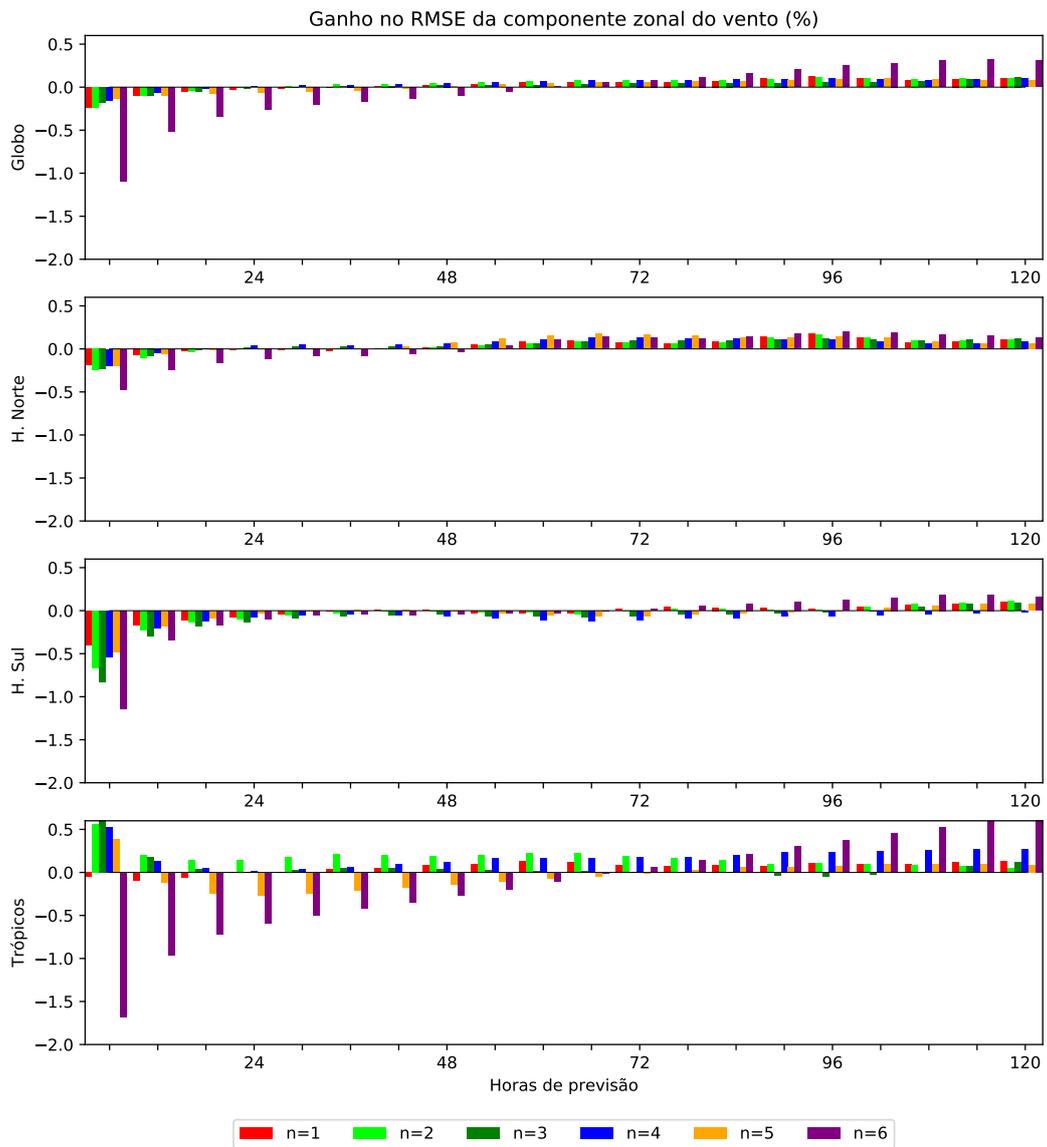
Fonte: Produção do autor.

Figura 5.34 - Idem à Figura 5.33, porém para a temperatura.



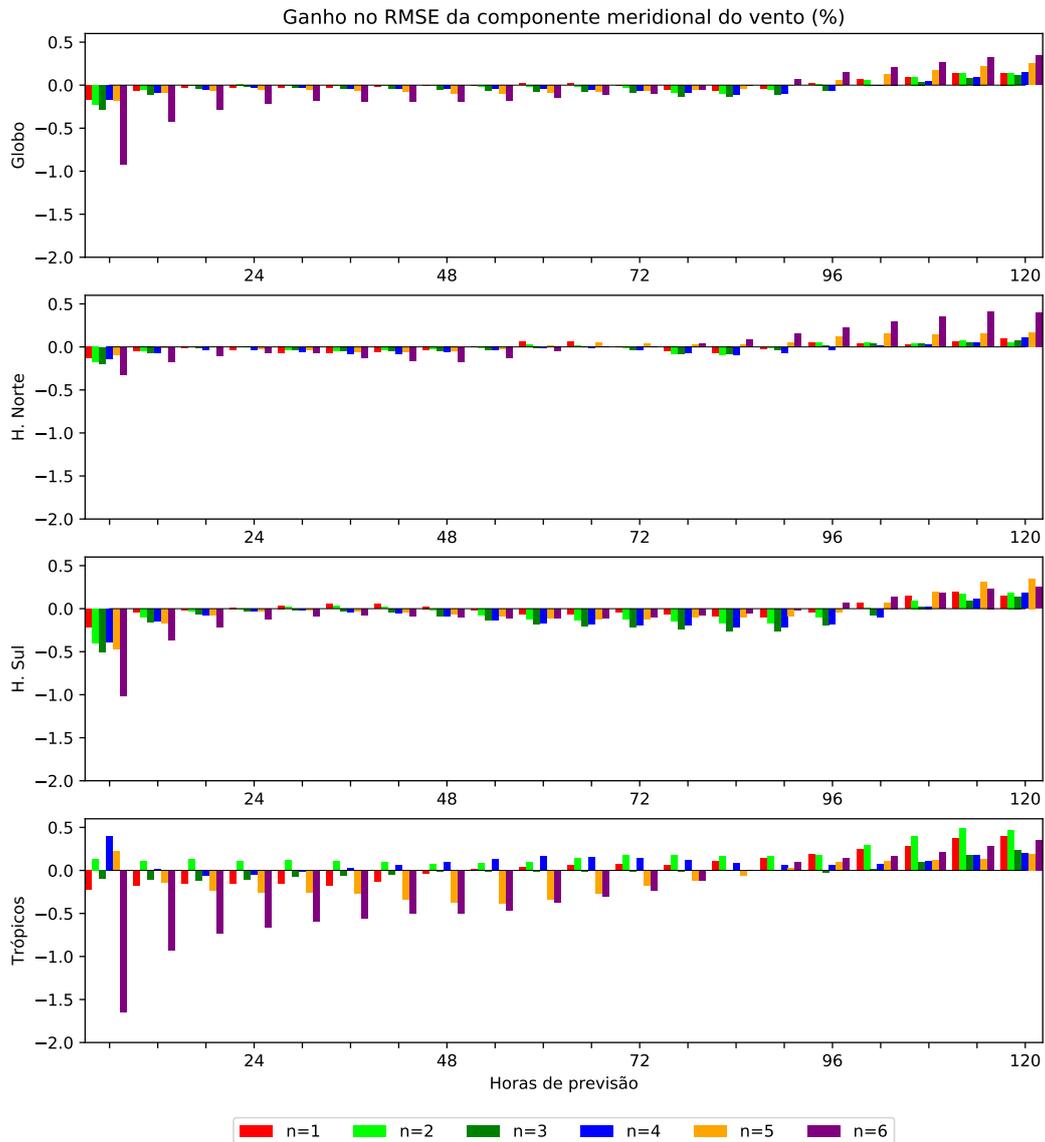
Fonte: Produção do autor.

Figura 5.35 - Idem à Figura 5.34, porém, para a componente zonal do vento.



Fonte: Produção do autor.

Figura 5.36 - Idem à Figura 5.34, porém, para a componente meridional do vento.



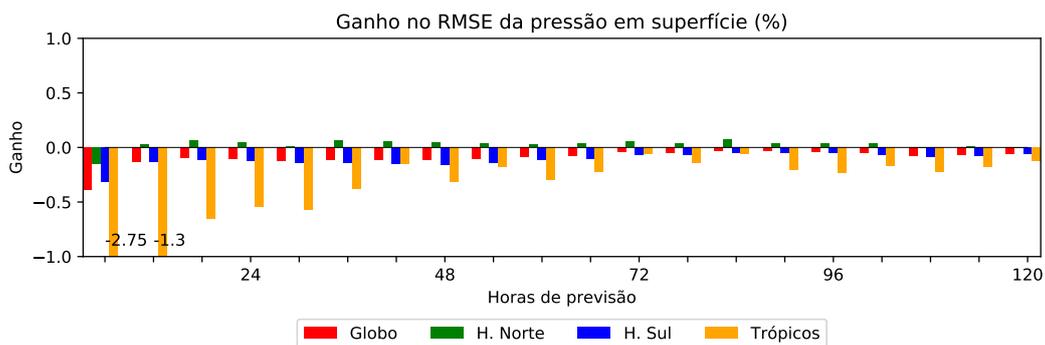
Fonte: Produção do autor.

Os resultados apresentados mostraram um padrão de melhora do experimento que utiliza o alfa dinâmico em relação ao experimento que utiliza  $\alpha = 0, 1$ . Em alguns casos, como a pressão em superfície e a componente zonal do vento, o experimento com o alfa dinâmico apresentou ganhos positivos na maioria dos casos. Quando comparados com os resultados das previsões geradas pelo LETKF e pelo 3DVar, nota-se aqui o mesmo padrão com relação à diferença dos erros de ambos os sistemas (Figura 5.13), que mostra o LETKF com um desempenho melhor que o 3DVar nas

primeiras horas de previsão. Quando é utilizado o  $\alpha = 0,1$ , aproximadamente 90% de LETKF é usado na confecção da análise final; já os valores do alfa dinâmico variam, utilizando, na média, mais 3DVar do que quando o alfa é fixo. Nota-se uma semelhança nos perfis da previsões, nas quais o sistema que utiliza mais o LETKF apresenta uma melhor performance nos primeiros horários. Acredita-se que devido a essa diminuição nos erros o sistema Híbrido Médio, utilizando o alfa dinâmico, tenha potencial para melhorar seus resultados.

Até o presente momento, foram discutidos os resultados para o período de dezembro, janeiro e fevereiro. A seguir, serão apresentados os resultados do ganho percentual do RMSE para o período de junho, julho e agosto, seguindo o mesmo padrão do período anterior. Na Figura 5.37 são apresentados os resultados do Ganho do RMSE nas previsões de até 120 horas, para a pressão em superfície nas quatro regiões de estudo: o Globo em vermelho, o Hemisfério Norte em verde, o Hemisfério Sul em azul e os Trópicos em amarelo. Observa-se, em sua grande maioria ganhos negativos, exceto para o Hemisfério Norte, que entre 12 horas de previsão e 102 horas apresenta, ainda que pequenos, ganhos positivos. Os ganhos negativos maiores foram vistos na região Tropical, chegando a quase 3% de ganho em 6 horas de previsão. O padrão visto para o trimestre DJF, se mantém no trimestre JJA. Nota-se uma diminuição no ganho negativo à medida que as horas de previsão avançam.

Figura 5.37 - Idem à Figura 5.31, porém, para o período de junho, julho e agosto.

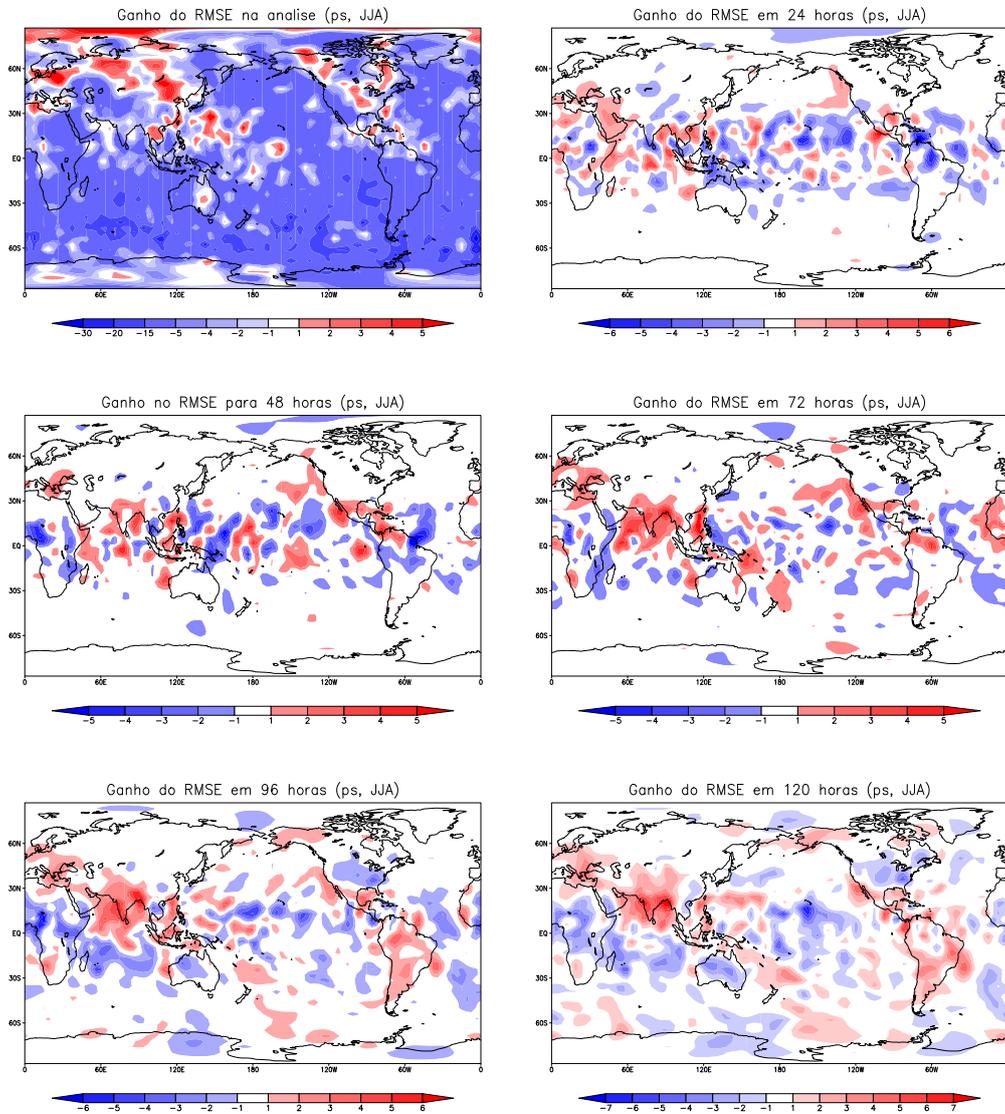


Fonte: Produção do autor.

Na Figura 5.38, são apresentadas as distribuições espaciais do Ganho percentual do RMSE médio no tempo, para a pressão em superfície na análise, e em 24, 48, 72, 96 e 120 horas de previsão. Nessa figura os tons em azul representam um ganho

negativo e os tons em vermelho representam ganhos do RMSE positivos. No campo de análise observam-se ganhos positivos mais localizados no Hemisfério Norte, o que corrobora o visto na Tabela 5.10, que mostra os menores erros na média para essa região. Observa-se também que os valores de ganho são muito maiores nos ganhos negativos do que nos positivos, chegando a seis vezes do valor máximo positivo. Quando se olha para o campo de ganho percentual em 24 horas de previsão, nota-se valores mais aproximados de ganhos positivos e negativos. Percebe-se ainda esses ganhos mais localizados na faixa entre 30°S e 30°N. No Hemisfério Sul são vistas poucas regiões com ganhos maiores que 1 (em módulo) e, dessas poucas regiões, vê-se mais ganhos negativos. Já o contrário é observado para o Hemisfério Norte, onde notam-se regiões com valores maiores de ganhos positivos. Esses resultados corroboram o que foi visto na Figura 5.37. Em 48 e 72 horas de previsão, nota-se que os valores de ganho foram intensificados, tanto os negativos quanto os positivos, porém continuam concentrados mais na região Tropical. Em 96 e 120 horas essas regiões de ganhos se espalham pelo Globo. Foi observado, a partir de 48 horas de previsão, um ganho positivo sobre a região do Tibete. Essa região foi aumentando e intensificando os valores de ganho até as 120 horas.

Figura 5.38 - Distribuição espacial do Ganho percentual do RMSE médio da pressão em superfície na análise e nas previsões de 24, 48, 72, 96 e 120 horas, para o trimestre JJA.

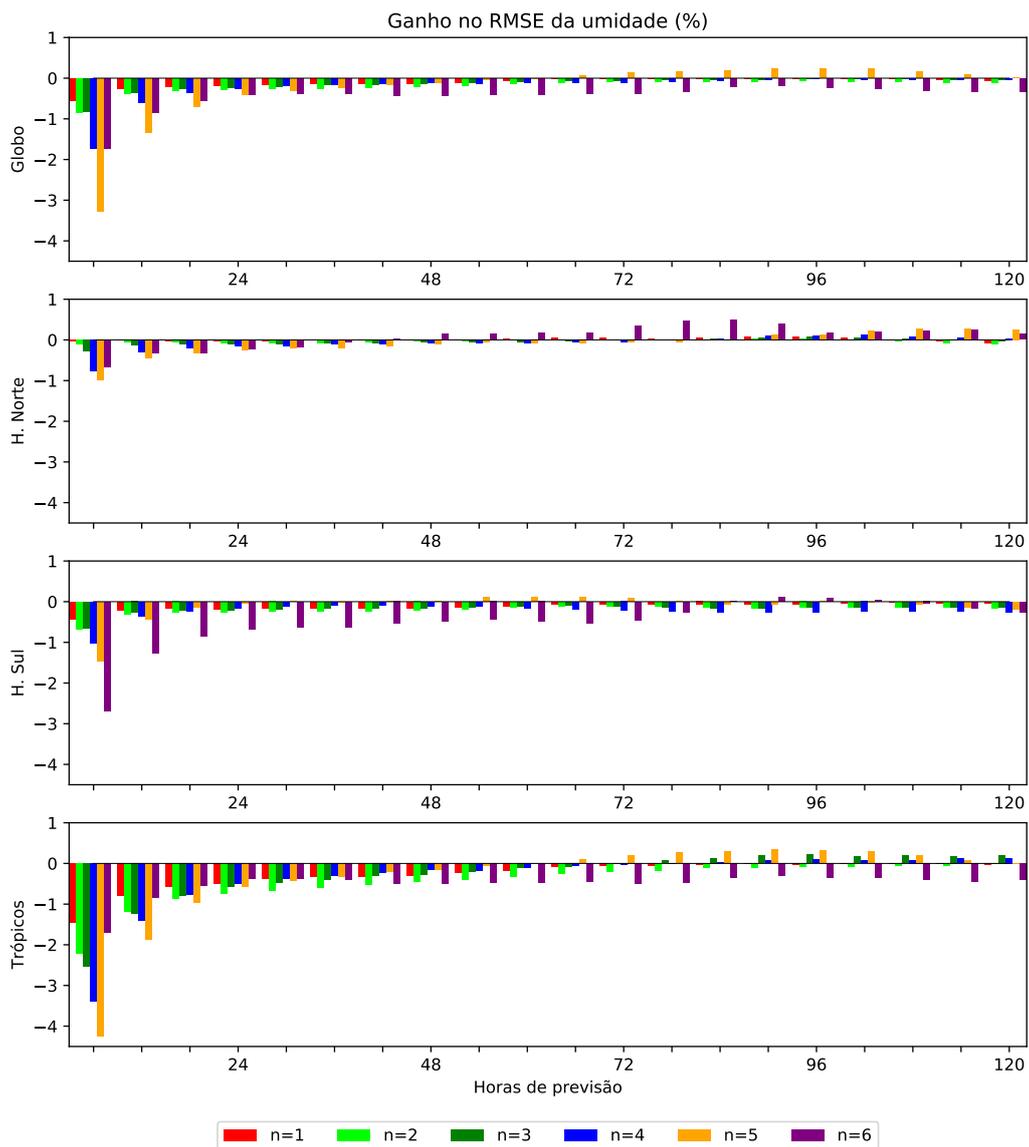


Fonte: Produção do autor.

A Figura 5.39 expõe os valores do ganho percentual do RMSE ao longo de 120 horas de previsão, da mesma forma que foi apresentado anteriormente, porém agora para o período de junho, julho e agosto. Nela é possível observar valores elevados (em módulo) nos primeiros horários de previsão. Quando comparado com o período anterior, esses valores são maiores (em módulo) para o trimestre JJA. Também é possível observar que a grande maioria dos resultados foram negativos, apresentando

alguns valores de ganhos positivos a partir de 66 horas de previsão no Globo e Trópico, 42 horas no Hemisfério Norte e a partir de 54 horas no Hemisfério Sul. Ao comparar com o período de verão austral, nota-se aqui valores maiores de ganhos negativos nas primeiras horas de previsão e, para Globo e Hemisfério Norte, por exemplo, mais ganhos positivos no inverno, ainda que com valores menores.

Figura 5.39 - Idem à Figura 5.33, porém para o período de junho, julho e agosto.



Fonte: Produção do autor.

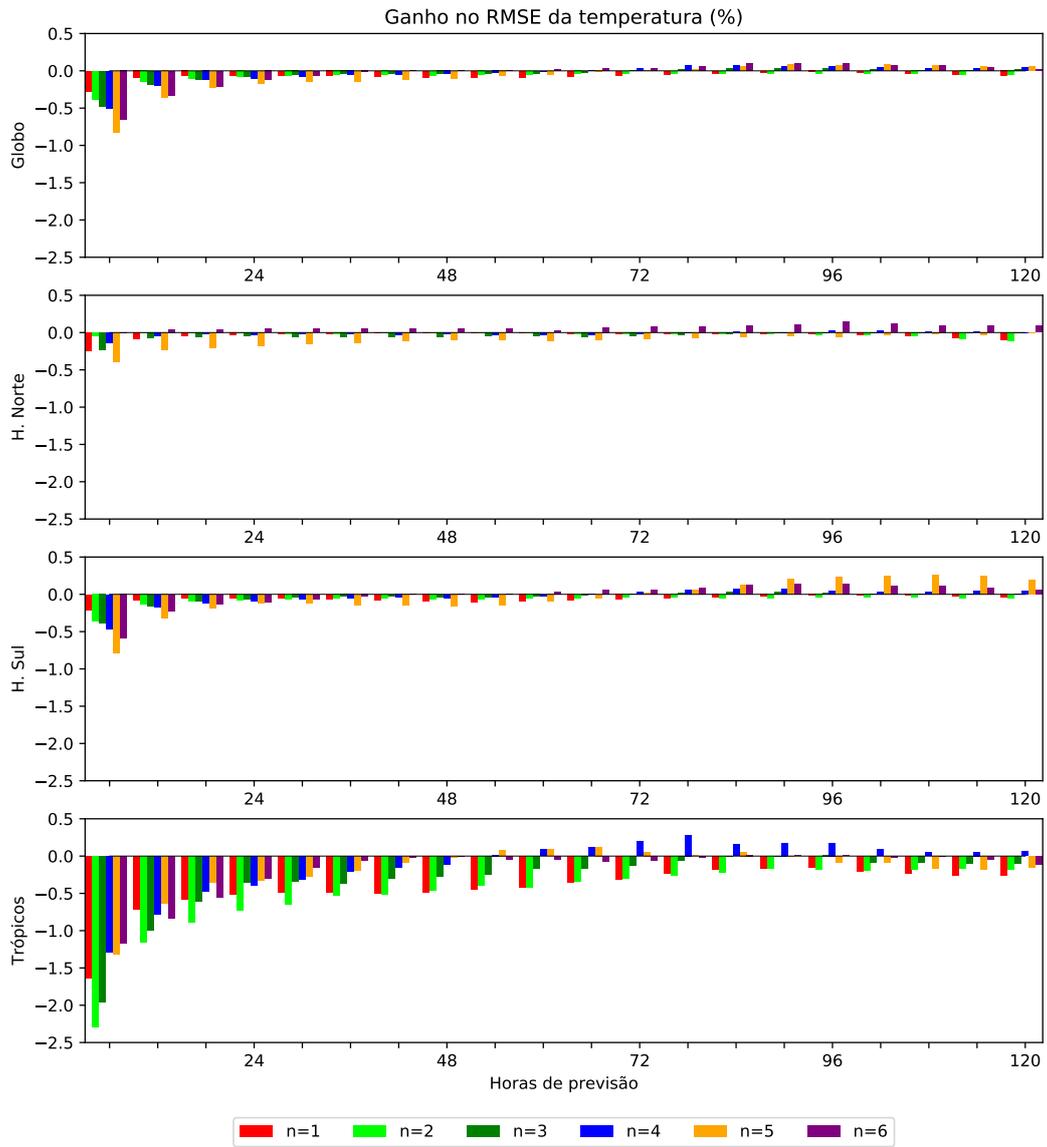
Na Figura 5.40 estão os valores de ganho para a temperatura. Observa-se os maiores

valores de ganhos na região tropical, e o padrão visto no período anterior, continua sendo visto aqui, os valores maiores (em módulo) são vistos nos primeiros horários de previsão, enquanto que, ao longo dos horários de previsão esse ganho negativo diminui, apresentando em alguns níveis ganhos positivos. Em relação ao período anterior, notam-se os maiores valores de ganhos negativos para o trimestre de JJA.

Os ganhos para as componentes zonal e meridional do vento são apresentados nas Figuras 5.41 e 5.42, respectivamente. Nelas, é possível observar, a grande maioria de ganhos negativos, principalmente para a componente meridional do vento, a qual apresentou ganhos positivos apenas no Hemisfério Norte e em quatro horários de previsão na região Tropical. Ao comparar com o período de dezembro, janeiro e fevereiro, nota-se uma grande diferença para essas variáveis, a começar pelos ganhos positivos que são vistos na maioria do Globo e Hemisfério Norte no vento zonal, para o trimestre de DJF, que no trimestre JJA tornaram-se negativos. O mesmo ocorre para o vento meridional no Globo, Trópicos e Hemisfério Sul. Além disso, em ambas as componentes, os valores de ganhos negativos foram maiores para o período de inverno austral.

Os resultados vistos nesta seção mostram que existe um padrão no comportamento das previsões. Nas primeiras horas de previsão se observa o experimento que utiliza  $\alpha = 0, 1$  com erros menores do que o experimento que utiliza o alfa dinâmico e, com o passar dos horários de previsão o ganho aumenta, chegando em alguns casos a resultados positivos. Este padrão pode estar relacionado à performance do LETKF e do 3DVar em suas previsões, conforme apresentado na Seção 5.1. Quando se utiliza o  $\alpha = 0, 1$  no Híbrido Médio, se utiliza aproximadamente 90% de LETKF. Então, acredita-se que esses 90% dominem a solução final nas primeiras horas de previsão, tornando os erros desse experimento menores do que os erros ao utilizar o alfa dinâmico. Porém, acredita-se que a melhora percebida no decorrer das previsões no experimento com alfa dinâmico, devido ao fato destas análises estarem melhor balanceadas. Para o caso de inverno não foram vistos os mesmos resultados em relação ao balanço, e suas previsões também não apresentaram bons resultados. A maioria dos casos em que o experimento com alfa dinâmico apresentou ganhos positivos foi durante o período de verão e, ainda, no período de inverno, os valores de ganhos negativos foram maiores. As variáveis que sofreram um maior impacto com a variação de período foram as componentes zonal e meridional do vento, apresentando majoritariamente ganhos negativos, mostrando que, nestes casos,  $\alpha = 0, 1$  foi mais importante.

Figura 5.40 - Idem à Figura 5.39, porém para a temperatura.



Fonte: Produção do autor.

Figura 5.41 - Idem à Figura 5.40, porém para a componente zonal do vento.

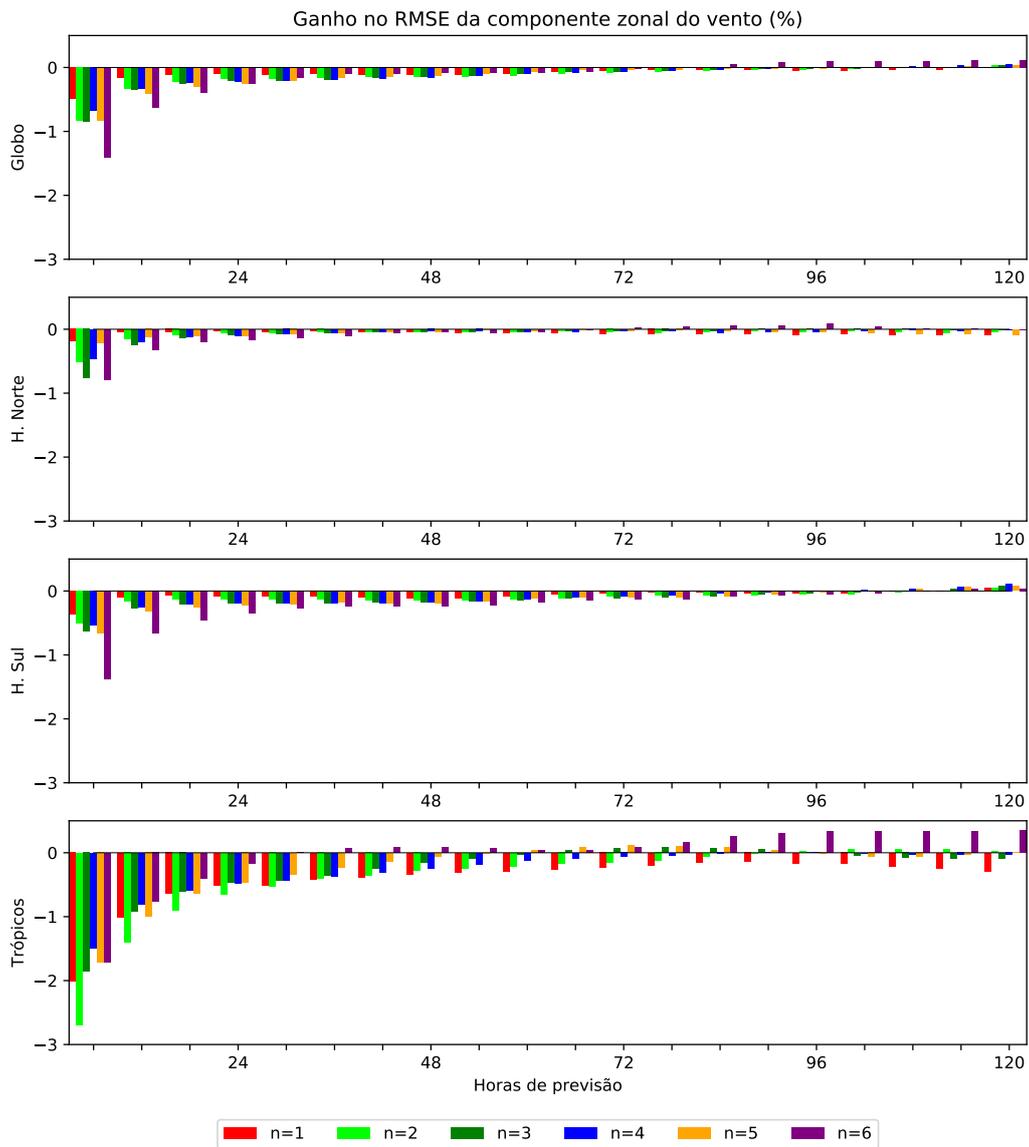
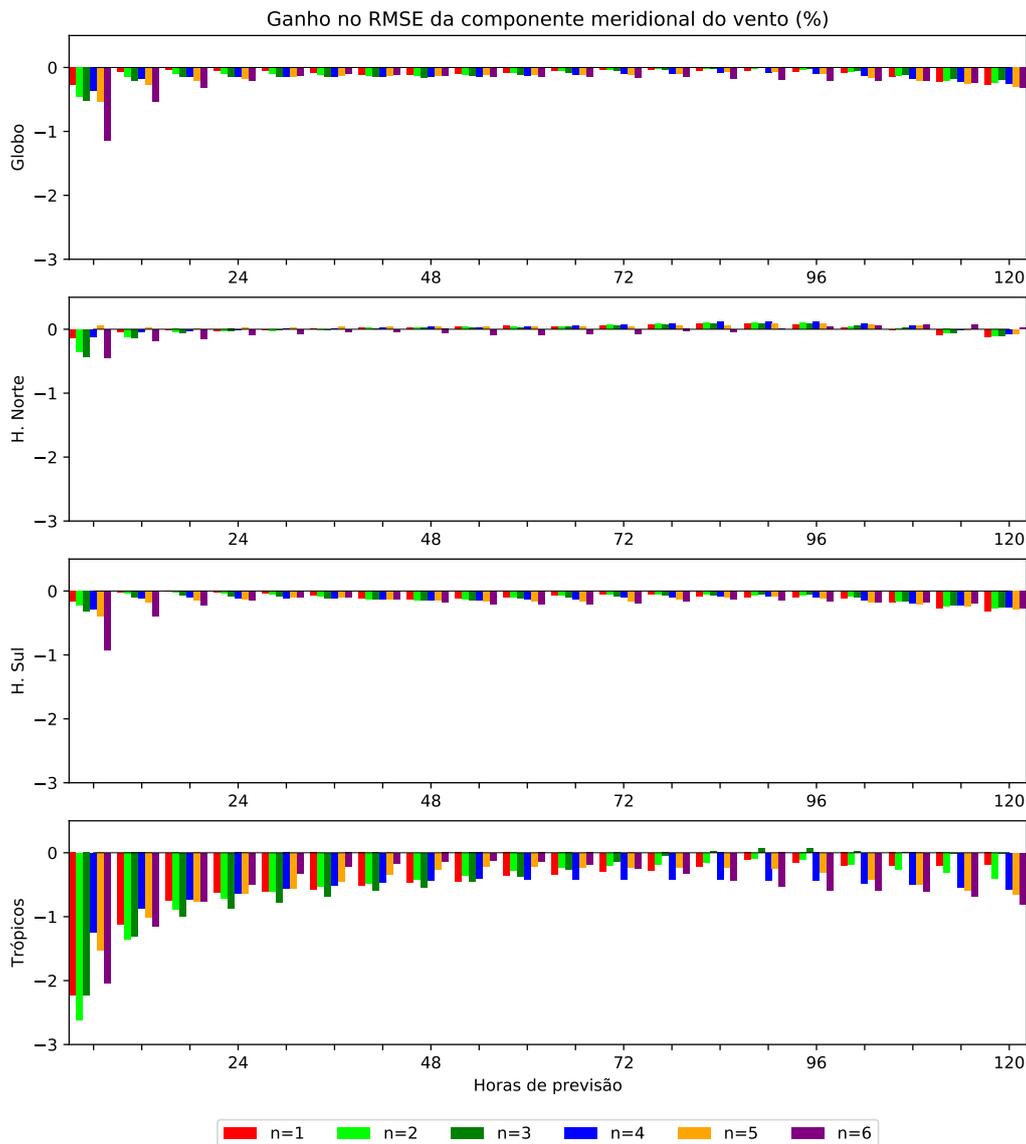


Figura 5.42 - Idem à Figura 5.40, porém para a componente meridional do vento.



Fonte: Produção do autor.

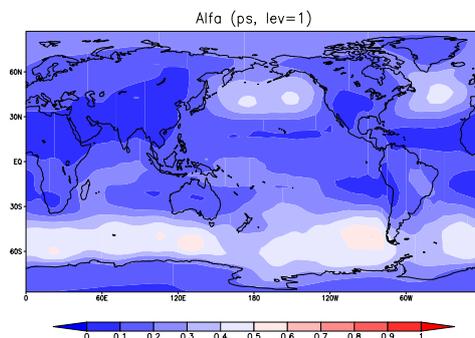
### 5.5 Sensibilidade ao número de membros

Nesta seção, são mostrados os resultados referentes ao aumento no número de membros do conjunto, passando de 16 para 40 membros, para os experimentos que utilizam o  $\alpha = 0, 1$  e o alfa dinâmico. São apresentados resultados sobre a estrutura do alfa dinâmico e resultados sobre as análises e ganhos percentuais das previsões para o período de dezembro, janeiro e fevereiro.

### 5.5.1 Estrutura média do Alfa Dinâmico

A Figura 5.43 mostra a distribuição espacial do alfa médio ao longo do trimestre DJF. Tons em azul mostram valores baixos de alfa quanto mais escuro, mais próximo de zero e, portanto, maior peso dado à análise do LETKF na composição da análise final. Já os tons em vermelho mais intenso apresentam valores maiores de alfa - quanto mais vermelho, mais próximo de um e, conseqüentemente, mais peso é dado à análise do 3DVar na composição da análise final. Comparando Figura 5.43 com a Figura 4.2, que é similar, porém utilizando 16 membros, nota-se uma coerência em relação às regiões onde se utiliza um alfa maior, como, por exemplo, as regiões de ciclones no Hemisfério Norte e no Hemisfério Sul. Existem algumas pequenas diferenças, como, por exemplo, no Polo Sul, sul do continente Africano, pontos sobre a Austrália e sobre o Oceano Pacífico Central, onde utilizando um número maior de membros diminuiu o valor de alfa, porém, existem também regiões onde esse valor aumentou, como, por exemplo, as regiões dos ciclones do Hemisfério Norte.

Figura 5.43 - Distribuição espacial do alfa médio da pressão em superfície para o período de dezembro, janeiro e fevereiro utilizando 40 membros.



Fonte: Produção do autor.

As demais variáveis e níveis não foram mostradas aqui, pois o padrão médio do alfa se mantém o mesmo do experimento utilizando 16 membros, e assim como para o campo de pressão em superfície, apresentam pequenas variações. Um dos motivos para que o alfa utilizando 40 membros tenha uma distribuição espacial similar ao alfa com 16 membros é que, embora o espalhamento varie conforme o número de membros, as regiões de maior incerteza do sistema continuam quase as mesmas, e como o alfa é uma normalização do espalhamento, independentemente do valor, onde

o espalhamento tem os maiores valores, o alfa vai ser maior, e onde o espalhamento é menor, o alfa terá valores menores. Dessa forma, foi possível concluir que a estrutura do alfa se mantém praticamente igual, não sofrendo alterações em decorrência do aumento no número de membros.

### 5.5.2 Impacto do número de membros nas análises

As análises foram avaliadas a partir do RMSE médio ao longo do período de dezembro, janeiro e fevereiro. Essas avaliações foram divididas em quatro regiões: Globo, Trópicos (compreendendo 20°S a 20°N), Hemisfério Norte (compreendendo 20°N a 80°N) e Hemisfério Sul (compreendendo 20°S a 80°S). Os resultados são apresentados na forma de tabela do RMSE da pressão em superfície e em perfis verticais das diferenças do RMSE dos experimentos com 16 e 40 membros para as demais variáveis.

A Tabela 5.11 mostra os valores de RMSE para a pressão em superfície de ambos os experimentos de 40 membros. Nota-se que os menores valores são relacionados ao experimento que utiliza o  $\alpha = 0, 1$ . Percebe-se ainda que o aumento no número de membros não gerou tanta diferença nos valores de RMSE para o experimento com o alfa dinâmico quando comparado com o período anterior (Tabela 5.9). Ao comparar os valores de RMSE para o experimento com  $\alpha = 0, 1$ , nota-se que o aumento no número de membros diminuiu os erros em até quase 2 Pa no Hemisfério Sul, por exemplo, e aproximadamente 1 Pa no Globo e região Tropical.

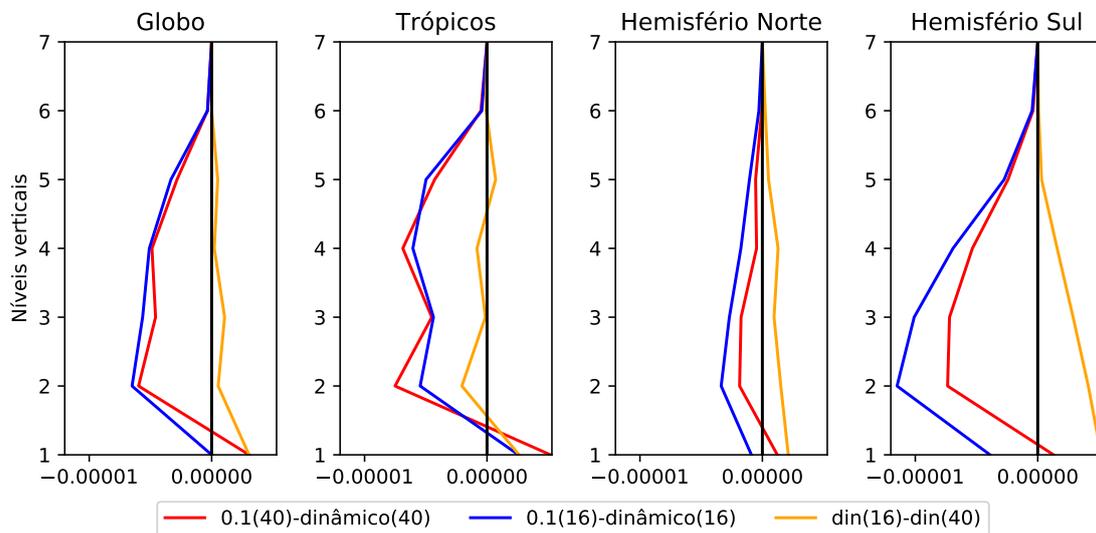
Tabela 5.11 - RMSE médio da pressão em superfície para as quatro regiões: Globo, Trópicos, Hemisfério Norte e Hemisfério Sul, para os meses de dezembro, janeiro e fevereiro utilizando 40 membros.

	Globo	Trópicos	H. Norte	H. Sul
$\alpha = 0, 1$	25,96	26,69	21,41	29,34
alfa dinâmico	29,30	29,28	23,95	34,13

Na Figura 5.44 são apresentados os perfis verticais das diferenças dos RMSE dos experimentos com 40 e 16 membros ( $\alpha = 0, 1$  e alfa dinâmico) para a umidade nas quatro regiões de estudo. As curvas em vermelho representam a diferença entre o experimento utilizando o  $\alpha = 0, 1$  e o experimento utilizando o alfa dinâmico, ambos com 40 membros. As curvas em azul representam essa mesma diferença, porém, para os experimentos com 16 membros. Aqui, quando os valores forem positivos,

significa que o experimento com alfa dinâmico apresentou erros menores; quando os valores forem negativos, o experimento utilizando  $\alpha = 0,1$  foi melhor. Por fim, as curvas em amarelo representam a diferença entre os experimentos utilizando o alfa dinâmico com 16 e 40 membros. Para essa diferença valores positivos, indicam que o experimento com 40 membros se sobressaiu em relação ao experimento com 16 membros e, quando a diferença for negativa, o experimento com 16 membros apresentou os melhores resultados.

Figura 5.44 - Perfil vertical da diferença do RMSE médio da umidade para quatro regiões: Globo, Trópicos, Hemisfério Norte e Hemisfério Sul.



Em vermelho está o experimento com  $\alpha = 0,1$  menos o alfa dinâmico, ambos com 40 membros; em azul o experimento com  $\alpha = 0,1$  menos alfa dinâmico, ambos com 16 membros, e, em amarelo a diferença entre os experimentos com alfa dinâmico, um com 16 membros e outro com 40 membros. Os eixos  $x$  são os valores do erro e nos eixos  $y$  os sete níveis do modelo.

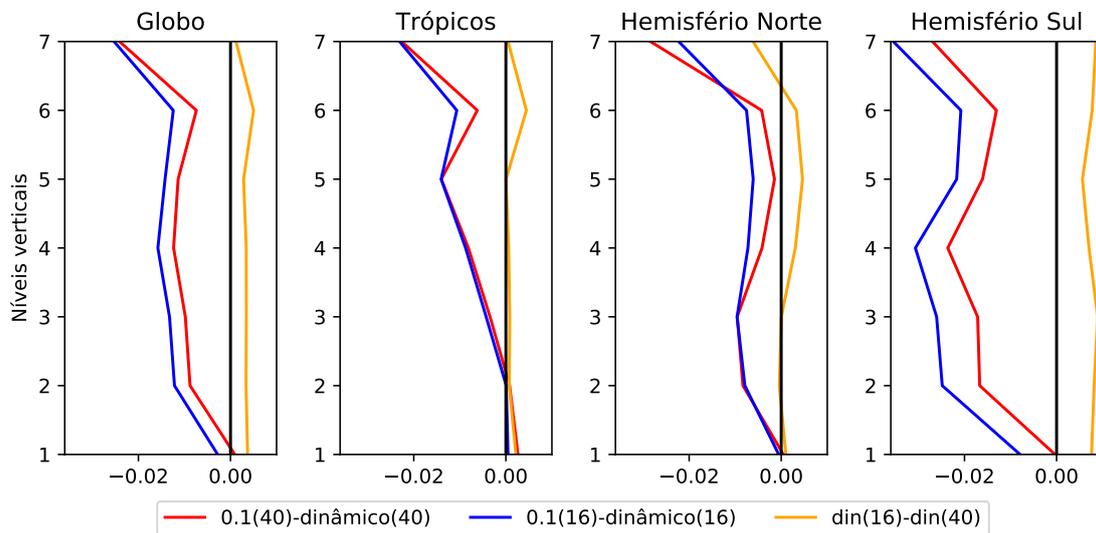
Fonte: Produção do autor.

Nessa figura, é notável que a diferença entre os erros do experimento que utilizou  $\alpha = 0,1$  em relação ao alfa dinâmico, com 16 membros é maior do que quando houve um aumento no número de membros. Em alguns casos, ainda, essa diferença torna-se positiva, mostrando que, de fato, o aumento no número de membros melhorou os resultados, tornando-os melhores que o experimento de alfa fixo. Isso é possível de observar para todas as regiões em superfície, quando a curva em vermelho fica

localizada no lado positivo do gráfico. Nota-se também que quando comparado os experimentos de alfa dinâmico, com 16 e 40 membros (curva amarela), o aumento no número de membros resulta em uma diminuição dos erros. Isso é observado em todo o perfil do Globo e dos Hemisférios, variando apenas na região Tropical, onde essa melhora é vista em superfície, e no quinto nível.

O perfil vertical da diferença do RMSE da temperatura é apresentado na Figura 5.45 e segue o mesmo padrão da figura anterior. Percebe-se, na grande maioria, diferenças negativas nos erros maiores para quando os experimentos utilizaram 16 membros. Isso não é visto apenas no Hemisfério Norte, nos níveis 2 e 7, quando se utilizou 40 membros, apresentou uma diferença maior. Nota-se ainda, que em alguns pontos utilizar mais membros tornou a análise melhor do que o experimento com alfa fixo. Foi o caso da temperatura em superfície no Globo, nos Trópicos e no Hemisfério Norte, que apresentaram valores positivos. Observa-se na curva amarela que, comparando os dois experimentos com alfa dinâmico, o aumento do número de membros foi benéfico na maioria dos casos, exceto no Hemisfério Norte nos níveis 2 e 7.

Figura 5.45 - Idem à Figura 5.44, porém, para a temperatura.

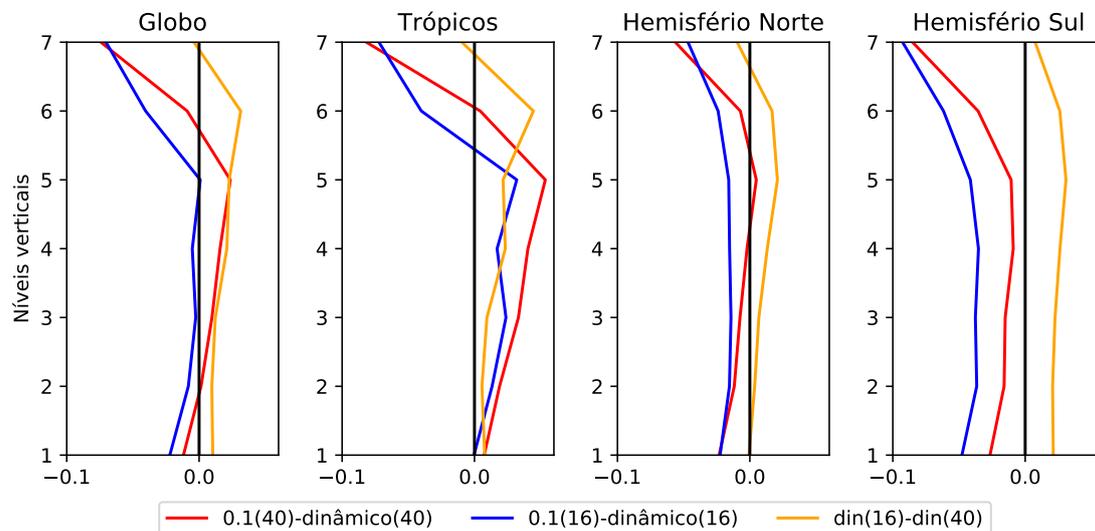


Fonte: Produção do autor.

O perfil vertical para a componente zonal do vento é mostrado na Figura 5.46. É possível observar que o aumento no número de membros resultou em análises com

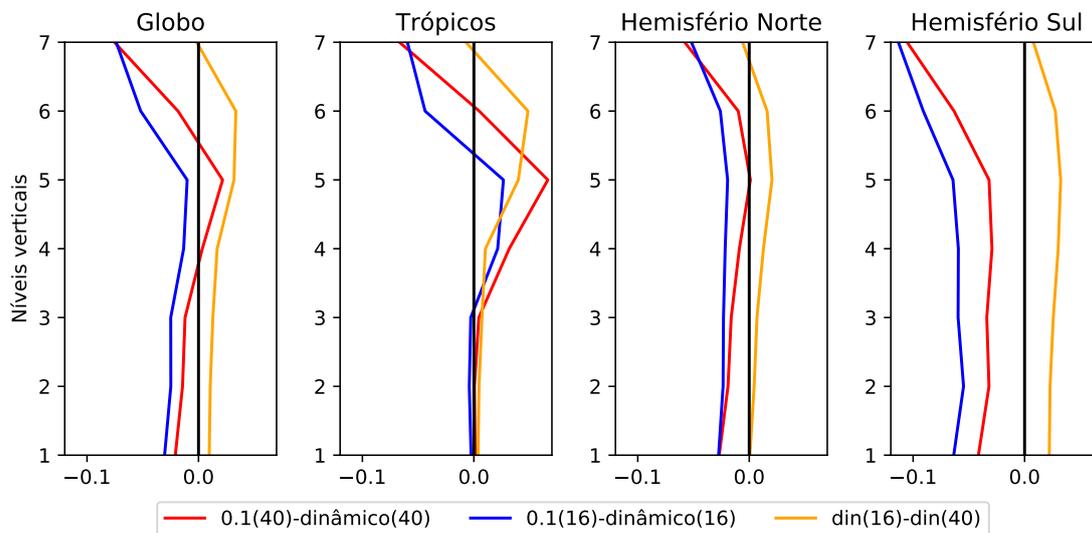
erros menores para o experimento com alfa dinâmico do que quando utilizado o alfa fixo com 40 membros. Essa melhora foi vista entre os níveis 2 e 5 do Globo, da superfície até o nível 6 dos Trópicos e no nível 5 do Hemisfério Norte. Para o Hemisfério Sul isso não foi visto, porém, a diferença entre os erros diminuiu. Observa-se, também, que há pontos em que, ainda que utilize 16 membros, o alfa dinâmico apresenta erro menor do que o alfa fixo também com 16 membros, como é o caso do nível 5 do Globo e os níveis 2, 3, 4, e 5 dos Trópicos. Quando comparados os experimentos de alfa dinâmico, nota-se que utilizar 40 membros apresenta uma melhora em relação à utilização de 16 membros - essa melhora só não é vista no último nível do Globo, Trópicos e Hemisfério Norte. Esse padrão é visto também para a componente meridional do vento, apresentado na Figura 5.47. Nessa figura, pode-se notar uma melhora do alfa dinâmico com 40 membros em relação ao alfa fixo com 40 membros, no nível 4 e 5 do Globo, da superfície até o nível 6 nos Trópicos e nível 5 no Hemisfério Norte. Percebe-se também uma melhora em relação ao experimento com 16 membros, exceto no último nível do Globo, Trópicos e Hemisfério Norte.

Figura 5.46 - Idem à Figura 5.44, porém, para a componente zonal do vento.



Fonte: Produção do autor.

Figura 5.47 - Idem à Figura 5.44, porém, para a componente meridional do vento.



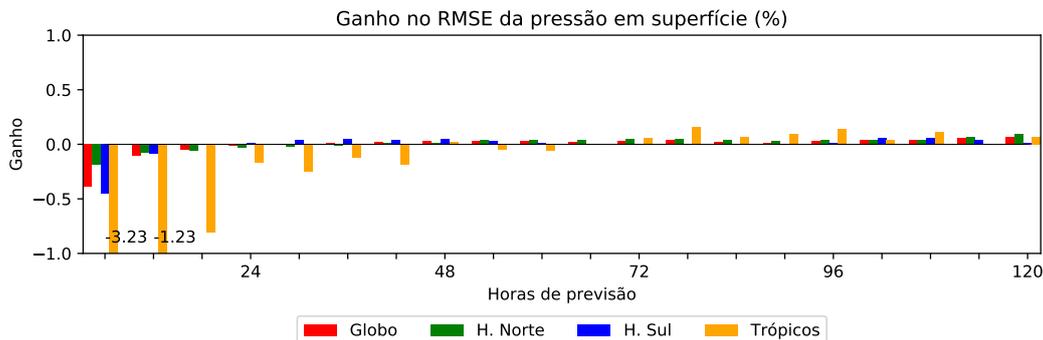
Fonte: Produção do autor.

Embora o erro da pressão em superfície tenha diminuído mais para o experimento com  $\alpha = 0, 1$ , como visto na Tabela 5.11, essa diferença entre os erros não influenciou os erros das demais variáveis, tanto em superfície quanto em médios e altos níveis. Pode-se observar que o aumento no número de membros fez com que diminuísse o erro do experimento com alfa dinâmico para a umidade e temperatura, tornando-o melhor que o experimento com alfa fixo em superfície. O vento também mostrou erros menores do que comparado aos experimentos com 16 membros e, ainda, em alguns pontos se mostrou melhor que o experimento com alfa fixo. Percebe-se também, que as maiores diferenças entre o alfa dinâmico e o alfa fixo com 40 membros encontram-se no Hemisfério Sul, podendo ter relação com a melhoria que o aumento no número de membros traz, diminuindo as incertezas em regiões onde há a ocorrência de sistemas transientes. Conclui-se que, quando comparados os experimentos com alfa dinâmico (16 e 40 membros), utilizar um número maior de membros apresenta análises com menores erros. Conforme os resultados, ao aumentar no número de membros as análises do alfa dinâmico melhoraram mais do que as análises geradas pelo experimento com alfa fixo, uma vez que a diferença dos erros diminuiu a medida que o número de membros aumentou. Ou seja, o número de membros teve um impacto maior no experimento com o alfa dinâmico do que com o alfa fixo. Infere-se, então, que haja uma tendência de melhora das análises geradas pelo Híbrido Médio utilizando o alfa dinâmico, quanto maior for o número de membros.

### 5.5.3 Impacto do número de membros nas previsões

A Figura 5.48 apresenta o ganho no RMSE das previsões do campo de pressão em superfície para até 120 horas quando utilizado 40 membros. Cada cor representa uma região de estudo: em vermelho o Globo, em verde o Hemisfério Norte, em azul o Hemisfério Sul e em amarelo os trópicos. As primeiras 24 horas de previsão apresentam apenas ganhos negativos, em contrapartida, a partir de 30 horas de previsão o Hemisfério Sul já passa a apresentar ganhos positivos. Os ganhos positivos no Globo apareceram a partir de 36 horas e, para os Trópicos apenas em 72 horas. Ao comparar com os experimentos com número menor de membros, o mesmo padrão é visto, com ganhos negativos nos primeiros horários e positivos no restante. É possível notar também que os valores do ganhos positivos foram maiores quando foram utilizados apenas 16 membros (mostrado na Figura 5.31).

Figura 5.48 - Ganho percentual nos valores de RMSE da pressão em superfície para as previsões de até 120 horas pertencentes ao período de dezembro, janeiro e fevereiro.



No eixo  $x$  estão os horários de previsão e no eixo  $y$  os valores de ganho relativos (%). Cada região está representada por uma cor: em vermelho o Globo, em verde o Hemisfério Norte, em azul o Hemisfério Sul e em amarelo os Trópicos.

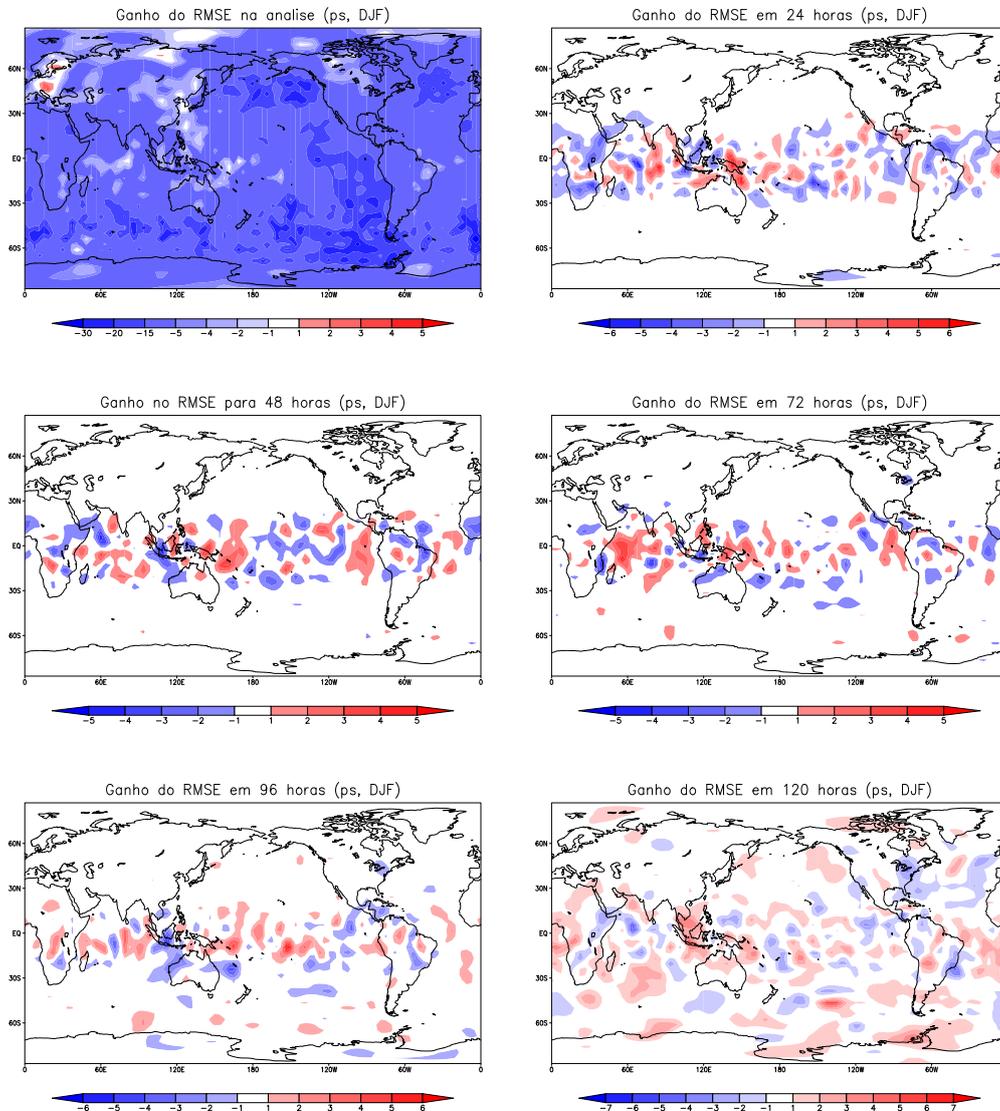
Fonte: Produção do autor.

O ganho percentual do RMSE nas previsões também é mostrado na Figura 5.49, através da distribuição espacial do ganho na análise e previsões de 24, 48, 72, 96 e 120 horas. Na análise observam-se ganhos negativos em grande parte do globo, com ganhos maiores (em módulo) no Hemisfério Sul e em pontos sobre o Hemisfério Norte. Os ganhos positivos foram vistos apenas sobre a faixa oeste da Europa e norte da Rússia. Isso corrobora o visto na Tabela 5.11, que mostra valores maiores para o

alfa dinâmico. Em relação ao número de membros, conclui-se que para o experimento com 40 membros houve uma piora nas análises de pressão. Nas primeiras 72 horas de previsão, o ganho ficou mais localizado na faixa entre 30°S e 30°N, sendo que em 24 horas apresenta apenas uma região sobre o Polo Sul de ganho negativo, em 48 horas alguns pontos ainda sobre o Hemisfério Sul com ganhos negativos e positivos, e em 72 horas aparecem mais regiões, no Hemisfério Sul, sendo a maioria positiva e no Hemisfério Norte um ponto negativo. Quando se compara com os ganhos vistos quando utilizado 16 membros (Figura 5.32), nota-se que a distribuição espacial está mais restrita à faixa entre 30°S e 30°N, e que ainda os valores absolutos estão maiores. Em 96 horas, os ganhos são vistos, na maioria, na região tropical e Hemisfério Sul, porém, ao comparar com o experimento com 16 membros, nota-se que diminuíram os ganhos no globo todo, principalmente no Hemisfério Norte, reduzindo valores e regiões. Já para 120 horas a distribuição espacial está mais próxima daquela com 16 membros, porém os valores, tanto negativos quanto positivos estão menores.

Tanto no gráfico do ganho percentual do RMSE nas previsões, quanto sua na distribuição espacial, nota-se o mesmo padrão visto quando utilizado 16 membros: uma análise com ganho negativo alto e, previsões mais curtas com ganho negativo havendo uma inversão para as previsões mais longas, tornando-os em ganhos positivos; além de ser observado também, que a medida em que as horas de previsão avançam, os ganhos positivos partem dos Hemisférios em direção a região Tropical.

Figura 5.49 - Distribuição espacial do Ganho do RMSE médio da pressão em superfície na análise e nas previsões de 24, 48, 72, 96 e 120 horas, para o trimestre DJF, para 40 membros.



Fonte: Produção do autor.

Através do que foi exposto aqui, pode-se concluir que, embora o sistema Híbrido Médio com 40 membros tenha apresentado uma melhora em relação ao mesmo sistema com 16 membros nas análises, essa melhora não foi propagada para as previsões. Além disso, ainda que o sistema Híbrido Médio utilizando o alfa dinâmico com 40 membros tenha apresentado uma melhora mais expressiva do que o sistema utilizando o alfa fixo com 40 membros, o sistema que utilizou  $\alpha = 0,1$  apresentou erros

menores nas análises e nas previsões de até 60 horas. Conclui-se, também, que o padrão apresentado nas previsões para o sistema Híbrido Médio com o alfa dinâmico se manteve o mesmo para 16 e 40 membros, mostrando uma melhora nas previsões mais longas. Conforme já discutido anteriormente, infere-se que essa melhora no ganho das previsões com o passar das horas de previsão, esteja relacionado às próprias previsões de ambos os sistemas: LETKF, que apresenta erros menores nos primeiros horários de previsão, e o 3DVar, que tem suas melhores previsões a partir de 60 horas aproximadamente. Vale lembrar que o ganho no RMSE do experimento com alfa dinâmico foi comparado com o experimento com  $\alpha = 0,1$ , que utiliza aproximadamente 90% de LETKF, justificando, assim, o resultado das primeiras horas de previsão.

## 5.6 Otimização do intervalo do alfa dinâmico

Com a finalidade de otimizar os resultados do sistema Híbrido Médio utilizando o alfa dinâmico, realizou-se um experimento reduzindo o intervalo de abrangência do alfa para que este ficasse em torno do alfa ótimo encontrado ( $\alpha = 0,1$ ). Dessa forma, os resultados que serão apresentados aqui, foram obtidos com o uso de um alfa dinâmico que varia entre 0,06 até 0,15.

### 5.6.1 Erros nas análises

Na Tabela 5.12 são mostrados os valores do RMSE da pressão em superfície para as quatro regiões de estudo, sendo a primeira linha para o experimento que utiliza  $\alpha = 0,1$ , a segunda linha para os resultados para o experimento que utiliza o alfa dinâmico que varia entre 0 e 1, e, por fim, a terceira linha para os resultados que utilizam o alfa dinâmico em torno de 0,1. As duas primeiras linhas são as mesmas da Tabela 5.9 e foram acrescentadas aqui para facilitar as comparações. Na Tabela 5.12 é possível observar que os erros ao se utilizar o alfa dinâmico em torno de 0,1 diminuem aproximadamente 2 Pa com relação ao alfa dinâmico entre 0 e 1, se aproximando assim dos resultados obtidos com o experimento que utiliza o alfa fixo.

Tabela 5.12 - RMSE médio da pressão em superfície para as quatro regiões: Globo, Trópicos, Hemisfério Norte e Hemisfério Sul.

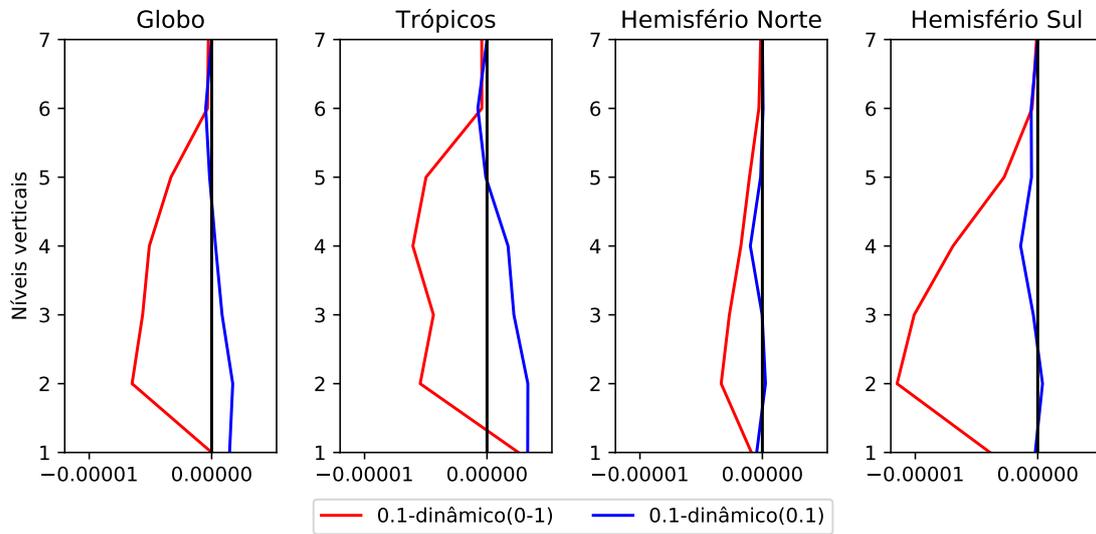
	Globo	Trópicos	H. Norte	H. Sul
$\alpha = 0,1$	27,08	27,81	21,87	31,03
alfa dinâmico [0;1]	29,75	29,41	23,84	35,27
alfa dinâmico (0,1)	27,21	27,97	21,98	31,14

Nas Figuras a seguir são apresentados perfis verticais da diferença do RMSE das análises geradas pelo experimento que utiliza  $\alpha = 0,1$  contra o experimento que utiliza o alfa dinâmico variando de 0 a 1 (vermelho) e, contra o experimento utilizando o alfa dinâmico em torno de 0,1 (azul) para as quatro regiões de estudo. Quando os valores forem positivos, os experimentos que utilizaram o alfa dinâmico apresentaram valores de RMSE menores do que o experimento com  $\alpha = 0,1$ ; quando os valores forem negativos, é visto o contrário.

A Figura 5.50, apresenta os valores para a umidade. Observa-se, quando comparado apenas os experimentos com alfa dinâmico, na maioria, valores maiores para o experimento em torno de 0,1, exceto para o sexto nível do modelo no Globo, Trópicos e Hemisfério Sul. No entanto, o restante dos perfis mostraram erros menores. Nota-se, ainda, que até o quarto nível do modelo a utilização do alfa dinâmico em torno de 0,1 apresentou resultados melhores do que o alfa fixo. Isso foi visto no Globo e Trópicos, enquanto para os Hemisférios ele foi melhor apenas no segundo nível. Vale ressaltar que, embora a maioria dos valores não tenham sido positivos para o experimento com alfa dinâmico em torno de 0,1, seus valores se aproximaram do zero, mostrando, com isso, uma diferença pequena entre os erros vistos no alfa fixo e nesse experimento com alfa dinâmico.

A Figura 5.51 mostra os perfis para a temperatura. Nota-se que o experimento com alfa dinâmico entre 0 e 1 se mostrou melhor que o experimento com alfa dinâmico em torno de 0,1 apenas no primeiro nível do Hemisfério Norte, ainda assim, sendo inferior ao experimento com alfa fixo. Em contra partida, nos Trópicos ambos os experimentos apresentaram uma performance melhor que o experimento com alfa fixo. O experimento utilizando o alfa dinâmico em torno de 0,1 se mostrou melhor que os demais nos níveis 1, 2 e 3 do Globo, do nível 1 ao 4 nos Trópicos, e no nível 3 do Hemisfério Sul. Embora o alfa fixo tenha sido melhor no restante dos perfis, nota-se uma melhora expressiva por conta do alfa dinâmico quando usada a normalização em torno de 0,1,

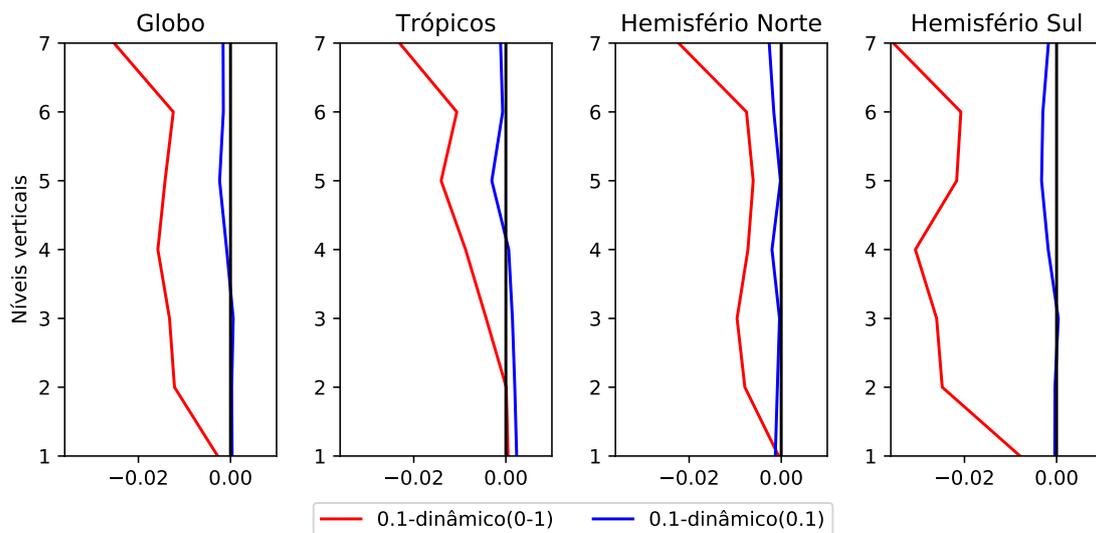
Figura 5.50 - Perfil vertical da diferença do RMSE médio da umidade para quatro regiões: Globo, Trópicos, Hemisfério Norte e Hemisfério Sul.



Em vermelho está o experimento com  $\alpha = 0,1$  menos o alfa dinâmico, variando de 0 a 1 e, em azul o experimento com  $\alpha = 0,1$  menos alfa dinâmico em torno de 0,1, Os eixos  $x$  são os valores do erro e nos eixos  $y$  os sete níveis do modelo.

Fonte: Produção do autor.

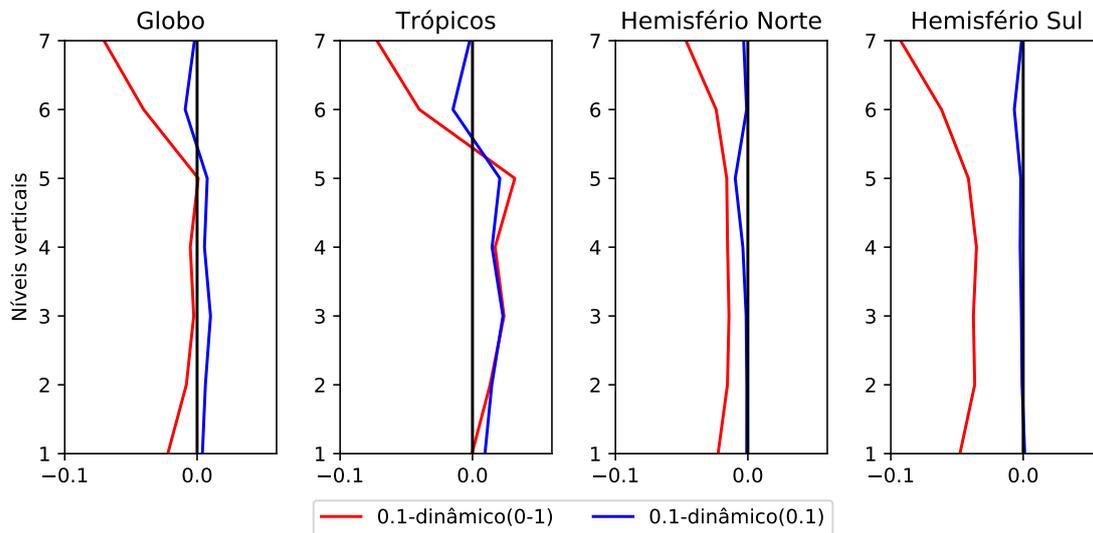
Figura 5.51 - Idem à Figura 5.50, porém, para a temperatura.



Fonte: Produção do autor.

Os perfis verticais da componente zonal do vento são apresentados na Figura 5.52. É possível observar que para a maioria dos níveis e regiões utilizar um alfa dinâmico em torno de 0,1 apresenta resultados melhores do que o alfa entre 0 e 1, Isso só não acontece para os níveis 4 e 5 dos trópicos, onde o experimento entre 0 e 1 apresenta-se com erros menores. Quando comparado o experimento com o alfa dinâmico em torno de 0,1 contra o  $\alpha = 0,1$  percebe-se que até o quinto nível do Globo e dos Trópicos, o experimento com o alfa dinâmico apresentou erros menores. Nos demais níveis e regiões o experimento com o alfa fixo teve os menores erros. No entanto, as diferenças entre os erros do experimento com alfa dinâmico em torno de 0,1 e do experimento com alfa fixo ficaram próximas de zero.

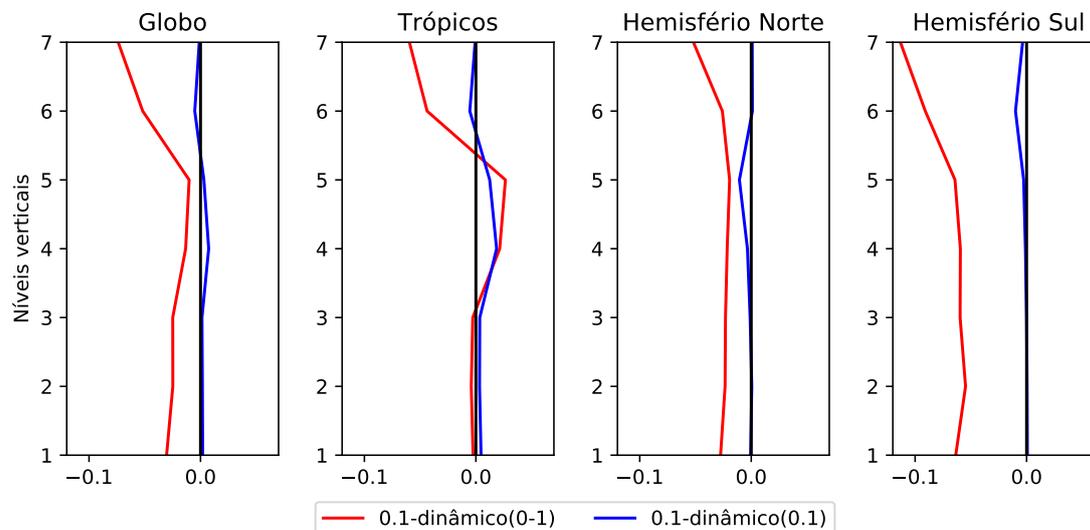
Figura 5.52 - Idem à Figura 5.50, porém, para a componente zonal do vento.



Fonte: Produção do autor.

A Figura 5.53 mostra o perfil vertical para a componente meridional do vento. Nota-se um padrão bem similar ao visto para a componente zonal do vento nos Hemisférios. Já na região Tropical, ainda que o experimento com alfa dinâmico entre 0 e 1 tenha se mostrado melhor que o em torno de 0,1, ele é melhor que o alfa fixo apenas nos níveis 4 e 5 e, no Globo, ele nunca é melhor. Comparando apenas o experimento com alfa fixo contra o experimento que utiliza o alfa dinâmico em torno de 0,1, nota-se que da superfície até o quinto nível, tanto no Globo quanto nos Trópicos o experimento com alfa dinâmico foi melhor.

Figura 5.53 - Idem à Figura 5.50, porém, para a componente meridional do vento.



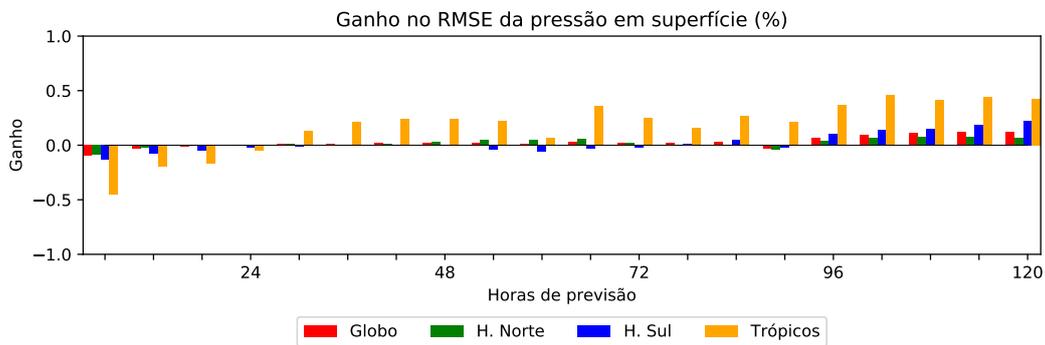
Fonte: Produção do autor.

Foi possível concluir aqui que, ao restringir os valores de alfa em torno do alfa ótimo, houve um ganho considerável quando comparado ao experimento que utiliza o alfa dinâmico variando entre 0 e 1 nas análises. Além disso, conclui-se que em vários casos, o experimento com o alfa dinâmico em torno de 0,1 apresentou valores de RMSE nas análises, menores do que quando utilizado  $\alpha = 0, 1$ .

### 5.6.2 Erros nas previsões

Assim como exposto nas Seções anteriores, esta mostra o comportamento do experimento que utiliza o alfa dinâmico em torno de 0,1 ao longo de 120 horas de previsão. Na Figura 5.54 verifica-se o ganho percentual do experimento com alfa dinâmico (em torno de 0,1) em relação ao experimento que utiliza  $\alpha = 0, 1$ . Observa-se na maioria, ganhos negativos nos primeiros horários de previsão e os ganhos tornam-se positivos ao longo das horas de previsão. Porém, entre 54 e 72 horas o Hemisfério Sul apresenta ganhos negativos e, ainda em 90 horas, a única região que apresenta ganho positivo são os trópicos. Comparando com os resultados do alfa dinâmico entre 0 e 1 (Figura 5.31), observam-se valores menores, tanto nos ganhos negativos quanto nos positivos, e ainda nota-se que, os ganhos negativos voltam a aparecer no decorrer das previsões, o que, quando comparado com o alfa dinâmico entre 0 e 1, não acontece.

Figura 5.54 - Ganho percentual nos valores de RMSE da pressão em superfície para as previsões de até 120 horas.



Fonte: Produção do autor.

Foi possível concluir que, quando houve uma alteração no intervalo do alfa dinâmico, restringindo a valores próximos do alfa ótimo, houve melhora nas análises geradas pelo sistema Híbrido Médio com o alfa dinâmico. No entanto, embora tenha sido observada essa melhora, não se observou a mesma nas previsões do campo de pressão. Houve uma diminuição no ganho, tanto negativo quanto positivo, e ainda apresentou mais ganhos negativos no Hemisfério Sul do que quando utilizado o intervalo entre 0 e 1.



## 6 CONCLUSÕES

Os sistemas de assimilação de dados Híbridos tem sido objeto de diversos estudos ao longo das últimas duas décadas. Porém, sistemas híbridos tradicionais, os quais visam o cálculo de uma nova matriz de covariância dos erros do *background*, a partir da combinação de matrizes oriundas de sistemas de assimilação independentes, são custosos computacionalmente. Esse foi um dos motivos que levou ao surgimento, em 2014, de uma nova metodologia de sistemas híbridos, o Híbrido Médio. Nela é realizada uma combinação linear entre duas análises: uma vinda de um sistema por conjunto e outra vinda de um sistema variacional, o qual utiliza a análise média do sistema por conjunto como seu *background*. Tais análises são ponderadas por um peso fixo que varia entre 0 e 1. Essa nova metodologia tem apresentado resultados competitivos quando comparada aos métodos tradicionais; uma de suas vantagens é o seu custo, equivalente ao do sistema por conjunto utilizado.

*Nesse trabalho foram estudados os efeitos da inclusão de um coeficiente de ajuste dinâmico, que tomou o lugar do coeficiente de ajuste estático da combinação linear, buscando balancear a combinação entre as duas análises de acordo com a estimativa da incerteza do conjunto, obtida a cada ciclo de assimilação. Essa medida da incerteza é dada pelo espalhamento do conjunto, que é normalizado entre 0 e 1 e em seguida entregue como o alfa dinâmico ao sistema.* Esse novo coeficiente foi implementado no Sistema Híbrido Médio, utilizando o LETKF e o 3DVar, e foi testado em experimentos no modelo SPEEDY. Foram realizados ao todo oito experimentos principais, sendo cinco deles testando valores fixos de alfa ( $\alpha = 0, 1$ ,  $\alpha = 0, 3$ ,  $\alpha = 0, 5$ ,  $\alpha = 0, 7$  e  $\alpha = 0, 9$ ), dois deles utilizando os sistemas envolvidos no Híbrido Médio de forma separada (LETKF e 3DVar) e um último experimento que utilizou o alfa dinâmico, todos utilizando 16 membros. Além desses, outros três experimentos secundários foram realizados, sendo dois deles utilizando 40 membros ( $\alpha = 0, 1$  e alfa dinâmico) e um último alterando o intervalo do alfa dinâmico. Foram realizados experimentos para período de verão e inverno.

A partir do desenvolvimento e execução das etapas apresentadas na Seção 1.1 na forma de objetivos para o presente trabalho, foi possível verificar o comportamento do LETKF e do 3DVar em um ambiente de modelo perfeito, bem como testar a sensibilidade do Sistema Híbrido Médio a diferentes valores de alfa. Foi possível também avaliar a inclusão do critério de ajuste dinâmico no sistema Híbrido Médio em suas análises e previsões, além de realizar testes de sensibilidade com relação ao

número de membros e com relação ao intervalo de alfa e avalia-los em suas análises e previsões.

No Capítulo 1, o problema desta tese foi introduzido, e os objetivos foram estabelecidos. Mostrou-se a trajetória percorrida pelo CPTEC/INPE no desenvolvimento da sua assimilação de dados e onde a instituição se encontra perante o restante do mundo. O surgimento do sistema de assimilação de dados, foco desse trabalho, foi brevemente apresentado, assim como a motivação para o uso dele. Por fim, foram expostos os objetivos e questões que foram respondidas ao longo dessa tese.

O Capítulo 2 apresentou a revisão bibliográfica sobre os assuntos abordados. Um histórico da assimilação de dados foi exposto, assim como seus últimos avanços. Realizou-se também uma revisão a respeito dos sistemas de assimilação de dados variacionais, sendo eles em três e quatro dimensões; sobre os sistemas baseados por conjunto, com ênfase maior nos sistemas baseados no Filtro de Kalman; por fim, sobre sistemas Híbridos.

A metodologia empregada neste trabalho, conforme apresentado no Capítulo 3, utiliza o modelo SPEEDY. Este foi escolhido a fim de diminuir possíveis complicações com o uso de um modelo realístico, como, por exemplo, tempo de uso de máquina e armazenamento de dados. Foram mostradas também as configurações para o sistema Híbrido Médio, 3DVar e LETKF. Na sequência, foi exposto o método de inclusão do coeficiente de ajuste dinâmico proposto nessa tese, sendo esse o ponto principal do trabalho. Nessa seção mostra-se como o espalhamento do conjunto é utilizado na forma de alfa dinâmico. Na última seção são mostradas as configurações dos diversos experimentos, bem como as observações sintéticas que foram utilizadas.

No Capítulo 4 resultados relacionados à estrutura do alfa dinâmico para os dois períodos de estudo foram detalhados, mostrando diversas características do novo parâmetro. foram apresentadas as distribuições espaciais do alfa médio ao longo dos trimestres, permitindo com isso identificar as regiões em que, na média, foi utilizado mais LETKF ou mais 3DVar na composição da análise final do sistema Híbrido Médio. A distribuição espacial permitiu também que uma melhor análise fosse realizada em relação à regiões de ocorrências de determinados sistemas sinóticos. ***Os resultados vistos nessa seção mostram que, em grande parte do Globo, houve predominância do LETKF em comparação ao 3DVar na composição da análise final do Híbrido Médio.*** Como o alfa dinâmico está relacionado com o espalhamento do LETKF e, este está relacionado ao grau de incerteza do sistema por conjunto; regiões onde atuam sistemas sinóticos, como ciclones, Alta da Bolívia

e ZCIT, apresentam uma incerteza maior, conseqüentemente, nessas regiões houve um uso maior do 3DVar. Também pôde-se relacionar os valores de alfa à densidade de observações. Embora este seja um trabalho utilizando modelo perfeito, as observações sintéticas utilizadas retratam distribuições espaciais e temporais realísticas: em regiões oceânicas, por exemplo, há uma densidade baixa de observações. Isso pode ser um fator decisivo para que a confiabilidade do LETKF nessas regiões seja menor.

Nesse mesmo capítulo é verificado o comportamento médio do alfa dinâmico ao longo dos períodos de estudo, possibilitando avaliar como o alfa se comporta com o passar do tempo em cada uma das 4 regiões avaliadas. Observa-se que os valores médios de alfa não ultrapassaram 0.5 para o período de dezembro, janeiro e fevereiro. Para o trimestre JJA foram encontrados valores levemente superiores. Por fim, nesse capítulo, mostra-se o desvio padrão do alfa dinâmico. Com ele torna-se possível avaliar a variação do alfa dinâmico espacialmente durante os trimestres. Foram observados valores de desvio padrão baixos em regiões onde foi visto o alfa médio baixo e, quando o alfa médio teve valores elevados, o desvio padrão também. Vale lembrar que, como o alfa dinâmico é uma normalização dos valores de espalhamento, sempre haverá no mínimo um ponto com  $\alpha = 1$  e no máximo um ponto com  $\alpha = 0$ .

No Capítulo 5 são apresentados e discutidos os resultados; as conclusões serão melhor apresentadas nos parágrafos que seguem. Primeiramente, realiza-se uma comparação da performance do LETKF e do 3DVar em suas análises e previsões. Em seguida são discutidos os resultados com relação as análises geradas pelo sistema Híbrido Médio e suas várias configurações. São mostrados seis experimentos, sendo cinco deles com valores fixos de alfa ( $\alpha = 0, 1$ ,  $\alpha = 0, 3$ ,  $\alpha = 0, 5$ ,  $\alpha = 0, 7$  e  $\alpha = 0, 9$ ) e um último utilizando o alfa dinâmico. A intenção aqui foi encontrar um valor ótimo de alfa fixo, já que, como mostrado na Fundamentação Teórica (Capítulo 2), os valores empregados a ele são empíricos, de forma que nunca foram testados mais do que três valores ( $\alpha = 0, 25$ ,  $\alpha = 0, 5$  e  $\alpha = 0, 75$ ) e, comparar esses resultados com os resultados obtidos com o sistema Híbrido Médio que utiliza o alfa dinâmico. A seção seguinte faz uma avaliação em relação ao balanço das análises vistas até então, a fim de atestar o balanço de quando se utiliza um alfa em forma de uma matriz. Foram avaliadas e discutidas estatísticas de O-A e O-F, assim como as previsões de até 120 horas geradas pelo experimento com alfa dinâmico e com  $\alpha = 0, 1$ . Além desses resultados, os resultados de outros três experimentos adicionais são mostrados nesse Capítulo. O primeiro mostra resultados sobre a sensibilidade do sistema quando utilizado 40 membros. Nessa seção são avaliados os resultados dos experimentos

com alfa dinâmico e com  $\alpha = 0, 1$  ambos com 40 membros. Esse impacto é mostrado nas análises e nas previsões de até 120 horas. Na seção seguinte é avaliado e discutido o impacto quando o alfa dinâmico sofre alterações no seu intervalo, variando de 0.06 a 1.15. Esses resultados são vistos também em análises e previsões.

Para avaliar a performance do LETKF e do 3DVar nas análises foram estimados os valores de RMSE médio para as quatro regiões de estudo e, ainda, esses erros foram avaliados na forma de distribuição espacial. Através desses resultados foi possível concluir que tanto para o sistema variacional quanto para o sistema por conjunto, os erros vistos foram mais localizados no Hemisfério Sul e região Tropical, acompanhando os sistemas sinóticos atuantes nessas regiões, como, por exemplo, ZCIT e ciclones. Infere-se que esses erros estejam associados à instabilidades dessas regiões e a falta de dados, em especial, no Hemisfério Sul. Embora este seja um trabalho realizado com observações sintéticas, a distribuição espacial delas tenta refletir a distribuição real. Observou-se também que a magnitude dos erros foi maior para os meses de junho, julho e agosto e, quando comparados os sistemas, o 3DVar apresentou os maiores erros. Para avaliar as previsões geradas pelas análises de cada um dos sistemas, foi calculada a diferença dos RMSE ao longo das horas de previsão. ***Conclui-se então, que os erros do LETKF são menores que os erros do 3DVar nas análises e nos primeiros horários de previsão, ao passo que nos últimos horários esse sinal inverte, sendo os erros do 3DVar menores que os do LETKF.*** Observou-se esse padrão para todas as variáveis, porém, para o período de inverno austral a inversão desse sinal demora mais para ocorrer.

A avaliação das análises geradas pelo sistema Híbrido Médio em suas diversas configurações mostrou que, para ambos os períodos (verão e inverno), quanto maior o valor do alfa, maior é o erro obtido, ou seja, quanto mais 3DVar é usado na geração da análise final do Híbrido Médio, pior é a qualidade da análise. Porém, usar  $\alpha = 0, 0$  (que é o LETKF puro) não necessariamente corresponde os melhores resultados. ***Dessa forma, conclui-se que, para esse cenário, dentre os valores de alfa fixo testados, o  $\alpha = 0, 1$  apresenta o melhor desempenho nas análises.*** O experimento utilizando o alfa dinâmico apresentou resultados inferiores aos resultados obtidos com o  $\alpha = 0, 1$ , e, em alguns casos inferiores aos resultados obtidos com o  $\alpha = 0, 3$ , porém, diante dos demais valores de alfa, os erros dele foram menores. ***Conclui-se que, em relação ao alfa dinâmico, o alfa ótimo apresentou, na grande maioria dos casos, erros menores nas análises.*** Foi observado que realizar a combinação linear entre o sistema variacional e o sistema

por conjunto beneficia a análise final do sistema Híbrido Médio, porém, quando são utilizados valores baixos de alfa ou quando utilizado o alfa dinâmico.

*Quando foram avaliadas as análises de todos os experimentos em relação ao balanço, conclui-se que, para o período de verão austral, o experimento que utilizou o alfa dinâmico foi melhor balanceado quando comparado com os demais. Para o período de inverno, o uso do  $\alpha = 0,3$  apresenta o melhor balanço.*

Ao serem avaliadas as estatísticas de O-A (observação menos análise) e O-F (observação menos estimativa inicial), foi possível concluir que em relação ao experimento utilizando o  $\alpha = 0,1$ , o experimento com alfa dinâmico apresentou melhores resultados para o período de dezembro, janeiro e fevereiro, exceto para o Hemisfério Sul. Já para o período de inverno o experimento com alfa dinâmico foi melhor no Globo e nos Trópicos, ao passo que o alfa fixo apresentou melhores resultados nos Hemisférios.

*Os resultados obtidos através do ganho percentual no RMSE das previsões de até 120 horas do experimento que utilizou o alfa dinâmico em relação ao experimento que utilizou o  $\alpha = 0,1$  mostraram que há um padrão no comportamento, apresentando ganhos negativos nas primeiras horas e, à medida que avançam as horas de previsão, o ganho inverte de sinal, passando a ficar positivo.* Acredita-se que esse padrão tenha sido observado em decorrência da performance do LETKF e do 3DVar, na qual o sistema por conjunto apresenta melhores resultados nas primeiras horas de previsão. Ao utilizar  $\alpha = 0,1$ , utiliza-se aproximadamente 90% de LETKF na solução final. Sendo assim, acredita-se que a solução nos primeiros horários esteja se aproximando da solução do LETKF. As variáveis que apresentaram um maior impacto foram as componentes zonal e meridional do vento. Para o período de inverno austral, os valores de ganhos negativos foram maiores, e em alguns casos, não apresentaram ganhos positivos.

*Em relação aos experimentos adicionais realizados, quando há um aumento no número de membros, foi observado que o experimento que utiliza o alfa dinâmico apresentou um maior impacto nas análises do que o experimento com alfa fixo ( $\alpha = 0,1$ ). Porém, esse impacto não se propagou para as previsões, apresentando ganhos negativos maiores do que quando utilizado 16 membros e ganhos positivos menores. Para o experimento alterando o intervalo do alfa dinâmico, quando o intervalo passa a variar entre 0,06 e 0,15, foi observado que os erros nas análises do experi-*

*mento com alfa dinâmico se aproximaram do experimento que utiliza o  $\alpha = 0,1$  e, em alguns casos, ele se mostrou melhor.* Já para os resultados nas previsões, foi visto o mesmo padrão dos outros experimentos, porém com ganhos negativos menores que os anteriores e, em algumas situações com ganhos negativos em previsões mais distantes da análise.

## 6.1 Trabalhos futuros

No decorrer deste trabalho, surgiram questões interessantes que não foram possíveis de abordar. Algumas não puderam ser exploradas por questão de tempo e outras por questões computacionais, como, por exemplo, o uso de um modelo realístico de melhor resolução.

Dentre os pontos a serem trabalhados futuramente, aqui destacam-se:

- Testar a hipótese proposta em sistemas de previsão numérica de tempo mais complexos e, de maior resolução, em experimentos utilizando dados reais, possibilitando assim a realização de estudos de caso mais realísticos;
- Considerar o parâmetro de inflação no cálculo do alfa dinâmico, uma vez que ele introduz uma alteração artificial no espalhamento do sistema por conjunto;
- Testar outros ranges do alfa dinâmico, como, por exemplo, valores entre 0 e 0.3, já que as análises até o  $\alpha = 0,3$  apresentaram melhoras significativas em relação ao alfa dinâmico;
- Fazer uma avaliação detalhada envolvendo as previsões de cada experimento de alfa fixo, a fim de avaliar o desempenho que cada um deles tem não somente nas análises, como também nas previsões.

## REFERÊNCIAS BIBLIOGRÁFICAS

ANDREOLI, R. V.; FERREIRA, S. H. S.; SAPUCCI, L. F.; SOUZA, R. A. F.; MENDONÇA, R. W. B.; HERDIES, D. L.; ARAVÉQUIA, J. A. Contribuição de diversos sistemas de observação na previsão de tempo no cptec/inpe. **Revista Brasileira de Meteorologia**, v. 23, p. 218–237, 2008. 2

ARAVÉQUIA, J. A.; SZUNYOGH, I.; FERTIG, E. J.; KALNAY, E.; KUHL, D.; KOSTELICH, E. J. Evaluation of a strategy for the assimilation of satellite radiance observations with the local ensemble transform kalman filter. **Monthly Weather Review**, v. 139, n. 6, p. 1932–1951, June 2011. ISSN 0027-0644.

Disponível em:

<<http://urlib.net/dpi.inpe.br/plutao/2011/06.11.03.38.28>>. Acesso em: 20 nov. 2015. 9

AZEVEDO, H. B. de; GONÇALVES, L. G. G. de; BASTARZ, C. F.; SILVEIRA, B. B. Observing system experiments in a 3dvar data assimilation system at cptec/inpe. **Weather and Forecasting**, v. 32, n. 3, p. 873–880, 2017. Disponível em: <<https://doi.org/10.1175/WAF-D-15-0168.1>>. 2

BARKER, D. M.; HUANG, W.; GUO, Y.-R.; BOURGEOIS, A. J.; XIAO, Q. N. A three-dimensional variational data assimilation system for mm5: implementation and initial results. **Monthly Weather Review**, v. 132, p. 897–914, 2004.

Disponível em:

<[https://doi.org/10.1175/1520-0493\(2004\)132<0897:ATVDAS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0897:ATVDAS>2.0.CO;2)>. 17

BASTARZ, C. F. **Assimilação de dados global híbrida por conjunto-variacional no CPTEC**. 275 p. Tese (Doutorado em Meteorologia) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2017-07-18 2017. Acesso em: 05 nov. 2018. 2, 10

BAÑOS, I. H. **Assimilação de perfis de rádio ocultação GPS no modelo global de PNT do CPTEC**. 145 p. Dissertação (Mestrado em Meteorologia) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2017-02-20 2017. Disponível em:

<<http://urlib.net/sid.inpe.br/mtc-m21b/2017/02.03.10.32>>. Acesso em: 04 maio 2018. 107

BISHOP, C. H.; ETHERTON, B. J.; MANJUNDAR, S. J. Adaptive sampling with the ensemble transform kalman filter. part i: theoretical aspects. **Monthly Weather Review**, v. 129, p. 420–436, 2001. 8, 9

BONAVITA, M.; HAMRUD, M.; ISAKSEN, L. Enkf and hybrid gain ensemble data assimilation. part ii: EnKF and hybrid gain results. **Monthly Weather Review**, v. 143, n. 12, p. 4865–4882, 2015. Disponível em: <<https://doi.org/10.1175/MWR-D-15-0071.1>>. 3, 11, 98

BOUQUIER, F.; COURTIER, P. **Data assimilation concepts and methods**. [S.l.: s.n.], 1999. 8

CHARNEY, J. G.; FJORTOFT, R.; NEUMAN, J. Numerical integration of the barotropic vorticity equation. **Tellus**, v. 2, p. 237–254, 1950. 5

CINTRA, R.; ARAVÉQUIA, J. A. Apresentação do método de geração de dados para assimilação do modelo ETA/CPTEC. In: FEDERAÇÃO LATINO-AMERICANA E IBÉRICA DE SOCIEDADES DE METEOROLOGIA., 10., 1998, Brasília, DF. **Congresso Brasileiro de Meteorologia**, 10. [S.l.], 1998. 1

CINTRA, R. S. C. **Preparação de dados de observações para o sistema de assimilação de dados PSAS do CPTEC**. São José dos Campos: [s.n.], 2005. 60 p. Disponível em: <<http://urlib.net/sid.inpe.br/mtc-m17@80/2006/12.04.16.01>>. 2

CLAYTON, A. M.; LORENC, A. C.; BARKER, D. M. Operational implementation of a hybrid ensemble/4d-var global data assimilation system at the met office. **Quarterly Journal of the Royal Meteorological Society**, v. 139, n. 675, p. 1445–1461, 2013. Disponível em: <<https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2054>>. 10

ERRICO, R. M. What is an adjoint model?. **Bulletin American Meteorology Society**, v. 78, p. 2577–2591, 1997. 7

ETHERTON, B. J.; BISHOP, C. H. Resilience of hybrid ensemble/3dvar analysis schemes to model error and ensemble covariance error. **Monthly Weather Review**, v. 132, n. 5, p. 1065–1080, 2004. 10

EVENSEN, G. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. **Journal of Geophysical Research**, v. 99, p. 10143–10162, 1994. 8, 9

\_\_\_\_\_. The ensemble kalman filter: theoretical formulation and practical implementation. **Ocean Dynamics**, v. 53, p. 343-367, 2003. 9

GANDIN, L. S. Objective analysis of meteorological fields. **Quarterly Journal of the Royal Meteorological Society**, v. 92, 1965. 1

GREYBUSH, S. J.; KALNAY, E.; MIYOSHI, T.; IDE, K.; HUNT, B. R. Balance and ensemble kalman filter localization techniques. **Monthly Weather Review**, v. 139, n. 2, p. 511-522, 2011. Disponível em: <<https://doi.org/10.1175/2010MWR3328.1>>. 97, 98

HAMILL, T. M.; SNYDER, C. A hybrid ensemble kalman filter-3d variational analysis scheme. **Monthly Weather Review**, v. 128, n. 8, p. 2905-2919, 2000. 2, 10, 96, 97

HARLIM, J.; HUNT, B. R. **Local Ensemble Transform Kalman Filter: an Efficient Scheme For Assimilating Atmospheric Data**. 2005. Disponível em: <[https://www.atmos.umd.edu/~ekalnay/pubs/harlim\\_hunt05.pdf](https://www.atmos.umd.edu/~ekalnay/pubs/harlim_hunt05.pdf)>. 9

HERDIES, D. L.; H., F. S.; BONATTI J. P. AND CINTRA, R.; SILVA, A. D. O sistema de assimilação de dados atmosféricos global do CPTEC/INPE. In: SOCIEDADE BRASILEIRA DE METEOROLOGIA, 12., 2002, Foz do Iguaçu. **Congresso Brasileiro de Meteorologia, 12**. [S.l.], 2002. 1

HOUTEKAMER, P. L.; ZHANG, F. Review of the ensemble kalman filter for atmospheric data assimilation. **Monthly Weather Review**, v. 144, n. 12, p. 4489-4532, 2016. Disponível em: <<https://doi.org/10.1175/MWR-D-15-0440.1>>. 19

HUNT, B. R.; KOSTELICH, E. J.; SZUNYOGH, I. Objective weather-map analysis. **Physica D Nonlinear Phenomena**, v. 230, n. 1-2, p. 112-126, 2007. Disponível em: <<http://www.sciencedirect.com/science/article/B6TVK-4MNHXX0-3/2/43ec2cf5dffe5aacdf538cdeb90cd480>>. 2, 9, 18

JULIAN, P. R.; THIEBAUX, H. J. On some properties of correlation functions used in optimum interpolation schemes. **Monthly Weather Review**, v. 103, p. 605-616, 1975. 1

KALMAN, R. E. A new approach to linear filtering and prediction problems. **Transactions of the ASME-Journal of Basic Engineering**, v. 82, n. Series D, p. 35-45, 1960. 8

KALNAY, E. **Atmospheric modeling, data assimilation and predictability**. [S.l.]: Cambridge University, 2003. 341 p. 5

KLEIST, D. T.; PARRISH, D. F.; DERBER, J. C.; TREADON, R.; WU W. ; LORD, S. Introduction of the gsi into de ncep global data assimilation system. **Weather and Forecasting**, v. 24, p. 1691–1705, 2009. 2

LI, H.; LIU, J.; FERTIG, E.; KALNAY, E.; ARAVÉQUIA, J. A.; SZUNYOGH, I.; KOSTELICH, E. J.; TODLING, R. Improved analyses and forecasts with airs retrievals using the local ensemble transform kalman filter. In: . [S.l.: s.n.], 2007. 9

LIMA, L. N. **Estudo das incertezas na simulação por conjuntos e no uso da assimilação de dados no oceano Atlântico Sudoeste**. 144 p. Tese (Tese (Doutorado em Meteorologia)) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2018-03-20 2018. Disponível em: <<http://urlib.net/sid.inpe.br/mtc-m21c/2018/03.09.19.53>>. Acesso em: 05 ago. 2018. 19

LIU, J.; EUGENIA, K.; TAKEMASA, M.; CARLA, C. Analysis sensitivity calculation in an ensemble kalman filter. **Quarterly Journal of the Royal Meteorological Society**, v. 135, n. 644, p. 1842–1851, 2009. Disponível em: <<https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.511>>. 22, 24

LIU, J.; FERTIG, E. J.; LI, H.; KALNAY, E.; HUNT, B. R.; KOSTELICH, E. J.; SZUNYOGH, I.; TODLING, R. Comparison between local ensemble transform kalman filter and psas in the nasa finite volume gcm - perfect model experiments. **Nonlinear Processes in Geophysics**, v. 15, p. 645–659, 2008. 9

MIYOSHI, T. **Ensemble Kalman Filter Experiments with a primitive-equation Global Model**. 226 p. Tese (Doutorado em Meteorologia) — University of Maryland, College Park, 2005. 13, 14, 19, 63, 79

MIYOSHI, T.; SATO, Y. Assimilating satellite radiances with a local ensemble transform kalman filter (LETKF) applied to the jma global model (GSM). **Online Letters on the Atmosphere**, v. 135, p. 37–40, 2007. 9

MOLTENI, L. Atmospheric simulations using a gcm with simplified physical parametrizations. i: model climatology and variability in multi-decadal experiments. **Climate Dynamics**, v. 20, p. 175–191, 2003. 13

OTT, E.; HUNT B. H.AND SZUNYOGH, I.; ZIMIN, A. V.; KOSTELICH, E. J. A local ensemble kalman filter for atmospheric data assimilation. **Tellus**, v. 56, p.

415–428, 2004. Disponível em:

<<http://www.tellusa.net/index.php/tellusa/article/view/14462>>. 9

PANOFSKY, H. Objective weather-map analysis. **Journal Applied Meteorology**, v. 6, p. 386–392, 1949. 5

PENNY, S. G. The hybrid Local Ensemble Transform Kalman Filter. **Monthly Weather Review**, v. 142, p. 2139–2149, 2014. 2, 3, 11, 14, 16, 19, 96

\_\_\_\_\_. Mathematical foundations of hybrid data assimilation from a synchronization perspective. **Chaos: An Interdisciplinary Journal of Nonlinear Science**, v. 27, n. 12, p. 126801, 2017. Disponível em: <<https://doi.org/10.1063/1.5001819>>. 12

PENNY, S. G.; BEHRINGER, D. W.; CARTON, J. A.; KALNAY, E. A hybrid global ocean data assimilation system at ncep. **Monthly Weather Review**, v. 143, n. 11, p. 4660–4677, 2015. Disponível em: <<https://doi.org/10.1175/MWR-D-14-00376.1>>. 3, 12, 98

PETRIE, R. E. **Localization in the ensemble Kalman Filter**. 80 p. Tese (Doutorado em Meteorologia) — University of Reading, Reading, 2008-08 2008. 9

RICHARDSON, L. F. **Weather prediction by numerical process**. Cambridge University Press, 1922. 236 p. Disponível em: <<http://archive.org/stream/weatherpredictio00richrich#page/n1/mode/1up>>. 5

SILVA, A. D.; PFAENDTNER, J.; GUO, J.; SIENKIEWICZ, M.; COHN, S. Assessing the effects of data selection with dao's physical-space statistical analysis system. **Monthly Weather Review**, p. 273–278, 1994. 1

SILVEIRA, B. B. **A emissividade da superfície continental em micro-ondas e suas implicações na assimilação de radiâncias**. 239 p. Tese (Doutorado em Meteorologia) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2017. Disponível em: <<http://urlib.net/sid.inpe.br/mtc-m21b/2017/08.30.13.23>>. 107

SZUNYOGH, I.; KOSTELICH, E. J.; GYARMATI, G.; KALNAY, E.; HUNT, B. R.; OTT, E.; SATTERFIELD, E.; YORKE, J. A. A local ensemble transform kalman filter data assimilation system for the ncep global model. **Tellus**, v. 60A, p. 113–130, 2008. 9

TALAGRAND, O. Assimilation of observations, an introduction. **Journal of the Meteorological Society of Japan Serie II**, v. 75, n. 1B, p. 191–209, 1997. 6

- WANG, X.; BISHOP, C. H.; JULIER, S. J. Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble. **Monthly Weather Review**, v. 132, p. 1590–1605, 2004. 9
- WANG, X.; HAMILL, T. M.; WHITAKER, J. S.; BISHOP, C. H. A comparison of hybrid ensemble transform kalman filter optimum interpolation and ensemble square root filter analysis schemes. **Monthly Weather Review**, v. 135, n. 3, p. 1055–1076, 2007. 10
- WANG, X.; PARRISH, D.; KLEIST, D.; WHITAKER, J. Gsi 3dvar-based ensemble–variational hybrid data assimilation for ncep global forecast system: single-resolution experiments. **Monthly Weather Review**, v. 141, n. 11, p. 4098–4117, 2013. Disponível em: <<https://doi.org/10.1175/MWR-D-12-00141.1>>. 10
- WESPETAL, M. E.; PENNY, S. G.; KALNAY, E.; MIYOSHI, T.; GREYBUSH, S. J. Estimating climatological forecast error covariance for 3d-var and hybrid data assimilation using ensemble perturbations, and a comparison with the nmc method. **Monthly Weather Review. Submitted.**, 2018. 12, 14, 17, 98
- WHITAKER, J. S.; HAMILL, T. M. Ensemble data assimilation without perturbed observations. **Monthly Weather Review**, v. 130, p. 1913–1924, 2002. 8
- WHITAKER, J. S. T. M. H.; WEI, X.; SONG, Y.; TOTH, Z. Ensemble data assimilation with the ncep global forecast system. **Monthly Weather Review**, v. 136, p. 463–482, 2008. 9
- WU, W. S.; PURSER, J.; PARRISH, D. Three-dimensional variational analysis with spatially inhomogeneous covariances. **Monthly Weather Review**, v. 130, p. 2905–2916, 2002. 2
- YANG, S.-C.; CORAZZA, M.; CARRASSI, A.; KALNAY, E.; MIYOSHI, T. Comparison of local ensemble transform kalman filter, 3dvar, and 4dvar in a quasigeostrophic model. **Monthly Weather Review**, v. 137, n. 2, p. 693–709, 2009. Disponível em: <<https://doi.org/10.1175/2008MWR2396.1>>. 63, 79
- ZHANG, F.; ZHANG, M.; HANSEN, J. Coupling ensemble kalman filter with four-dimensional variational data assimilation. **Advances in Atmospheric Sciences**, v. 26, n. 1, p. 1–8, 2009. 10

## APÊNDICE A - EQUIVALÊNCIA DA ALTERAÇÃO NA MATRIZ GANHO E O ALGORITMO DO HÍBRIDO MÉDIO

O  $\mathbf{x}^a$  é definido pela minimização da seguinte função custo:

$$J(\mathbf{x}^a) = (\mathbf{x}^a - \bar{\mathbf{x}}^a)^T \mathbf{B}^{-1}(\mathbf{x}^a - \bar{\mathbf{x}}^a) + (\mathbf{y}^o - \mathbf{H}(\mathbf{x}^a))^T \mathbf{R}^{-1}(\mathbf{y}^o - \mathbf{H}(\mathbf{x}^a)) \quad (\text{A.1})$$

Fazendo a derivada de  $J = 0$  podemos resolver a Equação A.1 para  $\mathbf{x}^a$ .

$$\frac{dJ}{d\mathbf{x}^a} = 2(\mathbf{x}^a - \bar{\mathbf{x}}^a)^T \mathbf{B}^{-1} - 2(\mathbf{y}^o - \mathbf{H}(\mathbf{x}^a))^T \mathbf{R}^{-1} \mathbf{H} = 0 \quad (\text{A.2})$$

$$\frac{1}{2} \left( \frac{dJ}{d\mathbf{x}^a} \right)^T = \mathbf{B}^{-1}(\mathbf{x}^a - \bar{\mathbf{x}}^a) - \mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y}^o - \mathbf{H}\mathbf{x}^a) = 0 \quad (\text{A.3})$$

$$(\mathbf{B}^{-1}\mathbf{x}^a - \mathbf{B}^{-1}\bar{\mathbf{x}}^a) - (\mathbf{H}^T \mathbf{R}^{-1}\mathbf{y}^o - \mathbf{H}^T \mathbf{R}^{-1}\mathbf{H}\mathbf{x}^a) = 0 \quad (\text{A.4})$$

$$(\mathbf{B}^{-1} - \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})\mathbf{x}^a = \mathbf{B}^{-1}\bar{\mathbf{x}}^a + \mathbf{H}^T \mathbf{R}^{-1}\mathbf{y}^o \quad (\text{A.5})$$

$$(\mathbf{I} + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})\mathbf{x}^a = \bar{\mathbf{x}}^a + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1}\mathbf{y}^o \quad (\text{A.6})$$

$$\mathbf{x}^a = (\mathbf{I} + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1}(\bar{\mathbf{x}}^a + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1}\mathbf{y}^o) \quad (\text{A.7})$$

Substituindo a Equação A.7 na Equação do Híbrido Médio (Equação 3.1)

$$\mathbf{x}_{HM}^a = (1 - \alpha)\bar{\mathbf{x}}^a + \alpha(\mathbf{I} + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1}(\bar{\mathbf{x}}^a + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1}\mathbf{y}^o) \quad (\text{A.8})$$

$$(\mathbf{I} + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})\mathbf{x}_{HM}^a = \alpha(\bar{\mathbf{x}}^a + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1}\mathbf{y}^o) + (\mathbf{I} + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})(1 - \alpha)\bar{\mathbf{x}}^a \quad (\text{A.9})$$

$$= (\alpha\bar{\mathbf{x}}^a + \alpha\mathbf{B}\mathbf{H}^T \mathbf{R}^{-1}\mathbf{y}^o) + (1 - \alpha)\bar{\mathbf{x}}^a + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}(1 - \alpha)\bar{\mathbf{x}}^a \quad (\text{A.10})$$

$$= \bar{\mathbf{x}}^a + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1}\alpha\mathbf{y}^o + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1}(1 - \alpha)\mathbf{H}\bar{\mathbf{x}}^a \quad (\text{A.11})$$

$$= \bar{\mathbf{x}}^a + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1}[\alpha\mathbf{y}^o + (1 - \alpha)\mathbf{H}\bar{\mathbf{x}}^a] \quad (\text{A.12})$$

$$= \bar{\mathbf{x}}^a + \alpha\mathbf{B}\mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^a) + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1}\mathbf{H}\bar{\mathbf{x}}^a \quad (\text{A.13})$$

$$= (\mathbf{I} + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})\bar{\mathbf{x}}^a + \alpha\mathbf{B}\mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^a) \quad (\text{A.14})$$

Então,

$$\mathbf{x}_{HM}^a = \bar{\mathbf{x}}^a + (\mathbf{I} + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1}\alpha\mathbf{B}\mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^a) \quad (\text{A.15})$$

Tendo que a matriz ganho ( $\mathbf{K}^B$ ) é escrita da seguinte forma:

$$\mathbf{K}^B = \mathbf{B}\mathbf{H}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \quad (\text{A.16})$$

$$= (\mathbf{I} + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1}\mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} \quad (\text{A.17})$$

Então,

$$\mathbf{x}_{HM}^a = \bar{\mathbf{x}}^a + \alpha \mathbf{K}^B (\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^a) \quad (\text{A.18})$$

Partindo do LETKF

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b) \quad (\text{A.19})$$

Substituindo da Equação do Híbrido Médio (3.1),

$$= (1 - \alpha)[\bar{\mathbf{x}}^b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b)] + \alpha[\mathbf{x}^b + \mathbf{K}^B(\mathbf{y}^o - \mathbf{H}\mathbf{x}^b)] \quad (\text{A.20})$$

$$= \bar{\mathbf{x}}^b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b) - \alpha\bar{\mathbf{x}}^b - \alpha\mathbf{K}\mathbf{y}^o + \alpha\mathbf{K}\mathbf{H}\bar{\mathbf{x}}^b + \alpha\mathbf{x}^b + \alpha\mathbf{K}^B\mathbf{y}^o - \alpha\mathbf{K}^B\mathbf{H}\mathbf{x}^b \quad (\text{A.21})$$

$$= \bar{\mathbf{x}}^b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b) + \alpha[-\bar{\mathbf{x}}^b - \mathbf{K}\mathbf{y}^o + \mathbf{K}\mathbf{H}\bar{\mathbf{x}}^b + \mathbf{x}^b + \mathbf{K}^B\mathbf{y}^o - \mathbf{K}^B\mathbf{H}\mathbf{x}^b] \quad (\text{A.22})$$

Como o  $\mathbf{x}^b$  nesse caso é o  $\bar{\mathbf{x}}^a$ , então substitui  $\mathbf{x}^b$  pela Equação A.19.

$$\begin{aligned} &= \bar{\mathbf{x}}^b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b) + \\ &+ \alpha \left\{ -\bar{\mathbf{x}}^b - \mathbf{K}\mathbf{y}^o + \mathbf{K}\mathbf{H}\bar{\mathbf{x}}^b + [\bar{\mathbf{x}}^b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b)] + \right. \\ &\quad \left. + \mathbf{K}^B\mathbf{y}^o - \mathbf{K}^B\mathbf{H}[\bar{\mathbf{x}}^b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b)] \right\} \end{aligned} \quad (\text{A.23})$$

$$\begin{aligned} &= \bar{\mathbf{x}}^b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b) + \\ &+ \alpha(-\bar{\mathbf{x}}^b - \mathbf{K}\mathbf{y}^o + \mathbf{K}\mathbf{H}\bar{\mathbf{x}}^b + \bar{\mathbf{x}}^b + \mathbf{K}\mathbf{y}^o - \mathbf{K}\mathbf{H}\bar{\mathbf{x}}^b + \\ &\quad + \mathbf{K}^B\mathbf{y}^o - \mathbf{K}^B\mathbf{H}\bar{\mathbf{x}}^b + \mathbf{K}^B\mathbf{H}\mathbf{K}\mathbf{y}^o - \mathbf{K}^B\mathbf{H}\mathbf{K}\mathbf{H}\bar{\mathbf{x}}^b) \end{aligned} \quad (\text{A.24})$$

$$= \bar{\mathbf{x}}^b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b) + \alpha\mathbf{K}^B[\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b + \mathbf{H}\mathbf{K}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b)] \quad (\text{A.25})$$

$$= \bar{\mathbf{x}}^b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b) + \alpha\mathbf{K}^B(\mathbf{I} - \mathbf{H}\mathbf{K})(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b) \quad (\text{A.26})$$

$$= \bar{\mathbf{x}}^b + [\mathbf{K} + \alpha\mathbf{K}^B(\mathbf{I} - \mathbf{H}\mathbf{K})](\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b) \quad (\text{A.27})$$

Conforme a Equação 3.4, então

$$\mathbf{x}_{HM}^a = \bar{\mathbf{x}}^b + \widehat{\mathbf{K}}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b) \quad (\text{A.28})$$

## PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

### **Teses e Dissertações (TDI)**

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

### **Manuais Técnicos (MAN)**

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

### **Notas Técnico-Científicas (NTC)**

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

### **Relatórios de Pesquisa (RPQ)**

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

### **Propostas e Relatórios de Projetos (PRP)**

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

### **Publicações Didáticas (PUD)**

Incluem apostilas, notas de aula e manuais didáticos.

### **Publicações Seriadas**

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Contam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

### **Programas de Computador (PDC)**

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.

### **Pré-publicações (PRE)**

Todos os artigos publicados em periódicos, anais e como capítulos de livros.