



Universidade do Estado do Rio de Janeiro

Centro de Tecnologia e Ciências

Faculdade de Engenharia

Alexandre Ribeiro Freitas

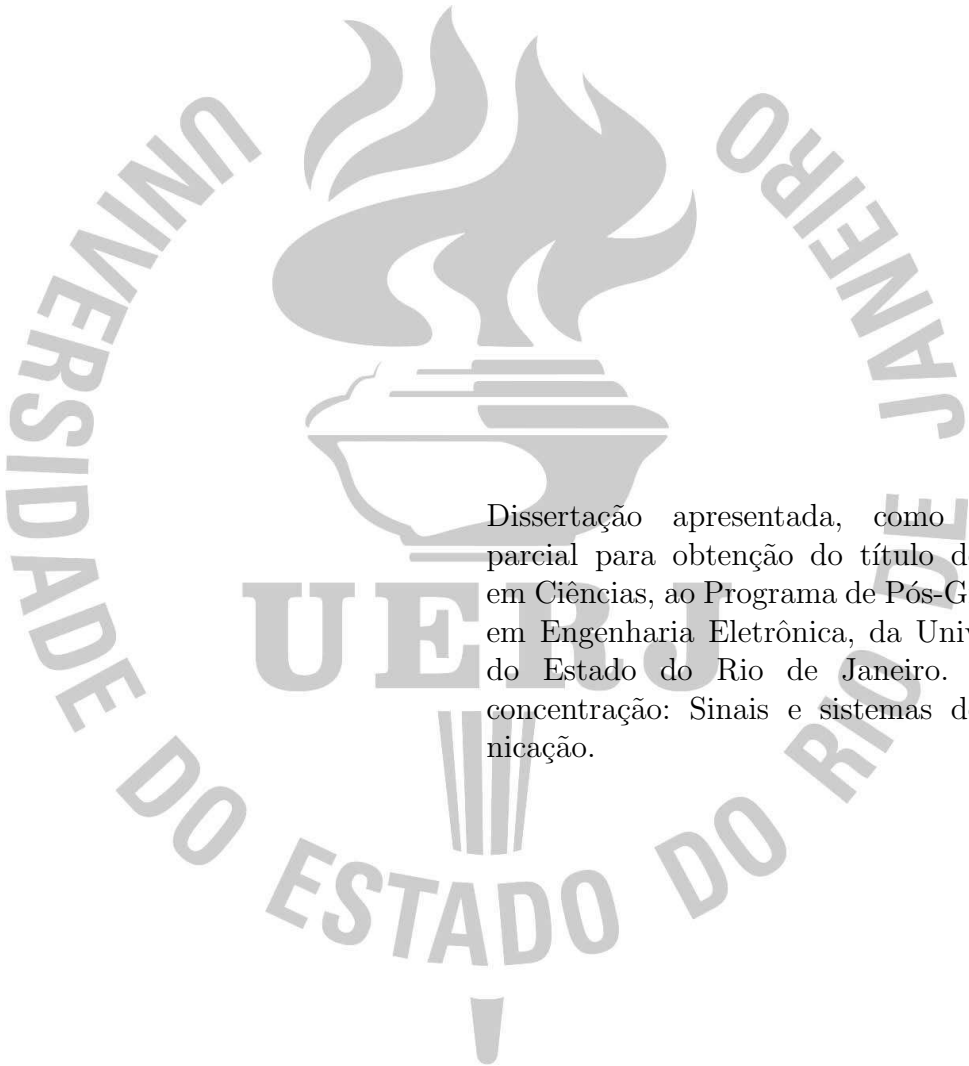
**Filtragem adaptativa aplicada à redução de desvios em séries
temporais de previsão numérica climática**

Rio de Janeiro

2015

Alexandre Ribeiro Freitas

**Filtragem adaptativa aplicada à redução de desvios em séries temporais de
previsão numérica climática**



Dissertação apresentada, como requisito parcial para obtenção do título de Mestre em Ciências, ao Programa de Pós-Graduação em Engenharia Eletrônica, da Universidade do Estado do Rio de Janeiro. Área de concentração: Sinais e sistemas de Comunicação.

Orientadores: Prof. Dr. Michel Pompeu Tcheou
Prof. Dr. Lisandro Lovisolo

Rio de Janeiro

2015

CATALOGAÇÃO NA FONTE

UERJ / REDE SIRIUS / BIBLIOTECA CTC/B

F866 Freitas, Alexandre Ribeiro.
Filtragem adaptativa aplicada à redução de desvios em
séries temporais de previsão numérica climática /
Alexandre Ribeiro Freitas. - 2015.
122 f.

Orientadores: Michel Pompeu Tcheou, Lisandro
Lovisolo.

Dissertação (Mestrado) – Universidade do Estado do
Rio de Janeiro, Faculdade de Engenharia.

1. Engenharia Eletrônica. 2. Filtragem – Dissertações.
3. Algoritmos adaptativos – Dissertações. 4. Previsão
numérica do tempo (Meteorologia) – Dissertações. I.
Tcheou, Michel Pompeu. II. Lovisolo, Lisandro. III.
Universidade do Estado do Rio de Janeiro. IV. Título.

CDU 004.421:551.509

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

Assinatura

Data

Alexandre Ribeiro Freitas

**Filtragem adaptativa aplicada à redução de desvios em séries temporais de
previsão numérica climática**

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre em Ciências, ao Programa de Pós-Graduação em Engenharia Eletrônica, da Universidade do Estado do Rio de Janeiro. Área de concentração: Sinais e sistemas de Comunicação.

Aprovado em 28 de Setembro de 2015.

Banca Examinadora:

Prof. Dr. Michel Pompeu Tcheou (Orientador)

Faculdade de Engenharia - UERJ

Prof. Dr. Lisandro Lovisolo (Orientador)

Faculdade de Engenharia - UERJ

Prof.^a Dra. Chou Sin Chan

Instituto Nacional de Pesquisas Espaciais - INPE

Prof. Dr. Tadeu Nagashima Ferreira

Universidade Federal Fluminense - UFF

Rio de Janeiro

2015

AGRADECIMENTO

Agradeço primeiramente a Deus por estar sempre comigo e colocar pessoas que me ajudaram a concluir mais este desafio.

Agradeço a minha esposa Lilian pelo amor. Sua compreensão nos momentos de ausência e suas palavras de incentivo nos momentos de desânimo foram essenciais para o término deste trabalho.

Agradeço aos meus amigos e familiares pelo apoio e compreensão. Em especial a minha mãe Maryuse pelas palavras de incentivo.

Agradeço ao meu orientador Michel P. Tcheou pois ele me deu a oportunidade de iniciar a pesquisa que posteriormente se tornou esta dissertação de mestrado. Agradeço pela paciência de mostrar os erros cometidos por mim, tanto na escrita quanto no desenvolvimento dos algoritmos e agradeço também pelo companheirismo em momentos difíceis.

Agradeço ao meu orientador Lisandro Lovisolo, que me ajudou de inúmeras formas. Agradeço pela disponibilidade de me orientar e por suas inspirações que ajudaram a nortear o caminho deste trabalho. Agradeço por todos os momentos em que me mostrou a melhor forma de expor as ideias e também pelas dicas e orientações de como alcançar os objetivos propostos por ele.

Aos colegas e professores Cleiton Magalhães, Ernesto Rodrigues, Flávio Avila e Luis Monteiro pelas valiosas conversas e pelo apoio no decorrer do mestrado.

Agradeço a Dra. Chou pelo incentivo, compreensão, disponibilidade e contribuições durante o desenvolvimento deste trabalho. Agradeço também por em esclarecer as dúvidas na área de meteorologia e pela oportunidade de participar da pesquisa em conjunto com o INPE.

Agradeço ao Programa de Pós-Graduação em Engenharia Eletrônica por permitir que eu pudesse avançar em conhecimento. Agradeço ao Laboratório de Processamento de Sinais, Aplicações Inteligentes e Comunicação (PROSAICO) pela infraestrutura computacional fornecida para a realização da pesquisa. Agradeço a CAPES e ao PCTEC/INCPE pelo apoio financeiro. Agradeço a todos do PCTEC/INPE pela disponibilização dos dados necessários para concluir este trabalho.

*”Devemos buscar a verdade, amá-la, defendê-la,
e entregá-la incorrupta para nossa posteridade”*

Felipe Malâncton

RESUMO

FREITAS, Alexandre Ribeiro. *Filtragem adaptativa aplicada à redução de desvios em séries temporais de previsão numérica climática*. 122 f. Dissertação (Mestrado em Engenharia Eletrônica) - Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro (UERJ), Rio de Janeiro, 2015.

Esta Dissertação apresenta resultados da aplicação de filtros adaptativos, utilizando os algoritmos NLMS (*Normalized Least Mean Square*) e RLS (*Recursive Least Square*), para a redução de desvios em previsões climáticas. As discrepâncias existentes entre o estado real da atmosfera e o previsto por um modelo numérico tendem a aumentar ao longo do período de integração. O modelo atmosférico Eta é utilizado operacionalmente para previsão numérica no CPTEC/INPE e como outros modelos atmosféricos, apresenta imprecisão nas previsões climáticas. Existem pesquisas que visam introduzir melhorias no modelo atmosférico Eta e outras que avaliam as previsões e identificam os erros do modelo para que seus produtos sejam utilizados de forma adequada. Dessa forma, neste trabalho pretende-se filtrar os dados provenientes do modelo Eta e ajustá-los, de modo a minimizar os erros entre os resultados fornecidos pelo modelo Eta e as reanálises do NCEP. Assim, empregamos técnicas de processamento digital de sinais e imagens com o intuito de reduzir os erros das previsões climáticas do modelo Eta. Os filtros adaptativos nesta dissertação ajustarão as séries ao longo do tempo de previsão. Para treinar os filtros foram utilizadas técnicas de agrupamento de regiões, como por exemplo o algoritmo de clusterização *k-means*, de modo a selecionar séries climáticas que apresentem comportamentos semelhantes entre si. As variáveis climáticas estudadas são o vento meridional e a altura geopotencial na região coberta pelo modelo de previsão atmosférico Eta com resolução de 40 km, a um nível de pressão de 250 hPa. Por fim, os resultados obtidos mostram que o filtro com 4 coeficientes, adaptado pelo algoritmo RLS em conjunto com o critério de seleção de regiões por meio do algoritmo *k-means* apresenta o melhor desempenho ao reduzir o erro médio e a dispersão do erro, tanto para a variável vento meridional quanto para a variável altura geopotencial.

Palavras-chave: Filtro adaptativo; Previsão atmosférica; *k-means*; NLMS; RLS.

ABSTRACT

FREITAS, Alexandre Ribeiro. *Adaptive filtering applied to reduction of deviations in numerical climate prediction time series*. 122 f. Dissertação (Mestrado em Engenharia Eletrônica) - Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro (UERJ), Rio de Janeiro, 2015.

This thesis presents results of application of adaptive filters, using the NLMS (Normalized Least Mean Square) and RLS (Recursive Least Square), to reduce deviations in weather forecasts. Discrepancies between the actual state of the atmosphere and predicted by a numerical model tend to increase over the integration period. The atmospheric model Eta is used operationally for numerical forecasting at CPTEC / INPE and as other atmospheric models, shows the accuracy of the weather forecasts. There is research aimed at introducing improvements in atmospheric model Eta 40km and others who evaluate the predictions and identify model errors for their products to be used properly. This work is intended to filter data from the Eta model and adjust them so as to minimize the error between the results provided by the Eta model and the NCEP. Thus, we employ digital signal processing techniques and images in order to improve weather forecasts Eta model. The adaptive filter in this paper will adapt the series throughout the forecast period. To train the filter regions, clustering techniques were used, such as clustering algorithm k-means so as to select climate series which have similar behavior to each other. Climate variables are the meridional wind and geopotential height in the area covered by atmospheric forecast model Eta with resolution of 40 km, to a pressure level of 250 hPa. Finally, the results show that the filter with 4 coefficients, adapted by RLS algorithm in conjunction with the criteria for selection of areas through k-means algorithm has the best performance by reducing the EQMn and dispersion of error, both the meridional wind variable and for the variable geopotential height.

Keywords: Adapted filters; weather forecast; *k-means*; NLMS; RLS.

LISTA DE FIGURAS

Figura 1	Região estudada representada pelo interior do retângulo.	22
Figura 2	Série do vento meridional de 2007 cujo primeiro dia de previsão é treze de Dezembro	23
Figura 3	Configuração geral do filtro adaptativo	25
Figura 4	Esquema de aninhamento do modelo Eta	28
Figura 5	Variável vertical σ representada pelas linhas vermelhas	29
Figura 6	Variável vertical Eta representada pelas linhas amarelas na horizontal .	32
Figura 7	Grade E de Arakawa	33
Figura 8	Resolução espacial e condições de contorno	34
Figura 9	Configuração geral do filtro adaptativo	38
Figura 10	Vorticidade (rotação) do vento visualizada no GraDS	45
Figura 11	Resolução inicial (a) e resolução final (b) do modelo Eta e do NCEP...	46
Figura 12	Filtro Passa Baixas bidimensional utilizado	47
Figura 13	Filtro Passa Baixas utilizado	47
Figura 14	Efeito da utilização do filtro passa baixas	48
Figura 15	Método utilizado para a implementação do filtro adaptativo	48
Figura 16	Clusterização realizada pelo algoritmo <i>k-means</i> sobre parâmetro vento meridional, onde cada isolinha separa um cluster	51
Figura 17	Seleção ao longo das abscissas e das ordenadas (Fora de escala).....	52
Figura 18	Seleção das bordas laterais do Modelo (Fora de escala)	53
Figura 19	Seleção em regiões circulares (Fora de escala).....	55
Figura 20	Seleção em anéis retangulares	56
Figura 21	Variável vento meridional (m/s) a 250 hPa previsto pelo modelo Eta no domínio de previsão	57
Figura 22	Variável altura geopotencial (mcp) a 250 hPa previsto pelo modelo Eta no domínio de previsão	58
Figura 23	Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo NLMS e com 500 séries de treinamento	63
Figura 24	Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo RLS e com 500 séries de treinamento	64
Figura 25	Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo NLMS e com 1000 séries de treinamento	64

Figura 26 Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo RLS e com 1000 séries de treinamento	65
Figura 27 Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo NLMS e com 500 séries de treinamento	65
Figura 28 Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo RLS e com 500 séries de treinamento	66
Figura 29 Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo NLMS e com 1000 séries de treinamento.....	66
Figura 30 Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo RLS e com 1000 séries de treinamento	67
Figura 31 Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo NLMS e com 500 séries de treinamento.....	72
Figura 32 Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo RLS e com 500 séries de treinamento	73
Figura 33 Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo NLMS e com 1000 séries de treinamento	73
Figura 34 Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo RLS e com 1000 séries de treinamento	74
Figura 35 Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo NLMS e com 500 séries de treinamento.....	74
Figura 36 Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo RLS e com 500 séries de treinamento	75
Figura 37 Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo NLMS e com 1000 séries de treinamento	75
Figura 38 Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo RLS e com 1000 séries de treinamento	76

Figura 39	Porcentagem de regiões com redução no EQMn da variável vento meridional, na aplicação do filtro de 4 coeficientes, adaptado pelo algoritmo NLMS	79
Figura 40	Porcentagem de regiões com redução no EQMn da variável vento meridional, na aplicação do filtro de 4 coeficientes, adaptado pelo algoritmo RLS	79
Figura 41	Porcentagem de regiões com redução no EQMn da variável vento meridional, na aplicação do filtro de 32 coeficientes, adaptado pelo algoritmo NLMS	80
Figura 42	Porcentagem de regiões com redução no EQMn da variável vento meridional, na aplicação do filtro de 32 coeficientes, adaptado pelo algoritmo RLS	80
Figura 43	Porcentagem de regiões com redução no EQMn da variável altura geopotencial, na aplicação do filtro de 4 coeficientes, adaptado pelo algoritmo NLMS	83
Figura 44	Porcentagem de regiões com redução no EQMn da variável altura geopotencial, na aplicação do filtro de 4 coeficientes, adaptado pelo algoritmo RLS	83
Figura 45	Porcentagem de regiões com redução do EQMn sobre a variável altura geopotencial, na aplicação do filtro de 32 coeficientes, adaptado pelo algoritmo NLMS	84
Figura 46	Porcentagem de regiões com redução do EQMn sobre a variável altura geopotencial, na aplicação do filtro de 32 coeficientes, adaptado pelo algoritmo RLS	84
Figura 47	EQMn antes e depois da aplicação do filtro de comprimento de um dia de previsões nas séries climáticas do Eta sobre a variável vento meridional	85
Figura 48	EQMn antes e depois da aplicação do filtro de comprimento de oito dias de previsões nas séries climáticas do Eta sobre a variável vento meridional	86
Figura 49	Média do EQMn por janela, na aplicação do filtro com comprimento de um dia de previsões sobre a variável vento meridional	86
Figura 50	Média do EQMn por janela, na aplicação do filtro com comprimento de oito dias de previsões sobre a variável vento meridional	87
Figura 51	Distribuição da média do EQMn por janela, na aplicação do filtro com um dia de previsões sobre a variável vento meridional	87
Figura 52	Distribuição da média do EQMn por janela, na aplicação do filtro com oito dias de previsões sobre a variável vento meridional	88
Figura 53	Série climática da variável vento meridional referente as dados do Eta, NCEP e Eta ajustado com filtro de 4 coeficientes	89
Figura 54	Série climática da variável vento meridional referente as dados do Eta, NCEP e Eta ajustado com filtro de 32 coeficientes	89

Figura 55	Erro entre a série climática da variável vento meridional referente as dados do Eta e série do NCEP em linha tracejada e erro entre a série do Eta ajustado com filtro de 4 coeficientes e a série do NCEP em linha cheia	90
Figura 56	Erro entre a série climática da variável vento meridional referente as dados do Eta e série do NCEP em linha tracejada e erro entre a série do Eta ajustado com filtro de 32 coeficientes e a série do NCEP em linha cheia.....	90
Figura 57	EQMn por janela, na aplicação do filtro com comprimento de um dia de previsões sobre a variável vento meridional	91
Figura 58	EQMn por janela, na aplicação do filtro com comprimento de oito dias de previsões sobre a variável vento meridional	91
Figura 59	Distribuição do EQMn por janela, na aplicação do filtro com comprimento de um dia de previsões sobre a variável vento meridional.....	92
Figura 60	Distribuição do EQMn por janela, na aplicação do filtro com comprimento de oito dias de previsões sobre a variável vento meridional	92
Figura 61	Série da janela que apresentou maior redução do EQMn na aplicação do filtro com comprimento de um dia de previsões sobre a variável vento meridional.....	93
Figura 62	Série da janela que apresentou maior redução do EQMn na aplicação do filtro com comprimento de oito dias de previsões sobre a variável vento meridional.....	93
Figura 63	EQMn antes e depois da aplicação do filtro de comprimento de um dia de previsões nas séries climáticas do Eta sobre a variável altura geopotencial ..	94
Figura 64	EQMn antes e depois da aplicação do filtro de comprimento de oito dias de previsões nas séries climáticas do Eta sobre a variável altura geopotencial	94
Figura 65	Média do EQMn por janela, na aplicação do filtro com comprimento de um dia de previsões sobre a variável altura geopotencial.....	95
Figura 66	Média do EQMn por janela, na aplicação do filtro com comprimento de oito dias de previsões sobre a variável altura geopotencial	95
Figura 67	Distribuição da média do EQMn por janela, na aplicação do filtro com um dia de previsões sobre a variável altura geopotencial	96
Figura 68	Distribuição da média do EQMn por janela, na aplicação do filtro com oito dias de previsões sobre a variável altura geopotencial	96
Figura 69	Série climática da variável altura geopotencial referente as dados do Eta, NCEP e Eta ajustado pelo filtro de 4 coeficientes	97
Figura 70	Série climática da variável altura geopotencial referente as dados do Eta, NCEP e Eta ajustado pelo filtro de 32 coeficientes.....	97
Figura 71	Erro entre a série climática da variável altura geopotencial referente as dados do Eta e série do NCEP em linha tracejada e erro entre a série do Eta ajustado com filtro de 4 coeficientes e a série do NCEP em linha cheia	98

Figura 72 Erro entre a série climática da variável altura geopotencial referente as dados do Eta e série do NCEP em linha tracejada e erro entre a série do Eta ajustado com filtro de 32 coeficientes e a série do NCEP em linha cheia...	98
Figura 73 EQMn por janela, na aplicação do filtro com comprimento de um dia de previsões sobre a variável altura geopotencial.....	99
Figura 74 EQMn por janela, na aplicação do filtro com comprimento de oito dias de previsões sobre a variável altura geopotencial.....	99
Figura 75 Distribuição do EQMn por janela, na aplicação do filtro com comprimento de um dia de previsões sobre a variável altura geopotencial	100
Figura 76 Distribuição do EQMn por janela, na aplicação do filtro com comprimento de oito dias de previsões sobre a variável altura geopotencial	100
Figura 77 Série da janela que apresentou maior redução do EQMn na aplicação do filtro com comprimento de um dia de previsões sobre a variável altura geopotencial.....	101
Figura 78 Série da janela que apresentou maior redução do EQMn na aplicação do filtro com comprimento de um dia de previsões sobre a variável altura geopotencial.....	101
Figura 79 Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro adaptado pelo algoritmo RLS	102
Figura 80 Porcentagem de regiões com redução no EQMn da altura geopotencial, na aplicação do filtro adaptado pelo algoritmo RLS	103
Figura 81 Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro adaptado pelo algoritmo RLS	103
Figura 82 Porcentagem de regiões com redução no EQMn da altura geopotencial, na aplicação do filtro adaptado pelo algoritmo RLS	104
Figura 83 Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro adaptado pelo algoritmo RLS	104
Figura 84 Porcentagem de regiões com redução no EQMn da altura geopotencial, na aplicação do filtro adaptado pelo algoritmo RLS	105
Figura 85 Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro adaptado pelo algoritmo RLS	106
Figura 86 Porcentagem de regiões com redução no EQMn da altura geopotencial, na aplicação do filtro adaptado pelo algoritmo RLS	106
Figura 87 Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro adaptado pelo algoritmo RLS	107
Figura 88 Porcentagem de regiões com redução no EQMn da altura geopotencial, na aplicação do filtro adaptado pelo algoritmo RLS	107

Figura 89 Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes do <i>k-means</i> , sobre a variável vento meridional	108
Figura 90 Distribuição do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes do <i>k-means</i> , sobre a variável vento meridional	109
Figura 91 Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção de regiões retangulares ao longo das ordenadas, sobre a variável vento meridional	109
Figura 92 Distribuição do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção em regiões retangulares ao longo das ordenadas, sobre a variável vento meridional	110
Figura 93 Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção de regiões retangulares ao longo das abscissas, sobre a variável vento meridional	110
Figura 94 Distribuição do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção em regiões retangulares ao longo das abscissas, sobre a variável vento meridional	111
Figura 95 Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões circulares, sobre a variável vento meridional	111
Figura 96 Distribuição do EQMn por janela na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões circulares, sobre a variável vento meridional	112
Figura 97 Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões anulares, sobre a variável vento meridional	112
Figura 98 Distribuição do EQMn por janela na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões anulares, sobre a variável vento meridional	113
Figura 99 Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes do <i>k-means</i> , sobre a variável altura geopotencial	114
Figura 100 Distribuição do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes do <i>k-means</i> , sobre a variável altura geopotencial	115

Figura 101 Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção em regiões retangulares ao longo das ordenadas, sobre a variável altura geopotencial	115
Figura 102 Distribuição do EQMn por janela do vento meridional, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção em regiões retangulares ao longo das ordenadas, sobre a variável altura geopotencial.....	116
Figura 103 Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção em regiões retangulares ao longo das abscissas, sobre a variável altura geopotencial	116
Figura 104 Distribuição do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção em regiões retangulares ao longo das abscissas, sobre a variável altura geopotencial.....	117
Figura 105 Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões circulares, sobre a variável altura geopotencial	117
Figura 106 Distribuição do EQMn por janela do vento meridional, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões circulares, sobre a variável altura geopotencial...	118
Figura 107 Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões anulares, sobre a variável altura geopotencial	118
Figura 108 Distribuição do EQMn por janela do vento meridional, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões anulares, sobre a variável altura geopotencial	119

LISTA DE TABELAS

Tabela 1	Complexidade computacional dos algoritmos NLMS e RLS.	43
Tabela 2	Porcentagem de regiões com redução no EQMn do vento meridional, aplicando filtros de 1 dia (4 coeficientes), com cardinalidade de treinamento determinada pelas bordas laterais	59
Tabela 3	Porcentagem de regiões com redução no EQMn do vento meridional, aplicando filtros de 8 dias (32 coeficientes), com cardinalidade de treinamento determinada pelas bordas laterais	60
Tabela 4	Porcentagem de regiões com redução no EQMn do vento meridional, aplicando filtros de 1 dia (4 coeficientes), com cardinalidade de treinamento determinada pela borda superior	61
Tabela 5	Porcentagem de regiões com redução no EQMn do vento meridional, aplicando filtros de 8 dias (32 coeficientes), com cardinalidade de treinamento determinada pela borda superior	62
Tabela 6	Porcentagem de regiões que se obtém redução no EQMn da altura geopotencial, aplicando filtros de 1 dia (4 coeficientes), com cardinalidade do treinamento determinado pelas forçantes horizontais	68
Tabela 7	Porcentagem de regiões que se obtém redução no EQMn da altura geopotencial aplicando filtros de 8 dias (32 coeficientes), com cardinalidade do treinamento determinado pelas forçantes horizontais	69
Tabela 8	Porcentagem de regiões que se obtém redução no EQMn da altura geopotencial, aplicando filtros de 1 dia (4 coeficientes), com cardinalidade do treinamento determinado pelas forçantes verticais	70
Tabela 9	Porcentagem de regiões que se obtém redução no EQMn da altura geopotencial, aplicando filtros de 8 dias (32 coeficientes), com cardinalidade do treinamento determinado pelas forçantes verticais	71
Tabela 10	Porcentagem de regiões que se obtém redução no EQMn da variável vento meridional aplicando filtros de 1 dia (4 coeficientes)	77
Tabela 11	Porcentagem de regiões que se obtém redução no EQMn da variável vento meridional aplicando filtros de 8 dias (32 coeficientes)	78
Tabela 12	Porcentagem de regiões que se obtém redução no EQMn da variável altura geopotencial aplicando filtros de 1 dia (4 coeficientes)	81
Tabela 13	Porcentagem de regiões que se obtém redução no EQMn da variável altura geopotencial aplicando filtros de 8 dias (32 coeficientes)	82
Tabela 14	Tabela comparativa dos critérios de seleção de regiões propostos sobre a variável vento meridional	113

Tabela 15 Tabela comparativa dos critérios de seleção de regiões propostos sobre a variável altura geopotencial	119
--	-----

LISTA DE SIGLAS

ASCII	<i>American Standard Code for Information Interchange</i>
COLA	<i>Center for Ocean, Land and Atmosphere Studies</i>
CPTEC	Centro de Previsão de Tempo e Estudos Climáticos
ECMWF	<i>European Centre for Medium Range Weather Forecast</i>
EQMn	Erro Quadrático Médio normalizado
GFDL	<i>Geophysical Fluid Dynamics Laboratory</i>
GFS	<i>Global Forecast System</i>
GRIB	<i>Gridded Binary or General Regularly-distributed Information in Binary form</i>
IC	Intervalo de Confiança
INPE	Instituto Nacional de Pesquisas Espaciais
LMS	<i>Least Mean Square</i>
MCGA	Modelo de Circulação Geral Atmosférico
NOAA	<i>National Oceanic and Atmospheric Administration</i>
MRF	<i>Medium Range Forecasting Model</i>
MSE	<i>Mean Square Error</i>
NeTCDF	<i>Network Common Data Form</i>
NLMS	<i>Normalized Least Mean Square</i>
RLS	<i>Recursive Least Square</i>

SUMÁRIO

	INTRODUÇÃO	21
1	MODELOS ATMOSFÉRICOS	27
1.1	Modelo de Circulação Geral Atmosférico (MCGA)	28
1.1.1	<u>Coordenada Vertical: sigma (σ)</u>	28
1.1.2	<u>Variáveis Prognósticas</u>	29
1.1.3	<u>Resolução Espacial</u>	30
1.1.4	<u>Condições Iniciais e Condições de Contorno</u>	30
1.2	Modelo GFS (<i>Global Forecast System</i>)	30
1.2.1	<u>Coordenada Vertical</u>	30
1.2.2	<u>Variáveis Prognósticas</u>	30
1.2.3	<u>Resolução Espacial</u>	31
1.3	Modelo Atmosférico Eta	31
1.3.1	<u>Coordenada Vertical: eta (η)</u>	31
1.3.2	<u>Variáveis Prognósticas</u>	32
1.3.3	<u>Grade horizontal</u>	33
1.3.4	<u>Condições Iniciais e Condições de contorno</u>	33
1.4	Reanálises do NCEP	34
2	ADAPTAÇÃO DE FILTROS PARA A CORREÇÃO DE DES- VIOS DE PREVISÃO	36
2.1	Filtros Digitais	36
2.2	Filtragem Adaptativa	38
2.2.1	<u>Algoritmos LMS e NLMS</u>	41
2.2.2	<u>Algoritmo RLS</u>	41
2.3	Métricas de desempenho	43
2.3.1	<u>Erro quadrático médio normalizado</u>	43
2.3.2	<u>Variância</u>	44
2.3.3	<u>Intervalo de Confiança</u>	44
3	METODOLOGIA	45
3.1	Filtragem adaptativa das previsões	48
3.1.1	<u>Referência Temporal</u>	49
3.1.2	<u>Critérios de Seleção de Regiões de Treinamento/Ajuste</u>	49
3.1.2.1	Seleção via <i>k-means</i> (primeiro critério)	49
3.1.2.2	Seleção via regiões retangulares (segundo critério)	51
3.1.2.3	Seleção via região circular (terceiro critério)	54
3.1.2.4	Seleção via anéis retangulares (quarto critério)	55

4	RESULTADOS	57
4.1	Configurações	57
4.2	Resultados encontrados	59
4.2.1	<u>Análises específicas</u>	85
4.2.2	<u>Avaliação dos melhores critérios de seleção de regiões</u>	101
4.2.2.1	Avaliação dos critérios propostos	108
	CONCLUSÃO	121
	REFERÊNCIAS	122

INTRODUÇÃO

A previsão climática possui inúmeros benefícios para a sociedade. Condições climáticas afetam a agricultura, o transporte de cargas, a indústria, o setor de energia elétrica, entre outras atividades econômicas (JONES et al., 2000). Tanto a identificação adequada das épocas para plantio e colheita, como uma boa tomada de decisão para a operação de reservatórios de usinas hidrelétricas dependem fortemente do clima a curto e longo prazos (BRAVO; COLLISCHONN; TUCCI, 2008). Assim, dispor de previsões meteorológicas cada vez mais acuradas traz benefícios inquestionáveis para a sociedade.

A previsão climática é o resultado (da estimativa) mais provável para evolução real do clima no futuro. Esta estimativa é obtida por meio de um conjunto de parâmetros baseados nas leis físicas da atmosfera que são tratados matematicamente por meio de modelos numéricos. Estes modelos integram, ou seja, calculam a solução de sistemas de equações atmosféricas tendo como dependência dados das condições iniciais fornecidas no início da integração e em alguns casos, das condições de contorno. As condições iniciais descrevem o estado observado na atmosfera, e servem como um ponto de partida para o modelo. Modelos regionais necessitam de contorno lateral. As condições de contorno são dados inseridos nas bordas do modelo, que se propagam para o interior do domínio do modelo induzindo a previsão, de modo que o modelo seja capaz de descrever eventos climatológicos de pequena escala (LAPRISE et al., 2008).

Os modelos de previsão têm desempenhos que podem variar conforme a estação do ano, da região, do horário inicial de integração, da resolução do modelo e também da quantidade de observações assimiladas pelas análises (BUSTAMANTE et al., 2005; GUIMARÃES; DEREZYNSKI; C., 2006; ROZANTE; CAVALCANTI, 2006). O prazo de integração de um modelo regional capaz de produzir previsões confiáveis depende, dentre outras coisas, da estabilidade numérica, da resolução, da qualidade das informações das condições de contorno utilizadas e da época do ano (ROZANTE, 2004; MOURA; HERDIES; ROZANTE, 2010).

Há diversos modelos numéricos em uso para realizar a previsão numérica atmosférica. Os modelos numéricos representam a atmosfera como sendo complexos conjuntos de equações matemáticas nas quais simula-se (prevê-se) o estado futuro da atmosfera a partir de condições iniciais específicas. Esses modelos dividem a atmosfera em camadas verticais e a superfície em uma malha horizontal, com diversos nós, formando assim uma malha tridimensional.

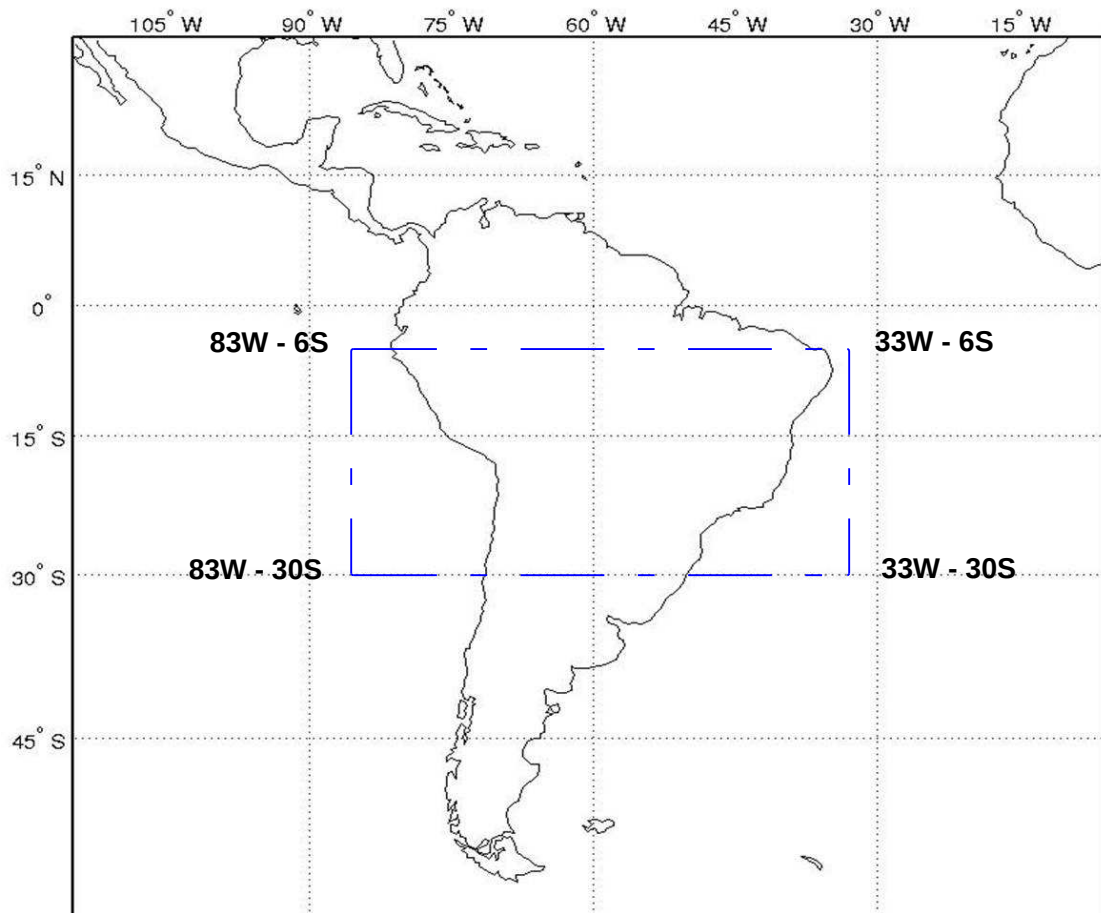
O modelo atmosférico Eta é utilizado operacionalmente para previsão numérica no CPTEC/INPE (Centro de Previsão de Tempo e Estudos Climáticos / Instituto Nacional de Pesquisas Espaciais). Ele foi originalmente desenvolvido pela Universidade de Belgrado em conjunto com o Instituto de Hidrometeorologia da Iugoslávia (BLACK, 1994). No CPTEC o Eta é empregado desde 1996 com o objetivo de complementar a previsão numérica

climática do modelo de circulação geral atmosférico (MCGA) que vinha sendo realizada desde o início de 1995. As variáveis prognósticas do modelo Eta são: temperatura do ar, componentes zonal e meridional do vento, umidade específica, pressão à superfície, energia cinética turbulenta, umidade e temperatura do solo, água líquida ou gelo das nuvens.

O modelo Eta aqui considerado, tem resolução espacial de 40 km e sua abrangência é regional. Ele utiliza como condição inicial os dados de análise do modelo GFS (*Global Forecast System*), fornecidos pelo NCEP (*National Center for Environmental Prediction*), e como condições de contorno, as previsões do modelo global de circulação geral atmosférico do CPTEC (BUSTAMANTE et al., 2005).

O interior do retângulo da Figura 1 representa a região estudada. Os lados do retângulo representam a região onde são inseridas as condições de contorno provenientes do MCGA, e seus vértices, se encontram nas coordenadas, 83W (West) e 6S (South), 83W e 30S, 33W e 6S, e 33W e 30S, e no interior do retângulo usa-se o GFS do NCEP como condição inicial.

Figura 1 – Região estudada representada pelo interior do retângulo.



Dentre as atividades do NCEP há o fornecimento de um conjunto de dados que representam com fidelidade as condições climáticas observadas. Esses dados são conhecidos como reanálises do NCEP/CFSR (*Climate Forecast System Reanalysis*). As reanálises

do NCEP são uma série de dados meteorológicos obtidos através de dados observados em todo o planeta e que posteriormente são reanalisados (PINTO; DINIZ; COSTA, 2008; MEIRES, 2012).

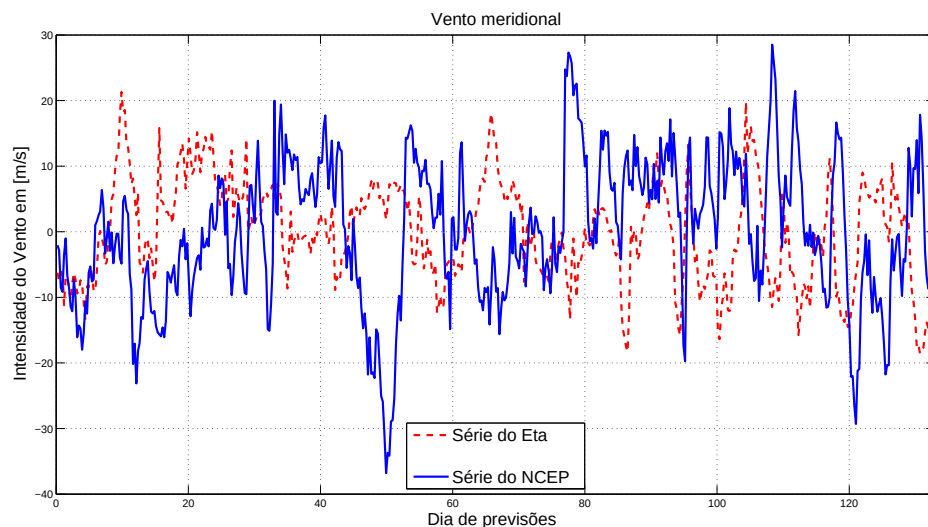
As discrepâncias existentes entre o estado real da atmosfera e o previsto por um modelo numérico tendem a aumentar ao longo do tempo de integração. Isso ocorre principalmente devido a dois tipos de erros: i) erros devidos às aproximações físicas do modelo numérico e ii) erros nas condições iniciais, que são amplificados com o tempo de integração do modelo, devido a não-linearidades nas equações do mesmo.

As condições de contorno lateral impactam fortemente as integrações climáticas de prazos longos (meses a décadas). Os erros oriundos desses contornos apresentam-se como um fator limitante para a utilização do modelo Eta além de um determinado prazo de previsão, como por exemplo previsões para os próximos 100 anos.

Já as condições iniciais afetam as integrações de curto a médio prazo, de alguns dias a algumas semanas (BUSTAMANTE et al., 2005; LAPRISE et al., 2008). Esses erros em geral crescem ao longo do tempo e levam ao decaimento da qualidade da previsão (BUSTAMANTE et al., 2005; LAPRISE et al., 2008). Para previsões sazonais (quatro meses e meio), os modelos numéricos apresentam erros na previsão atmosférica inerentes à modelagem (BUSTAMANTE et al., 2005).

O nível de qualidade da condição inicial utilizada impacta a acurácia dos resultados obtidos com o modelo. Há também erros ocasionados pelas condições de contorno lateral que podem propagar-se para o interior do domínio e comprometer as previsões (BUSTAMANTE et al., 2005; ROZANTE; CAVALCANTI, 2006).

Figura 2 – Série do vento meridional de 2007 cujo primeiro dia de previsão é treze de Dezembro. Em linha pontilhada tem-se a série dos dados do modelo Eta e na linha cheia tem-se a série dos dados provenientes das reanálises do NCEP



A Figura 2 apresenta um exemplo de previsão atmosférica da variável vento meri-

dional num horizonte de 13 de Dezembro de 2007 a 30 de Abril de 2008 que correspondem a 138 dias de previsões no total. A previsão fornecida pelo modelo atmosférico Eta é representada pela linha tracejada, enquanto a série proveniente dos dados de reanálise do NCEP estão em linha cheia. Percebe-se que existe discrepância entre as duas séries. Assim, pesquisas têm sido realizadas no sentido de verificar os desempenhos de modelos numéricos em variadas resoluções e com diferentes condições iniciais. Existem pesquisas que visam introduzir melhorias no modelo Eta 40 km para que ele sirva como boa condição de contorno para o Eta de menor resolução (Eta 15 km) e outras que avaliam as previsões e identificam os erros do modelo para que seus produtos sejam utilizados de forma adequada (ROZANTE, 2004; CHOU et al., 2006). Outros avanços foram alcançados por meio da interação entre áreas de conhecimento que visam compreender os padrões de ocorrência das variáveis meteorológicas no espaço e no tempo (CHOU et al., 2006). Dessa forma, neste trabalho pretende-se filtrar os dados provenientes do modelo Eta e ajustá-los, de modo a minimizar o erro entre os resultados fornecidos pelo modelo Eta e às reanálises do NCEP. Assim, empregamos técnicas de processamento digital de sinais e imagens com o intuito de melhorar as previsões climáticas do modelo Eta.

Como o Eta fornece um conjunto de dados climáticos com intervalos de seis horas, podemos tratá-los como sinais volumétricos discretos (em função da latitude, longitude e altitude) no tempo. Em cada altitude tem-se uma matriz na qual cada elemento traz variáveis meteorológicas correspondentes à coordenada geográfica (Latitude/Longitude).

Um sinal discreto é representado por uma sequência $\{x(k)\}$, $k \in \mathbb{Z}$. Cada $x(k)$ corresponde à amplitude do sinal em kT , e T a princípio pode descrever qualquer unidade de medida, como tempo, espaço, temperatura, entre outros (DINIZ; SILVA; NETTO, 2004; COHEN, 1995).

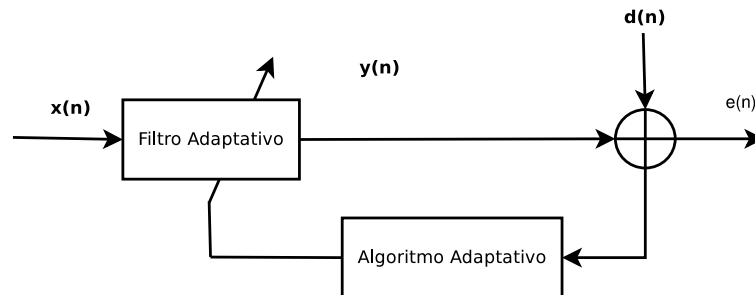
A abordagem de considerar dados meteorológicos como sinais de tempo discreto nos permite considerar a aplicação de sistemas de tempo discreto a esses dados. Um sistema, ou filtro, de tempo discreto mapeia a sequência de entrada $\{x(k)\}$ em uma sequência de saída $\{y(k)\}$ de forma que $y(k) = H\{x(k)\}$ onde o operador $H\{\cdot\}$ é o filtro no tempo discreto (DINIZ; SILVA; NETTO, 2004).

Entre as diversas aplicações de filtros discretos há a filtragem adaptativa. Nesse caso, os coeficientes do filtro são ajustados em função de um sinal de referência conhecido ao qual o sinal deveria corresponder. Assim um filtro, quando aplicado ao sinal, resulta na saída em um sinal mais próximo ao sinal de referência.

A capacidade de adaptar filtros e aplicá-los aos sinais faz da filtragem adaptativa uma interessante ferramenta para a correção de erros de previsão meteorológica. A Figura 3 mostra a configuração geral para a adaptação de filtros (DINIZ, 2008). Nela, $x(n)$ é o sinal de entrada, $y(n)$ é a saída do filtro adaptativo e $d(n)$ é definido como o sinal de referência ou desejado. O sinal de erro $e(n)$ é definido como $d(n) - y(n)$. O algoritmo adaptativo é usado para ajustar os coeficientes do filtro adaptativo de modo que sinal de

erro $e(n)$ seja minimizado (DINIZ, 2008). Pode-se considerar os dados climáticos do Eta como o sinal de entrada, e os parâmetros climáticos das reanálises do NCEP como o sinal de referência.

Figura 3 – Configuração geral do filtro adaptativo



Objetivo

Neste contexto, nesta dissertação, pretende-se melhorar a qualidade da previsão climática do modelo Eta para previsão sazonal (quatro meses e meio), e para isso, pretende-se minimizar os erros entre as previsões do modelo Eta de 40 km e os dados de reanálise do NCEP. Para atingir este objetivo será utilizada como ferramenta a filtragem adaptativa, que adapta os coeficientes de um determinado filtro, que ao ser aplicado sobre as séries do modelo Eta deve produzir uma série mais próxima aos dados de reanálise do NCEP. Assim, pretende-se investigar o uso de filtros adaptativos na correção de séries temporais de previsão climática.

Estrutura do Texto

No Capítulo 1 diferentes modelos atmosféricos utilizados neste trabalho são descritos. São eles: Modelo Eta, MCGA, Modelo GFS (*Global Forecast System*) e as reanálises do NCEP; Citam-se as principais características desses modelos.

No Capítulo 2 são apresentados alguns conceitos sobre filtros digitais, bem como suas aplicações, entre elas a filtragem adaptativa. Nesse capítulo discute-se sobre as principais características de algoritmos adaptativos, elegendo alguns aspectos decisivos para a proposta deste trabalho. Nesse capítulo também são apresentadas as métricas de desempenho empregadas. Essas métricas visam estabelecer critérios objetivos para a avaliação do desempenho da aplicação dos filtros adaptativos para a redução do erro, ou seja, da melhora da qualidade da previsão frente à reanálise do NCEP.

No Capítulo 3 é apresentada a implementação proposta. Alguns critérios de seleção de variáveis para formar os conjuntos de sinais de referência para a adaptação foram estudados, e esses são detalhados nesse capítulo.

No Capítulo 4 são apresentados os resultados obtidos de modo a avaliar o desempenho da implementação proposta. Apresenta-se ainda um estudo de caso com o objetivo de se avaliar com mais detalhes os efeitos da filtragem adaptativa sobre as séries climáticas.

No capítulo 5 apresentam-se as conclusões e propostas de pesquisas futuras.

1 MODELOS ATMOSFÉRICOS

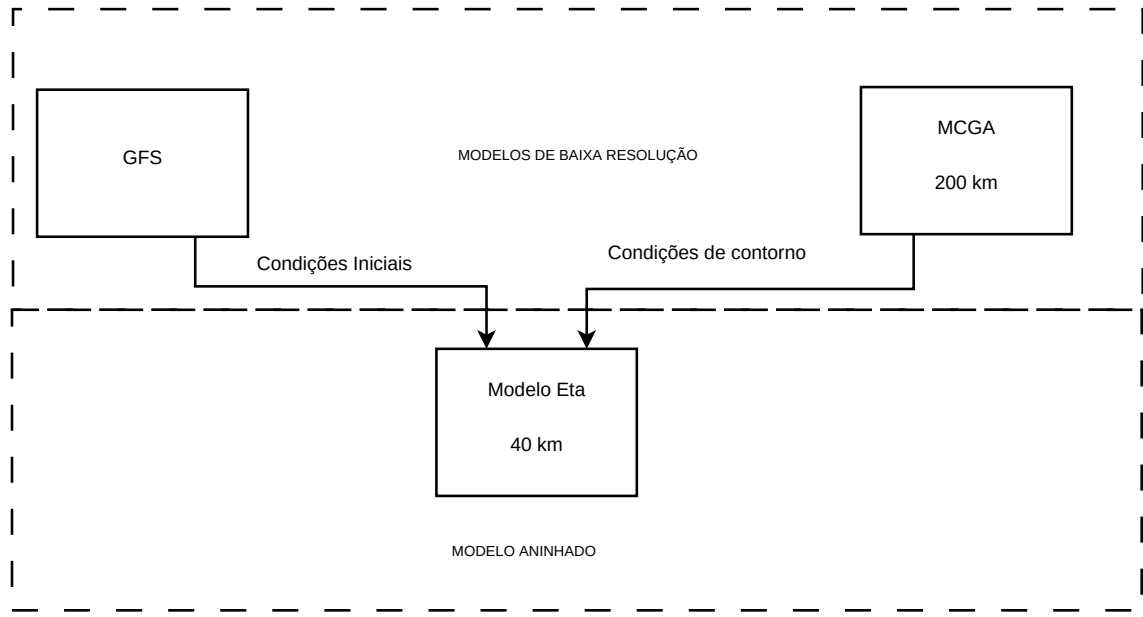
Modelos Atmosféricos de Circulação Global (MCGs), simulam o comportamento atmosférico em todo o planeta enquanto que os Modelos de Circulação Regional (MCRs), simulam o comportamento atmosférico em uma área específica. MCGs têm como objetivo representar o comportamento de fenômenos de grande escala e possuem baixa resolução espacial, na ordem de algumas centenas de quilômetros. Enquanto os MCRs objetivam representar comportamentos de meso escala em uma área específica, características locais e dados mais detalhados da superfície, e possuem resolução maior que do MCGs, poucos quilômetros ou poucas dezenas de quilômetros.

Para realizar a previsão atmosférica, modelos MCRs necessitam de condições iniciais e de condições de contorno. As condições iniciais descrevem o estado atual observado na atmosfera. As equações que modelam o comportamento atmosférico são resolvidas no tempo para obtenção de informações do comportamento da atmosfera num tempo futuro. Nesses modelos, as condições de contorno são os dados inseridos nas bordas do modelo, geralmente provenientes de previsões realizadas por um MCG de maior escala. A transferência de informação da maior escala (MCG) para a menor escala (MCR) é geralmente realizada através do *downscaling* dinâmico (LAPRISE et al., 2008), ou seja, através do aninhamento de resultados da modelagem em grande escala (MCG) para o MCR de baixa escala. A técnica de *downscaling* é utilizada para transformar modelos de alta escala (baixa resolução) dos modelos de circulação global para escalas mais refinadas (alta resolução) usando modelos climáticos regionais de alta resolução sobre uma determinada área de interesse (LAPRISE et al., 2008).

Neste trabalho o GFS (*Global Forecast System*) e MCGA (Modelo de Circulação Geral da Atmosfera) do CPTEC são modelos de grande escala, ou seja, são modelos atmosféricos de circulação global. O modelo Eta é um modelo de circulação regional e como tal é aninhado em modelos de grande escala. A Figura 4 mostra o esquema de aninhamento do modelo Eta. Ele recebe as condições iniciais do GFS e a condição de contorno do MCGA do INPE, ambos modelos de grande escala. O MCGA do CPTEC também utiliza o GFS como condição inicial.

Os modelos atmosféricos apresentam algumas características básicas que os diferenciam entre si. São elas a configuração horizontal, configuração vertical, variáveis meteorológicas e as condições iniciais e de contorno. Vejamos a seguir as características básicas do MCGA, GFS e Eta cujos resultados/previsões são considerados neste trabalho.

Figura 4 – Esquema de aninhamento do modelo Eta



1.1 Modelo de Circulação Geral Atmosférico (MCGA)

O Modelo de circulação geral atmosférico atualmente em uso no CPTEC tem sua origem no MRF (*Medium Range Forecasting Model*), usado para previsão do tempo no NCEP. O MRF era uma combinação do código espectral global desenvolvido pelo NCEP e as parametrizações físicas desenvolvidas no *Geophysical Fluid Dynamics Laboratory* (GFDL) do NOAA (*National Oceanic and Atmospheric Administration*) na Universidade de Princeton, EUA. Desde sua implantação no COLA (*Center for Ocean, Land and Atmosphere Studies*), o modelo MRF tem sofrido atualizações significativas de ambos os grupos. Entretanto, algumas mudanças de natureza experimental feitas pelo COLA, tornaram inviável suas adoções pelo NCEP. Assim, os modelos do NOAA e do COLA têm divergido substancialmente e assim o MCGA do COLA tem sido utilizado pelo CPTEC para previsão climática. O CPTEC adotou a versão 1.7 do COLA, sendo assim utilizada neste trabalho.

1.1.1 Coordenada Vertical: sigma (σ)

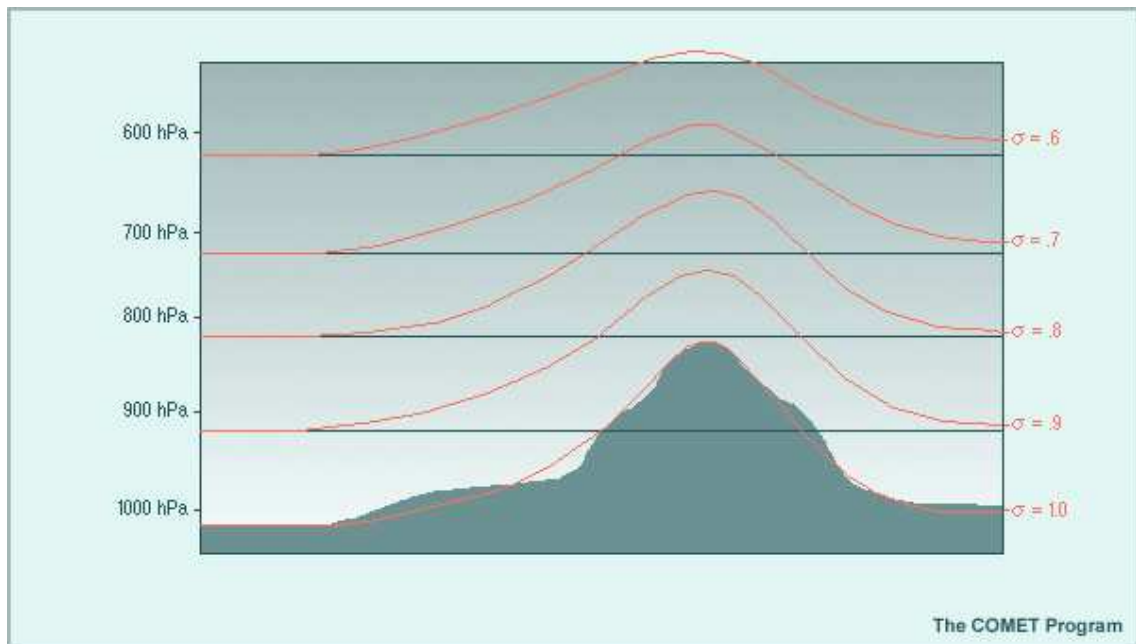
A coordenada vertical utilizada pelo MCGA é a coordenada sigma (JAMES et al., 1997; BONATTI, 2015). Esta coordenada define a forma como serão calculadas as equações de previsão em diversas alturas diferentes. Assim, o MCGA utiliza o sistema de coordenada sigma (σ) que é definida como,

$$\sigma = \frac{p - p_t}{p_s - p_t}, \quad (1)$$

p_t é a pressão no topo do modelo, p é a pressão no nível da previsão na atmosfera, e

p_s é a pressão na superfície da terra (mas não necessariamente no nível médio do mar). A Figura 5 mostra o esquema do sistema de coordenada sigma. Percebe-se que as inclinações do modelo sobre a montanha são suavizadas conforme a pressão diminui até serem aproximadamente paralelas a superfícies de pressão constante no topo do modelo ($\sigma = 0$)

Figura 5 – Variável vertical σ representada pelas linhas vermelhas



O uso da coordenada sigma permite formulações relativamente simples para tratar das condições de fronteira inferior do modelo atmosférico e se ajusta naturalmente a terrenos inclinados. Entretanto esta coordenada traz algumas limitações. Para realizar a previsão do vento, o modelo precisa calcular de forma correta o gradiente de pressão, que é simples de ser calculado em coordenada de pressão, porém quando a superfície sigma se inclina, o cálculo do gradiente de pressão gera erros. Estes erros podem se tornar muito grandes em regiões íngremes, comprometendo a previsão do vento. Outro problema é que para inclinações abruptas, este sistema acaba por representar de forma errônea a altitude verdadeira, causando erros nas previsões de pressão e temperatura em locais próximos a montanhas.

1.1.2 Variáveis Prognósticas

As variáveis prognósticas do MCGA são: o logaritmo da pressão à superfície, a vorticidade (componente vertical do rotacional do vento), a divergência do vento horizontal, a temperatura virtual e a umidade específica. Outras variáveis estão previstas nas parametrizações presentes no modelo, como temperatura de superfície, do interior do solo e do interior do dossel e umidade do solo, etc. As equações do modelo são escritas na

forma espectral e resolvidas numericamente.

O modelo MCGA do INPE aqui estudado, apresenta resultados de previsão quatro vezes ao dia, às 00, 06, 12 e 18 horas UTC e o período de previsão se estende de dias a décadas.

1.1.3 Resolução Espacial

O MCGA é empregado em resolução horizontal de 200x200 km e 100x100 km.

1.1.4 Condições Iniciais e Condições de Contorno

As condições iniciais do MCGA consistem de valores das variáveis: logaritmo da pressão na superfície, temperatura virtual, divergência horizontal, vorticidade vertical e umidade específica. Estes dados geralmente provêm dos produtos operacionais disponibilizados pelo NCEP. Estes centros de pesquisa têm dados de reanálise e algumas destas reanálises são utilizadas pelo MCGA.

A topografia é representada por um série truncada, sendo sujeita ao efeito de Gibbs. Ou seja, a soma das componentes da série de Fourier gera ondas no espaço físico.

O MCGA necessita de condições de contorno inferior da superfície da terra. Esta condição visa representar processos atmosféricos na escala do tempo de um mês ou mais.

1.2 **Modelo GFS** (*Global Forecast System*)

O GFS é um modelo de previsão global produzido pelo NCEP, que pertence ao NOAA. O NCEP está localizado nos Estados Unidos e é composto por nove centros de pesquisa, cada um responsável por diferentes serviços e produtos específicos. Este modelo apresenta resultados às 0, 06, 12 e 18 UTC e disponibiliza suas informações sobre todo o globo.

1.2.1 Coordenada Vertical

O modelo GFS utiliza como coordenada vertical, a coordenada sigma com 64 níveis, irregularmente espaçados (NOAA, 2015).

1.2.2 Variáveis Prognósticas

As variáveis prognósticas deste modelo são a vorticidade, divergência, temperatura virtual, pressão atmosférica e umidade específica (NOAA, 2015). O período de previsão se estende de dias a décadas.

1.2.3 Resolução Espacial

A resolução espacial é espectral triangular, o que equivale a aproximadamente 70 quilômetros (NOAA, 2015).

1.3 **Modelo Atmosférico Eta**

O modelo regional Eta é utilizado operacionalmente para previsão numérica no CPTEC/INPE (Centro de Previsão de Tempo e Estudos Climáticos/Instituto Nacional de Pesquisas Espaciais) (CHOU; BUSTAMANTE; GOMES, 2005). Ele foi inicialmente desenvolvido pela Universidade de Belgrado em conjunto com o Instituto de Hidrometeorologia da Iugoslávia (MESINGER et al., 1988). O Eta é um modelo de circulação regional e tem por objetivo prever com detalhes sistemas organizados em meso escala, como por exemplo, fenômenos associados a frentes, brisa marítima, tempestades severas, entre outras. No modelo Eta aqui considerado, as previsões cobrem a maior parte da América do sul.

1.3.1 Coordenada Vertical: eta (η)

O nome do modelo Eta deriva da letra grega denominada eta (η), esse é o nome da coordenada vertical η e uma das principais características deste modelo. A coordenada vertical η é definida como (MESINGER, 1984)

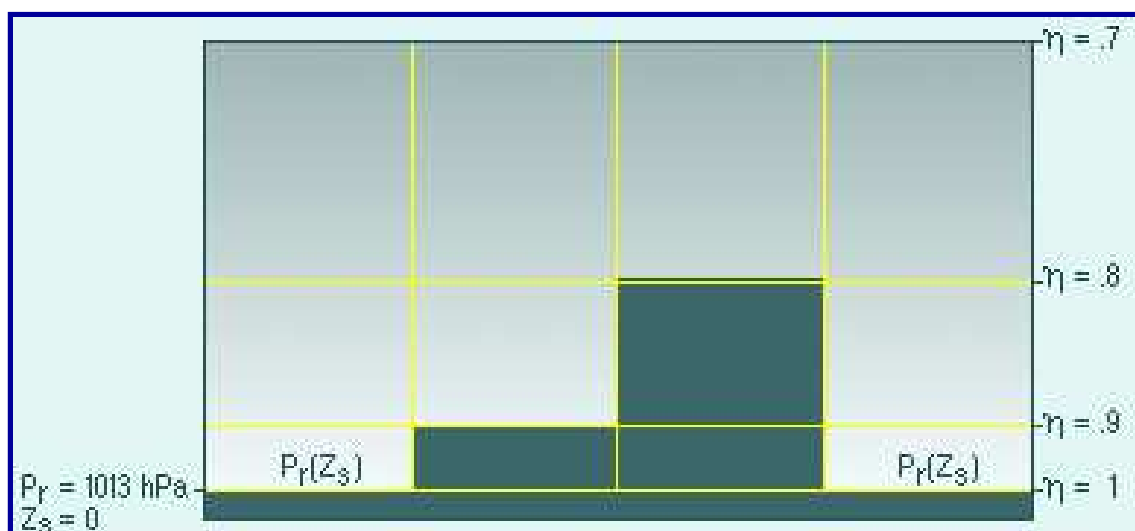
$$\eta = \frac{p - p_t}{p_s - p_t} \left[\frac{p_{ref}(z_s) - p_t}{p_{ref}(0) - p_t} \right], \quad (2)$$

na qual p é a pressão atmosférica em hPa (hectopascal), p_s é a pressão ao nível médio do mar igual a 1013,25 hPa, p_t é a pressão no topo do modelo, que se encontra a 50 hPa, p_{ref} é uma pressão de referência adequadamente definida e z_s é a altura da superfície em m. Tal coordenada foi criada com o intuito de reduzir os erros decorrentes do cálculo da força do gradiente de pressão usado em modelos que utilizam a coordenada sigma (BLACK, 1994).

A Figura 6 apresenta a divisão de uma montanha utilizando a variável vertical η . Percebe-se que a espessura entre as superfícies (linhas amarelas) aumentam com o aumento da altura ao longo da vertical.

Uma das vantagens da coordenada η é que suas superfícies são aproximadamente horizontais, o que reduz os erros nos cálculos de variáveis obtidas a partir de derivadas horizontais, fazendo com que o erro seja reduzido em regiões montanhosas. Outra vantagem está no fato de que modelo eta não necessita realizar interpolações verticais para calcular a FGP (Força do Gradiente de Pressão), o que é necessário em modelos sigma. Por fim, com o erro reduzido no cálculo da FGP, há melhora na previsão dos ventos, temperatura e umidade em áreas de terreno muito inclinado.

Figura 6 – Variável vertical Eta representada pelas linhas amarelas na horizontal



Neste modelo, o valor do vento é fixado em zero nas laterais dos degraus das montanhas. Dessa forma, um vale com extensão de dois espaçamentos de grade tem apenas um ponto de previsão de vento.

1.3.2 Variáveis Prognósticas

As variáveis prognósticas do modelo Eta são: temperatura do ar, componentes zonal e meridional do vento, umidade específica, pressão à superfície, energia cinética turbulenta, umidade e temperatura do solo e água líquida ou gelo das nuvens. O Eta utiliza a abordagem de divisão explícita que produz previsão baseada nessas variáveis. Após cada processo ser calculado, cada umas das variáveis são atualizadas e a integração prossegue. A precipitação convectiva é produzida pelo esquema Betts-Miller-Janjic (JANJIC, 1994). Os processos de superfície são resolvidos pelo esquema em. O pacote de radiação foi desenvolvido pelo *Geophysical Fluid Dynamics Laboratory*. O esquema *Lacis e Hansen* (1974) trata a radiação de onda longa, enquanto a radiação de onda curta é tratada pelo esquema de *Fels e Scharzkopf* (1975) (FELS; SCHWARZKOPF, 1975).

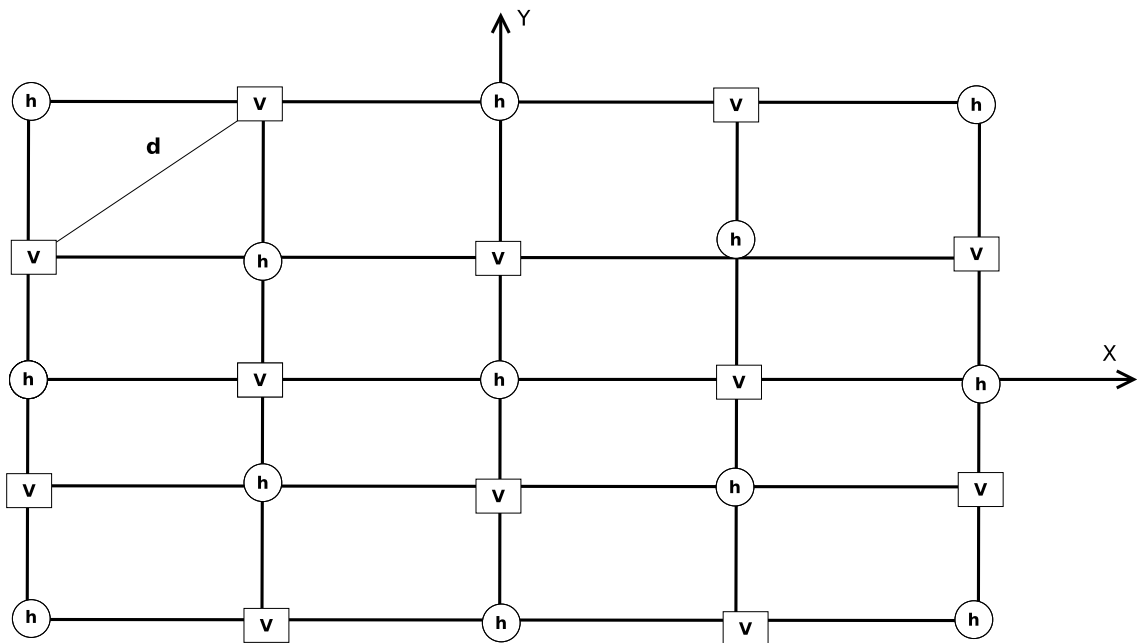
As previsões de velocidade de vento do modelo Eta apresentam limitações numéricas e truncamento. Por exemplo, ao utilizar células horizontais de 40 por 40 km, dentro da área correspondente a cada célula pode haver variações de vegetação, altitude ou áreas urbanas que não são percebidas na resolução espacial específica.

O modelo Eta aqui estudado, apresenta resultados de previsão quatro vezes ao dia, às 0, 06, 12 e 18 UTC. O período de previsão se estende por dias, meses, décadas e até séculos.

1.3.3 Grade horizontal

O modelo Eta estudado, apresenta resolução espacial de 40 km. A base para a estrutura horizontal do modelo Eta é a Grade “E” de Arakawa (MESINGER; ARAKAWA, 1976), a Figura 7 mostra essa grade horizontal. Nos pontos denominados **h** são definidas as variáveis de pressão, temperatura e umidade específica. Nos pontos **v** são definidas as componentes horizontais do vento. O comprimento **d** indica o espaçamento entre os dois pontos **h** e **v**. A magnitude desse comprimento é geralmente usada para definir a resolução horizontal do modelo, no caso, $d = 40$ km. A grade “E” é regular em coordenadas esféricas e o ponto de interseção entre o meridiano de zero grau e o equador é transferido para o centro do domínio do modelo, para minimizar a convergência entre os paralelos e meridianos na área central do domínio do modelo.

Figura 7 – Grade E de Arakawa



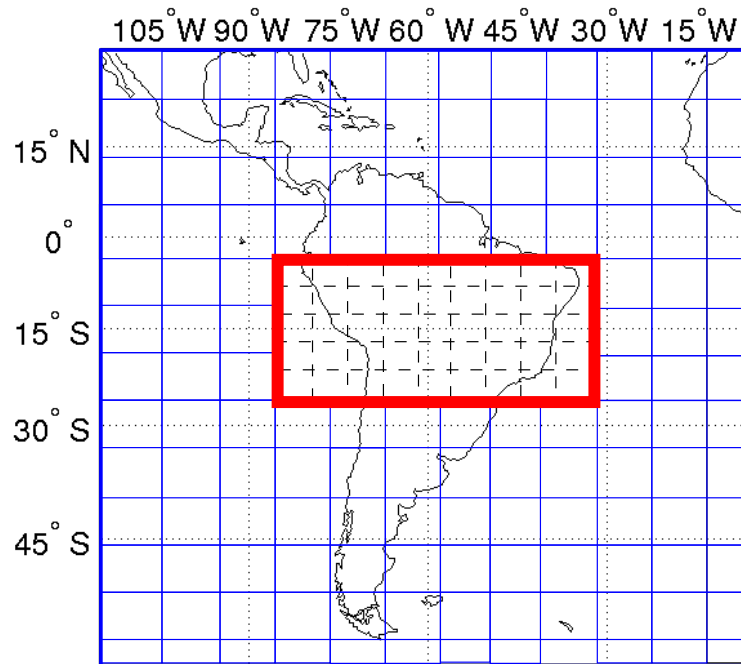
1.3.4 Condições Iniciais e Condições de contorno

O modelo regional atmosférico Eta utiliza como condição inicial os dados de análise do modelo atmosférico GFS (*Global Forecast System*), fornecidos pelo NCEP e como condições de contorno, as previsões do modelo global de circulação geral atmosférico (MCGA) do CPTEC (BUSTAMANTE et al., 2005). Esta condição inicial é obtida por meio de uma análise estática, onde a estimativa inicial é ajustada de acordo com as observações do horário de análise. Nas bordas laterais, o modelo é atualizado a cada 6 horas com as previsões do modelo global.

A Figura 8 mostra em linha tracejada as grades do modelo de maior resolução espacial (Eta de 40 km); com linhas contínuas, as grades do modelo global GFS de menor

resolução e a região sombreada as condições de contorno lateral provenientes do MCGA do CPTEC.

Figura 8 – Resolução espacial e condições de contorno, em linha tracejada estão as grades do modelo de maior resolução espacial (Eta de 40 km); com linhas contínuas, as grades do modelo global GFS de menor resolução e a região sombreada as condições de contorno lateral provenientes do MCGA do CPTEC



1.4 Reanálises do NCEP

Como já mencionado, além do Modelo GFS, o NCEP fornece um outro conjunto de dados que representam com maior fidelidade as condições climáticas observadas. Esses dados são conhecidos como reanálises do NCEP/CFSR (*Climate Forecast System Reanalysis*). As reanálises do NCEP são uma série de dados meteorológicos obtidos através da assimilação e reanálise de dados observados em todo o planeta (SAHA et al., 2010).

Em geral, os dados observacionais provêm de diversas fontes, entre estas estão: radiossonda, observações de superfície através de equipamentos e sensores, micro-ondas, imagens de satélite, entre outros. Os dados do CFSR iniciam em 1948 e por se tratar de períodos de dados razoavelmente longos com inclusão máxima de observações do planeta estes dados são importantes para o estudo de condições gerais da atmosfera em áreas isoladas. Estes dados também representam com uma maior fidelidade as condições climáticas atmosféricas, servindo como parâmetro de comparação de desempenho de outros modelos atmosféricos.

Como mencionado durante o processo de inserção das condições iniciais e de contorno no modelo atmosférico Eta, erros são gerados e crescem ao longo da integração

do modelo. Há também outras fontes de erros de modelagem a serem considerados, que contribuem para o erro de previsão existente no modelo Eta. Dessa forma, como os dados de reanálise do NCEP são próximos das observações e largamente utilizados como dados de referência eles podem ser utilizados para auxiliar no ajuste dos dados do modelo Eta através dos filtros adaptados.

2 ADAPTAÇÃO DE FILTROS PARA A CORREÇÃO DE DESVIOS DE PREVISÃO

Este capítulo trata dos conceitos dos filtros digitais e de filtros adaptativos. Inicialmente as principais características de filtros digitais e suas principais divisões são apresentadas. A seguir é apresentado a filtragem adaptativa, suas características e implementações. Por fim os algoritmos adaptativos são abordados, bem como suas diferenças e suas vantagens e desvantagens.

2.1 Filtros Digitais

Um filtro digital processa sinais digitais de forma a transformar o sinal de entrada em um outro sinal de saída. O filtro pode ser caracterizado em função de sua resposta $y(n)$ quando se aplica na entrada do filtro um sinal do tipo impulso. Quando a resposta ao impulso apresenta duração finita o filtro é chamado de filtro FIR (*Finite Impulse Response*) e quando a resposta ao impulso apresenta duração infinita o filtro é chamado de filtro IIR (*Infinite Impulse Response*).

Nos filtros FIR a saída depende somente da entrada presente e de um passado finito da entrada. Pelo fato dos valores passados da saída não influenciarem no cálculo dos valores de saída presentes, este filtro também é chamado de filtro não-recursivo.

Na maioria das aplicações, os sistemas no tempo discreto podem ser escritos por equações de diferenças. A entrada e a saída de um sistema descrito por uma equação de diferenças linear se relacionam genericamente por (DINIZ; SILVA; NETTO, 2004)

$$\sum_{i=0}^N a_i y(n-i) - \sum_{l=0}^M b_l x(n-l) = 0, \quad (3)$$

onde $y(n)$ é a saída do filtro, $x(n)$ é o sinal de entrada, sendo a_i e b_l os coeficientes do filtro, N é o número de coeficientes da saída, e M é o número de coeficientes do sinal de entrada. A equação 3 pode ser reescrita, sem perda de generalidade, considerando que $a_0 = 1$, assim

$$y(n) = - \sum_{i=1}^N a_i y(n-i) + \sum_{l=0}^M b_l x(n-l), \quad (4)$$

quando $a_1 = a_2 = a_3 = \dots = a_N = 0$, então a amostra n da saída depende somente dos valores do sinal de entrada e os coeficientes b_l se relacionam diretamente com a resposta ao impulso do sistema, isto é, $b_l = h(l)$. Assim, o sistema é chamado de não-recursivo ou filtro FIR. Estes filtros são caracterizados por uma equação de diferenças dada por (DINIZ; SILVA; NETTO, 2004; SILVA et al., 2012)

$$y(n) = \sum_{l=1}^M h(l)x(n-l), \quad (5)$$

aplicando-se a transformada Z à Equação 5, o que permite obter a função de transferência a partir da resposta ao impulso $h(l)$, que é dada por

$$H(z) = \sum_{l=0}^M h(l)z^{-l}. \quad (6)$$

No caso de filtros IIR a saída depende dos sinais de entrada presentes e passados, e também dos sinais de saídas passados. Pelo fato dos sinais de saída passados influenciarem os sinais de saída atuais, esses filtros são chamados de filtros recursivos (DINIZ; SILVA; NETTO, 2004). A implementação de um filtro IIR em geral é feita pela transformação de um filtro analógico em um filtro digital, com as mesmas especificações. Utiliza-se essa abordagem pois o projeto de filtros analógicos está estabelecido e excelentes resultados são assim alcançados. Outro motivo é que métodos de projetos de filtros analógicos têm fórmulas relativamente simples. Desta maneira, métodos de projetos de filtros digitais baseados nesses são também simples de obter (HAYKIN; VEEN, 2001). Estes filtros são caracterizados por uma equação de diferenças dada pela Equação 4 e através da transformada Z sobre esta equação obtém-se a função de transferência do filtro IIR dada por

$$H(z) = \frac{\sum_{i=0}^M b_i z^{-i}}{1 + \sum_{i=1}^N a_i z^{-i}}. \quad (7)$$

Filtros FIR implementam respostas desejadas para o módulo das componentes espectrais com fase linear, ou seja, sem nenhuma distorção de fase no sinal filtrado. Essa característica é importante em algumas aplicações como processamento de imagem, biomedicina e alguns esquemas de transmissão de dados (DINIZ; SILVA; NETTO, 2004; SILVA et al., 2012). Filtros FIR são sempre BIBO (*bounded input-bounded output*) estáveis, ou seja, qualquer entrada limitada em amplitude resulta em uma saída limitada em amplitude, e com alguma sensibilidade a erros de quantização. Já filtros IIR necessitam de menos coeficientes que os filtros FIR para atender a uma mesma especificação de projeto, contudo essa melhoria é obtida em detrimento da distorção de fase. Além disso podem não apresentar sempre a BIBO estabilidade e apresentam maior sensibilidade a quantização de coeficientes (DINIZ; SILVA; NETTO, 2004; SILVA et al., 2012).

Dessa forma, para melhorar a previsão sazonal do Eta 40km relativamente às reanálises do NCEP, aplica-se um filtro cujos coeficientes são ajustados por meio da abordagem conhecida como filtragem adaptativa. Esta abordagem é escolhida pela sua aplicabilidade em diversas áreas envolvendo otimização e problemas de identificação de sistemas e equalização (DINIZ, 2008).

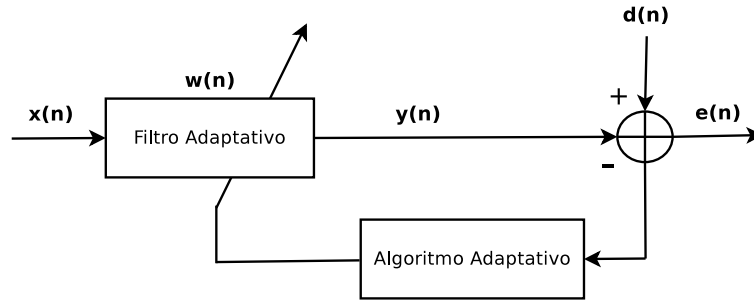
Neste trabalho usamos filtros FIR, pois a princípio a quantidade de coeficientes é pequena e a distorção de fase é indesejada (SILVA et al., 2012; HAYKIN; VEEN, 2001).

2.2 Filtragem Adaptativa

A utilização de filtros adaptativos compreende dois estágios. O primeiro é de regime transitório e outro é de regime permanente. No regime transitório o filtro adaptativo é iniciado com um conjunto qualquer de coeficientes, que vão sendo adaptados para uma solução em função de uma critério de otimização. Após esse estágio, os coeficientes do filtro podem ser usados para aplicar o filtro em regime permanente, ou seja os coeficientes otimizados são mantidos (KUHN, 2012).

Geralmente, a adaptação requer a existência de um sinal de referência que rege a adaptação dos coeficientes do filtro. A Figura 9 mostra a configuração geral empregada para adaptar filtros (DINIZ, 2008), em que $x(n)$ é o sinal de entrada, $y(n)$ é a saída do filtro adaptativo, $w(n)$ é o filtro adaptado, e $d(n)$ é o sinal de referência ou desejado. O sinal de erro $e(n)$ é definido como $d(n) - y(n)$.

Figura 9 – Configuração geral do filtro adaptativo



Um algoritmo adaptativo é usado para ajustar os coeficientes do filtro de modo que o sinal de erro $e(n)$ seja minimizado (DINIZ, 2008). Assim, num instante n , uma amostra do sinal de entrada $x(n)$ é submetida ao filtro adaptado gerando a amostra do sinal de saída $y(n)$. Esta é comparada a amostra correspondente do sinal de referência $d(n)$ gerando a amostra do sinal de erro $e(n)$. O sinal do erro é utilizado no procedimento de adaptação dos coeficientes do filtro do instante n para o instante $n + 1$. Espera-se que no fim desse processo (de adaptação) a saída do filtro adaptativo torne-se a mais próxima possível da resposta desejada e a magnitude de $e(n)$ seja reduzida.

Em geral, se minimiza o erro médio quadrático (MSE - *Mean Square Error*) de $e(n)$ ao longo do processo de adaptação. O MSE é definido como

$$\varepsilon(n) = E[e^2(n)] = E[(d(n) - y(n))^2] = E[d^2(n) - 2d(n)y(n) + y^2(n)], \quad (8)$$

assim,

$$\varepsilon(n) = E[d^2(n)] - 2E[d(n)y(n)] + E[y^2(n)]. \quad (9)$$

A saída do filtro adaptativo é

$$y(n) = \sum_{i=0}^N w_i(n)x(n-i) = \mathbf{w}^T(n)\mathbf{x}(n), \quad (10)$$

onde $\mathbf{x}(n) = [x(n) \ x(n-1) \ x(n-2) \ \dots \ x(n-N)]^T = [x_0(n) \ x_1(n) \ \dots \ x_i(n)]$ é o sinal de entrada e $x_i(n) = x(n-i)$, e $\mathbf{w}(n) = [w_0(n) \ w_1(n) \ \dots \ w_N(n)]^T$ é o vetor de coeficientes do filtro adaptativo (DINIZ, 2008), T indica a transposição da matriz. Substituindo a Equação 10 na Equação 9 tem-se

$$\varepsilon(n) = E[d^2(n)] - 2E[d(n)\mathbf{w}^T(n)\mathbf{x}(n)] + E[\mathbf{w}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{w}(n)]. \quad (11)$$

Se os coeficientes de $\mathbf{w}^T(n)$ são fixos, isto é, o filtro é estático, $\mathbf{w}^T(n) = \mathbf{w}^T$, a função do MSE em ambiente estacionário é dada por

$$\varepsilon = E[d^2(n)] - 2\mathbf{w}^T E[d(n)\mathbf{x}(n)] + \mathbf{w}^T E[\mathbf{x}(n)\mathbf{x}^T(n)]\mathbf{w} \quad (12)$$

temos que $\mathbf{p} = E[d(n)\mathbf{x}(n)]$ é o vetor de correlação cruzada entre o sinal desejado e o sinal de entrada, e $\mathbf{R} = E[\mathbf{x}(n)\mathbf{x}^T(n)]$ é a matriz de correlação do sinal de entrada, assim teremos

$$\varepsilon = E[d^2(n)] - 2\mathbf{w}^T\mathbf{p} + \mathbf{w}^T\mathbf{R}\mathbf{w} \quad (13)$$

Para encontrarmos os valores dos coeficientes de $\mathbf{w}$, devemos encontrar o ponto de mínimo da função ε em função de $\mathbf{w}$. Para isso, calcula-se primeiro o vetor gradiente da função MSE, em relação $\mathbf{w}$

$$\mathbf{g}_{\mathbf{w}} = \frac{\partial \varepsilon}{\partial \mathbf{w}} = -2\mathbf{p} + 2\mathbf{R}\mathbf{w} \quad (14)$$

como queremos minimizar $\mathbf{g}_{\mathbf{w}}$, fazemos $\mathbf{g}_{\mathbf{w}} = \mathbf{0}$, e assim

$$\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p}. \quad (15)$$

Essa solução é chamada solução de Wiener (DINIZ, 2008; HAYKIN, 1995). Percebe-se a partir da Equação 15 que a solução de Wiener requer o conhecimento das características estatísticas dos sinais envolvidos. Na prática, realizar uma estimação precisa de $\mathbf{R}$ e $\mathbf{p}$ é muito difícil (DINIZ, 2008). Além disso a equação (15) requer o cálculo da inversa de uma matriz cuja solução na maioria das vezes não pode ser obtida analiticamente, o que gera elevada complexidade computacional na solução. Dessa forma, visando contornar esses problemas e encontrar uma solução viável do ponto de vista de implementação prática, é

necessário encontrar soluções alternativas. Neste contexto os algoritmos adaptativos visam contornar o problema da solução de Wiener.

Um aspecto importante sobre filtragem adaptativa está na escolha do algoritmo de adaptação, pois ele influencia uma série de aspectos, como por exemplo a complexidade computacional e a velocidade de convergência. Dessa forma, tendo em vista que há uma vasta gama de estruturas de filtragem e algoritmos adaptativos presentes na literatura, a escolha do algoritmo é determinada por um ou mais fatores, como velocidade de convergência, desajuste, *tracking* ou rastreamento, robustez, requisitos computacionais, estrutura e propriedades numéricas (HAYKIN, 1995; KUHN, 2012).

A velocidade de convergência pode ser definida como o número de iterações necessárias para que o algoritmo, em resposta a sinais estacionários, convirja para a solução ótima de Wiener, no sentido de minimizar o erro médio quadrático. Uma alta velocidade de convergência permite ao filtro adaptativo se ajustar rapidamente a um ambiente com características estatísticas desconhecidas.

O desajuste é um parâmetro adimensional que fornece uma medida quantitativa da diferença entre o erro médio quadrático residual (considerando a aplicação do algoritmo de adaptação) e o erro médio quadrático mínimo, que é produzido pelo filtro de Wiener.

O *tracking* ou rastreamento, trata da capacidade do algoritmo em acompanhar as variações das características estatísticas dos sinais em um meio não-estacionário. O desempenho deste rastreamento é influenciado pela taxa de convergência e pelas flutuações em regime permanente

Para que um filtro adaptativo seja robusto, pequenos distúrbios não podem resultar em erros elevados. Por sua vez, os requisitos computacionais estão ligados à quantidade de operações (adições, multiplicações, divisões e etc) requeridas para realizar uma iteração do algoritmo adaptativo.

A estrutura se refere à maneira como a informação flui no algoritmo. Neste quesito, certas estruturas apresentam vantagens devido, principalmente, a suas formas modulares, e implementações e robustez numérica (HAYKIN, 1995; KUHN, 2012).

Podem ocorrer erros de quantização quando filtros são implementados. As propriedades numéricas estão relacionadas com a forma como o algoritmo reage ao erros de quantização. Se esses erros se acumulam e levam a um comportamento anômalo do algoritmo ele é dito numericamente instável, entretanto se o algoritmo é pouco sensível às variações no comprimento da palavra (quantidade de bits) utilizada em sua implementação digital, então o algoritmo é dito numericamente estável.

Neste contexto, os algoritmos utilizados para realizar a adaptação dos filtros foram o NLMS (*Normalized Least Mean Square*) e o RLS (*Recursive Least Square*). Ambos foram escolhidos por sua simplicidade computacional e pela sua larga aplicação em inúmeras pesquisas (HAYKIN, 1995; KUHN, 2012; DINIZ; SILVA; NETTO, 2004).

2.2.1 Algoritmos LMS e NLMS

O algoritmo LMS (*Least Mean Square*) é um dos algoritmos mais utilizados em filtragem adaptativa devido a sua baixa complexidade computacional, garantia de convergência em ambientes estacionários e por apresentar comportamento estável quando implementado com aritmética de precisão finita. O algoritmo LMS tem por objetivo minimizar o erro médio quadrático entre as amostras do sinal de entrada e o sinal de referência. Assim, a equação de atualização no LMS é dada por

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 2\mu e(n)\mathbf{x}(n) \quad (16)$$

onde μ é o passo que é escolhido de forma a garantir a convergência.

Entretanto, o LMS apresenta baixa velocidade de convergência e é sensível à potência do sinal de entrada. Para contornar estas desvantagens e não perder a baixa complexidade computacional, decidiu-se aqui utilizar o algoritmo NLMS (*Normalized Least Mean Square*) para realizar a adaptação dos filtros (DINIZ; SILVA; NETTO, 2004; HAYKIN, 1995). Como o LMS, o NLMS também visa minimizar o erro médio quadrático entre o sinal de entrada e o sinal de referência. O NLMS apresenta tamanho de passo variável, possibilitando assim uma convergência mais rápida em comparação ao LMS. O NLMS é imune a alterações da potência do sinal de entrada, pois realiza a normalização do sinal. Para tanto, o algoritmo NLMS emprega a equação de atualização dos coeficientes,

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu_n}{\gamma + \mathbf{x}^T(n)\mathbf{x}(n)} e(n)\mathbf{x}(n), \quad (17)$$

onde $e(n) = d(n) - \mathbf{x}^T(n)\mathbf{w}(n)$, μ_n é o passo, γ é uma constante pequena introduzida para impedir que o tamanho do passo seja muito grande quando $\mathbf{x}^T(n)\mathbf{x}(n)$ for muito pequeno e também para evitar divisão por zero. De modo a garantir a estabilidade, o fator de convergência deve ser escolhido no intervalo $0 < \mu_n < 2$ (DINIZ, 2008). Assim o algoritmo NLMS atualiza os coeficientes do filtro de modo a minimizar erro médio quadrático dado por $\varepsilon(n) = E[e^2(n)]$.

2.2.2 Algoritmo RLS

O RLS é um algoritmo de estimação paramétrica que procura minimizar o erro quadrático na saída do filtro adaptativo. Os coeficientes do filtro são atualizados em função do conjunto de valores calculados na iteração anterior. Este algoritmo possui alta velocidade de convergência quando os autovalores da matriz de correlação do sinal de entrada estão espalhados (distantes do seu valor esperado). O algoritmo RLS também apresenta bom desempenho para sistemas sujeitos a rápidas variações na amplitude do sinal (DINIZ, 2008). No algoritmo RLS, a função objetivo $\xi^d(n)$ a ser minimizada é dada

por

$$\xi^d(n) = \sum_{i=0}^k \lambda^{k-i} \varepsilon^2(i), \quad (18)$$

$\varepsilon(i)$ é o erro de saída a *posteriori* no instante i . Este erro é computado depois que o vetor de coeficiente é atualizado e utilizado na função do custo para o novo vetor de dados, qual ele,

$$\varepsilon(n) = E[(d(n) - y(n))^2], \quad (19)$$

e assim teremos

$$\xi^d(n) = \sum_{i=0}^n \lambda^{n-i} [d(i) - \mathbf{x}^T(i) \mathbf{w}(n)]^2, \quad (20)$$

a variável λ é a “memória” do algoritmo, ou constante de esquecimento, sendo um caso especial de $\lambda = 1$, um filtro de memória infinita, assim o valor desta constante deve estar entre $0 \leq \lambda \leq 1$. No desenvolvimento do LMS e do NLMS utilizou-se o erro a priori $e(n)$. No algoritmo RLS o erro a *posteriori* $\varepsilon(n)$ é utilizado. Para obter a atualização, como se pretende minimizar o erro a *posteriori*, deriva-se em função de $\mathbf{w}(n)$ e iguala-se a zero a Equação 20, com que se obtém

$$\sum_{i=0}^n \lambda^{n-i} \mathbf{x}(i) \mathbf{x}^T(i) \mathbf{w}(n) = \sum_{i=0}^n \lambda^{n-i} \mathbf{x}(i) d(i). \quad (21)$$

Assim temos a equação de atualização dos coeficientes no algoritmo RLS,

$$\mathbf{w}(n) = \mathbf{R}_D^{-1}(n) \mathbf{p}_D(n), \quad (22)$$

na qual

$$\mathbf{R}_D^{-1}(n) = \left[\sum_{i=0}^n \lambda^{n-i} \mathbf{x}(i) \mathbf{x}^T(i) \right] \quad (23)$$

e

$$\mathbf{p}_D(n) = \sum_{i=0}^n \lambda^{n-i} \mathbf{x}(i) d(i). \quad (24)$$

$\mathbf{R}_D$ é a matriz de correlação do sinal de entrada e $\mathbf{p}_D(n)$ é a matriz de correlação cruzada determinística entre o vetor de entrada e o sinal desejado. Assume-se que o sinal de entrada é zero para $n < 0$ e então $\mathbf{R}_D$ será sempre singular, ou seja, seu determinante é sempre zero. O custo computacional para obter $\mathbf{R}_D^{-1}$ é muito elevado, desta forma utiliza-se o lema da inversão das matrizes para simplificar os cálculos (DINIZ, 2008), o que fornece

$$\mathbf{S}_D(n) = \mathbf{R}^{-1}_D = \frac{1}{\lambda} \left[\mathbf{S}_D(n-1) - \frac{\mathbf{S}_D(n-1)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{S}_D(n-1)}{\lambda + \mathbf{S}_D(n-1)\mathbf{x}^T(n)\mathbf{x}(n)} \right] \quad (25)$$

A Tabela 1 compara as complexidades computacionais dos algoritmos NLMS e RLS. A partir da Tabela 1 percebe-se que o algoritmo NLMS requer menos operações de multiplicação e de adição que o algoritmo RLS, tornando este último computacionalmente mais custoso. Entretanto a maior complexidade do RLS não é proibitiva.

Tabela 1 – Complexidade computacional dos algoritmos NLMS e RLS.

Complexidade computacional	NLMS	RLS
Multiplicações	$3N + 1$	$2N^2 + 11N + 8$
Adições	$3N$	$3N^2 + 7N + 4$
Divisões	1	1

Dessa forma tanto o algoritmo NLMS e RLS são aplicados na redução de erros na previsão atmosférica. Primeiramente os filtros devem ser treinados, as séries atmosféricas provenientes do modelo Eta são utilizadas para realizar o treinamento do filtro, isto é, o modelo Eta é o sinal de entrada $\mathbf{x}(n)$. Para tanto, o sinal de referência dado por $\mathbf{d}(n)$ representa as séries provenientes dos dados de reanálise do NCEP. Assim o erro médio quadrático $\varepsilon(n)$ é a função objetivo deste sistema aqui utilizado.

Neste trabalho filtro adaptativo corresponde ao sistema adaptativo que realiza a adaptação dos coeficientes de um determinado filtro, e define-se filtro adaptado como um filtro que teve seus coeficientes adaptados por meio de um sistema adaptativo.

2.3 Métricas de desempenho

Algumas métricas de desempenho foram utilizadas para avaliar os resultados encontrados na aplicação dos filtros adaptados sobre as séries climáticas do modelo atmosférico Eta. Para avaliarmos a qualidade dos resultados encontrados é necessário definir parâmetros quantitativos de desempenho. As métricas de desempenho utilizadas para validar os resultados são: o erro quadrático médio normalizado (EQMn) e a variância (var) do EQMn.

2.3.1 Erro quadrático médio normalizado

O erro quadrático médio normalizado (EQMn) entre o sinal desejado (reanálises do NCEP) e os sinais do modelo Eta (adaptado ou não) avalia se o filtro encontrado resultou em uma melhor predição. O erro quadrático médio normalizado é dado por

$$\text{EQMn} = \frac{1}{N} \frac{\sum_{n=1}^N [\mathbf{d}(n) - \mathbf{x}(n)]^2}{\sum_{n=1}^N [\mathbf{d}(n)]^2}, \quad (26)$$

na qual $\mathbf{x}(n)$ é a série utilizada e $\mathbf{d}(n)$ a série de referência, N é a quantidade de pontos da série.

Como discutimos anteriormente o objetivo principal do filtro adaptado é diminuir o erro médio quadrático. Assim, o objetivo é reduzir o erro do ano de interesse ou ano de ajuste, quando comparado aos dados de reanálise do NCEP. Assim, devemos verificar se de fato o filtro adaptado obtém melhor previsão do ano de ajuste. Um possível critério para isso é avaliar a quantidade de coordenadas (pontos de grade) nas quais o erro foi reduzido após a aplicação do filtro adaptado.

2.3.2 Variância

A variância é uma medida de dispersão dos valores de uma amostra em relação à sua média, dada por

$$\text{var}(\text{EQMn}(n)) = \frac{1}{T} \sum_{n=1}^T (\text{EQMn}(n) - \mu_e)^2, \quad (27)$$

na qual T é o tamanho da série e μ_e é a média de EQMn . Assim, calcula-se a variância para verificar se houve redução na dispersão do erro de previsão existente entre os dados de previsão climáticos do Eta e as reanálises do NCEP.

2.3.3 Intervalo de Confiança

De modo a verificar o nível de variação dos resultados encontrados nos testes utilizou-se o intervalo de confiança. Ele também funciona como parâmetro de quão estável são os resultados encontrados, que é a medida de quão perto os resultados estarão do resultado original, se os testes forem refeitos. O intervalo de confiança da média de n realizações pode ser definido como

$$IC = m_e \pm t_{\alpha/2} \frac{S}{\sqrt{c}}, \quad (28)$$

onde m_e é a média de EQMn ; t é um valor obtido através de uma tabela de quantis da distribuição, chamado valor crítico; α o intervalo de confiança escolhido, que neste trabalho é 95%; S é o desvio das amostras; e c é a quantidade de amostras, que neste caso é dez.

Assim estas métricas de desempenho são aplicadas nos testes realizados a partir da metodologia escolhida. A metodologia diz respeito ao método utilizado na aplicação do filtro adaptativo.

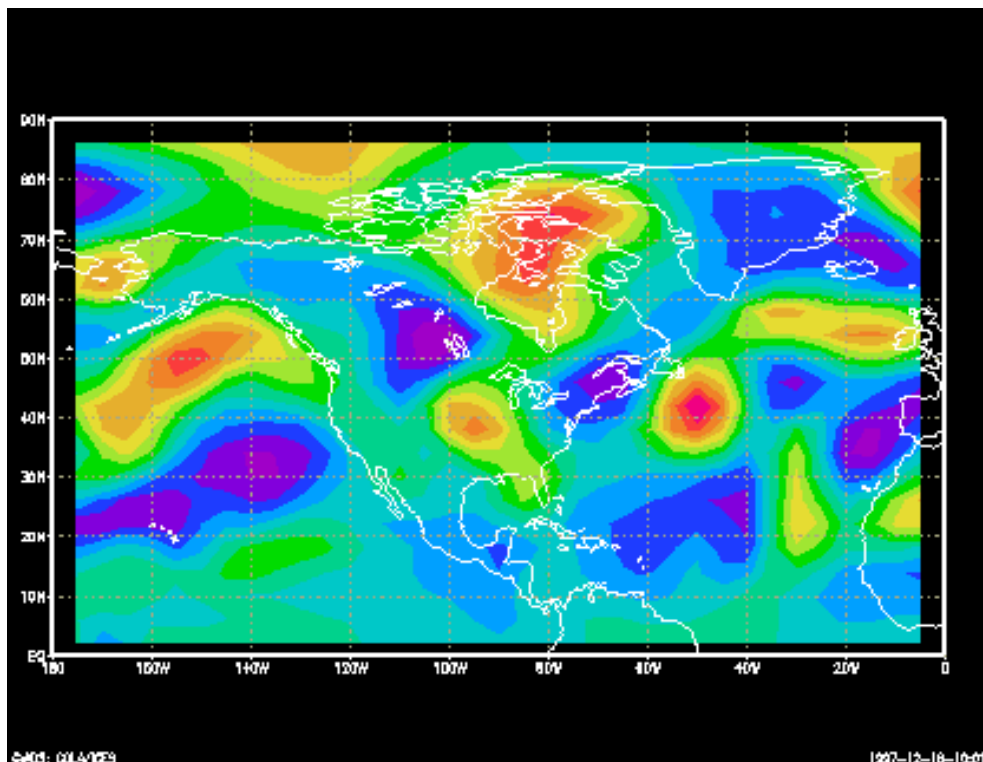
3 METODOLOGIA

Os dados de previsão atmosférica provenientes do modelo Eta são disponibilizados pelo CPTEC/INPE e as reanálises do NCEP estão disponíveis em (WORLEY, 2015). Tanto os dados provenientes do modelo Eta quanto os dados de reanálise do NCEP estão em formato de “Ponto de Grade”. Uma forma de acessar esses dados é utilizar um *software* denominado GrADS (Sistema de Visualização e Análise de Dados em Pontos de Grade) (COLA, 2015a).

O GrADS é um software utilizado no acesso, manipulação e visualização de dados geofísicos. Atualmente, o GrADS é o *software* mais utilizado nos centros operacionais e de pesquisa meteorológica espalhados pelo mundo, inclusive no Brasil (COLA, 2015a). Dados no formato BINÁRIO, GRIB, NetCDF ou HDF-SDS, são compatíveis com o GrADS. As variáveis podem possuir 4 dimensões (longitude, latitude, níveis verticais e tempo). No GrADS, os dados dispostos em forma de matriz podem conter uma ou mais variáveis dispostas sobre uma grade regular, ou não linear, ou gaussiana, ou em pontos de estações ou de resolução variável (COLA, 2015a). Pode-se visualizar essas variáveis por meio de vários tipos de gráficos que podem ser salvos em vários formatos de imagens gráficas.

A Figura 10 mostra um exemplo da vorticidade (rotação) do vento visualizada no GrADS (COLA, 2015b).

Figura 10 – Vorticidade (rotação) do vento visualizada no GrADS



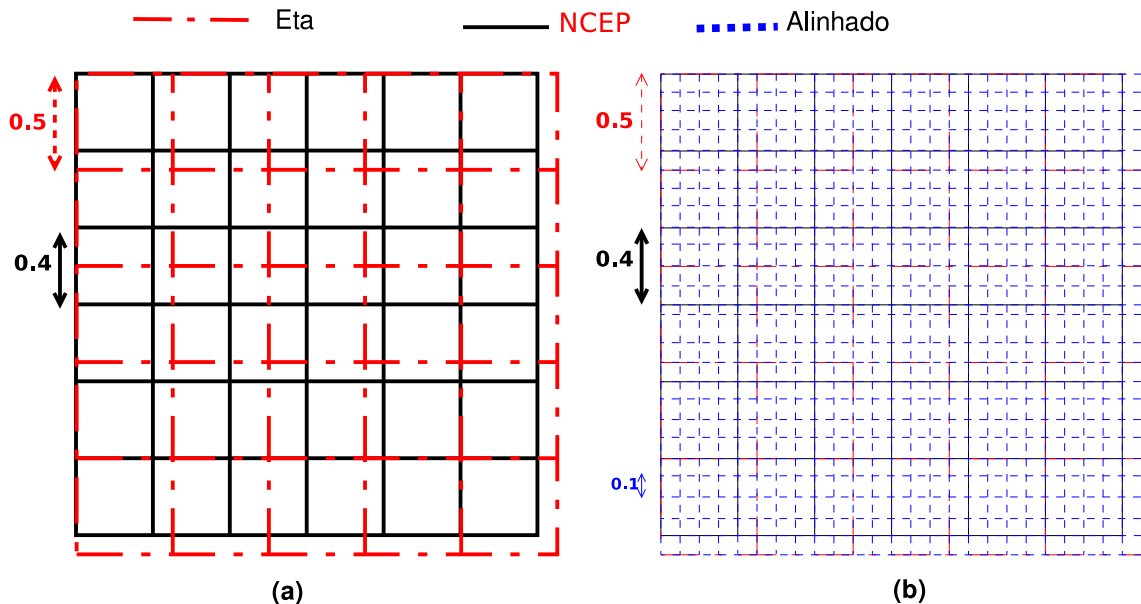
Por meio da interface programável do GrADS podem-se desenvolver diversas análises, calcular variáveis atmosféricas e aplicar visualizações automáticas. *Scripts*

GrADS podem ser criados para desempenhar inúmeras funções. Assim, criaram-se *scripts* para converter os dados fornecidos pelo modelo Eta e as reanálises do NCEP para um arquivo em formato ASCII (*American Standard Code for Information Interchange*). Este arquivo é lido através do *software* Matlab, no qual são realizadas todas as análises e testes desta dissertação.

Devido a diferenças entre as grades e resolução espacial (latitude e longitude) do Eta e do NCEP, faz-se necessário o ajuste dos dados para uma grade comum. O Eta possui $0,4^\circ$ de resolução, enquanto que o NCEP possui $0,5^\circ$. Assim, para a adequação das grades de resolução espacial, ambas são redefinidas com $0,1^\circ$, gerando um aumento no número de pontos.

As resoluções em latitude e longitude correspondem a comprimentos de arco. Digitalmente esses arcos em latitude e longitude correspondem a elementos de matrizes. As previsões para uma dada altura no Eta é uma matriz 61×126 e no NCEP tem-se uma matriz de 49×101 . Após o ajuste decimal das grades, as duas matrizes apresentam dimensões de 244×504 , ou seja, 122976 elementos no total cada. Este ajuste é feito por meio da realocação dos índices das matrizes do Eta. Os índices das linhas e colunas do Eta são multiplicados por 4, alterando-os para 244×504 . Assim, os valores para cada linha e coluna do modelo Eta são replicados por quatro enquanto que para as reanálises do NCEP eles são replicados cinco vezes. A Figura 11 (a) mostra a resolução espacial inicial e a Figura 11 (b) a resolução espacial final do modelo Eta e do NCEP.

Figura 11 – Resolução inicial (a) e resolução final (b) do modelo Eta e do NCEP



O aumento da resolução provocado pelo alinhamento requer a aplicação de um filtro espacial passa baixa de modo a eliminar as possíveis variações bruscas dos valores entre as posições geográficas. A Figura 12 mostra a resposta em frequência do filtro passa baixas bidimensional utilizado neste trabalho e que é aplicado no domínio da frequência,

com ganho em dB. Na Figura 13 o mesmo filtro passa baixas em função da frequência.

O filtro passa baixas foi projetado utilizando o algoritmo *Weighted Chebyshev approximation* (CHEN, 2009) tendo como parâmetros, $\delta_p = 1$, $\delta_s = 40$, sendo δ_p a amplitude do *ripple* de banda de passagem, δ_s a amplitude do *ripple* de rejeita-banda. As amplitudes na banda de passagem e na banda de rejeição são respectivamente 1 e zero. Assim, pelas Figuras 12 e 13 percebe-se que há ganho de zero *dB* em baixas frequências e a partir da frequência de corte normalizada de 0,4 a atenuação está na faixa de -40dB . A frequência de corte foi escolhida de modo a não causar perda severa das altas frequências do sinal e ao mesmo tempo permitir uma transição suave entre as linhas/colunas dos dados do Eta e do NCEP.

Figura 12 – Filtro Passa Baixas bidimensional utilizado

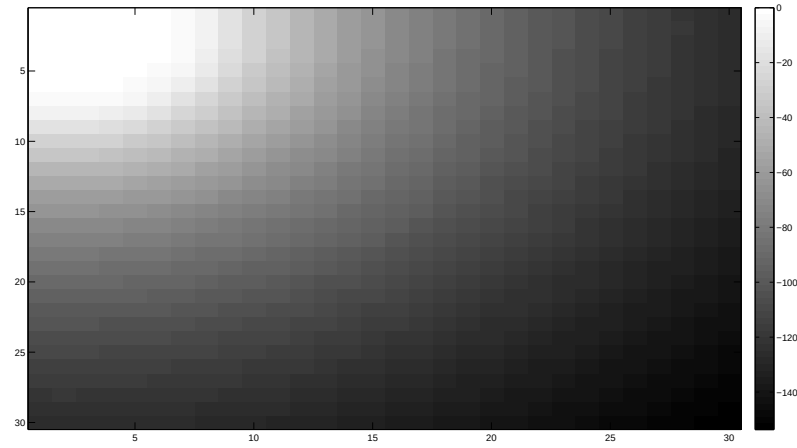
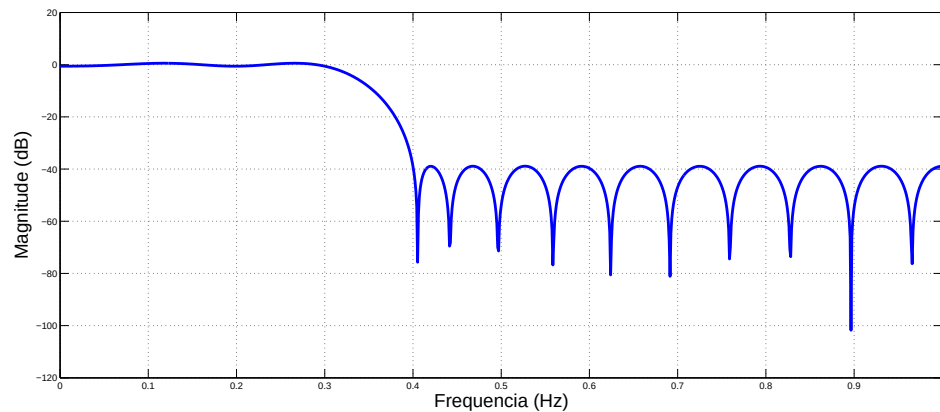


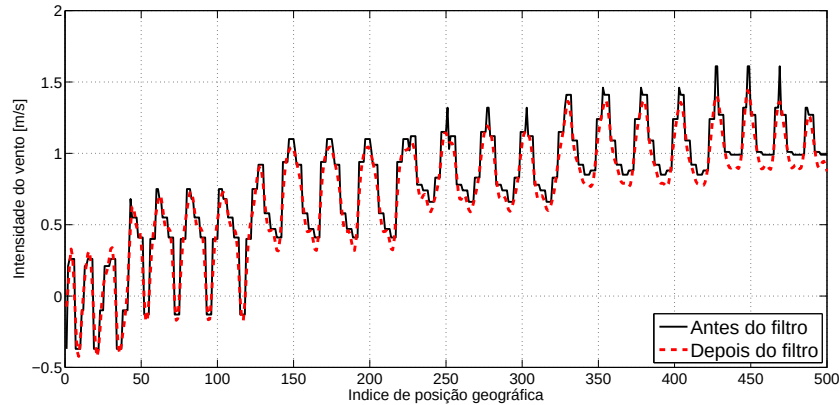
Figura 13 – Filtro Passa Baixas utilizado



Com o objetivo de ilustrar o efeito do filtro passa-baixas sobre os dados do Eta e NCEP, na Figura 14 a matriz de dados foi vetorizada e com isso aplicou-se o filtro

ao longo da horizontal. Assim, na Figura 14 tem-se em linha cheia o parâmetro meteorológico vento meridional ao longo das coordenadas geográficas, e em linha tracejada este mesmo parâmetro após a aplicação do filtro passa baixas. Percebe-se que este segundo está suavizado em relação à série antes do filtro passa baixas.

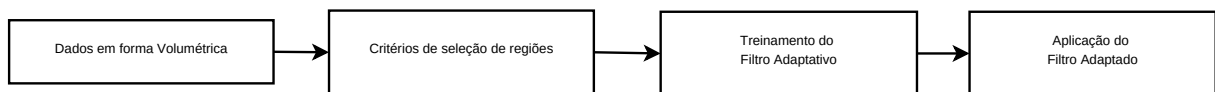
Figura 14 – Efeito da utilização do filtro passa baixas



3.1 Filtragem adaptativa das previsões

Os filtros adaptativos nesta dissertação realizarão o ajuste das séries ao longo do tempo de previsão. Assim, para melhorar o desempenho dos filtros adaptativos é importante que as séries apresentem características estatísticas semelhantes entre si. Com isso percebe-se a necessidade de agrupar pixels (em função da latitude e longitude), que representam as séries ao longo do tempo.

Figura 15 – Método utilizado para a implementação do filtro adaptativo



Assim, o diagrama de blocos da Figura 15 mostra o método utilizado para a aplicação do filtro adaptativo.

Inicialmente os dados estão dispostos volumetricamente, com os eixos horizontal (latitude) e vertical (longitude) representando as posições geográficas das séries climáticas ao longo do tempo (profundidade). Em seguida, são selecionadas as regiões que fornecem o conjunto de séries climáticas para realizar o treinamento do filtro. O tamanho dessas séries depende do horizonte de previsão, podendo ser sazonal ou diário. Na etapa seguinte as séries climáticas são utilizadas para adaptar os coeficientes do filtro. Os coeficientes do filtro são otimizados para cada série climática e assim o filtro mantém seus coeficientes atualizados para continuar a adaptação com a próxima série. Por fim o filtro encontrado poderá ser aplicado a séries de previsão do ano seguinte ao ano de treinamento.

3.1.1 Referência Temporal

O dados disponíveis para realizar o treinamento do filtro e o ajuste das séries, vão do ano de 2001 até o ano de 2010. O filtro encontrado após o seu treinamento, é aplicado no mesmo horizonte de previsão, na mesma região do ano seguinte ao ano de treinamento. Por exemplo, se queremos ajustar o ano de 2003, usamos o ano de 2002 para gerar o conjunto de treinamento. Tal abordagem supõe a hipótese que o ano anterior ao ano de ajuste é capaz de fornecer um bom conjunto de dados para o treinamento do filtro, supondo mudanças pequenas de um ano para o outro.

3.1.2 CrITÉRIOS de Seleção de Regiões de Treinamento/Ajuste

Inicialmente, utiliza-se toda a região para o treinamento do filtro e mesmo com um grande número de séries, os resultados não foram bons. Neste caso, o algoritmo não convergiu, ou seja, a aplicação do filtro adaptado às previsões não geraram séries com erro reduzido. Isso aconteceu pois algumas séries climáticas estão em pontos geográficos distantes uns dos outros e apresentam características estatísticas e climáticas distintas, comprometendo a capacidade da filtragem adaptativa de obter um filtro genérico capaz de reduzir o erro da sequência filtrada. Outro ponto negativo é a complexidade computacional requerida para obter filtros com essa montagem.

Percebeu-se assim a necessidade de agrupar as regiões que possuam alguma similaridade entre si. Dessa forma, utilizaram-se quatro critérios de seleção de regiões para agrupar os dados. Isso é feito em função de características semelhantes entre os elementos, ou no espaço de dados ou no espaço geográfico físico. Utilizaram-se para isso o *k-means*, regiões retangulares, regiões circulares e regiões anulares.

3.1.2.1 Seleção via *k-means* (primeiro critério)

O primeiro critério de seleção de regiões é a clusterização, feita pelo algoritmo *k-means* (HAN; KAMBER; PEI, 2012; SEMAAN et al., 2012). O *k-means* permite classificar as informações. Essa classificação é realizada por meio de comparações entre os valores numéricos dos dados, sem a necessidade de nenhuma supervisão humana. Para realizar a classificação, uma função objetivo avalia a qualidade do particionamento de modo que os objetos dentro de um grupo sejam semelhantes entre si, mas diferentes de objetos de outros grupo. Ou seja, a função objetivo visa uma alta similaridade dentro dos clusters e uma baixa similaridade entre os clusters (HAN; KAMBER; PEI, 2012). Assim, esse critério tem por objetivo separar as regiões em função do comportamento dos dados.

Seja um conjunto de dados $\mathcal{D}$, contendo q objetos. O *k-means* particiona os objetos em função de suas distâncias aos centróides dos clusters. O algoritmo *k-means* distribui

os objetos de $\mathcal{D}$ em k clusters, $C_1, \dots, C_k, C \subset \mathcal{D}$, $C_i \cap C_j = \emptyset$ para $(1 \leq i, j \leq k)$. O centróide pode ser definido como a média dos objetos (ou pontos atribuídos ao cluster). A distância entre o objeto $\mathbf{v} \in C_i$ e $\mathbf{c}_i$ como centro de massa ou centróide do cluster é medida através de $dist(\mathbf{c}_i, \mathbf{v})$, onde $\mathbf{c}_i = (c_x, c_y)$ e $\mathbf{v}_i = (v_x, v_y)$. Por exemplo, $dist(\mathbf{c}_i, \mathbf{v})$ pode ser a distância euclidiana entre o objeto $\mathbf{v}$, e o centroide $\mathbf{c}_i$, que é dada por

$$dist(\mathbf{c}_i, \mathbf{v}) = \sqrt{(c_x - v_x)^2 + (c_y - v_y)^2}, \quad (29)$$

onde c_x, v_x, c_y, v_y são os elementos respectivamente dos vetores $\mathbf{v}$ e $\mathbf{c}_i$. Assim, a função objetivo é dada como a soma dos erros para todos os objetos no conjunto de dados (HAN; KAMBER; PEI, 2012)

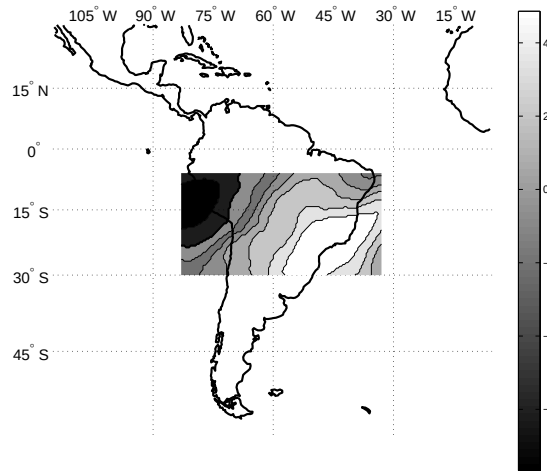
$$E = \sum_{i=1}^k \sum_{v \in c_i; v \in D} dist(\mathbf{c}_i, \mathbf{v})^2, \quad (30)$$

o parâmetro k é o número de clusters, $\mathbf{v}$ é um ponto no espaço que representa um dado objeto, e $\mathbf{c}_i$ é o centróide de um cluster C_i (tanto $\mathbf{v}$ e $\mathbf{c}_i$ são multidimensionais) e $dist(\mathbf{c}_i, \mathbf{v})$ é a distância entre o objeto $\mathbf{v}$, e o centróide mais próximo a ele, $\mathbf{c}_i$.

Inicialmente escolhe-se aleatoriamente k distintos objetos em $\mathcal{D}$ para representar inicialmente os centros dos clusters. Depois associa-se cada objeto de $\mathcal{D}$ ao cluster de centro mais próximo, com base na distância euclidiana. Iterativamente reduz-se E em cada cluster. Para cada cluster, calcula-se um novo centróide usando os objetos atribuídos ao cluster na iteração anterior. Todos os objetos são, então, redistribuídos utilizando os centróides atualizados. As iterações continuam até que a redistribuição fique estável, isto é, os clusters formados na rodada atual são os mesmos formados na rodada anterior ou uma quantidade de iterações pré-determinadas seja realizada (HAN; KAMBER; PEI, 2012). Neste trabalho definiu-se previamente um total de 200 iterações.

A Figura 16 mostra um exemplo de clusterização feito pelo algoritmo *k-means* sobre a média da variável vento meridional.

Figura 16 – Clusterização realizada pelo algoritmo *k-means* sobre parâmetro vento meridional, onde cada isolinha separa um cluster



Diferentemente dos outros critérios de seleção de regiões que separam as regiões pela sua forma, o *k-means* separa as regiões com base no comportamento dos dados. O *k-means* é utilizado para agrupar regiões referente a séries climáticas que apresentam comportamento estatístico semelhantes. Agrupando assim os dados da região estudada em função da evolução da grandeza escolhida. Neste trabalho a grandeza escolhida foi a média da série, tanto para a variável do vento meridional quanto para a variável altura geopotencial.

3.1.2.2 Seleção via regiões retangulares (segundo critério)

Este critério de seleção de região consiste em agrupar dados para o processo de adaptação de filtro por meio de regiões retangulares. Primeiramente, todas as regiões retangulares selecionadas apresentam largura de dois pixels. Este critério apresenta algumas variações de como essa região retangular é escolhida, podendo ser ao longo da coordenada horizontal (abscissa), ao longo da coordenada vertical (ordenada) e ao longo das bordas.

A seleção da região ao longo da abscissa, consiste em sortear aleatoriamente duas linhas (uma subsequente a outra), e selecionar todas as colunas referentes aquelas linhas. Este critério seleciona 1008 séries climáticas.

A seleção da região ao longo da ordenada, consiste em sortear aleatoriamente duas colunas (uma subsequente a outra), e selecionar todas as linhas referentes aquelas colunas. Este critério seleciona 488 séries climáticas.

A seleção ao longo das bordas, consiste em selecionar distintamente as bordas do modelo, são elas, borda superior, borda inferior e bordas laterais. No processo de

adaptação, cada uma das bordas é utilizada separadamente das outras. As bordas laterais selecionam 488 séries climáticas, enquanto que as bordas superior e inferior selecionam 1008 séries climáticas.

A Figura 17 mostra o critério de seleção retangular, no caso de uma seleção ao longo da abscissa e outra ao longo da ordenada. A Figura 18 mostra o critério de seleção retangular, no caso de seleção das bordas.

Figura 17 – Seleção ao longo das abscissas e das ordenadas (Fora de escala)

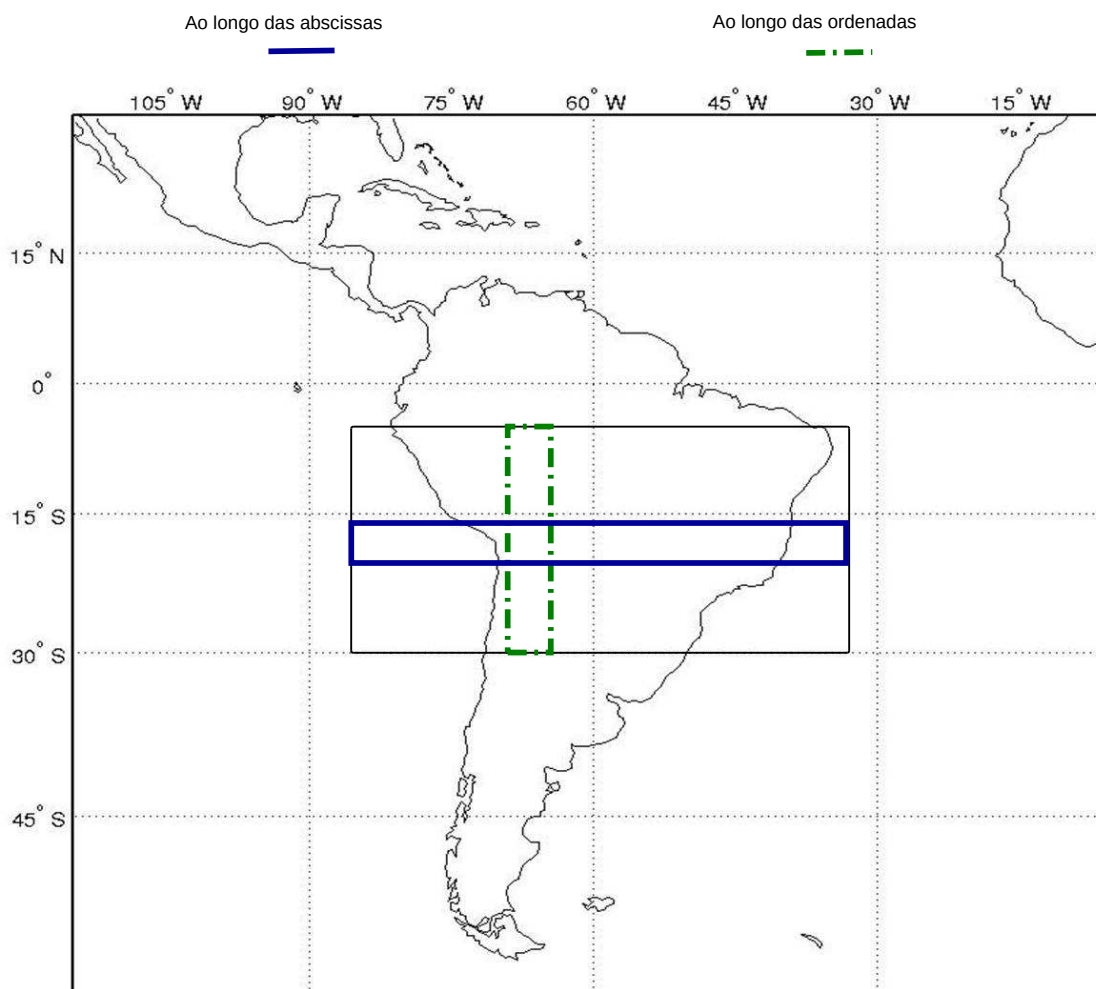
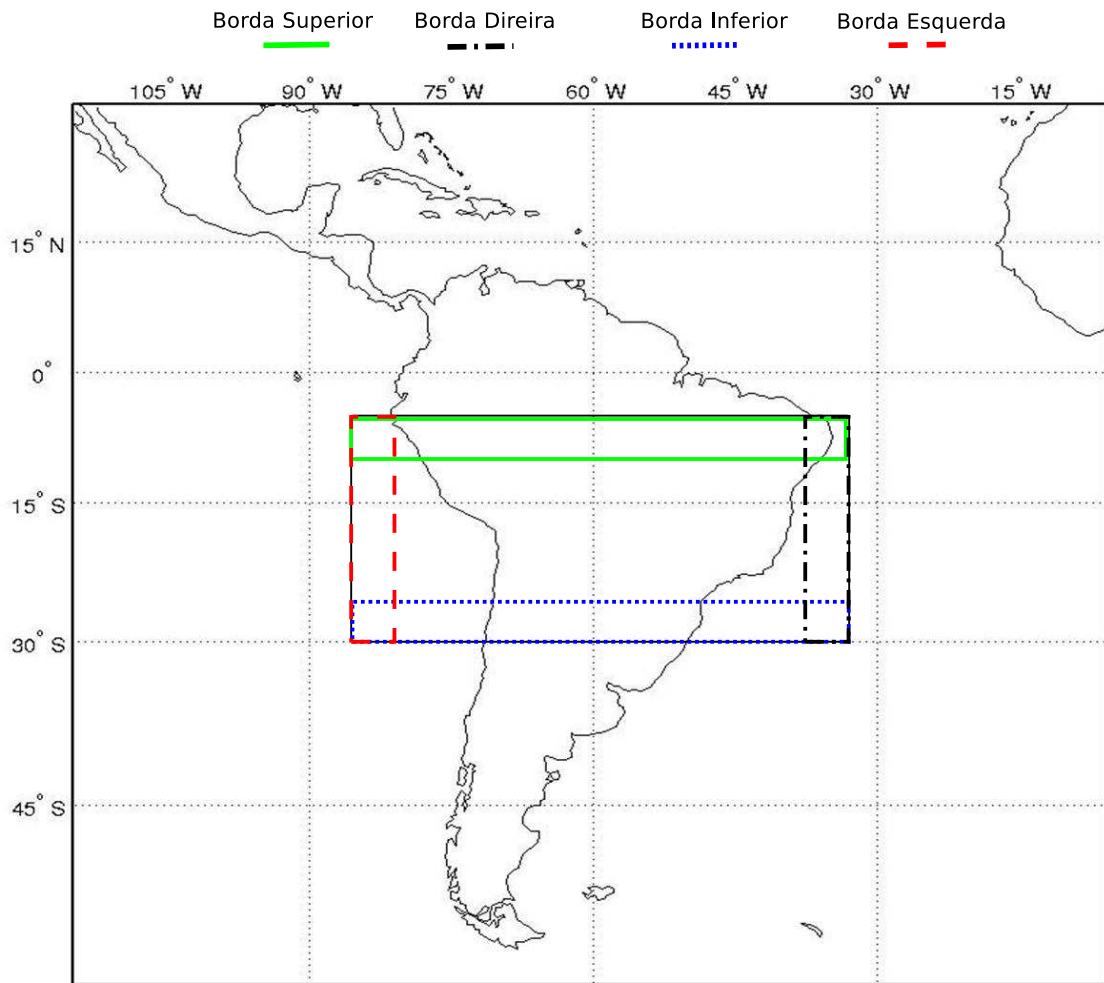


Figura 18 – Seleção das bordas laterais do Modelo (Fora de escala)



As condições de contorno inserem os dados nas bordas do modelo, tal qual, tem-se dados inseridos ao longo do eixo das ordenadas e ao longo do eixo das abscissas. Assim por meio deste critério pretende-se avaliar a dependência das séries sobre esses dados inseridos no modelo, se eles são mais dependentes dos dados inseridos ao longo das abscissas ou das ordenadas. Dessa forma, com critério de seleção de regiões retangulares ao longo do eixo das ordenadas e bordas laterais, pretende-se verificar se a escolha de regiões de uma mesma latitude, independente de um longitude formam um bom conjunto estatístico para treinamento das séries. E por sua vez com o critério de seleção de regiões retangulares ao longo do eixo das abscissas e bordas superior e inferior, pretende-se verificar se a escolha de regiões de uma mesma longitude, independente de uma latitude formam um bom conjunto estatístico para treinamento das séries.

3.1.2.3 Seleção via região circular (terceiro critério)

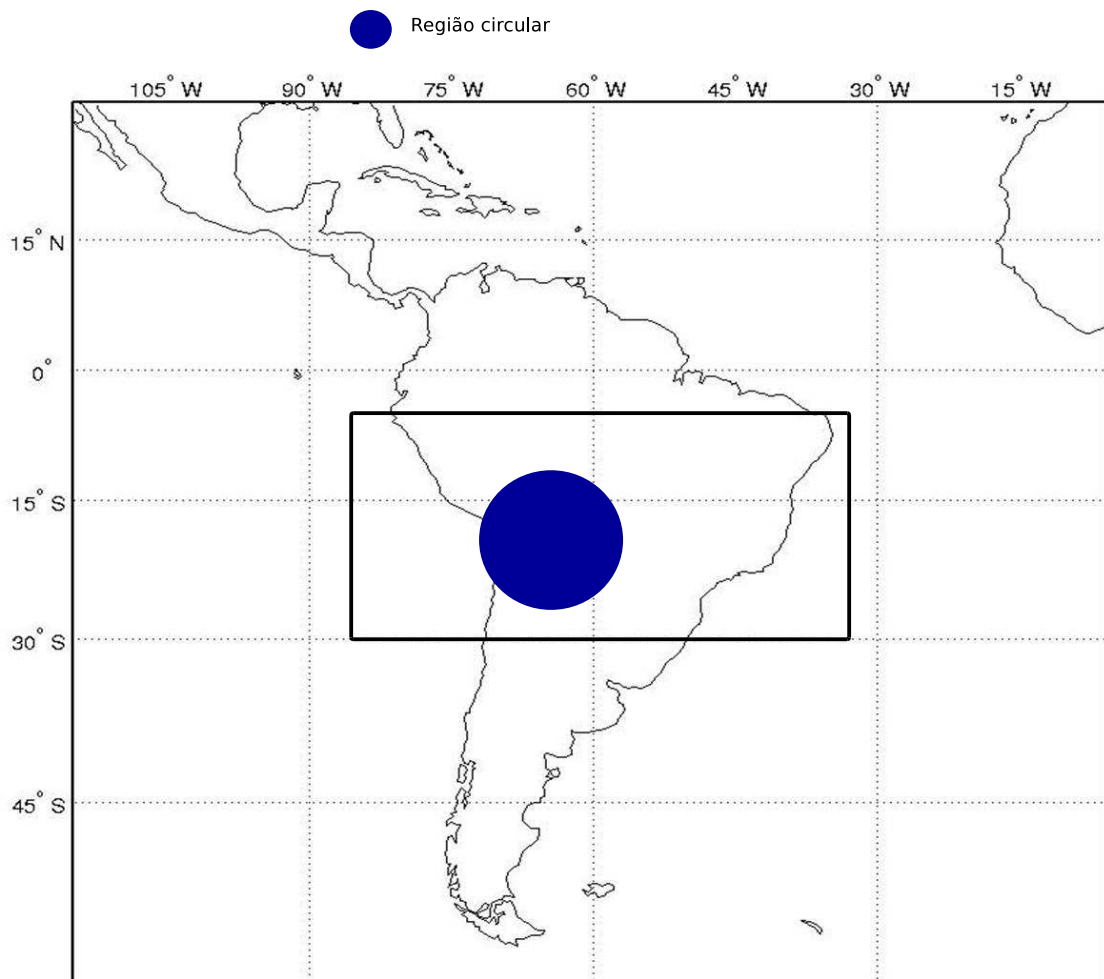
O terceiro critério de seleção foram regiões circulares em torno de um determinado pixel (coordenada geográfica), obedecendo a

$$r^2 = (x - a)^2 + (y - b)^2, \quad (31)$$

onde a é a coordenada sobre o eixo x (Latitude), b a coordenada sobre o eixo y (Longitude), o par ordenado (a, b) é o centro e r é o raio da circunferência. Os raios das circunferências foram escolhidos para que tivessem aproximadamente o mesmo número de séries que o critério de seleção retangular. Assim, no primeiro caso, o raio é igual a 13 pixels, totalizando 529 séries climáticas, no segundo caso, o raio é igual a 18 pixels, totalizando 1009 séries.

Neste critério de seleção, pretende-se analisar se regiões circulares são um bom conjunto para o treinamento do filtro. Sabe-se que o comportamento das séries climáticas não tem forma definida, dessa forma pretende-se avaliar se este formato de região engloba séries com características estatísticas semelhantes. A Figura 19 apresenta uma região circular, a região em azul escuro é um exemplo de região selecionada para o treinamento do filtro.

Figura 19 – Seleção em regiões circulares (Fora de escala)

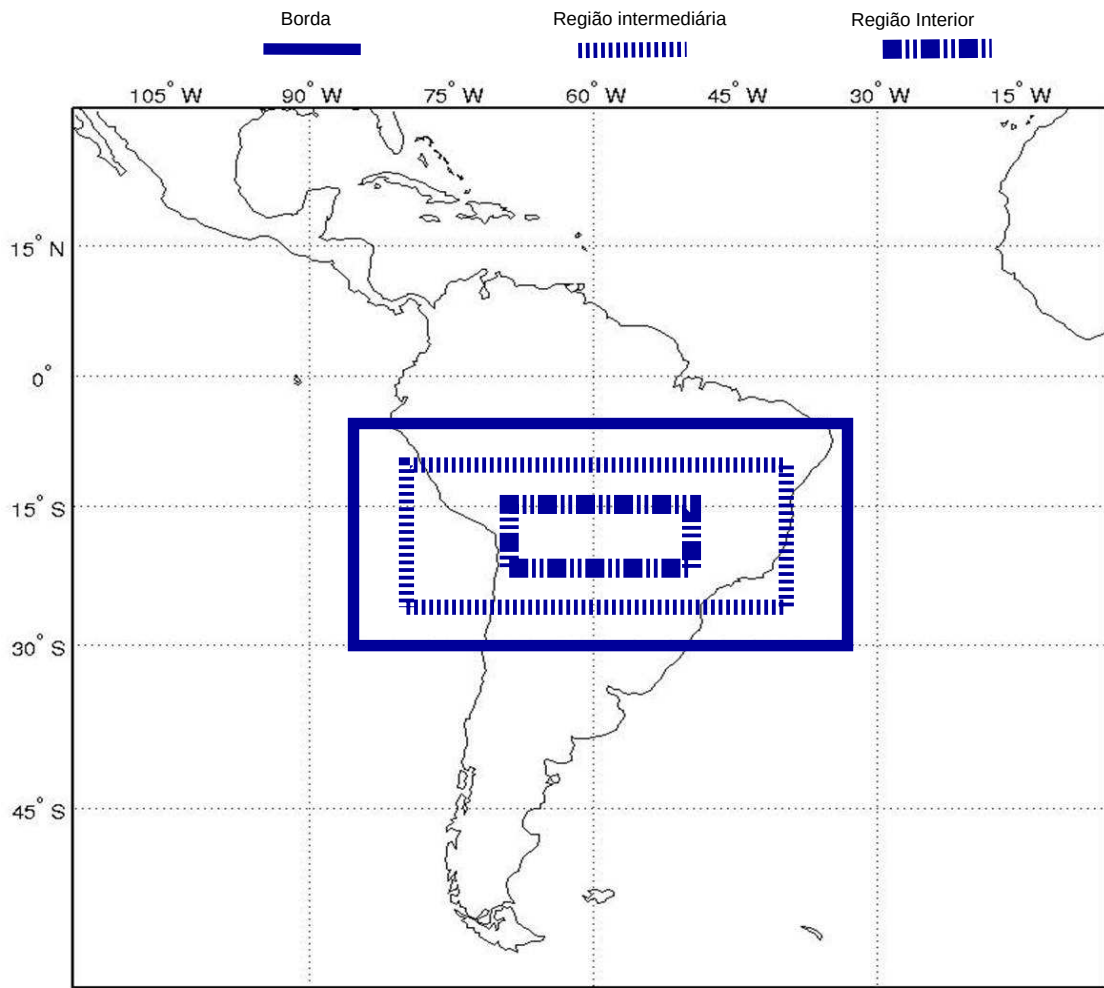


3.1.2.4 Seleção via anéis retangulares (quarto critério)

Este critério de seleção de região consiste em agrupar dados para o processo de adaptação de filtro por meio de anéis concêntricos retangulares.

A Figura 20 apresenta as regiões utilizadas neste critério. A linha cheia representa a região anular nas bordas do modelo. A região anular intermediária é representada por barras. A região anular no interior do domínio de previsão é representada pela linha ora cheia ora em barras. Percebe-se que temos três anéis concêntricos retangulares, um na borda do domínio de previsão, um na região intermediária do domínio de previsão, e outro na região mais interna do domínio de previsão. A região anular da borda do modelo tem largura de um pixel e apresenta 1492 séries climáticas. Para que as outras regiões apresentassem aproximadamente o mesmo número de séries climáticas foi necessário alterar a largura do anel utilizado. Assim, a região anular na região intermediária tem largura de dois pixels e apresenta 1665 séries climáticas e por fim, a região anular no interior da região apresenta 1489 séries climáticas.

Figura 20 – Seleção em anéis retangulares



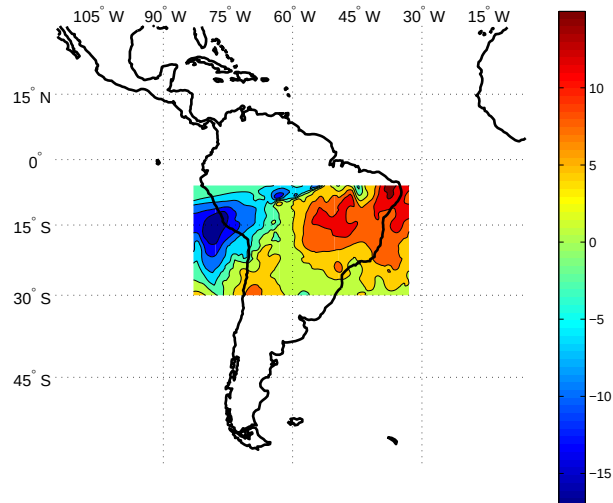
Neste critério pretende-se analisar se regiões no formato de anéis concêntricos retangulares são um bom conjunto de treinamento dos filtros adaptativos. Se temos forçantes laterais tanto ao longo das abscissas e ao longo das ordenadas, neste critério pretende-se avaliar se existe dependência mútua das características estatísticas das séries, ao longo destes anéis concêntricos retangulares.

4 RESULTADOS

4.1 Configurações

Os parâmetros climáticos utilizados nos testes foram o vento meridional, medido em m/s e a altura geopotencial, medida em metro geopotencial (mgp), ambas a uma pressão de 250 mbar. As Figuras 21 e 22 mostram o retrato no tempo de previsão as variáveis vento meridional e altura geopotencial, respectivamente, sobre o domínio de previsão estudado.

Figura 21 – Variável vento meridional (m/s) a 250 hPa previsto pelo modelo Eta no domínio de previsão

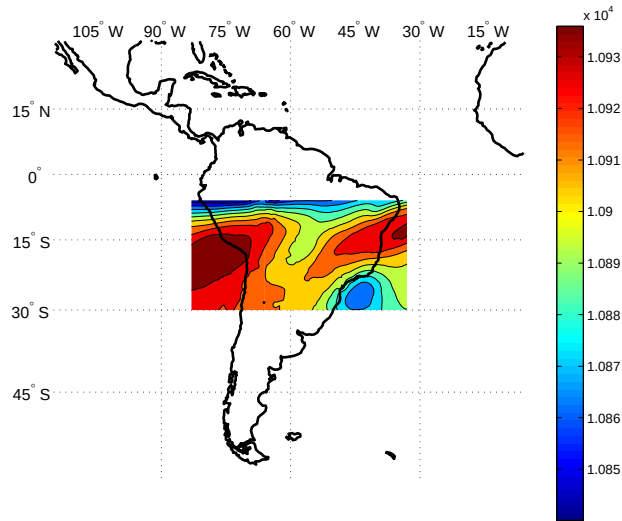


O período do conjunto de previsão considerado é de 2001 a 2010, compreendendo sempre intervalos de previsão entre 13 de dezembro a 30 de abril do ano seguinte. As séries utilizadas apresentam assim 138 dias de previsão, isto é horizonte sazonal. As variáveis vento meridional e a altura geopotencial foram escolhidas por serem variáveis de previsão que influenciam na modelagem atmosférica (ROZANTE, 2004). As previsões do modelo Eta 40km e as reanálises do NCEP apresentam resolução temporal de seis horas, com isso, para cada dia há quatro saída de previsões e com isso para cada coordenada há 552 pontos.

Testes preliminares foram realizados de modo a tentar determinar os parâmetros dos algoritmos adaptativos NLMS e RLS. Os parâmetros foram escolhidos com base no desempenho da adaptação realizada por estes algoritmos.

Utilizaram-se filtros adaptados com quatro coeficientes (equivalente a um dia de previsões) e com trinta e dois coeficientes (oito dias de previsões) para realizar o ajuste dos dados meteorológicos do modelo Eta. Essas ordens de filtro foram definidas de modo a

Figura 22 – Variável altura geopotencial (mcp) a 250 hPa previsto pelo modelo Eta no domínio de previsão



avaliar o impacto do comprimento do filtro sobre o ajuste dos dados do Eta em relação aos do NCEP. Isto é, filtros correspondentes a apenas um dia de previsões são suficientes para ajustar os dados meteorológicos ou é necessário utilizar filtros de comprimentos maiores?

Os parâmetros que devem ser escolhidos no caso do NLMS são o filtro $\mathbf{w}(n)$ inicial, e o fator de convergência. O fator de convergência escolhido foi $\mu_n = 0.1$, e o filtro $\mathbf{w}(n)$ foi inicializado com zeros. No caso do RLS os parâmetros que devem ser escolhidos são o fator de esquecimento e o filtro inicial. O fator de esquecimento escolhido foi $\lambda = 0.999$ enquanto o filtro $\mathbf{w}(n)$ também foi inicializado com zeros.

Um dos fatores importantes quando se trata de filtragem adaptativa é a quantidade de séries usadas para adaptar o filtro. Essa quantidade foi definida pelo critério de seleção por bordas. Como mencionado, as bordas superior e inferior selecionadas possuem 1008 séries cada, enquanto que as bordas laterais possuem 488 séries. Para se ter o mesmo critério de comparação com relação ao número de séries optou-se, por utilizar aproximadamente as quantidades de 500 e 1000 séries, nos critérios de seleção de regiões retangulares, regiões pelo algoritmo *k-means* e regiões circulares.

As estratégias do diagrama de blocos da Figura 15 foram aplicadas dez vezes a cada um dos anos estudados de modo a fornecer uma estatística confiável de resultados, considerando um intervalo de confiança de 95%. Ainda, as séries treinaram os filtros em sequência aleatória e, em particular, no critério de seleção utilizando regiões circulares, a posição ao longo da latitude e longitude do centro da circunferência também foi escolhido aleatoriamente.

4.2 Resultados encontrados

A. Critérios de seleção: *k-means*, regiões retangulares e região circular

As colunas das Tabelas 2 a 5 apresentam os resultados obtidos com os diferentes critérios de seleção de regiões para o conjunto de treinamento na aplicação dos algoritmos adaptativos NLMS e RLS.

A Tabela 2 mostra resultados da porcentagem de regiões com redução no EQMn da variável vento meridional e aproximadamente 500 séries de treinamento e com filtro de ordem correspondente a um dia de previsões. Nesta tabela pode-se ver que os filtros adaptados pelo algoritmo NLMS reduziu em média o EQMn em 98% da região selecionada pelo algoritmo *k-means*, em 70% da região selecionada na borda esquerda, em 55% da região selecionada na borda direita, em 68% da região selecionada ao longo do eixo das ordenadas e reduziu na média o EQMn em 44% da região circular. Em contrapartida os filtros adaptados pelo algoritmo RLS reduziram na média o EQMn em 100% da região selecionada pelo algoritmo *k-means*, em 98% da região selecionada na borda esquerda, em 92% da região selecionada na borda direita, em 99% da região selecionada ao longo do eixo das ordenadas e reduziu na média o EQMn em 93% da região circular. Assim percebe-se que os filtros adaptados pelo algoritmo RLS, em todos os critérios de seleção de regiões, reduziram na média o EQMn em uma porcentagem maior de regiões do que filtros adaptados pelo algoritmo NLMS.

Tabela 2 – Porcentagem de regiões com redução no EQMn do vento meridional, aplicando filtros de 1 dia (4 coeficientes), com cardinalidade de treinamento determinada pelas bordas laterais

Ano de aplicação	<i>k-means</i>		Borda Esquerda		Borda Direita		Eixo das Ordenadas		Região Circular	
	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	100	100	82	93	82	92	97	100	100	100
2003	100	100	56	100	80	100	49	100	1	100
2004	95	100	96	100	47	100	50	91	21	35
2005	100	100	72	100	48	91	56	100	10	100
2006	100	100	72	100	28	87	74	100	30	100
2007	99	100	58	100	100	100	79	98	80	100
2008	87	100	62	100	5	90	54	100	82	100
2009	100	100	97	93	98	100	90	100	70	100
2010	100	100	40	100	10	70	63	100	0	100
Média	98	100	70	98	55	92	68	99	44	93

A Tabela 3 mostra resultados da porcentagem de regiões com redução no EQMn da variável vento meridional e aproximadamente 500 séries de treinamento e com filtro de ordem correspondente a oito dias de previsões. Nesta tabela, comparando a média da redução do EQMn dos filtros adaptados pelo algoritmo NLMS e pelo algoritmo RLS, percebe-se que os filtros adaptados pelo algoritmo RLS reduzem o EQMn em uma porcentagem superior de regiões que filtros adaptados pelo algoritmo NLMS. Um exemplo disso é que o filtro adaptado pelo algoritmo NLMS tendo como critério de seleção de regiões circulares, reduziu o erro em 59% da região, enquanto que o filtro adaptado pelo algoritmo RLS reduziu o erro em 93% da região. É importante observar ainda que somente o filtro adaptado pelo algoritmo RLS e com região de treinamento selecionada pelo algoritmo *k-means* reduziu o EQMn em 100% da região.

Tabela 3 – Porcentagem de regiões com redução no EQMn do vento meridional, aplicando filtros de 8 dias (32 coeficientes), com cardinalidade de treinamento determinada pelas bordas laterais

Ano de aplicação	<i>k-means</i>		Borda Esquerda		Borda Direita		Eixo das Ordenadas		Região Circular	
	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	100	100	35	100	59	59	59	90	100	100
2003	100	100	80	74	46	100	93	100	100	100
2004	0	100	38	23	50	98	41	100	30	35
2005	39	100	84	100	59	90	80	100	12	100
2006	100	100	95	100	28	81	58	100	100	100
2007	100	100	94	100	68	99	27	93	34	100
2008	64	100	56	100	63	80	53	100	33	100
2009	100	100	77	100	19	96	67	100	100	100
2010	100	100	10	100	52	53	84	92	29	100
Média	78	100	63	89	49	84	63	97	59	93

A Tabela 4 mostra resultados da porcentagem de regiões com redução no EQMn da variável vento meridional e aproximadamente 1000 séries de treinamento e com filtro de ordem correspondente a um dia de previsões. Nesta tabela pode-se observar que os filtros adaptados pelo algoritmo RLS ainda apresentam resultados superiores aos filtros adaptados pelo algoritmo NLMS. Entretanto, para as configurações dispostas nesta tabela, nenhum conjunto de critérios reduziu o EQMn em toda a região selecionada. Assim, dentre estas configurações aquelas que apresentaram melhores resultados foram o filtro adaptado pelo algoritmo RLS utilizando os critérios de seleção de regiões da borda superior, e da borda inferior que reduziram o EQMn em 97% da região cada. Por outro lado o filtro

adaptado pelo algoritmo NLMS, utilizando o critério de seleção de regiões da borda inferior reduziu apenas 38% da região, seu desempenho foi ruim principalmente nos anos de 2008 e 2009, tendo reduzido o EQMn em apenas 8% e 9% da região. Percebe-se então que o algoritmo RLS obteve resultado superior ao algoritmo NLMS.

Tabela 4 – Porcentagem de regiões com redução no EQMn do vento meridional, aplicando filtros de 1 dia (4 coeficientes), com cardinalidade de treinamento determinada pela borda superior

Ano de aplicação	k-means		Borda Superior		Borda Inferior		Eixo das Abscissas		Região Circular	
	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	100	100	70	100	30	95	85	99	100	100
2003	100	100	79	100	90	100	65	100	0	100
2004	23	48	42	70	43	100	47	92	3	19
2005	100	100	52	100	33	98	46	100	20	100
2006	100	100	82	100	18	100	62	99	40	100
2007	100	83	40	100	88	94	97	92	85	100
2008	98	100	60	100	8	100	40	87	85	100
2009	100	100	65	100	9	100	43	100	100	100
2010	99	100	33	100	22	90	15	100	20	100
Média	91	92	58	97	38	97	55	97	50	91

A Tabela 5 mostra resultados da porcentagem de regiões com redução no EQMn da variável vento meridional e aproximadamente 1000 séries de treinamento e com filtro de ordem correspondente a oito dias de previsões. Nesta tabela percebe-se que o filtro adaptado pelo RLS utilizando o critério de seleção de região pela borda inferior apresentou resultado superior as demais configurações, reduzindo o EQMn em 98% da região. Como pior configuração temos o filtro adaptado pelo algoritmo NLMS utilizando o critério de seleção de regiões por meio do algoritmo k-means, que reduziu o EQMn em apenas 42% da região, os piores anos da aplicação do filtro foram os anos de 2005, 2006, 2009 e 2010 em que não houve redução do EQMn nas regiões em que se aplicou o filtro.

Tabela 5 – Porcentagem de regiões com redução no EQMn do vento meridional, aplicando filtros de 8 dias (32 coeficientes), com cardinalidade de treinamento determinada pela borda superior

Ano de aplicação	<i>k-means</i>		Borda Superior		Borda Inferior		Eixo das Abscissas		Região Circular	
	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	100	100	68	100	49	95	54	94	98	100
2003	100	100	96	100	52	100	65	94	76	100
2004	72	48	30	70	93	100	85	91	40	20
2005	0	100	83	100	11	97	23	78	37	100
2006	0	100	97	100	35	100	67	78	98	100
2007	53	83	55	100	77	94	47	94	12	100
2008	57	100	40	100	40	100	18	62	15	100
2009	0	100	61	100	42	100	60	76	98	100
2010	0	100	50	100	32	100	50	90	33	100
Média	42	92	64	97	48	98	52	84	56	91

A partir dos dados das Tabelas 2 a 5 sabe-se que o uso de filtros com comprimento equivalente a um dia de previsão gera na média a redução no EQMn em uma fração de regiões maior do que se obteve com filtros de comprimento equivalente a oito dias de previsões. Percebe-se que os filtros adaptados usando algoritmo RLS reduzem o EQMn de séries ajustadas em mais regiões que os filtros adaptados pelo algoritmo NLMS. Observa-se ainda, que a seleção de regiões utilizando o *k-means* mostra-se como a melhor estratégia, pois com ela obteve-se redução do EQMn em toda a região. Por fim, percebe-se que é melhor utilizar 500 séries no treinamento do filtro, pois a quantidade de séries não influenciou significativamente os resultados.

Especificamente no ano de 2004, pelas Tabelas 3, 4 e 5 percebe-se que o uso do algoritmo NLMS utilizado na adaptação dos filtros não contribuiu de forma satisfatória para a redução no EQMn das séries climáticas. Em (ALMEIDA; BEZERRA; MEDEIROS, 2009) mostra-se que o ano de 2004 foi um ano atípico, e com isso o ano de 2003 não apresenta um comportamento estatístico consistente com o ano seguinte. Assim os filtros adaptados pelo algoritmo NLMS não foram capazes de reduzir esses erros de previsão. Contudo, pela Tabela 2, o uso do filtro adaptado pelo algoritmo NLMS na borda esquerda do modelo e nas regiões separadas pelo *k-means*, reduz o EQMn no ano de 2004 em 96% e 95% das regiões. Na Tabela 5 vê-se que o filtro com o uso do filtro adaptado pelo algoritmo NLMS na borda superior do modelo e nas regiões separadas pelo *k-means*, reduz o EQMn no ano de 2004 em 93% e 73% das regiões. Dessa forma percebe-se que nestas configurações

o NLMS conseguiu contornar o problema de modelagem do ano de 2003 para o ano de 2004.

As Figuras 23 e 24 mostram gráficos dos resultados da porcentagem de regiões com redução no EQMn da variável vento meridional e aproximadamente 500 séries de treinamento e com filtro de ordem correspondente a um dia de previsões treinado pelos algoritmos NLMS e RLS respectivamente. Pelos gráficos destas figuras percebe-se que o algoritmo RLS não só reduz na média o EQMn em uma porcentagem superior ao NLMS como em geral o intervalo de confiança é menor, denotando assim maior confiabilidade nos resultados obtidos. Percebe-se no gráfico da figura 24 a discrepância da porcentagem de regiões com redução no EQMn sobre o ano de 2004, para o caso de regiões circulares, em relação aos demais critérios. O critério de seleção em regiões circulares reduziu o EQMn em 35% da região. Assim percebe-se que nestas configurações este critério não conseguiu contornar o problema de modelagem do ano de 2003 para o ano de 2004.

Figura 23 – Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo NLMS e com 500 séries de treinamento

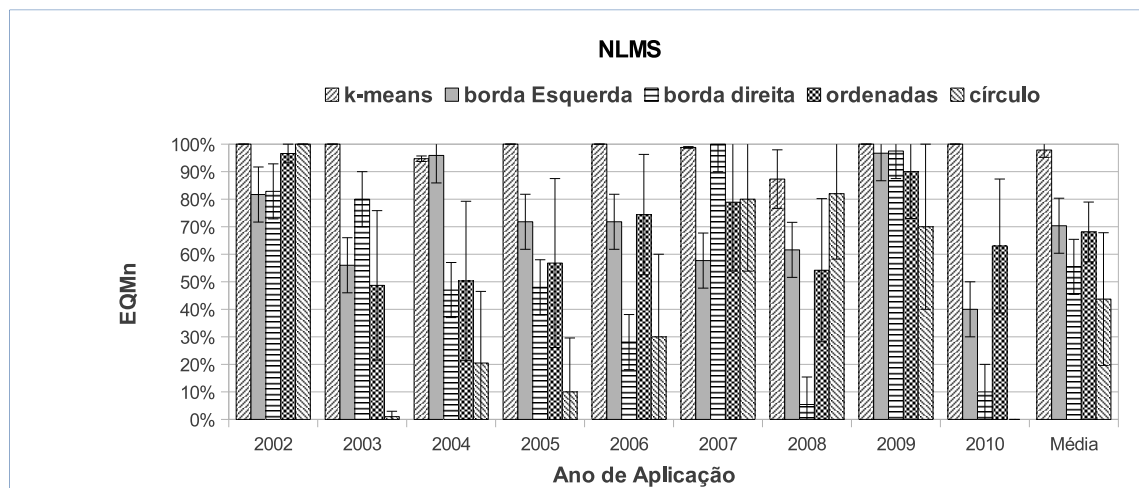
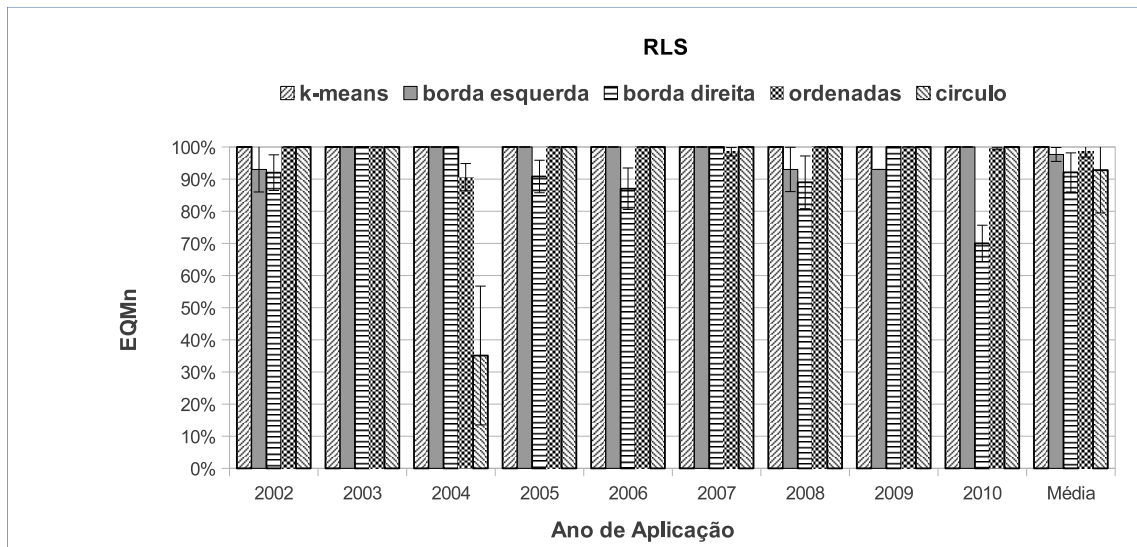


Figura 24 – Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo RLS e com 500 séries de treinamento



As Figuras 25 e 26 mostram gráficos dos resultados da porcentagem de regiões com redução no EQMn da variável vento meridional e aproximadamente 1000 séries de treinamento e com filtro de ordem correspondente a um dia de previsões treinado pelos algoritmos NLMS e RLS respectivamente. Tem-se que sobre estas configurações o filtro adaptado pelo algoritmo RLS reduziu na média o EQMn em mais regiões que o filtro treinado pelo algoritmo NLMS, assim como apresentou um menor intervalo de confiança. Observando o ano de 2004 no gráfico da Figura 26 percebe-se que somente o critério de seleção da borda inferior contornou o problema de modelagem do ano de 2003 para o ano de 2004.

Figura 25 – Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo NLMS e com 1000 séries de treinamento

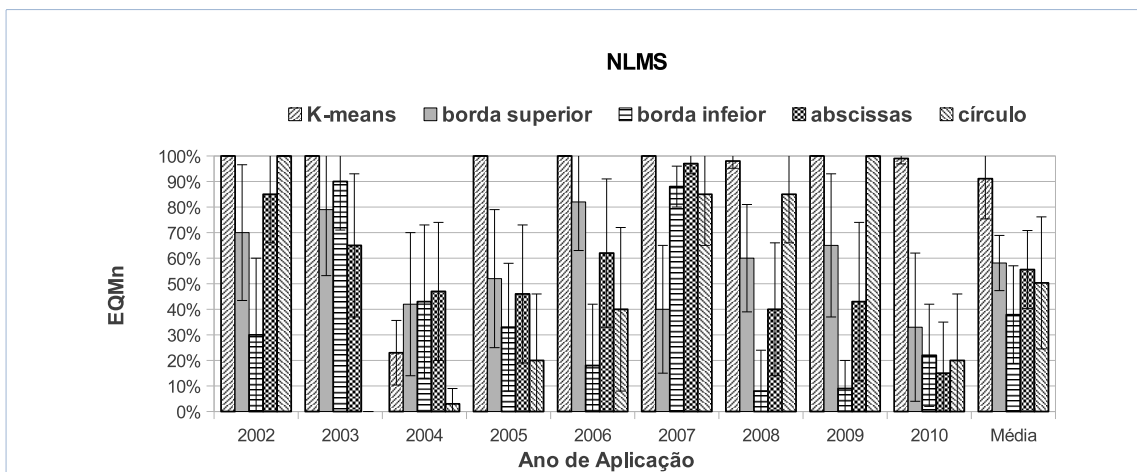
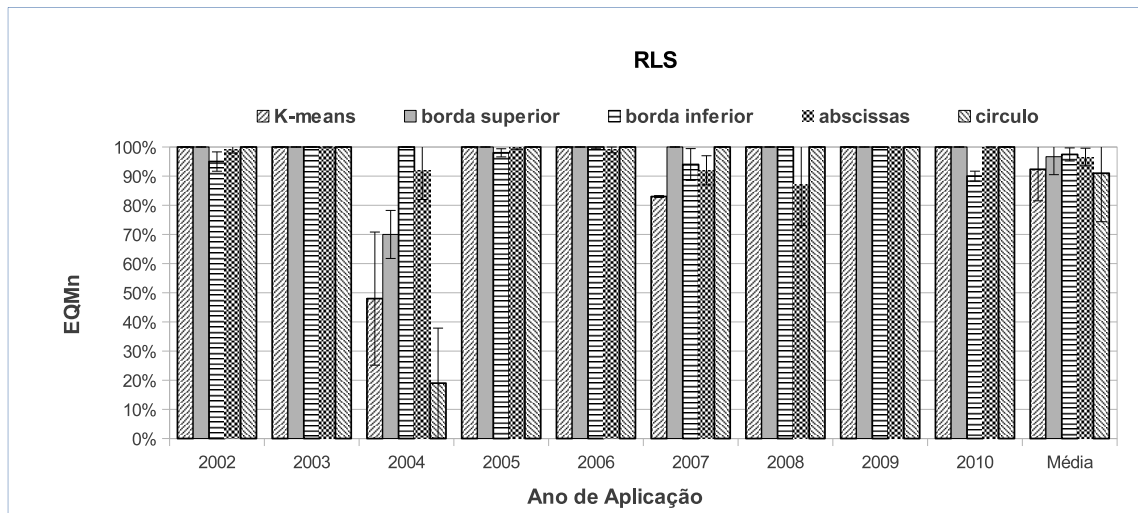


Figura 26 – Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo RLS e com 1000 séries de treinamento



As Figuras 27 e 28 mostram gráficos dos resultados da porcentagem de regiões com redução no EQMn da variável vento meridional e aproximadamente 500 séries de treinamento e com filtro de ordem correspondente a oito dias de previsões treinado pelos algoritmos NLMS e RLS respectivamente. Pelos gráficos, percebe-se que o filtro adaptado pelo algoritmo NLMS apresenta um comportamento instável na redução do EQMn, percebe-se isso pelos altos valores do intervalo de confiança. Em contra partida os filtros adaptados pelo algoritmo RLS são mais estáveis na redução na média do EQMn.

Figura 27 – Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo NLMS e com 500 séries de treinamento

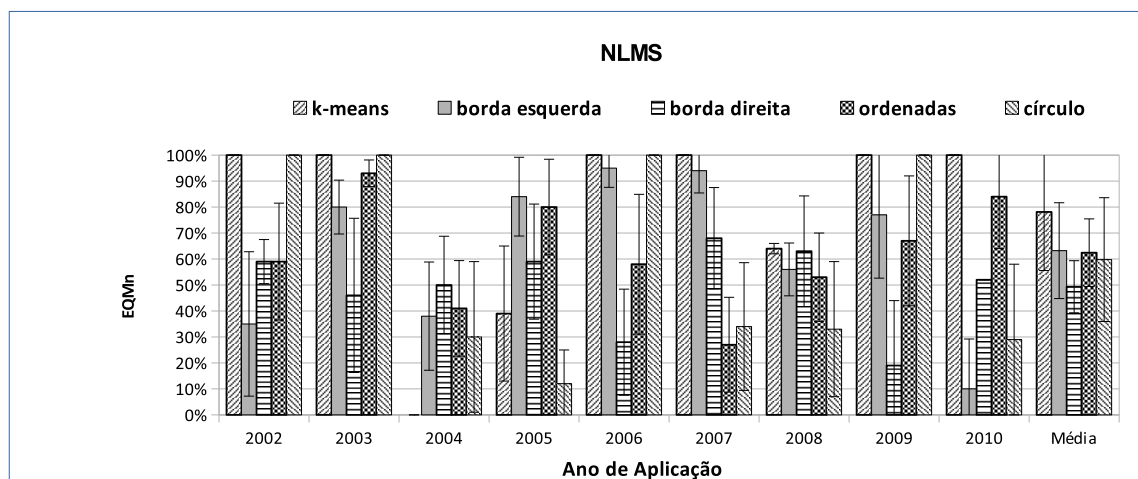
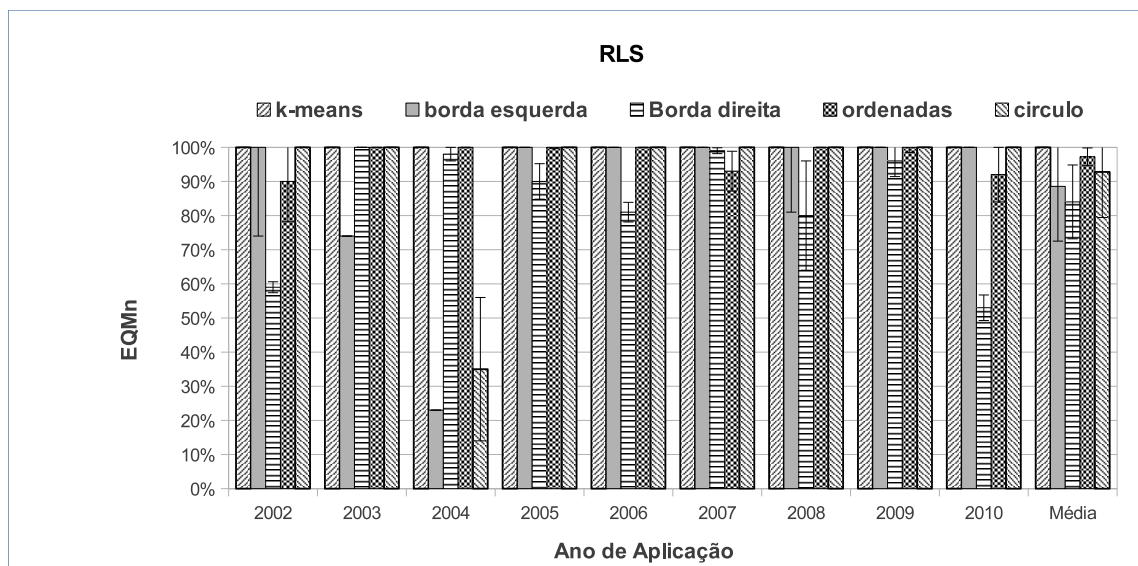


Figura 28 – Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo RLS e com 500 séries de treinamento



As Figuras 29 e 30 mostram gráficos dos resultados da porcentagem de regiões com redução no EQMn da variável vento meridional e aproximadamente 1000 séries de treinamento e com filtro de ordem correspondente a oito dias de previsões treinado pelos algoritmos NLMS e RLS respectivamente. Para esta configuração, se tem o filtro adaptado pelo algoritmo RLS com redução do EQMn em mais regiões que o filtro adaptado pelo algoritmo NLMS. Além disso o NLMS apresenta elevados valores de intervalo de confiança, denotando a grande variabilidade nos resultados.

Figura 29 – Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo NLMS e com 1000 séries de treinamento

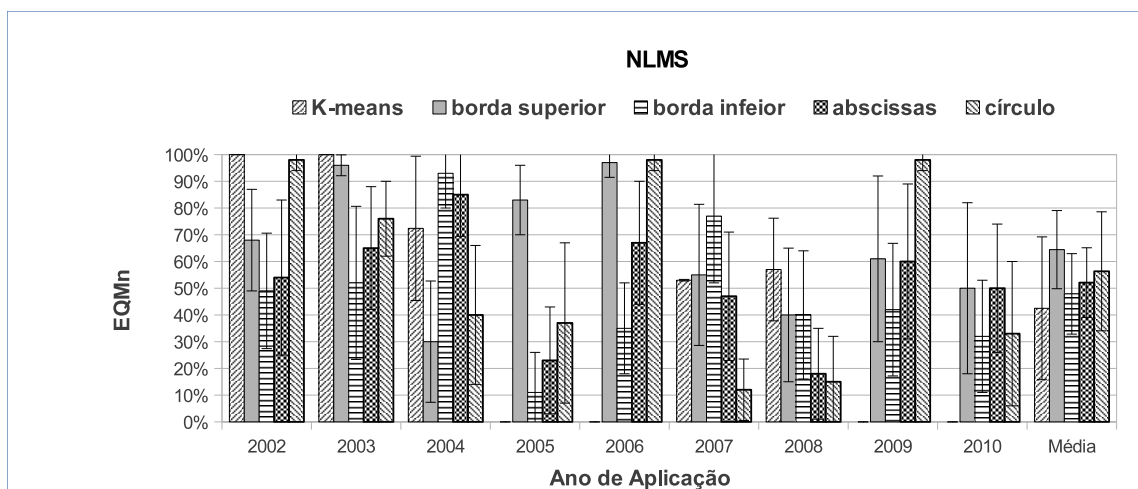
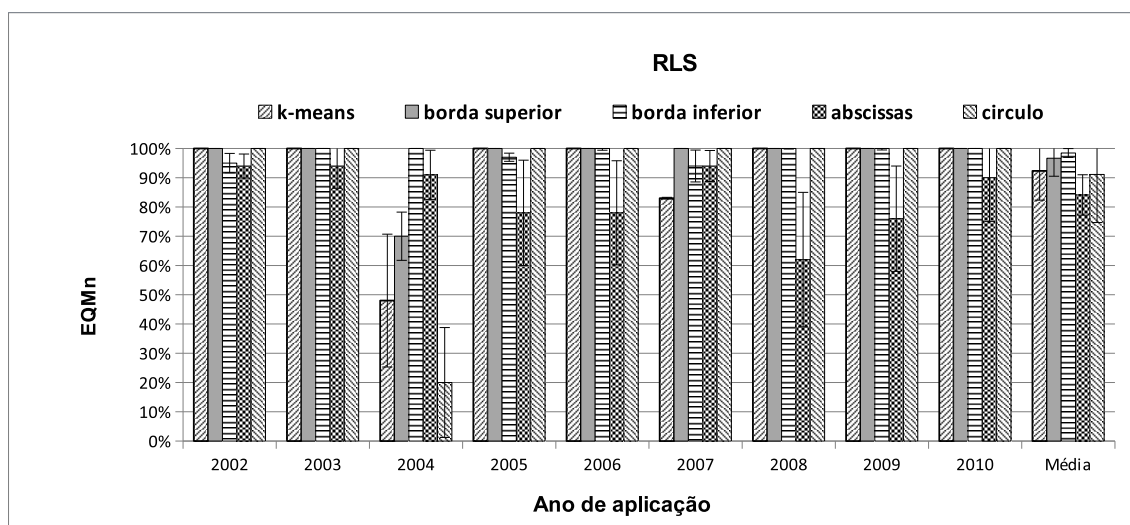


Figura 30 – Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo RLS e com 1000 séries de treinamento



Pelos gráficos das Figuras 23, a 30 pode-se perceber que o algoritmo RLS não só reduz o EQMn em uma porcentagem de regiões superior ao NLMS, como na média o intervalo de confiança é menor, denotando assim maior confiabilidade nos resultados obtidos. Percebe-se ainda que diante de outros critérios de seleção de regiões o método utilizando o algoritmo *k-means* se destaca por apresentar resultados melhores, tanto no uso do algoritmo de adaptação NLMS quanto no RLS. Por fim pode-se ver que filtros com comprimento de 1 dia de previsões (4 coeficientes) apresentaram na média resultados superiores a filtro com 8 dias de previsões (32 coeficientes).

De forma geral percebe-se que pelas Tabelas 2 a 5 e pelos gráficos das Figuras 23 a 30 os filtros adaptados utilizando algoritmo RLS reduziram o erro em porcentagens maiores de regiões do que os obtidos via NLMS. Ou seja, dentre as estratégias utilizadas no treinamento do filtro, o algoritmo RLS consegue adaptar melhor os filtros no intuito de reduzir o erro de previsão do vento meridional.

A Tabela 6 mostra resultados da porcentagem de regiões com redução no EQMn da variável altura geopotencial e aproximadamente 500 séries de treinamento e com filtro de ordem correspondente a um dia de previsões. A Tabela 6 mostra que os filtros adaptados pelo algoritmo RLS apresentam resultados superiores que os filtros adaptados pelo algoritmo NLMS. O pior caso no uso do NLMS foi utilizando o critério de seleção de regiões retangulares ao longo das ordenadas, que reduziu na média o EQMn em apenas 19% das regiões. Ainda, dentre as configurações apresentadas, o filtro adaptado pelo algoritmo RLS e com critério de seleção de região feito pelo algoritmo *k-means* apresentou o melhor resultado reduzindo na média o EQMn em 100% da região.

Tabela 6 – Porcentagem de regiões que se obtém redução no EQMn da altura geopotencial, aplicando filtros de 1 dia (4 coeficientes), com cardinalidade do treinamento determinado pelas forças horizontais

Ano de aplicação	<i>k-means</i>		Borda Esquerda		Borda direita		Eixo das ordenadas		Região circular	
	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	0	100	61	91	66	100	10	100	70	100
2003	100	100	63	93	86	98	0	100	72	82
2004	0	100	24	80	73	86	5	18	50	78
2005	82	100	8	35	37	90	16	73	31	90
2006	74	100	10	65	25	82	80	76	30	100
2007	100	100	20	76	52	80	0	98	49	99
2008	10	100	10	83	0	50	0	88	42	90
2009	10	100	38	73	13	68	1	70	10	88
2010	95	100	61	77	81	5	56	57	72	49
Média	52	100	32	75	48	73	19	75	47	86

A Tabela 7 mostra resultados da porcentagem de regiões com redução no EQMn da variável altura geopotencial e aproximadamente 500 séries de treinamento e com filtro de ordem correspondente a oito dias de previsões. Percebe-se que de todas as configurações presentes nesta Tabela, aquelas que utilizam o algoritmo RLS para treinar o filtro apresentaram desempenho superiores aquelas que utilizam o algoritmo NLMS para treinar o filtro. Tem-se então que a melhor configuração neste caso foi aquela cujo o filtro foi adaptado pelo algoritmo RLS e com critério de seleção de regiões feito pelo algoritmo *k-means* que reduziu na média o EQMn em 81% da região.

Tabela 7 – Porcentagem de regiões que se obtém redução no EQMn da altura geopotencial aplicando filtros de 8 dias (32 coeficientes), com cardinalidade do treinamento determinado pelas forçantes horizontais

Ano de aplicação	k-means		Borda Esquerda		Borda direita		Eixo das Ordenadas		Região circular	
	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	0	100	79	67	41	100	49	99	40	93
2003	11	100	25	71	43	95	0	84	31	65
2004	0	100	13	74	41	84	4	9	30	82
2005	26	91	57	23	70	83	2	52	20	100
2006	50	100	85	62	26	65	17	70	40	87
2007	1	65	8	37	55	69	23	80	31	73
2008	1	35	15	64	3	46	1	15	26	50
2009	90	97	70	78	27	58	31	79	10	60
2010	38	47	22	61	35	5	67	81	70	31
Média	24	81	41	60	39	71	22	63	33	71

A Tabela 8 mostra resultados da porcentagem de regiões com redução no EQMn da variável altura geopotencial e aproximadamente 1000 séries de treinamento e com filtro de ordem correspondente a um dia de previsões. Por esta tabela percebe-se que algumas configurações apresentaram anos com zero regiões com redução no EQMn, como por exemplo no caso do filtro adaptado pelo algoritmo NLMS com critério de seleção de regiões utilizando o k-means, sobre os anos de 2004, 2007 e 2008. Por outro lado, utilizando o mesmo critério de seleção de regiões k-means, mas adaptando o filtro com o algoritmo RLS, tem-se a redução na média do EQMn em 95% da região, sendo a melhor configuração apresentada nesta Tabela.

Tabela 8 – Porcentagem de regiões que se obtém redução no EQMn da altura geopotencial, aplicando filtros de 1 dia (4 coeficientes), com cardinalidade do treinamento determinado pelas forçantes verticais

Ano de aplicação	<i>k-means</i>		Borda Superior		Borda inferior		Eixo das Abscissas		Região circular	
	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	0	100	47,7	100	73	98	78	100	0	100
2003	80	100	57	100	27	31	49	70	50	90
2004	0	100	43	47	0	0	35	90	58	79
2005	46	96	42	80	0	63	46	79	30	90
2006	59	100	58	100	68	97	42	92	51	100
2007	0	100	57	99	10	37	36	73	34	99
2008	0	100	0	88	0	60	15	85	30	91
2009	10	73	0	100	0	86	15	78	8	93
2010	72	88	65	70	97	48	84	52	49	35
Média	30	95	41	87	31	58	44	80	34	83

A Tabela 9 mostra resultados da porcentagem de regiões com redução no EQMn da variável altura geopotencial e aproximadamente 1000 séries de treinamento e com filtro de ordem correspondente a oito dias de previsões. Neste caso o critério de seleção de regiões que apresentou o melhor resultado foi a seleção de regiões circulares, com o filtro adaptado pelo algoritmo RLS, que reduziu na média o EQMn em 76% das regiões.

Tabela 9 – Porcentagem de regiões que se obtém redução no EQMn da altura geopotencial, aplicando filtros de 8 dias (32 coeficientes), com cardinalidade do treinamento determinado pelas forçantes verticais

Ano de aplicação	<i>k-means</i>		Borda Superior		Borda inferior		Eixo das Abscissas		Região circular	
	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	0	100	55	99	70	73	45	87	64	100
2003	11	100	21	94	15	26	48	65	50	100
2004	0	100	6	41	6	13	31	70	35	89
2005	26	45	9	41	61	57	44	76	28	80
2006	50	100	47	90	60	85	53	80	76	84
2007	1	7	31	80	0	8	44	80	43	86
2008	1	5	3	33	0	15	32	44	29	70
2009	90	29	70	68	2	70	8	70	10	37
2010	38	25	57	64	23	24	54	48	42	41
Média	24	57	33	68	26	41	40	69	42	76

As Tabelas 6 a 9 mostraram que o uso de filtros com comprimento equivalente a um dia de previsão gera na média a redução no EQMn em uma fração de regiões maior do que se obteve com filtros de comprimento equivalente a oito dias de previsões. Percebe-se que os filtros adaptados usando algoritmo RLS reduzem o EQMn de séries ajustadas em mais regiões que os filtros adaptados pelo algoritmo NLMS. Observa-se ainda, que a seleção de regiões utilizando o *k-means* mostra-se a melhor estratégia, pois com ela obteve-se redução do EQMn em mais regiões e todos os anos de ajuste, como pode ser visto na Tabela 6, em que esta estratégia reduziu o EQMn em toda a região. Por fim, percebe-se que como na variável vento meridional é melhor utilizar 500 séries no treinamento do filtro, pois a quantidade de séries não influenciou significativamente os resultados.

Especificamente no ano de 2004, demonstra-se que o uso do algoritmo NLMS utilizado na adaptação dos filtros não contribui bem para a redução no EQMn das séries climáticas nas séries da variável altura geopotencial. Como mencionado o ano de 2004 foi um ano atípico, e com isso o ano de 2003 não apresentam um comportamento estatístico consistente com o ano seguinte. Assim os filtros adaptados pelo algoritmo NLMS não conseguiram corrigir esses erros de previsão. Aliás, o desempenho do algoritmo NLMS foi muito ruim quando aplicado a altura geopotencial e se comparado ao algoritmo RLS aplicado a mesma variável.

Para esta variável somente o filtro adaptado pelo algoritmo RLS conseguiu reduzir o EQMn das séries do ano de 2004. Dessa forma percebe-se que nestas configurações somente

o filtro adaptado pelo algoritmo RLS conseguiu contornar o problema de modelagem do ano de 2003 para o ano de 2004.

Percebe-se que a única combinação que conseguiu reduzir o EQMn em toda a região selecionada foi a seleção utilizando o algoritmo *k-means* com adaptação dos coeficientes do filtro com 4 coeficientes pelo algoritmo RLS.

As Figuras 31 e 32 mostram a porcentagem de regiões com redução no EQMn da variável altura geopotencial e aproximadamente 500 séries de treinamento e com filtro de ordem correspondente a um dia de previsões treinado pelos algoritmos NLMS e RLS respectivamente. Por estes gráficos percebe-se que os filtros adaptados pelo algoritmo NLMS apresentaram resultados muito inferiores aos dos filtros adaptados pelo algoritmo RLS, tanto na porcentagem de regiões com redução no EQMn quanto na intensidade do valor do intervalo de confiança. O intervalo de confiança dos resultados do filtro com NLMS é pior que os filtros com RLS.

Figura 31 – Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo NLMS e com 500 séries de treinamento

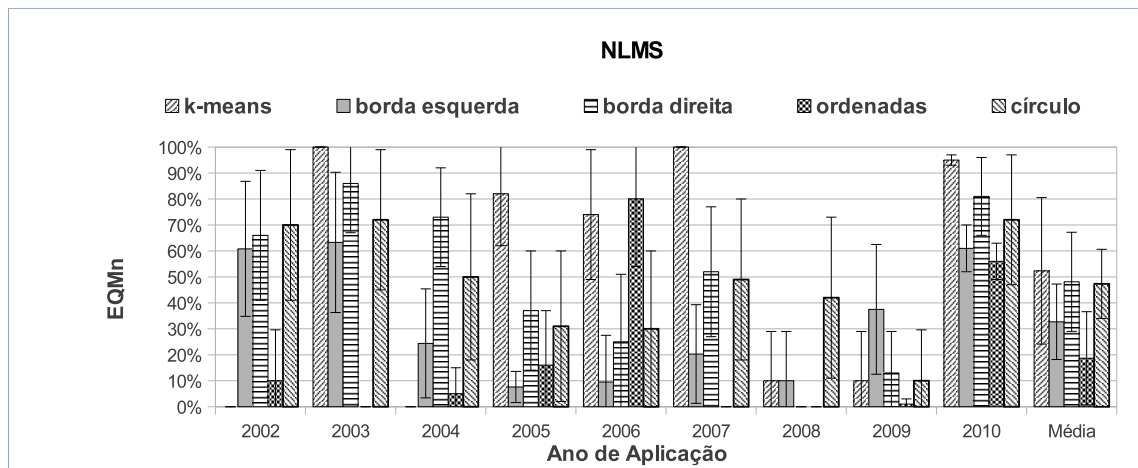
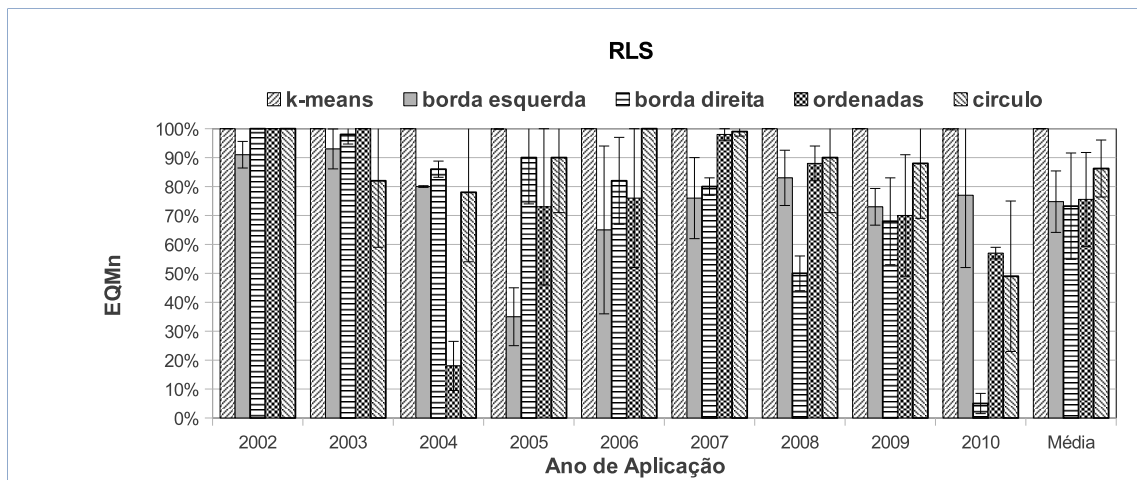


Figura 32 – Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo RLS e com 500 séries de treinamento



As Figuras 33 e 34 mostram gráficos dos resultados da porcentagem de regiões com redução no EQMn da variável altura geopotencial e aproximadamente 1000 séries de treinamento e com filtro de ordem correspondente a um dia de previsões treinado pelos algoritmos NLMS e RLS respectivamente. Ainda pelos gráficos percebe-se que os intervalos de confiança dos resultados dos filtros adaptados pelo algoritmo RLS são menores que os intervalos de confiança dos resultados dos filtros adaptados pelo algoritmo NLMS. Percebe-se ainda queda na redução do EQMn nos anos de 2003, 2004, 2007 e 2010 no caso dos filtros adaptados pelo algoritmo RLS. Nestes anos o critério de seleção de regiões da borda inferior não apresentou bons resultados.

Figura 33 – Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo NLMS e com 1000 séries de treinamento

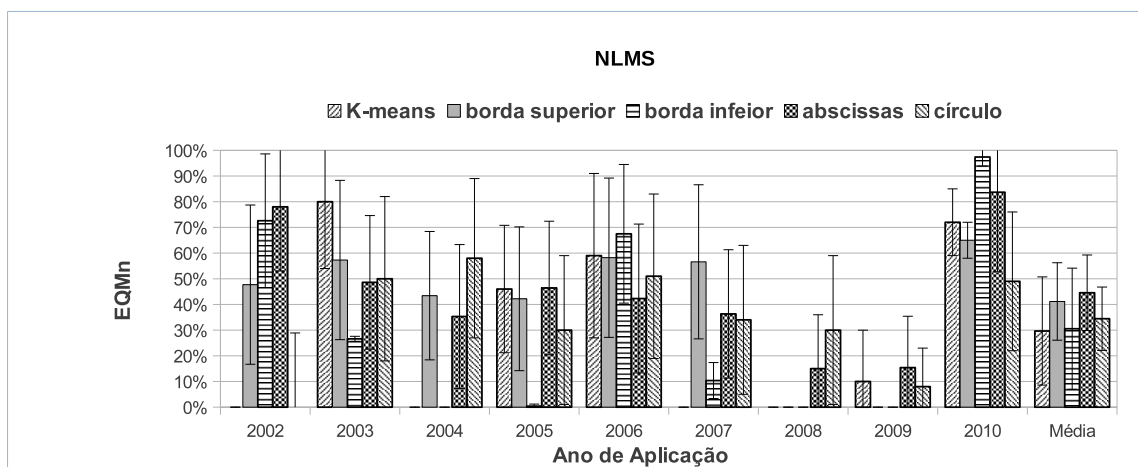
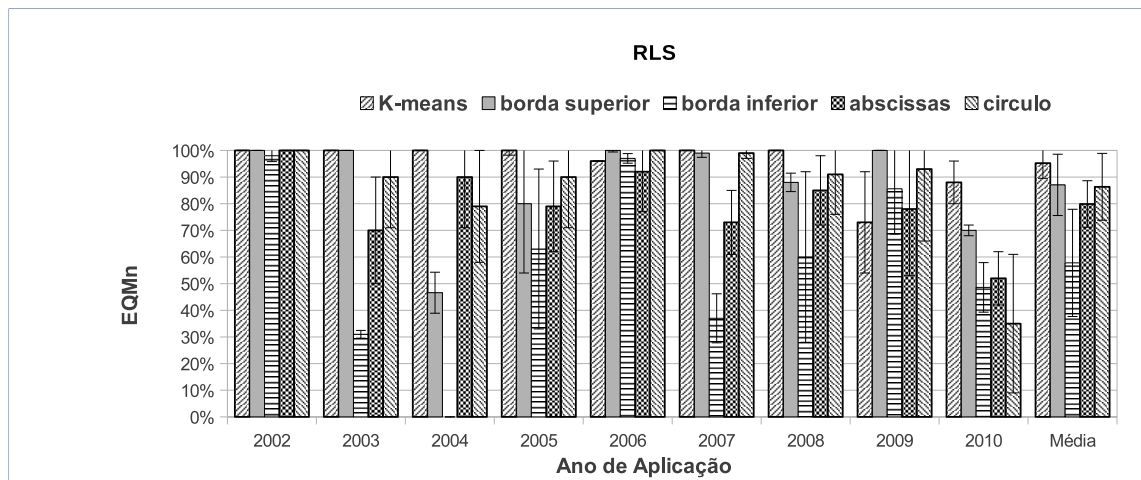


Figura 34 – Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 4 coeficientes, adaptado pelo algoritmo RLS e com 1000 séries de treinamento



As Figuras 35 e 36 mostram gráficos dos resultados da porcentagem de regiões com redução no EQMn da variável altura geopotencial e aproximadamente 500 séries de treinamento e com filtro de ordem correspondente a oito dias de previsões treinado pelos algoritmos NLMS e RLS respectivamente. Por estes gráficos, percebe-se que os filtros adaptados pelo algoritmo RLS apresentam melhor desempenho do que os filtros adaptados pelo algoritmo NLMS. Percebe-se inclusive, que para estas configurações os resultados dos filtros adaptados pelo NLMS são ruins, pois a redução na média do EQMn não ultrapassa 45% de regiões.

Figura 35 – Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo NLMS e com 500 séries de treinamento

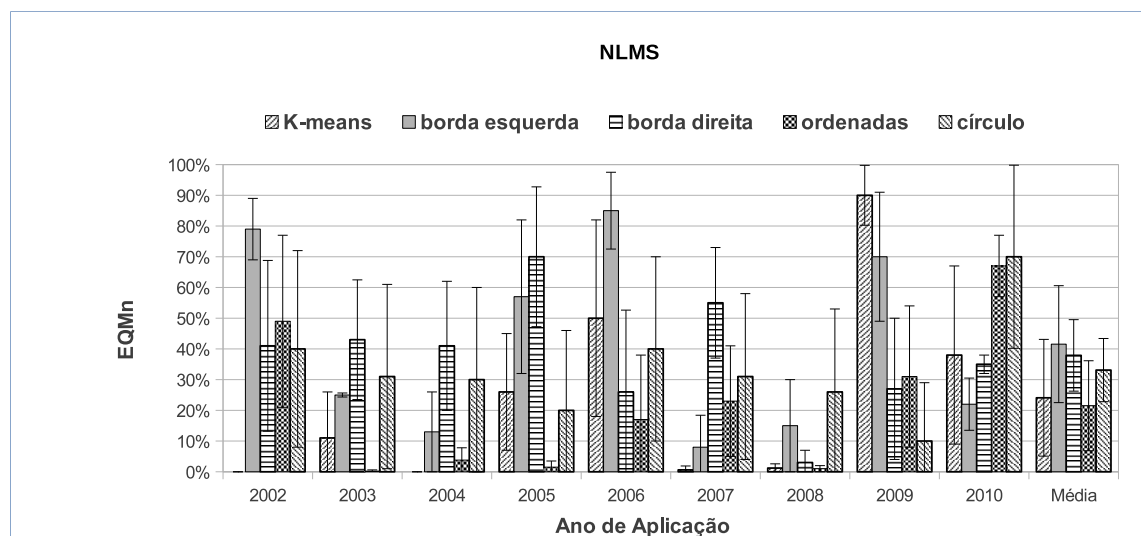
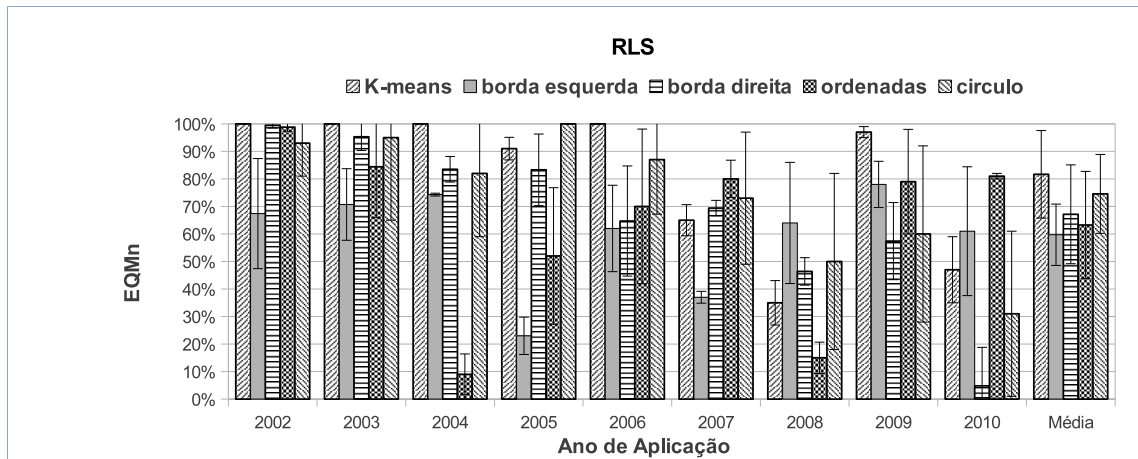


Figura 36 – Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo RLS e com 500 séries de treinamento



As Figuras 37 e 38 mostram gráficos dos resultados da porcentagem de regiões com redução no EQMn da variável altura geopotencial e aproximadamente 1000 séries de treinamento e com filtro de ordem correspondente a oito dias de previsões treinado pelos algoritmos NLMS e RLS respectivamente. Por estes gráficos percebe-se tanto os filtros adaptados pelo algoritmo RLS apresentam na média resultados superiores aos filtros adaptados pelo algoritmo NLMS. Entretanto, os filtros adaptados pelo algoritmo RLS também apresentam resultados com certa variação na porcentagem de redução do EQMn. Na média, o critério de seleção de regiões circulares foi o que apresentou o melhor resultado com redução no EQMn de 76% da região.

Figura 37 – Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo NLMS e com 1000 séries de treinamento

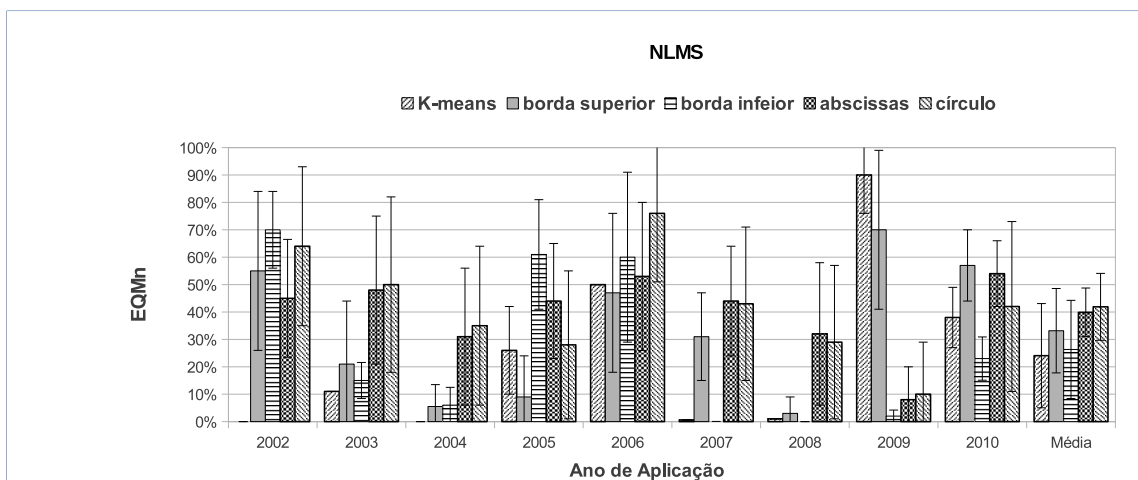
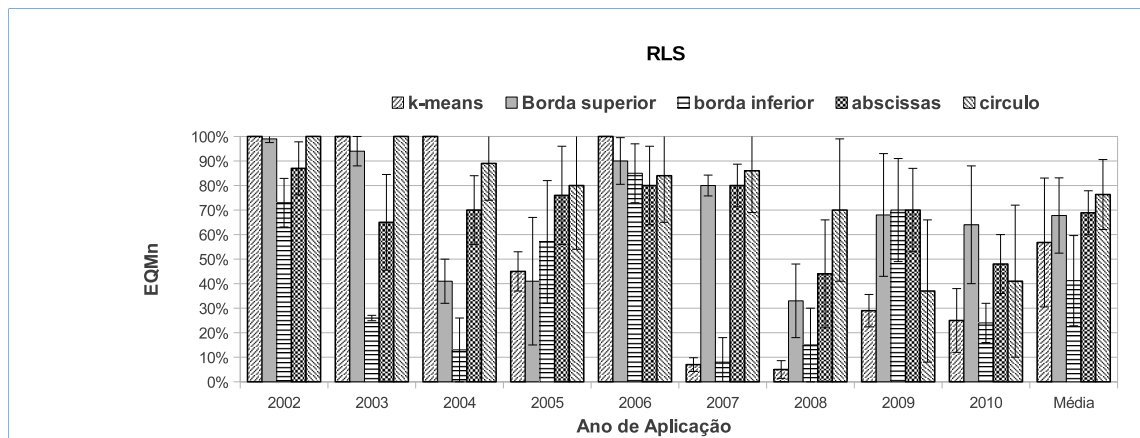


Figura 38 – Porcentagem de regiões com redução no EQMn na variável altura geopotencial, na aplicação do filtro, de 32 coeficientes, adaptado pelo algoritmo RLS e com 1000 séries de treinamento



Pelos gráficos das Figuras 31 a 38 pode-se perceber que o algoritmo RLS não só reduz o EQMn em uma porcentagem superior ao NLMS como na média o intervalo de confiança é menor, denotando assim maior confiabilidade nos resultados obtidos. Percebe-se ainda que diante de outros critérios de seleção de regiões o método utilizado o algoritmo *k-means* se destaca por apresentar resultados melhores, tanto no uso do algoritmo de adaptação NLMS quanto no RLS. Por fim pode-se ver que filtros com comprimento de 1 dia de previsões (4 coeficientes) apresentaram na média resultados superiores a filtro com 8 dias de previsões (32 coeficientes).

De forma geral, assim como foi na variável vento meridional, percebe-se que pelas Tabelas 2 a 5 e pelos gráficos das Figuras 31 a 38 que filtros adaptados utilizando algoritmo RLS reduziram o erro em porcentagens maiores de regiões do que os obtidos via NLMS. Ou seja, dentre as estratégias utilizadas no treinamento do filtro, o algoritmo RLS consegue adaptar melhor os filtros no intuito de reduzir o erro de previsão da variável altura geopotencial.

B. Critério de seleção: Anéis retangulares

As colunas das Tabelas 10 e 11 apresentam os resultados obtidos com o critério de seleção de regiões por anéis retangulares concêntricos, para o conjunto de treinamento na aplicação dos algoritmos adaptativos NLMS e RLS.

A Tabela 10 mostra resultados da porcentagem de regiões com redução no EQMn da variável vento meridional com regiões selecionadas por anéis retangulares e com filtro de ordem correspondente a um dia de previsões. De acordo com os resultados da tabela, percebe-se que o uso de filtros adaptados com o algoritmo RLS reduziu na média o EQMn em 97% da região anular da borda, 98% da região anular intermediária e 98% da região anular interior. Assim percebe-se que o RLS obteve na média a redução do EQMn em

mais regiões que o NLMS.

Tabela 10 – Porcentagem de regiões que se obtém redução no EQMn da variável vento meridional aplicando filtros de 1 dia (4 coeficientes)

Ano de aplicação	Bordas		Intermediário		Interior	
	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	66	95	52	100	69	100
2003	33	100	48	100	98	100
2004	49	93	63	95	80	92
2005	29	99	36	100	25	100
2006	34	98	26	100	90	100
2007	76	96	96	96	87	95
2008	34	99	25	93	9	92
2009	51	100	50	100	31	100
2010	33	95	50	100	80	100
Média	45	97	50	98	63	98

A Tabela 11 mostra resultados da porcentagem de regiões com redução no EQMn da variável vento meridional com regiões selecionadas por anéis retangulares e com filtro de ordem correspondente a oito dias de previsões. Pela tabela, percebe-se que o uso de filtros adaptados com o algoritmo RLS reduziu na média o EQMn em 91% da região anular da borda, 94% da região anular intermediária e 88% da região anular interior. Percebe-se que o RLS obteve na média a redução do EQMn em mais regiões que o NLMS, entretanto, os resultados da redução do EQMn para o caso do RLS foram na média ligeiramente inferiores aos resultados da Tabela 10

Tabela 11 – Porcentagem de regiões que se obtém redução no EQMn da variável vento meridional aplicando filtros de 8 dias (32 coeficientes)

Ano de aplicação	Bordas		Intermediário		Interior	
	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	54	94	59	97	48	100
2003	57	96	20	97	25	100
2004	66	90	52	91	43	90
2005	46	94	50	91	1	94
2006	69	89	64	97	98	99
2007	70	93	73	97	88	94
2008	70	83	30	84	31	49
2009	66	96	38	99	40	66
2010	31	87	25	90	38	100
Média	59	91	46	94	46	88

A partir desses dados sabe-se que o uso de filtros com comprimento equivalente a um dia de previsão gera na média a redução no EQMn em uma fração de regiões maior do que se obteve com filtros de comprimento equivalente a oito dias de previsões, principalmente se compararmos aqueles filtros adaptados através do algoritmo RLS. Percebe-se que os filtros adaptados usando algoritmo RLS reduzem o EQMn de séries ajustadas em mais regiões que os filtros adaptados pelo algoritmo NLMS. Pela Tabela 10 observa-se ainda, que na média, as três regiões anulares concêntricas (Borda, Intermediária e Interior) reduziram na média o EQMn em porcentagem semelhantes, indicando que ao se aplicar este método de modo que toda a região do modelo seja contemplada, da borda até o centro, os resultados da redução do EQMn serão numericamente semelhantes entre si, ou seja, em torno de 98%.

As Figuras 39 e 40, mostram gráficos dos resultados da porcentagem de regiões com redução no EQMn da variável vento meridional com filtro de ordem correspondente a um dia de previsões treinado pelos algoritmos NLMS e RLS, respectivamente. A Figura 39 mostra que o filtro adaptado pelo algoritmo NLMS reduz o EQMn de forma muito variável pois o intervalo de confiança é grande. Na Figura 40 a redução do EQMn é feita de forma mais estável, próximo de 100%, além disso o intervalo de confiança é pequeno.

Figura 39 – Porcentagem de regiões com redução no EQMn da variável vento meridional, na aplicação do filtro de 4 coeficientes, adaptado pelo algoritmo NLMS

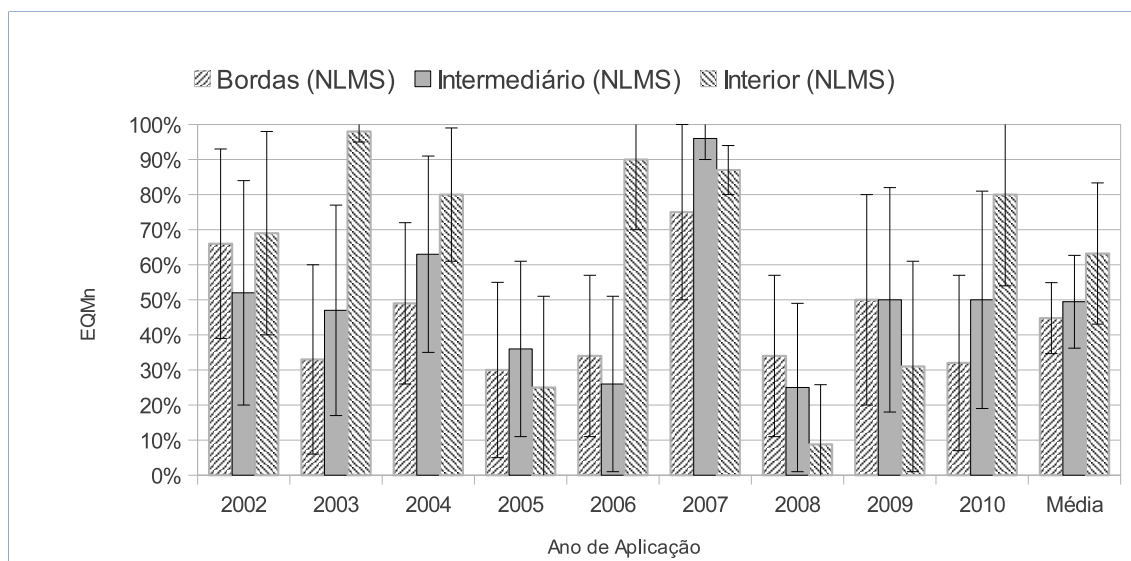
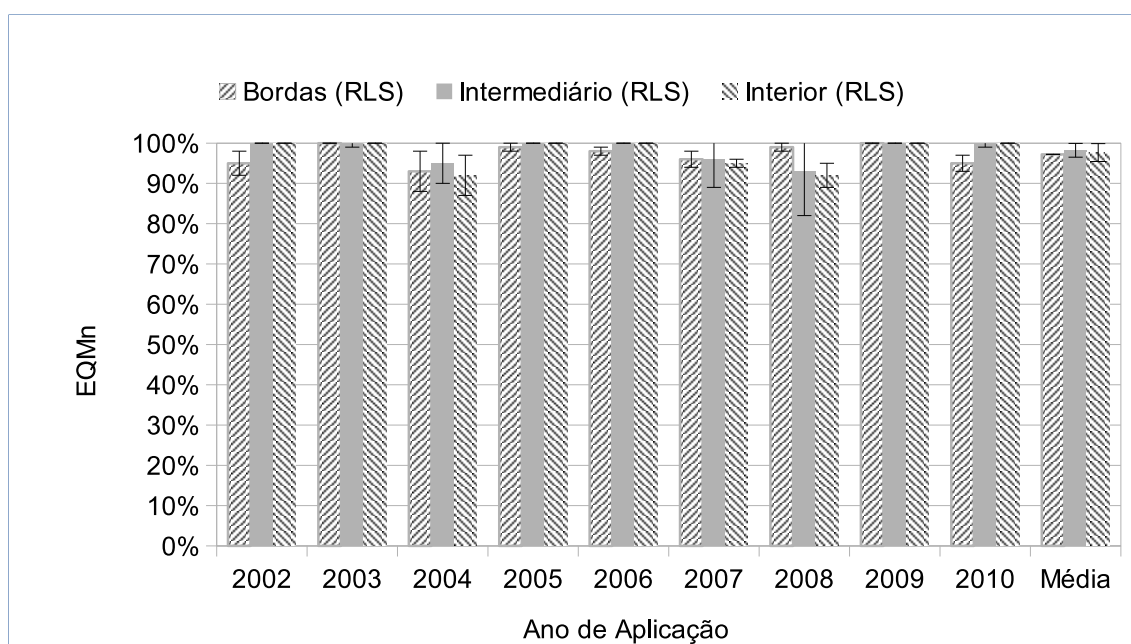


Figura 40 – Porcentagem de regiões com redução no EQMn da variável vento meridional, na aplicação do filtro de 4 coeficientes, adaptado pelo algoritmo RLS



As Figuras 41, 42, mostram gráficos dos resultados da porcentagem de regiões com redução no EQMn da variável vento meridional com filtro de ordem correspondente a oito dias de previsões treinado pelos algoritmos NLMS e RLS, respectivamente. A Figura 41, mostra que o filtro adaptado pelo algoritmo NLMS apresenta grande intervalo de confiança na redução do EQMn sobre as regiões. A Figura 42 mostra que a redução do EQMn é feita de forma mais estável na maioria dos casos, além disso o intervalo de confiança na média é menor que no caso do NLMS.

Figura 41 – Porcentagem de regiões com redução no EQMn da variável vento meridional, na aplicação do filtro de 32 coeficientes, adaptado pelo algoritmo NLMS

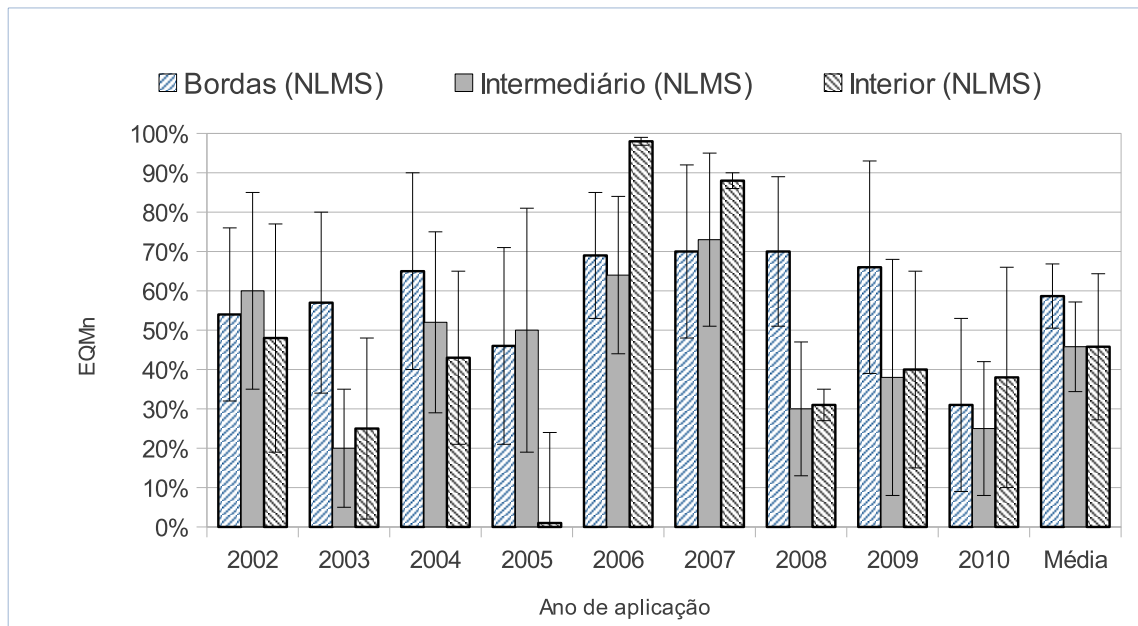
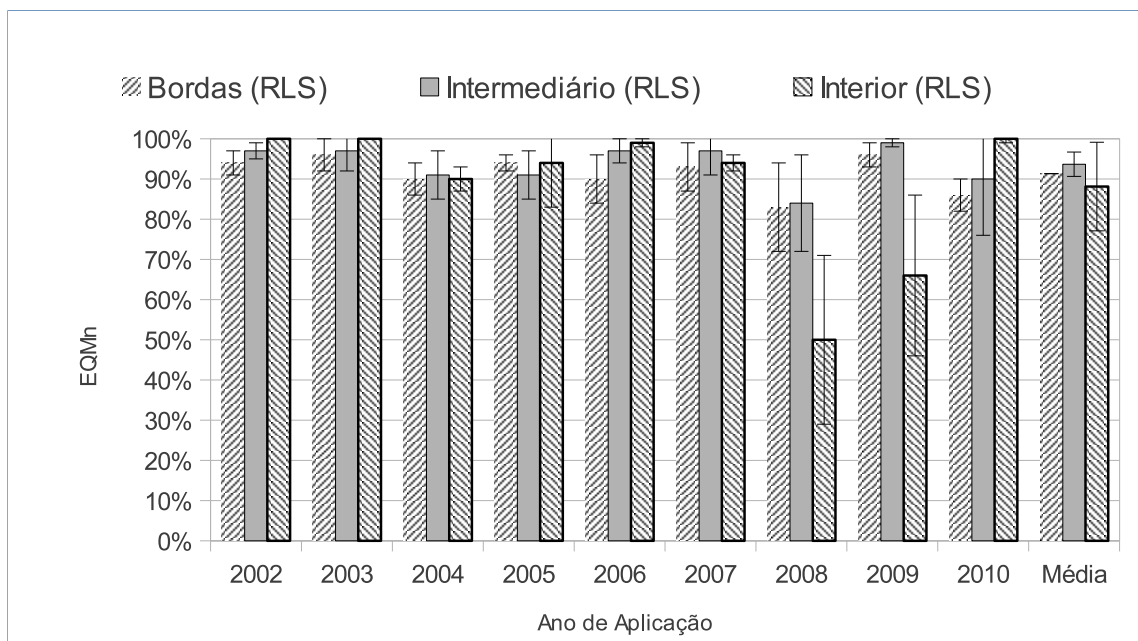


Figura 42 – Porcentagem de regiões com redução no EQMn da variável vento meridional, na aplicação do filtro de 32 coeficientes, adaptado pelo algoritmo RLS



Comparando os resultados dos gráficos da Figuras 40 com os resultados das Figuras 39, 41 e 42 percebe-se que o uso do algoritmo RLS para treinar o filtro com 4 coeficientes gerou a redução do EQMn de mais regiões do que as outras configurações. Nota-se ainda que o intervalo de confiança é menor para o caso da Figura 40 se comparado com os outros resultados.

As colunas das Tabelas 12 e 13 apresentam os resultados obtidos com o critério de

seleção de regiões por Anéis, para o conjunto de treinamento na aplicação dos algoritmos adaptativos NLMS e RLS sobre a variável altura geopotencial. Pela tabela 12, percebe-se que o uso de filtros adaptados com o algoritmo RLS reduziu na média o EQMn em 66% da região anular da borda, 71% da região anular intermediária e 86% da região anular interior. Percebe-se que o RLS obteve na média a redução do EQMn em mais regiões que o NLMS.

Tabela 12 – Porcentagem de regiões que se obtém redução no EQMn da variável altura geopotencial aplicando filtros de 1 dia (4 coeficientes)

Ano de aplicação	Bordas		Intermediário		Interior	
	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	37	88	75	95	60	100
2003	60	71	65	70	69	99
2004	36	46	45	67	42	96
2005	28	51	48	79	40	100
2006	32	83	55	96	65	100
2007	27	57	40	66	63	68
2008	0	65	16	78	47	82
2009	10	81	0	47	30	92
2010	52	48	43	40	75	36
Média	31	66	43	71	55	86

A Tabela 13 mostra resultados da porcentagem de regiões com redução no EQMn da variável altura geopotencial com regiões selecionadas por anéis retangulares e com filtro de ordem correspondente a oito dias de previsões. Pela tabela, percebe-se que o uso de filtros adaptados com o algoritmo RLS reduziu na média o EQMn em 59% da região anular da borda, 60% da região anular intermediária e 88% da região anular interior. Percebe-se que o RLS obteve na média a redução do EQMn em mais regiões que o NLMS.

Tabela 13 – Porcentagem de regiões que se obtém redução no EQMn da variável altura geopotencial aplicando filtros de 8 dias (32 coeficientes)

Ano de aplicação	Bordas		Intermediário		Interior	
	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	63	88	66	79	66	100
2003	40	70	36	61	63	95
2004	23	50	22	48	45	90
2005	38	56	54	72	34	94
2006	66	91	48	74	34	100
2007	16	44	53	61	48	74
2008	14	32	11	67	44	62
2009	27	59	34	36	44	90
2010	47	42	32	40	65	86
Média	37	59	40	60	49	88

A partir desses dados sabe-se que o uso de filtros com comprimento equivalente a um dia de previsão gera na média a redução no EQMn em uma fração de regiões maior do que se obteve com filtros de comprimento equivalente a oito dias de previsões. Percebe-se que os filtros adaptados usando algoritmo RLS reduzem o EQMn de séries ajustadas em mais regiões que os filtros adaptados pelo algoritmo NLMS. Observa-se ainda, que na média, conforme as regiões se aproximam do centro os resultados melhoram. Entretanto na média os resultados foram inferiores aos da variável vento meridional. Por exemplo, na Tabela 10 o filtro adaptado pelo algoritmo RLS reduziu na média o EQMn em 98% da região intermediária, enquanto que pela Tabela 12 o filtro adaptado pelo algoritmo RLS reduziu na média o EQMn em 71% da região intermediária.

As Figuras 43 e 44 mostram a porcentagem de regiões com redução no EQMn da variável altura geopotencial com filtro de ordem correspondente a um dia de previsões treinado pelos algoritmos NLMS e RLS, respectivamente.

Percebe-se pelo gráfico da Figura 43 que o filtro adaptado pelo algoritmo NLMS apresenta grande intervalo de confiança na redução do EQMn sobre as regiões. Percebe-se ainda que os anos de 2008 e 2009 foram ruins. Na Figura 44 a redução do EQMn é feita de forma mais estável na maioria dos casos, além disso o intervalo de confiança na média é menor que no caso do NLMS. Percebe-se ainda uma melhora nos resultados para o ano de 2008 e 2009.

Figura 43 – Porcentagem de regiões com redução no EQMn da variável altura geopotencial, na aplicação do filtro de 4 coeficientes, adaptado pelo algoritmo NLMS

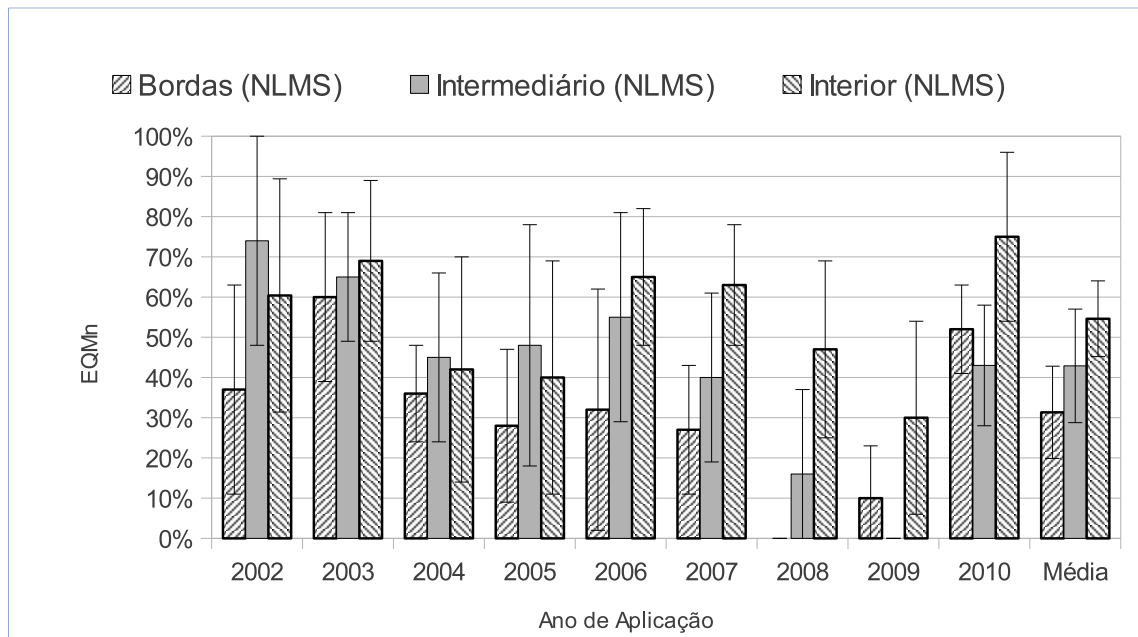
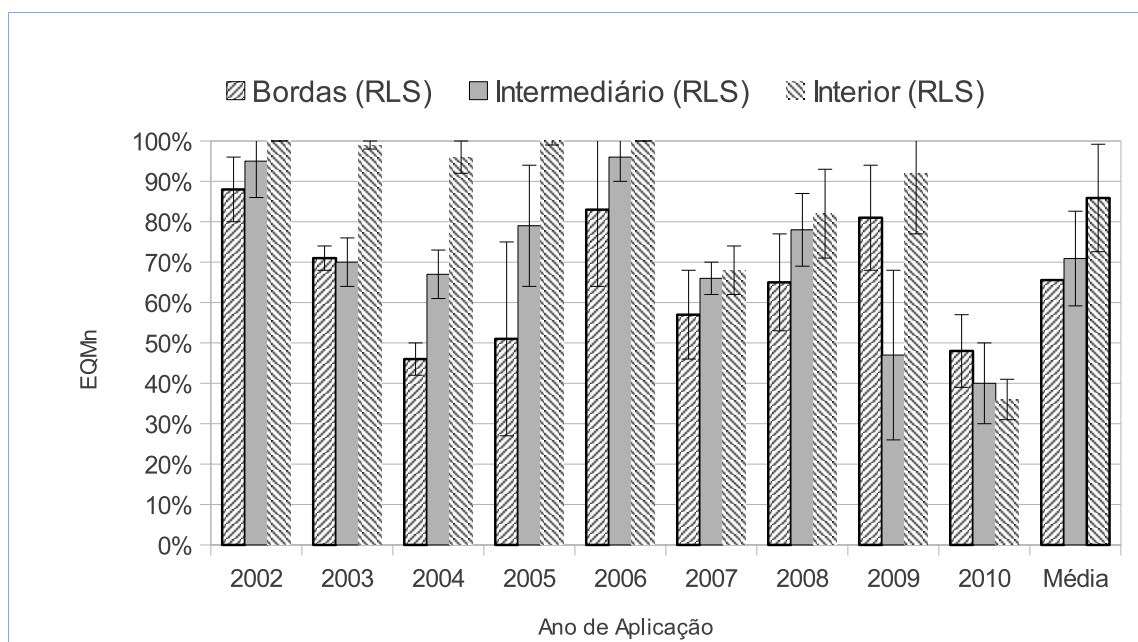


Figura 44 – Porcentagem de regiões com redução no EQMn da variável altura geopotencial, na aplicação do filtro de 4 coeficientes, adaptado pelo algoritmo RLS



As Figuras 45 e 46, mostram gráficos dos resultados da porcentagem de regiões com redução no EQMn da variável altura geopotencial com filtro de ordem correspondente a um dia de previsões treinado pelos algoritmos NLMS e RLS respectivamente. Percebe-se pelo gráfico da Figura 45 que o filtro adaptado pelo algoritmo NLMS apresenta grande intervalo de confiança na redução do EQMn sobre as regiões. Percebe-se ainda que na média para as três regiões anulares não se reduz o EQMn em mais de 50% das regiões.

Na Figura 46 a redução do EQMn é feita de forma mais estável na maioria dos casos, principalmente na região anular interior, além disso o intervalo de confiança na média é menor que no caso do NLMS.

Figura 45 – Porcentagem de regiões com redução do EQMn sobre a variável altura geopotencial, na aplicação do filtro de 32 coeficientes, adaptado pelo algoritmo NLMS

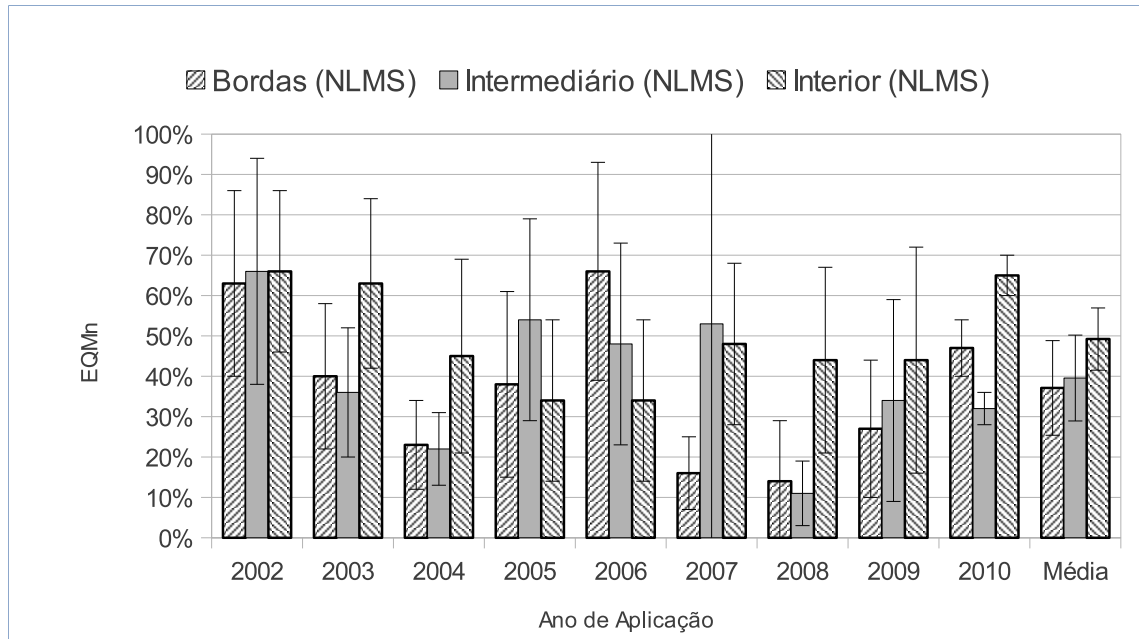
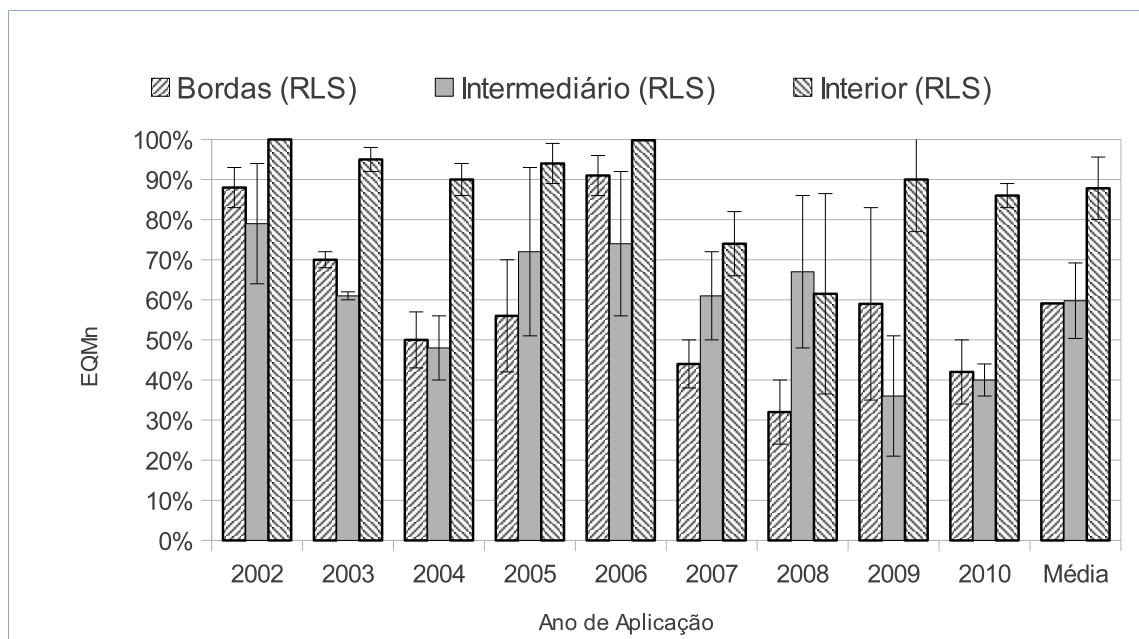


Figura 46 – Porcentagem de regiões com redução do EQMn sobre a variável altura geopotencial, na aplicação do filtro de 32 coeficientes, adaptado pelo algoritmo RLS



Percebe-se que pelos gráficos das Figuras 44 a 46 que o uso do algoritmo RLS para

treinar o filtro resultou na redução do EQMn em mais regiões do que o uso do algoritmo NLMS. Entretanto percebe-se que na média o intervalo de confiança no caso de filtros adaptados pelo RLS é menor do que o caso se filtros adaptados pelo NLMS.

4.2.1 Análises específicas

Com o objetivo de avaliar com mais detalhes o efeito do filtro sobre as séries ajustadas, selecionaram-se séries que obtiveram redução do EQMn em relação ao NCEP. Com esta seleção pretende-se verificar a intensidade de redução do EQMn em cada caso estudado. Outro ponto importante é verificar os efeitos nas séries do uso de um filtro adaptado com um dia de previsão e o filtro adaptado com oito dias de previsão. O ano de escolhido para o treinamento do filtro foi o ano de 2007, e aplicou-se o filtro adaptado sobre o ano de 2008. O algoritmo utilizado foi o algoritmo RLS utilizando o critério de seleção de regiões dado pelo *k-means*.

A. Componente meridional do vento

As Figuras 47 e 48 mostram o EQMn ao longo da posição geográfica, referente ao ajuste das séries do modelo Eta após a aplicação do filtro com comprimento de um dia de previsões e com oito dias de previsões. Percebe-se que ambos os filtros reduziram o EQMn para toda a região. Entretanto ao se comparar as duas figuras, percebe-se que a intensidade de redução do EQMn para o filtro de um dia de previsão foi maior.

Figura 47 – EQMn antes e depois da aplicação do filtro de comprimento de um dia de previsões nas séries climáticas do Eta sobre a variável vento meridional

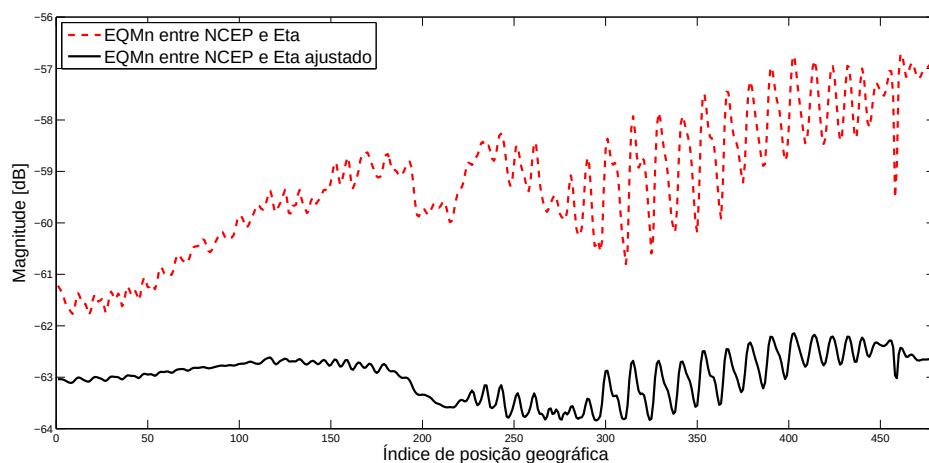
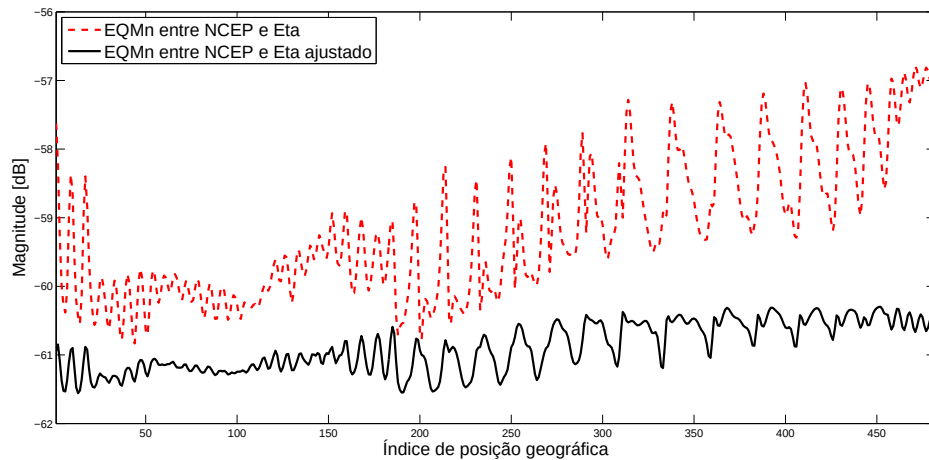


Figura 48 – EQMn antes e depois da aplicação do filtro de comprimento de oito dias de previsões nas séries climáticas do Eta sobre a variável vento meridional



Para identificar a distribuição do EQMn ao longo dos dias de previsões, utilizou-se uma janela deslizante com 4 coeficientes para filtros de 4 coeficientes e uma janela com 32 coeficientes para filtros com 32 coeficientes. Assim, se calculou a média do EQMn de todas as séries climáticas.

Nas Figuras 49 e 50 o EQMn é o resultado da média de todas as realizações na aplicação de filtros com 4 coeficientes e com 32 coeficientes respectivamente sobre a variável vento meridional. Dessa forma percebe-se que ao longo do tempo o uso do filtro com 4 coeficientes teve o EQMn reduzido em 95,89% do tempo de previsão enquanto que para o filtro de 32 coeficientes o EQMn foi reduzido em 75,89% do tempo de previsão.

Figura 49 – Média do EQMn por janela, na aplicação do filtro com comprimento de um dia de previsões sobre a variável vento meridional

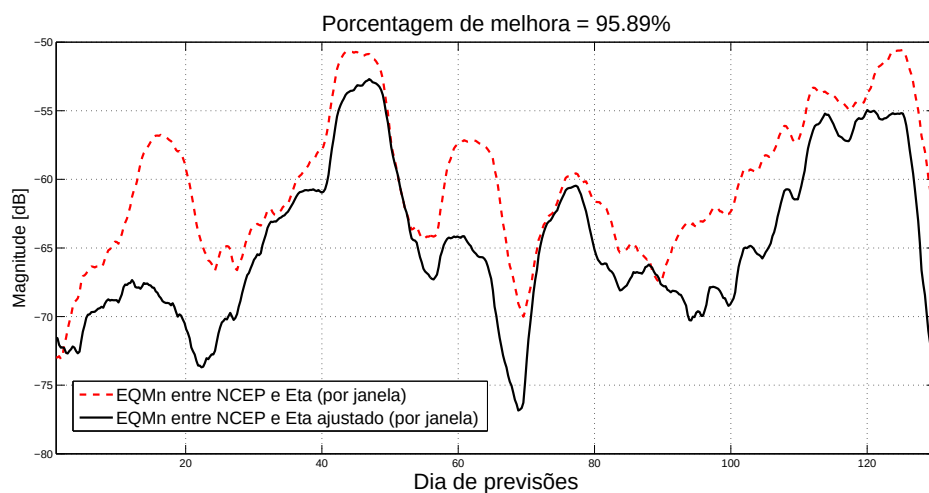
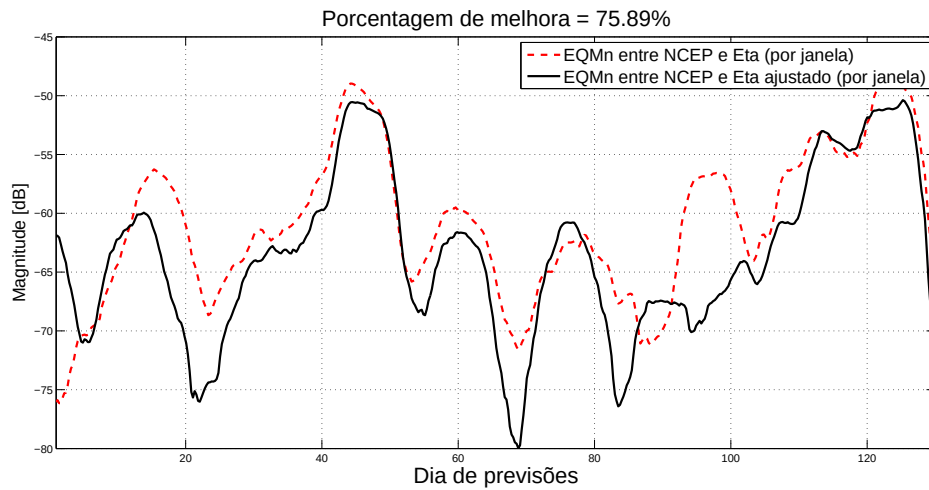


Figura 50 – Média do EQMn por janela, na aplicação do filtro com comprimento de oito dias de previsões sobre a variável vento meridional



As Figuras 51 e 52 mostram a distribuição estatística do EQMn na aplicação de filtros com 4 coeficientes e com 32 coeficientes respectivamente sobre a variável vento meridional. No caso das Figuras 51 e 52 percebe-se que houve redução na média do EQMn e redução da dispersão do EQMn em ambos os casos. Pela Figura 51 percebe-se que na média o EQMn reduziu em $3,87dB$, enquanto que pela Figura 52 percebe-se que na média o EQMn reduziu em $2,07dB$. Assim, neste caso, o filtro com 4 coeficiente reduziu na média o EQMn em maior intensidade que o filtro com 32 coeficientes.

Figura 51 – Distribuição da média do EQMn por janela, na aplicação do filtro com um dia de previsões sobre a variável vento meridional

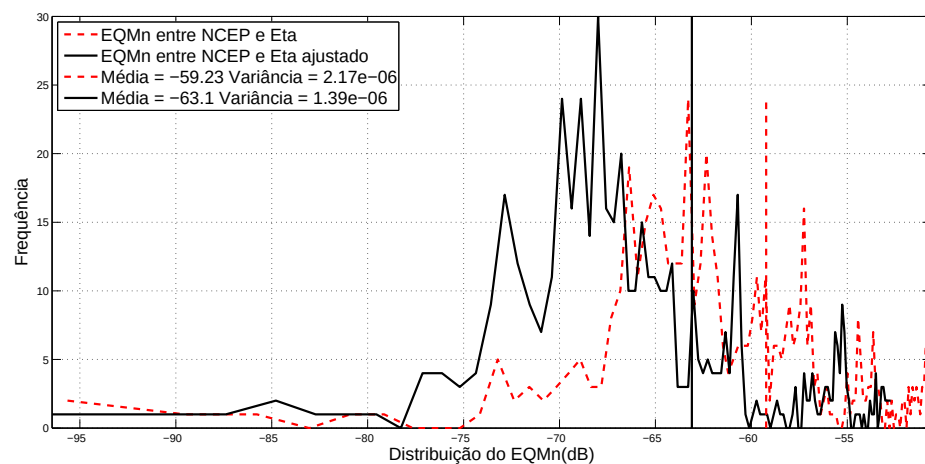
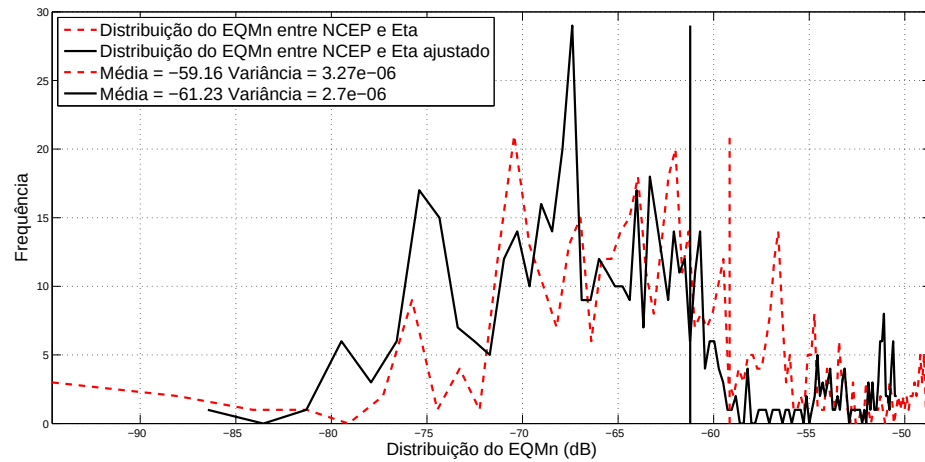


Figura 52 – Distribuição da média do EQMn por janela, na aplicação do filtro com oito dias de previsões sobre a variável vento meridional



Com base nos resultados apresentados, percebe-se que na média a intensidade de redução EQMn provocada pela aplicação do filtro com 32 coeficientes, isto é 8 dias de previsões, é superior à intensidade de redução do EQMn provocada pela aplicação do filtro com 4 coeficientes, isto é um dia de previsões. Ainda a aplicação de ambos os filtros resulta na redução do EQMn em 100% da região.

Para avaliar o efeito do ajuste dos filtros sobre as séries climáticas, colheram-se duas séries temporais do estudo de caso apresentado, uma para a aplicação do filtro com 4 coeficientes e outra para a aplicações do filtro com 32 coeficientes.

A Figura 53 mostra a série que teve seu ajuste feito por um filtro com 4 coeficientes enquanto que a Figura 54 mostra a série que teve seu ajuste feito com um filtro com 32 coeficientes. As Figuras 53 e 54 também mostram a série de referência do NCEP e a série Eta. Pela Figura 53 próximo ao vigésimo segundo dia de previsão, percebe-se que a série do Eta ajustado, se aproximou mais da série de referência do NCEP. Ainda, pela Figura 54, alguns dias após o vigésimo dia de previsão, percebe-se que a série do Eta ajustado, também se aproximou mais da série de referência do NCEP.

Figura 53 – Série climática da variável vento meridional referente as dados do Eta, NCEP e Eta ajustado com filtro de 4 coeficientes

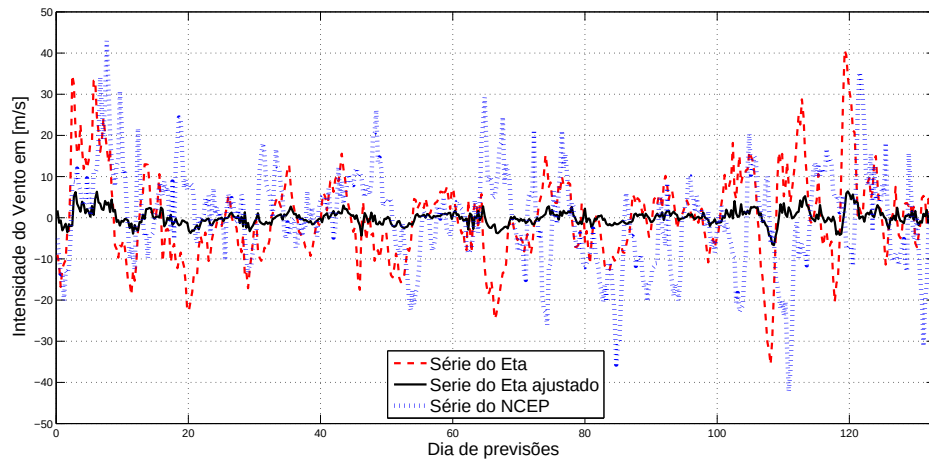
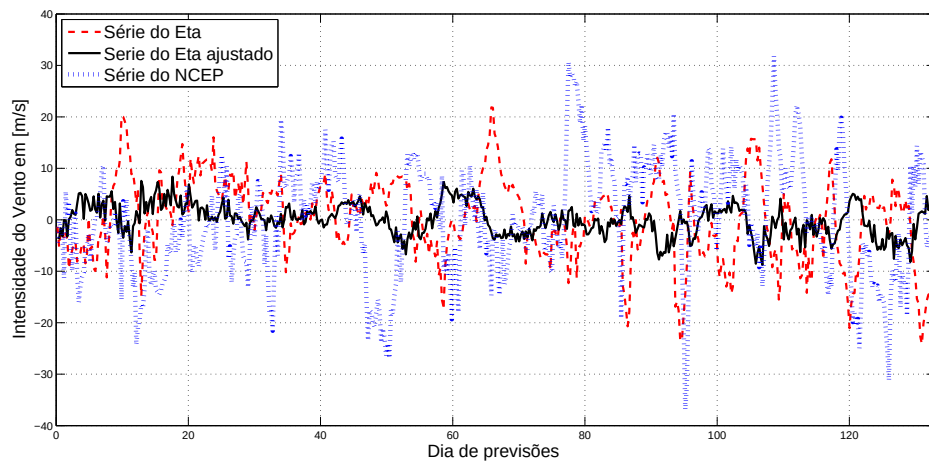


Figura 54 – Série climática da variável vento meridional referente as dados do Eta, NCEP e Eta ajustado com filtro de 32 coeficientes



As Figuras 55 mostram o erro entre a série do Eta e a séries do NCEP em linha tracejada e o erro entre a série do Eta ajustado por um filtro de 4 coeficientes e a série do NCEP em linha cheia. A Figura 56 mostra o erro entre a série do Eta e a séries do NCEP em linha tracejada e o erro entre a série do Eta ajustado por um filtro com comprimento de 32 coeficientes e a série do NCEP em linha cheia. Em ambas pode-se ver que após o dia 20 há um período em que há redução no erro.

Figura 55 – Erro entre a série climática da variável vento meridional referente as dados do Eta e série do NCEP em linha tracejada e erro entre a série do Eta ajustado com filtro de 4 coeficientes e a série do NCEP em linha cheia

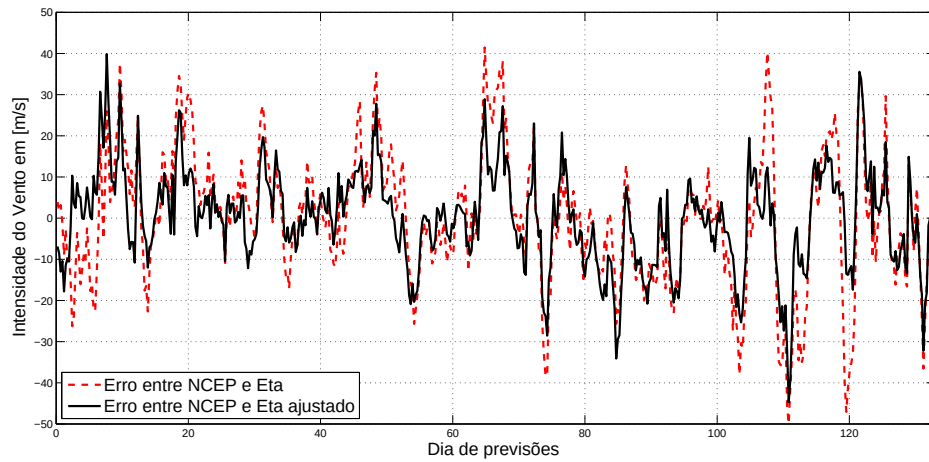
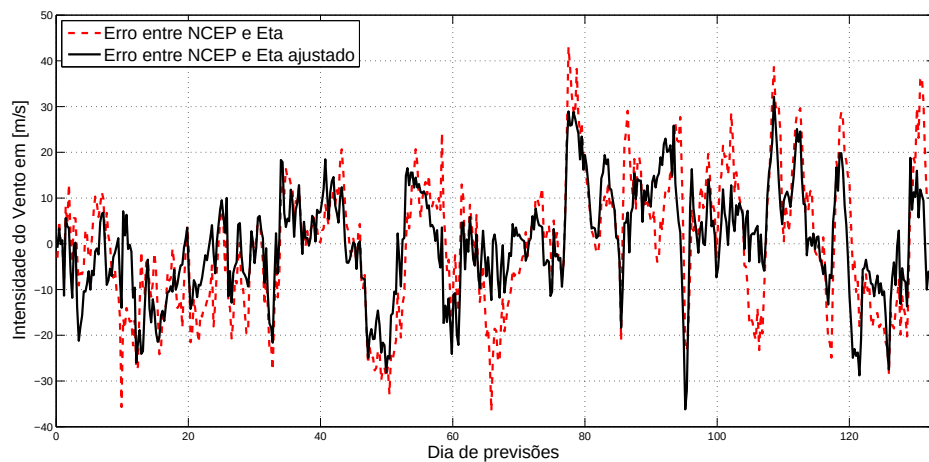


Figura 56 – Erro entre a série climática da variável vento meridional referente as dados do Eta e série do NCEP em linha tracejada e erro entre a série do Eta ajustado com filtro de 32 coeficientes e a série do NCEP em linha cheia



Com o objetivo de identificar a adesão da série ajustada à série do NCEP, com relação aos dados do Eta original, para estas séries calculou-se o EQMn entre as séries do Eta e o NCEP, e a série do Eta ajustado e o NCEP. Para identificar a distribuição do EQMn utilizou-se uma janela deslizante (móvel) com 4 coeficientes (um dia de previsões) e com 32 coeficientes (oito dias de previsões).

A Figura 57 mostra o EQMn por uma janela de 4 coeficientes ao longo dos dias de previsões e a Figura 58 mostra o EQMn por uma janela de 32 coeficientes ao longo dos dias de previsões. Pela Figura 57 percebe-se que o filtro com 4 coeficientes reduziu o EQMn em 69,72% do tempo de previsão, e pela Figura 58 pode-se ver que o filtro com

32 coeficientes reduziu o EQMn em 86,73% do tempo de previsão. Neste caso percebe-se que o filtro com 32 coeficientes apresentou melhores resultados.

Figura 57 – EQMn por janela, na aplicação do filtro com comprimento de um dia de previsões sobre a variável vento meridional

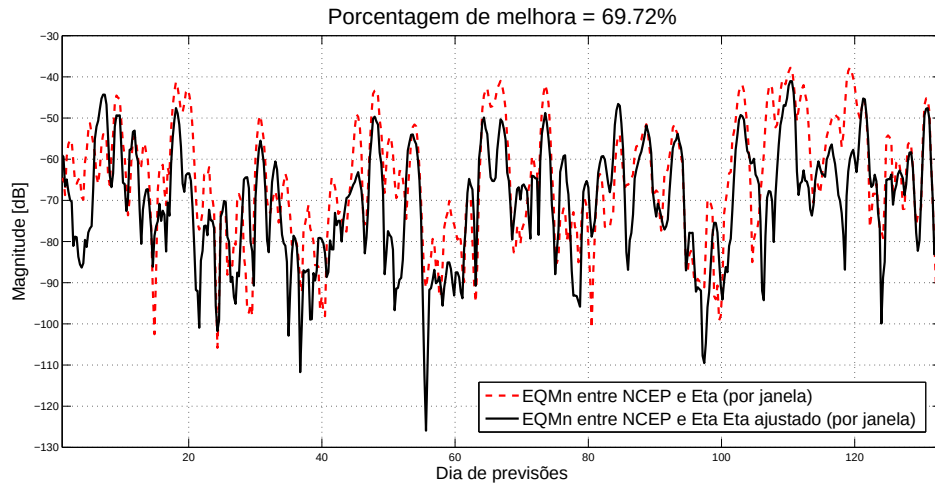
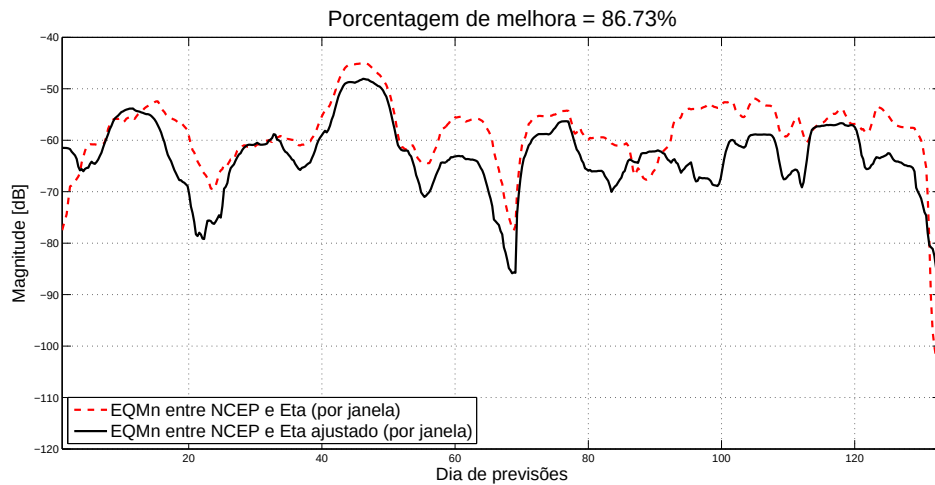


Figura 58 – EQMn por janela, na aplicação do filtro com comprimento de oito dias de previsões sobre a variável vento meridional



As Figuras 59 e 60 mostram a distribuição estatística do EQMn na aplicação do filtro com 4 coeficientes e com 32 coeficientes respectivamente sobre a variável vento meridional. No caso das Figuras 59 e 60 percebe-se que houve redução na média do EQMn e redução da dispersão do EQMn em ambos os casos. Pela Figura 59 percebe-se que na média o EQMn reduziu em 5,89dB, enquanto que pela Figura 60 percebe-se que na média o EQMn reduziu em 4,13dB. Assim, neste caso, o filtro com 4 coeficiente reduziu na média o EQMn em maior intensidade que o filtro com 32 coeficientes.

Figura 59 – Distribuição do EQMn por janela, na aplicação do filtro com comprimento de um dia de previsões sobre a variável vento meridional

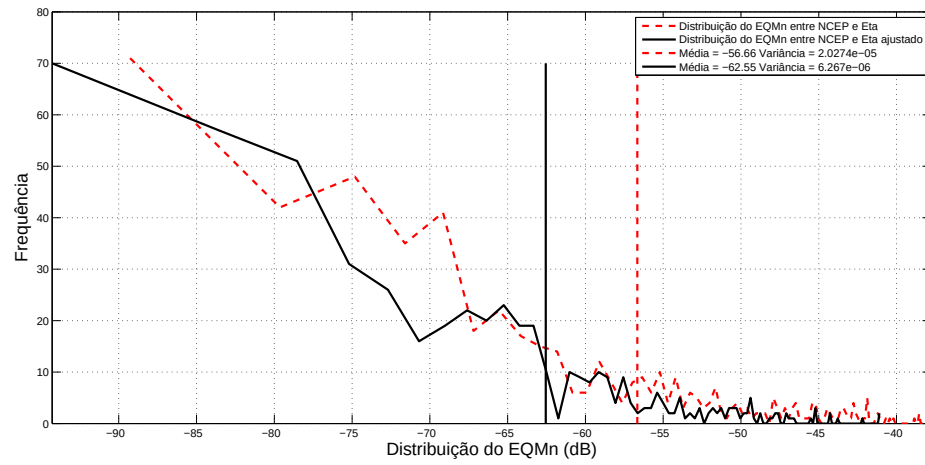
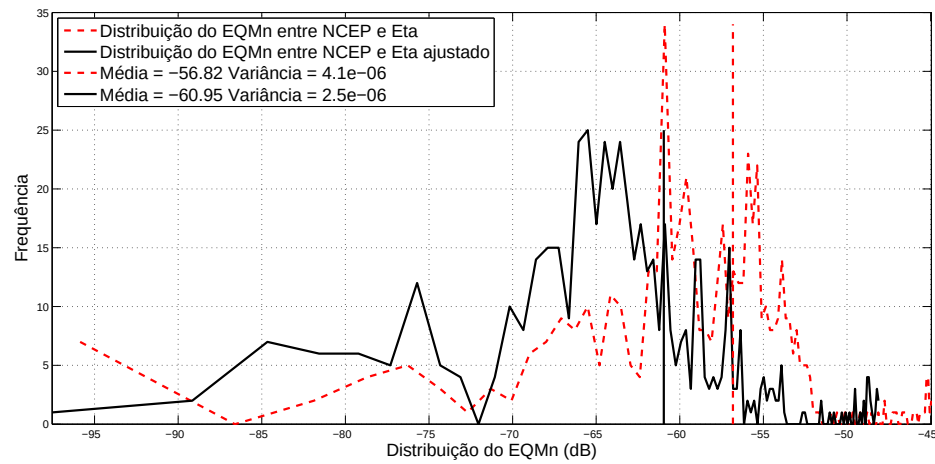


Figura 60 – Distribuição do EQMn por janela, na aplicação do filtro com comprimento de oito dias de previsões sobre a variável vento meridional



As Figuras 61 e 62 apresenta um retrato das séries na janela. Em ambos os casos percebe-se que a série do Eta ajustado se aproximou da série de referência do NCEP, denotando melhora na previsão atmosférica em ambos os casos.

Figura 61 – Série da janela que apresentou maior redução do EQMn na aplicação do filtro com comprimento de um dia de previsões sobre a variável vento meridional

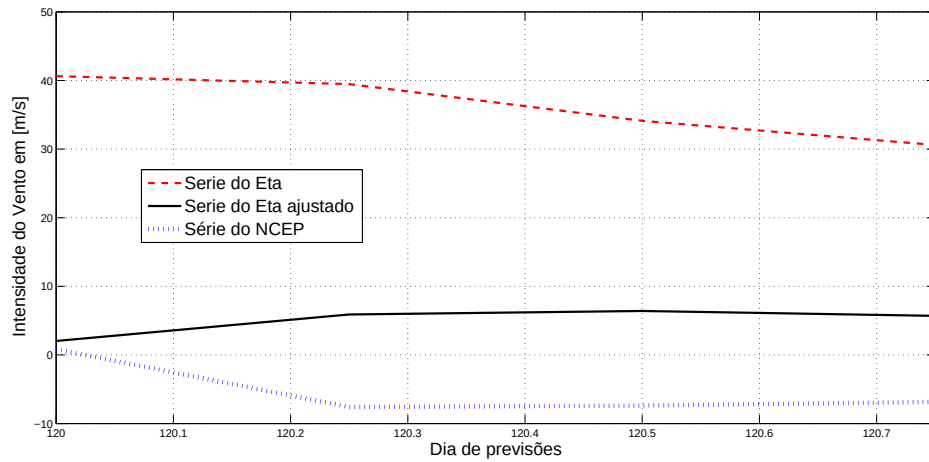
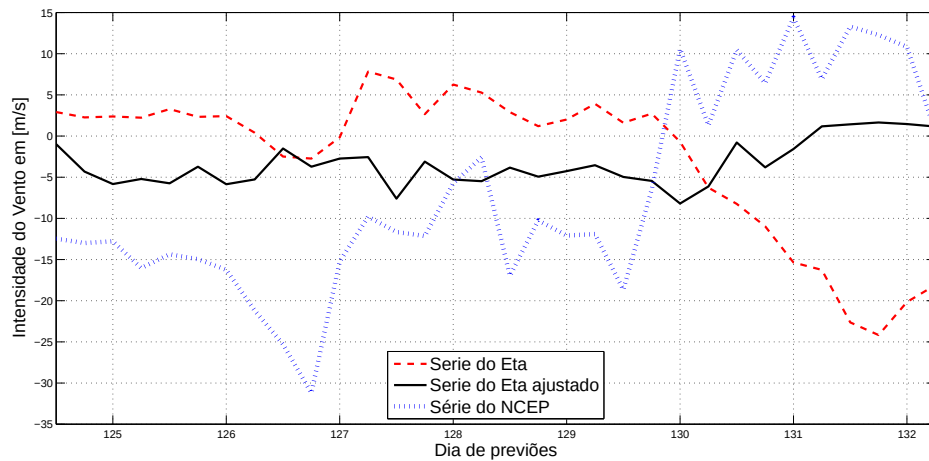


Figura 62 – Série da janela que apresentou maior redução do EQMn na aplicação do filtro com comprimento de oito dias de previsões sobre a variável vento meridional



B. Altura geopotencial

Para verificar o efeito dos filtros adaptado na variável altura geopotencial, foram realizados os mesmos procedimentos como no caso da variável vento meridional.

As Figuras 63 e 64 mostram o EQMn ao longo da posição geográfica, referente ao ajuste das séries do modelo Eta realizado pela aplicação do filtro com um dia (4 coeficientes) de previsões e com oito dias (32 coeficientes) de previsões. Pela Figura 63 observa-se que houve redução do EQMn em todas as séries ajustadas pelo filtro. Pela Figura 64 observa-se que não houve redução do EQMn nas 150 primeiras posições geográficas.

Figura 63 – EQMn antes e depois da aplicação do filtro de comprimento de um dia de previsões nas séries climáticas do Eta sobre a variável altura geopotencial

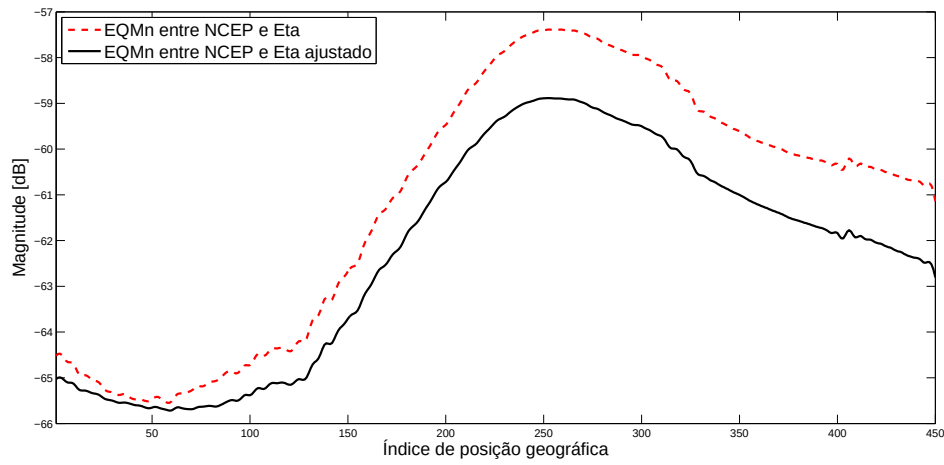
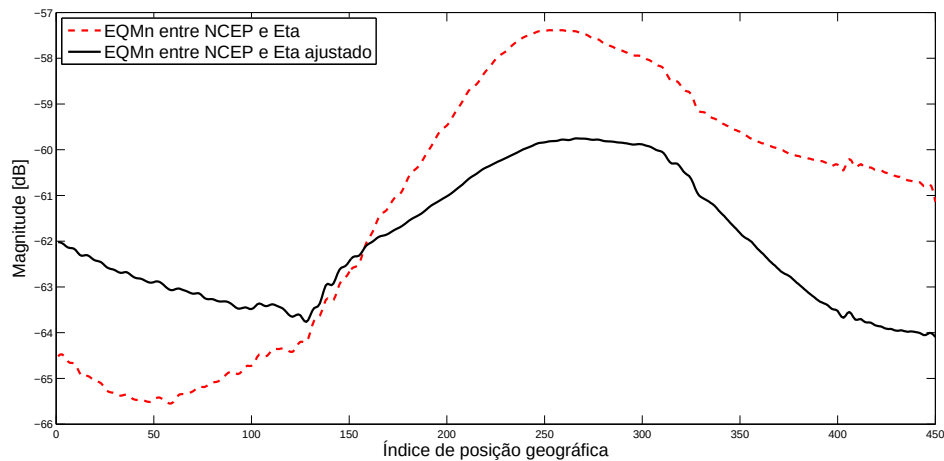


Figura 64 – EQMn antes e depois da aplicação do filtro de comprimento de oito dias de previsões nas séries climáticas do Eta sobre a variável altura geopotencial



Para identificar a distribuição do EQMn ao longo dos dias de previsões, utilizou-se uma janela deslizante com 4 coeficientes para filtros de 4 coeficientes e uma janela com 32 coeficientes para filtros com 32 coeficientes. Assim, se calculou a média do EQMn de todas as séries climáticas.

Nas Figuras 65 e 66 o EQMn é o resultado da média de todas as realizações na aplicação de filtros com 4 coeficientes e com 32 coeficientes respectivamente sobre a variável altura geopotencial. Dessa forma percebe-se que ao longo do tempo o uso do filtro com 4 coeficientes teve o EQMn reduzido em 87,85% do tempo de previsão enquanto que para o filtro de 32 coeficientes o EQMn foi reduzido em 62,24% do tempo de previsão.

Figura 65 – Média do EQMn por janela, na aplicação do filtro com comprimento de um dia de previsões sobre a variável altura geopotencial

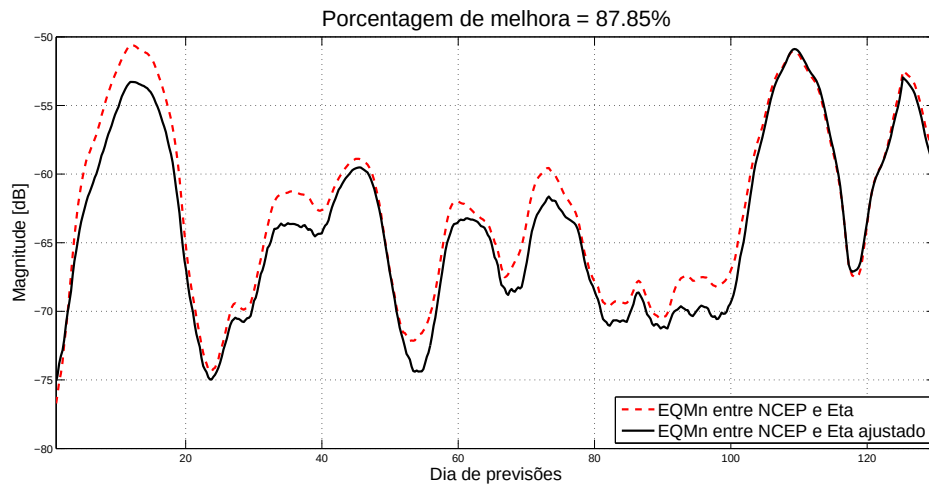
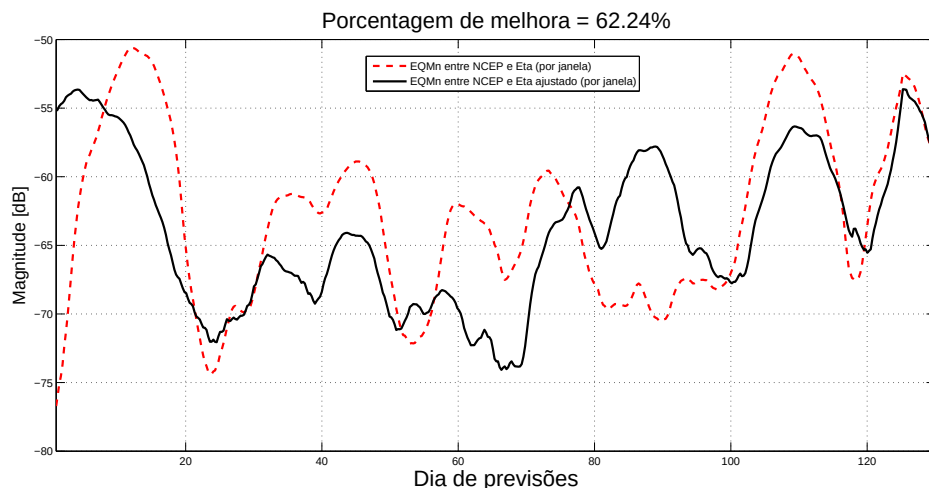


Figura 66 – Média do EQMn por janela, na aplicação do filtro com comprimento de oito dias de previsões sobre a variável altura geopotencial



As Figuras 67 e 68 mostram a distribuição estatística do EQMn na aplicação de filtros com 4 coeficientes e com 32 coeficientes respectivamente sobre a variável altura geopotencial. No caso das Figuras 67 e 68 percebe-se que houve redução na média do EQMn e redução da dispersão do EQMn em ambos os casos. Pela Figura 67 percebe-se que na média o EQMn reduziu em $1,25dB$, enquanto que pela Figura 52 percebe-se que na média o EQMn reduziu em $1,55dB$. Assim, neste caso, o filtro com 4 coeficiente reduziu na média o EQMn em maior intensidade que o filtro com 32 coeficientes.

Figura 67 – Distribuição da média do EQMn por janela, na aplicação do filtro com um dia de previsões sobre a variável altura geopotencial

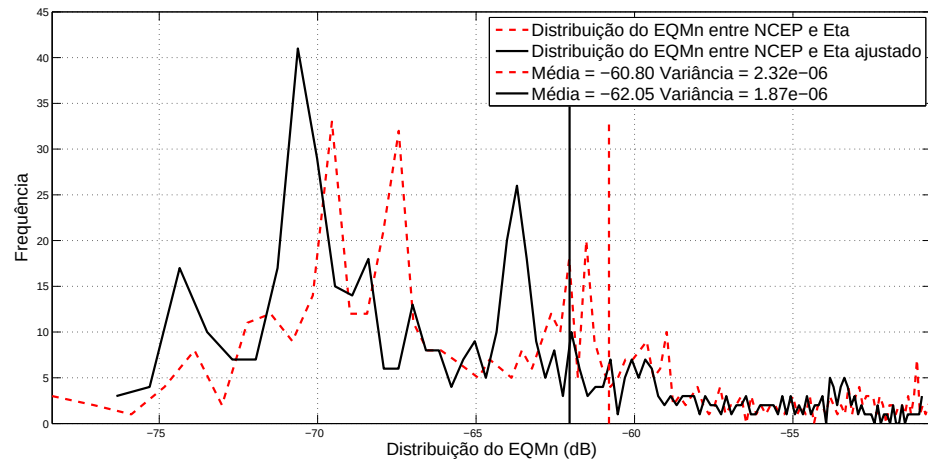
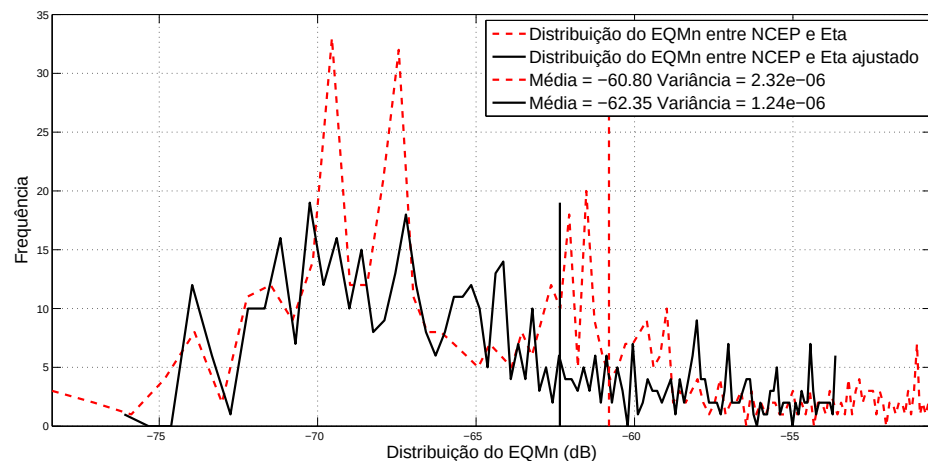


Figura 68 – Distribuição da média do EQMn por janela, na aplicação do filtro com oito dias de previsões sobre a variável altura geopotencial



Mostra-se que na média a intensidade de redução EQMn provocada pela aplicação do filtro com 32 coeficientes é superior à intensidade de redução do EQMn provocada pela aplicação do filtro com 4 coeficientes. Em contra partida, o filtro com 4 coeficientes resulta na redução do EQMn em 100% da região, o que não ocorre na aplicação do filtro de 32 coeficientes.

Para avaliar o efeito do ajuste dos filtros sobre as séries climáticas, colheu-se duas séries temporais do estudo de caso apresentado, uma para a aplicação do filtro com 4 coeficientes e outra para a aplicações do filtro com 32 coeficientes.

A Figura 69 mostra a série que teve seu ajuste feito por um filtro com 4 coeficientes enquanto que a Figura 70 mostra a série que teve seu ajuste feito com um filtro com 32 coeficientes. As Figuras 69 e 70 também mostram a série de referência do NCEP e a série

Eta. Ainda pela Figura 69 é difícil perceber uma melhora significativa da série do Eta ajustado. Pela Figura 70 pode-se ver que próximo ao quadragésimo dia de previsão a série do Eta ajustado se aproxima da série de referência do NCEP.

Figura 69 – Série climática da variável altura geopotencial referente as dados do Eta, NCEP e Eta ajustado pelo filtro de 4 coeficientes

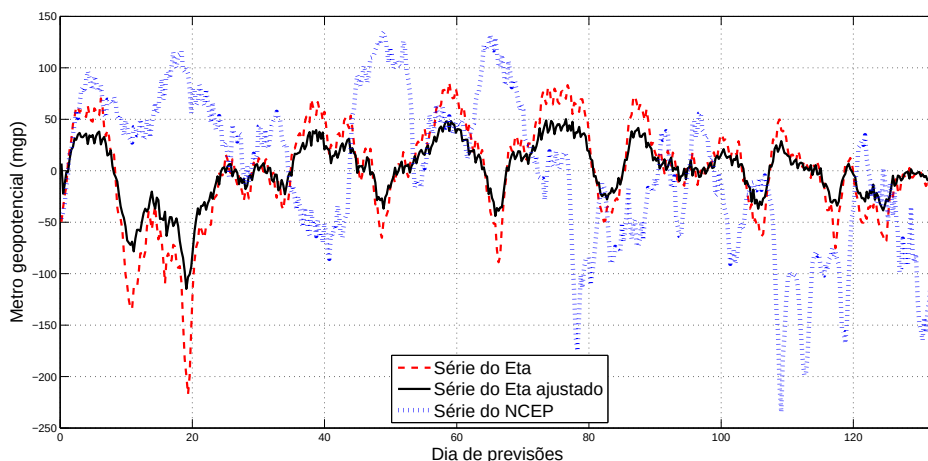
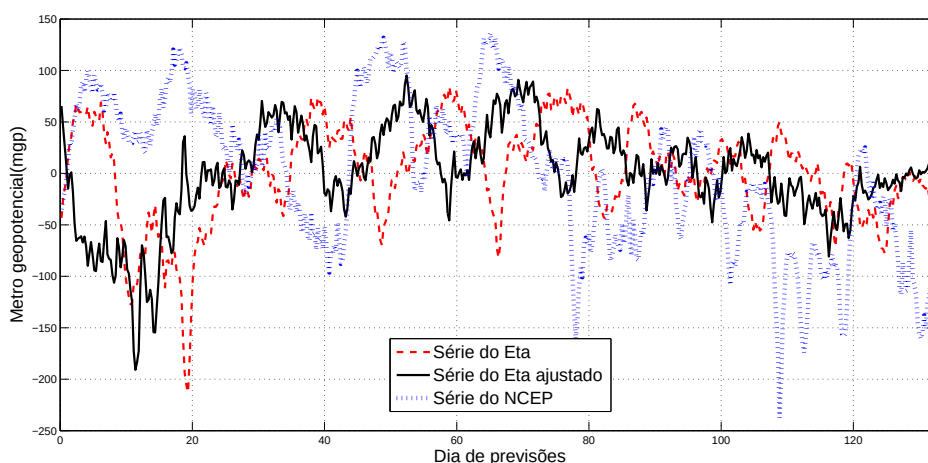


Figura 70 – Série climática da variável altura geopotencial referente as dados do Eta, NCEP e Eta ajustado pelo filtro de 32 coeficientes



As Figuras 71 mostram o erro entre a série do Eta e a séries do NCEP em linha tracejada e o erro entre a série do Eta ajustado por um filtro de 4 coeficientes e a série do NCEP em linha cheia. Nesta figura pode-se ver que no sexagésimo dia há redução no erro. A Figura 56 mostra o erro entre a série do Eta e a séries do NCEP em linha tracejada e o erro entre a série do Eta ajustado por um filtro com comprimento de 32 coeficientes e a série do NCEP em linha cheia. Pode-se ver que no vigésimo dia de previsão há redução no erro.

Figura 71 – Erro entre a série climática da variável altura geopotencial referente as dados do Eta e série do NCEP em linha tracejada e erro entre a série do Eta ajustado com filtro de 4 coeficientes e a série do NCEP em linha cheia

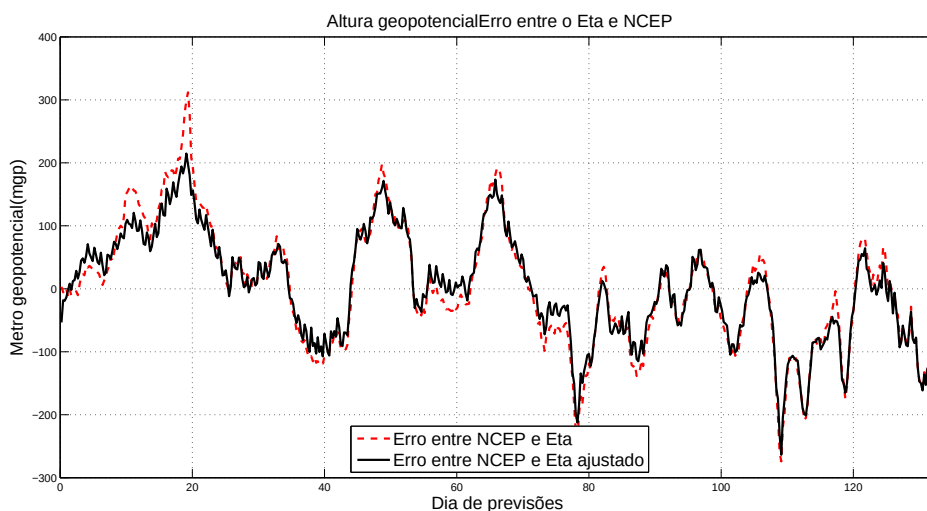
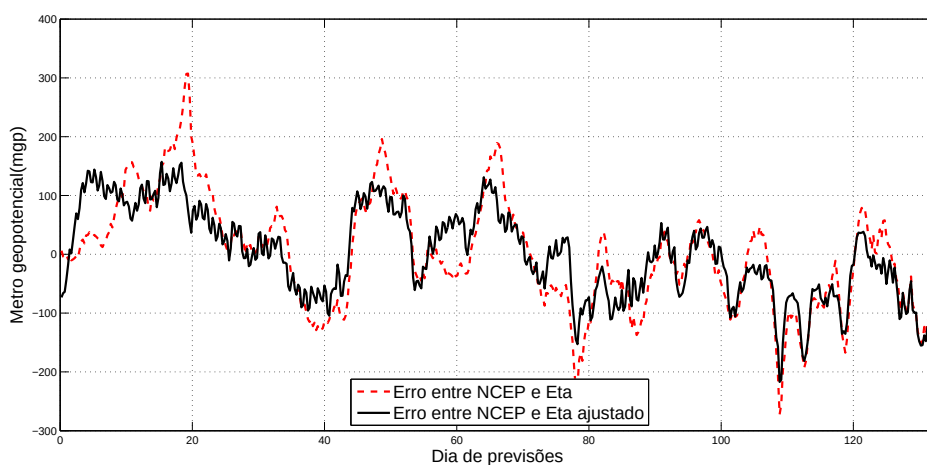


Figura 72 – Erro entre a série climática da variável altura geopotencial referente as dados do Eta e série do NCEP em linha tracejada e erro entre a série do Eta ajustado com filtro de 32 coeficientes e a série do NCEP em linha cheia



Com o objetivo de identificar mais precisamente a melhora da série do Eta ajustado com relação aos dados do Eta original, para estas séries calculou-se o EQMn entre as séries do Eta e o NCEP, e a série do Eta ajustado e o NCEP. Para identificar a distribuição do EQMn utilizou-se uma janela deslizante com 4 coeficientes e com 32 coeficientes.

A Figura 73 mostra o EQMn por uma janela de 4 coeficientes ao longo dos dias de previsões e a Figura 74 mostra o EQMn por uma janela de 32 coeficientes ao longo dos dias de previsões. Pela Figura 69 percebe-se que o filtro com 4 coeficientes reduziu o EQMn em 65,42% do tempo de previsão, e pela Figura 74 pode-se ver que o filtro com 32 coeficientes reduziu o EQMn em 66,54% do tempo de previsão. Neste caso percebe-se

que o filtro com 32 coeficientes apresentou melhores resultados.

Figura 73 – EQMn por janela, na aplicação do filtro com comprimento de um dia de previsões sobre a variável altura geopotencial

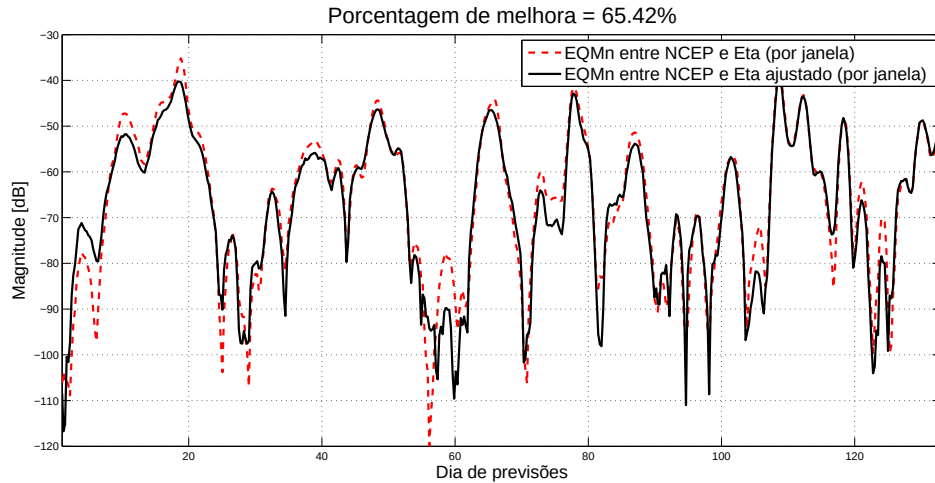
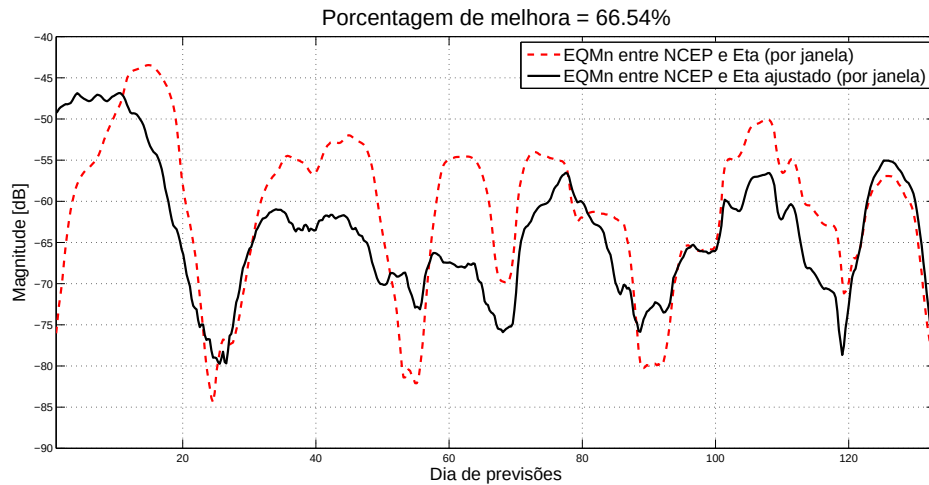


Figura 74 – EQMn por janela, na aplicação do filtro com comprimento de oito dias de previsões sobre a variável altura geopotencial



As Figuras 75 e 76 mostra a distribuição estatística do erro. As Figuras 75 e 76 mostram a distribuição estatística do EQMn na aplicação do filtro com 4 coeficientes e com 32 coeficientes respectivamente sobre a variável altura geopotencial. No caso das Figuras 75 e 76 percebe-se que houve redução na média do EQMn e redução da dispersão do EQMn em ambos os casos. Pela Figura 75 percebe-se que na média o EQMn reduziu em $1,5dB$, enquanto que pela Figura 76 percebe-se que na média o EQMn reduziu em $3,02dB$. Assim, neste caso, o filtro com 32 coeficiente reduziu na média o EQMn em maior intensidade que o filtro com 4 coeficientes.

Figura 75 – Distribuição do EQMn por janela, na aplicação do filtro com comprimento de um dia de previsões sobre a variável altura geopotencial

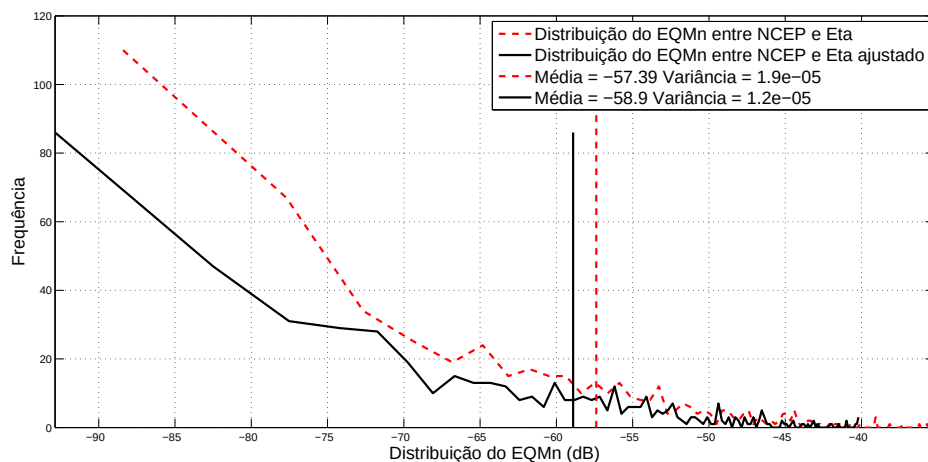
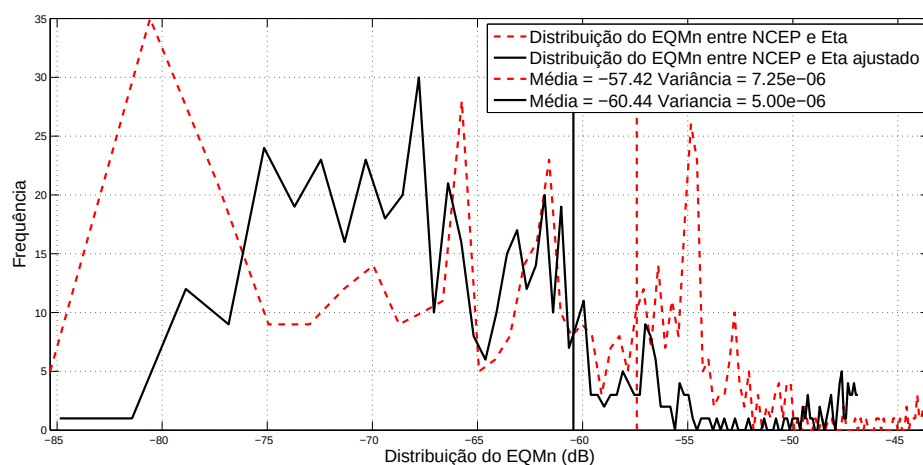


Figura 76 – Distribuição do EQMn por janela, na aplicação do filtro com comprimento de oito dias de previsões sobre a variável altura geopotencial



As Figuras 77 e 78 apresentam um retrato das séries na janela. Em ambos os casos percebe-se que a série do Eta ajustado se aproximou da séries de referência do NCEP, denotando melhora na previsão atmosférica.

Figura 77 – Série da janela que apresentou maior redução do EQMn na aplicação do filtro com comprimento de um dia de previsões sobre a variável altura geopotencial

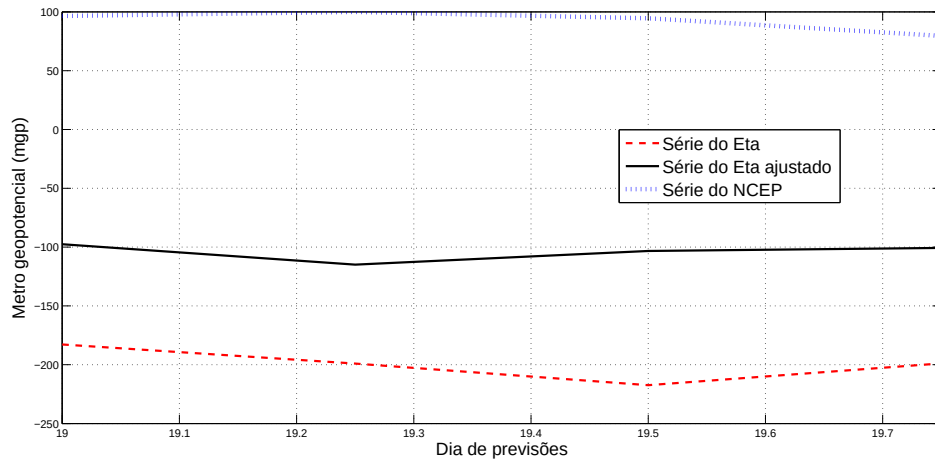
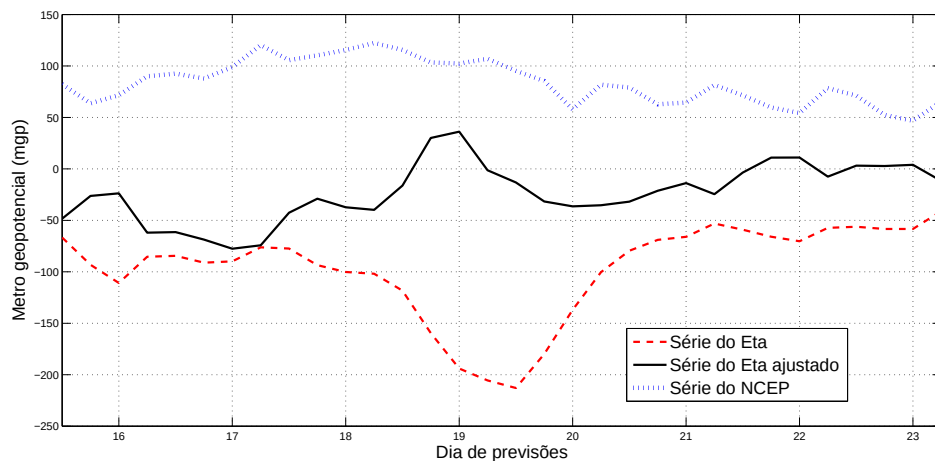


Figura 78 – Série da janela que apresentou maior redução do EQMn na aplicação do filtro com comprimento de um dia de previsões sobre a variável altura geopotencial



4.2.2 Avaliação dos melhores critérios de seleção de regiões

Com o objetivo de propor o melhor método no ajuste de erros de previsões escolheram-se aqueles que apresentaram melhores resultados. Nesta avaliação utilizou-se somente o RLS, pois com base nos resultados apresentados foi quem apresenta melhor desempenho. Os critérios de seleção de regiões são em i) *k-means* (500 séries de treinamento), regiões em ii) Anéis (anel intermediário), regiões retangulares iii) ao longo das ordenadas e iv) ao longo das abscissas e por fim o de v) regiões circulares (500 séries de treinamento).

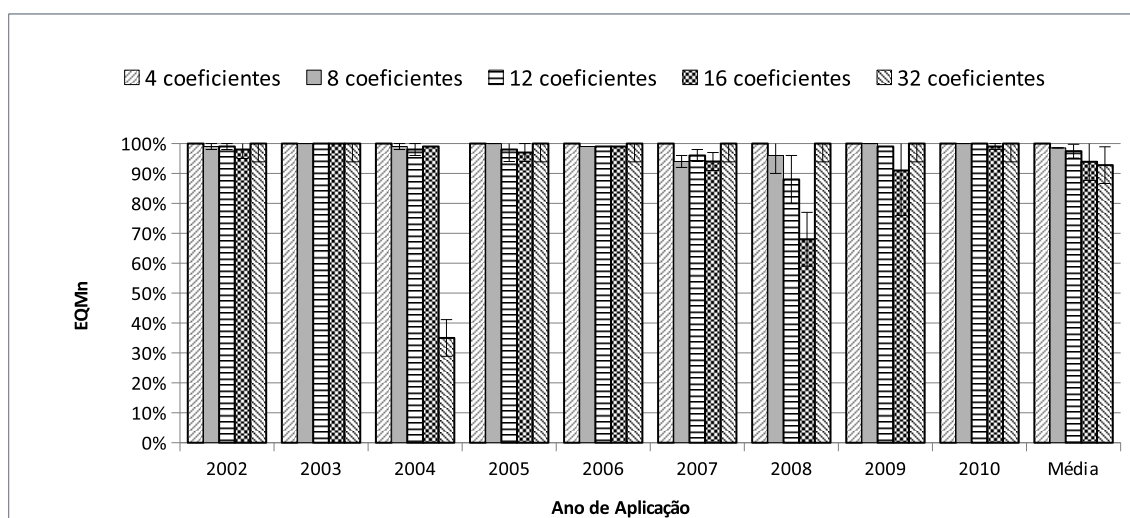
Para verificar qual o comprimento de filtro que apresenta o melhor resultado possível, testes foram feitos utilizando filtros com comprimento de 1 dia (4 coeficien-

tes), 2 dias (8 coeficientes), 3 dias (12 coeficientes), 4 dias (16 coeficientes) e 8 dias de previsões (32 coeficientes). Os testes foram feitos tanto para a variável vento meridional quanto para a variável altura geopotencial.

I. Filtro adaptado pelo algoritmo RLS e com conjunto de treinamento dado pela seleção de regiões utilizando o *k-means*

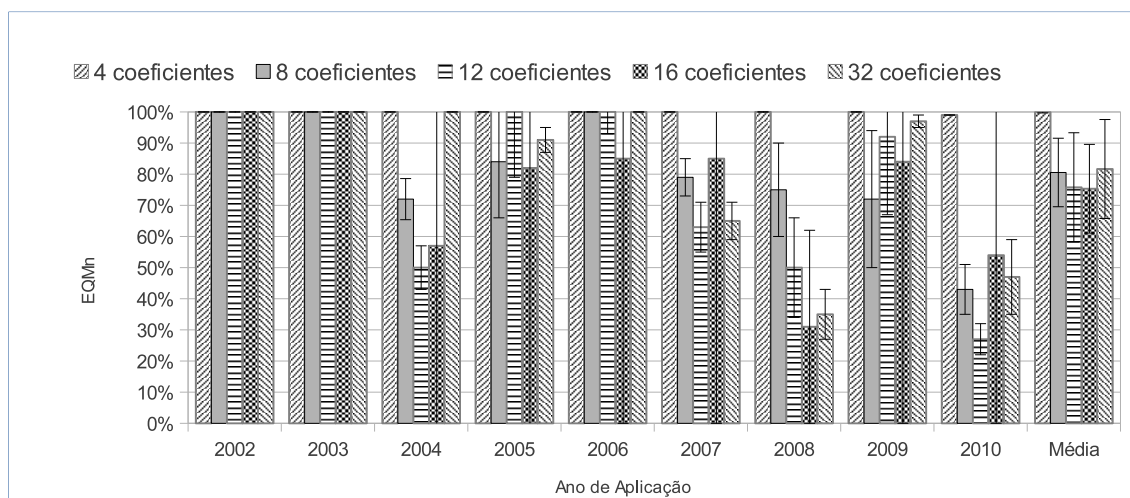
A Figura 79 mostra o gráfico dos resultados da porcentagem de regiões agrupadas pelo *k-means* com redução no EQMn da variável vento meridional, com filtros de 4, 8, 12, 16 e 32, coeficientes e aproximadamente 500 séries de treinamento, com 10 realizações. Pelo gráfico percebe-se que o filtro com 4 coeficientes de comprimento é o único que reduziu na média o EQMn de toda a região selecionada em todos os anos.

Figura 79 – Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro adaptado pelo algoritmo RLS



A Figura 80 mostra o gráfico dos resultados da porcentagem de regiões com redução no EQMn da variável altura geopotencial, com filtros de ordens iguais a 4, 8, 12, 16 e 32, e aproximadamente 500 séries de treinamento. A Figura 80 mostra que a seleção de região feita pelo algoritmo *k-means* o filtro com 4 coeficientes (1 dia de previsões) foi o que apresentou o melhor resultado, pois reduziu na média o EQMn de toda a região selecionada em todos os anos.

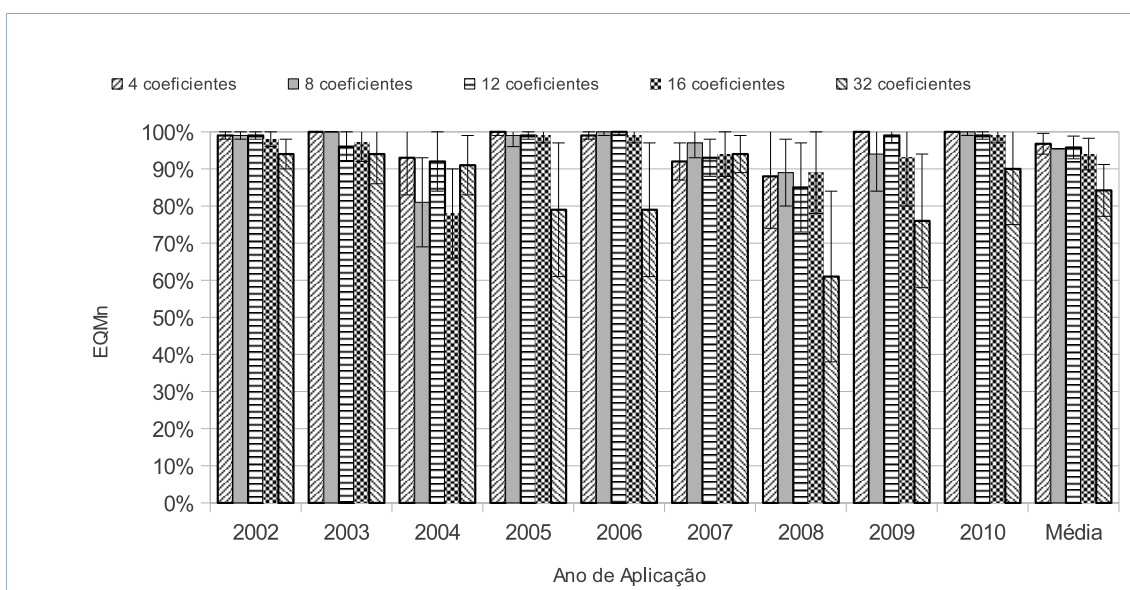
Figura 80 – Porcentagem de regiões com redução no EQMn da altura geopotencial, na aplicação do filtro adaptado pelo algoritmo RLS



II. Filtro adaptado pelo algoritmo RLS e com conjunto de treinamento dado pela seleção de regiões retangulares (ao longo das abscissas)

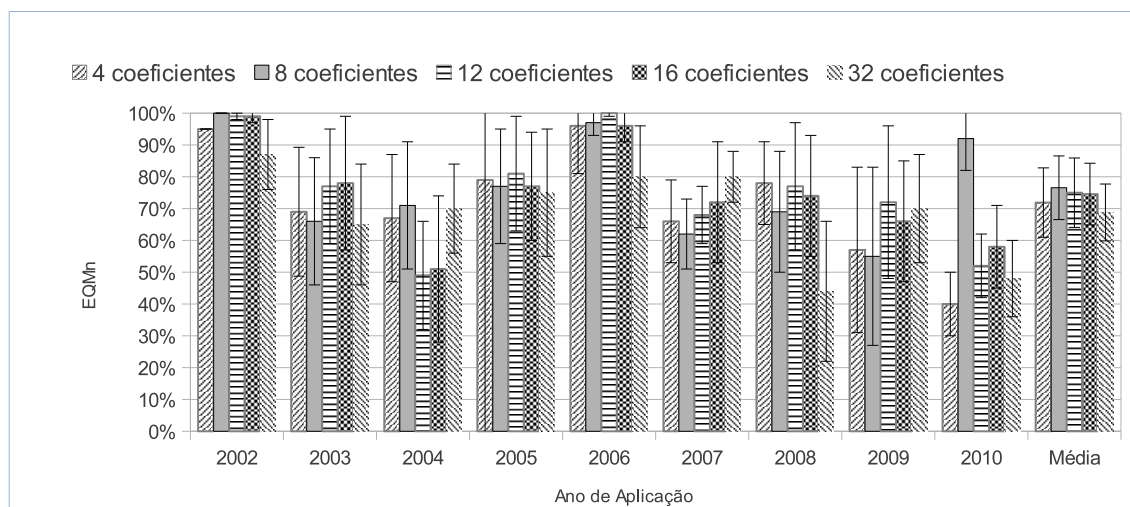
A Figura 81 mostra a porcentagem de regiões retangulares ao longo das abscissas com redução no EQMn da variável vento meridional, após a aplicação de filtros de 4, 8, 12, 16 e 32 coeficientes. O filtro com comprimento de 4 coeficientes reduziu na média o EQMn em 97% das regiões, obtendo assim um desempenho superior aos filtros com comprimentos superiores.

Figura 81 – Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro adaptado pelo algoritmo RLS



A Figura 82 mostra a porcentagem de regiões com redução no EQMn da variável altura geopotencial, com filtros de 4, 8, 12, 16 e 32 coeficientes.

Figura 82 – Porcentagem de regiões com redução no EQMn da altura geopotencial, na aplicação do filtro adaptado pelo algoritmo RLS

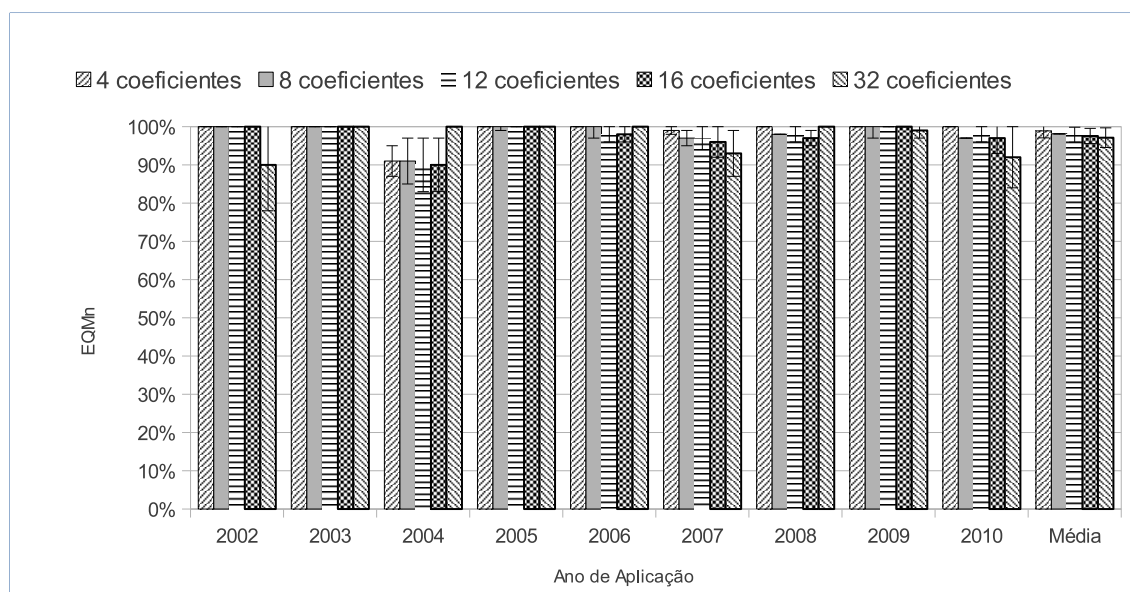


Pela Figura 82 percebe-se que entre a seleção de regiões retangulares ao longo das abscissas o filtro com 8 coeficientes (2 dias de previsões) foi o que apresentou o melhor resultado, reduzindo na média o EQMn em 77% das regiões sobre a variável altura geopotencial.

III. Filtro adaptado pelo algoritmo RLS e com conjunto de treinamento dado pela seleção de regiões retangulares (Ao longo das ordenadas)

A Figura 83 mostra a porcentagem de regiões retangulares ao longo das ordenadas, com redução no EQMn da variável vento meridional, com filtros de 4, 8, 12, 16 e 32 coeficientes.

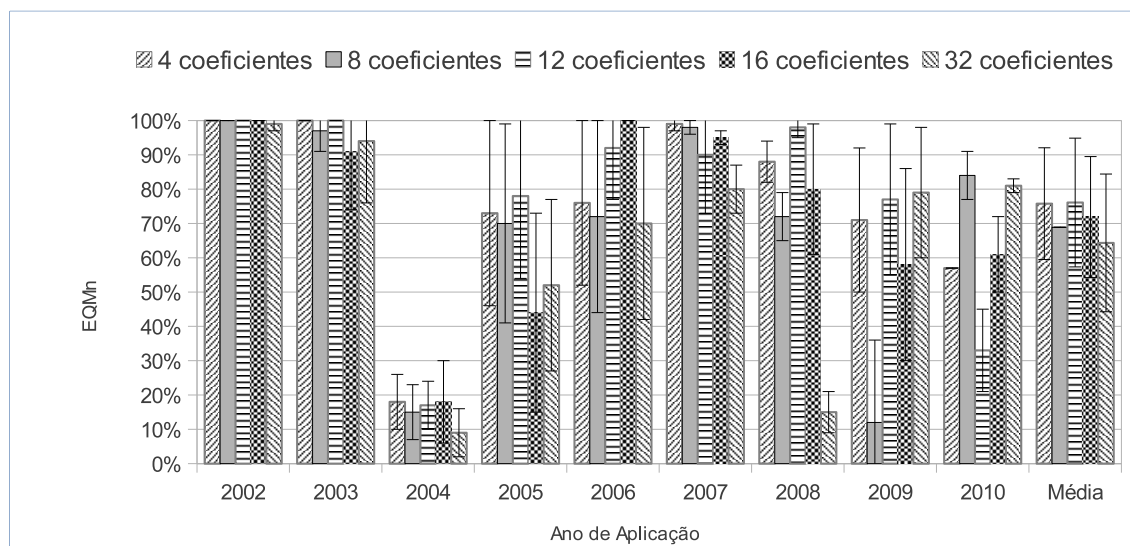
Figura 83 – Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro adaptado pelo algoritmo RLS



Pela Figura 83 percebe-se que entre a seleção de regiões retangulares ao longo das ordenadas o filtro com 4 coeficientes (1 dia de previsões) foi o que apresentou o melhor resultado reduzindo na média o EQMn em 99% da região.

A Figura 84 mostra o gráfico dos resultados da porcentagem de regiões com redução no EQMn da variável altura geopotencial, com filtros de ordens iguais a 4, 8, 12, 16 e 32.

Figura 84 – Porcentagem de regiões com redução no EQMn da altura geopotencial, na aplicação do filtro adaptado pelo algoritmo RLS

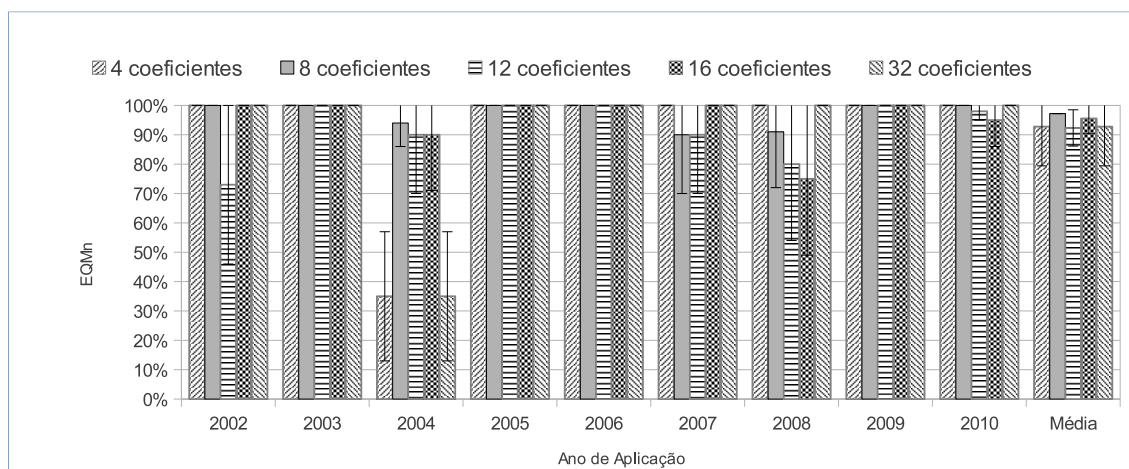


Pela Figura 84 percebe-se que entre a seleção de regiões retangulares ao longo das ordenadas o filtro com 4 coeficientes (1 dia de previsões) e com 12 coeficientes foram os que apresentaram os melhores resultados ambos reduzindo na média EQMn em 76% da região.

IV. Filtro adaptado pelo algoritmo RLS e com conjunto de treinamento dado pela seleção de regiões Circulares

A Figura 85 mostra a porcentagem de regiões circulares com redução no EQMn da variável vento meridional, com filtros de 4, 8, 12, 16 e 32 coeficientes.

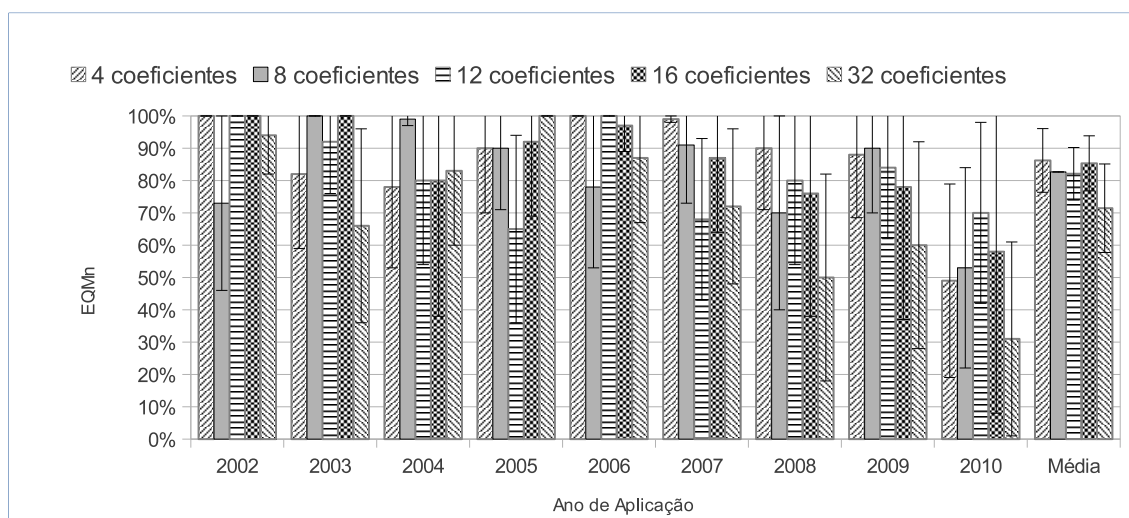
Figura 85 – Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro adaptado pelo algoritmo RLS



Pelo gráfico da Figura 85 percebe-se que entre a seleção de regiões em anéis retangulares o filtro com 8 coeficientes (1 dia de previsões) foi o que apresentou o melhor resultado, pois reduziu na média o EQMn em 97% da região.

A Figura 86 mostra o gráfico dos resultados da porcentagem de regiões com redução no EQMn da variável altura geopotencial, com filtros de ordens iguais a 4, 8, 12, 16 e 32.

Figura 86 – Porcentagem de regiões com redução no EQMn da altura geopotencial, na aplicação do filtro adaptado pelo algoritmo RLS

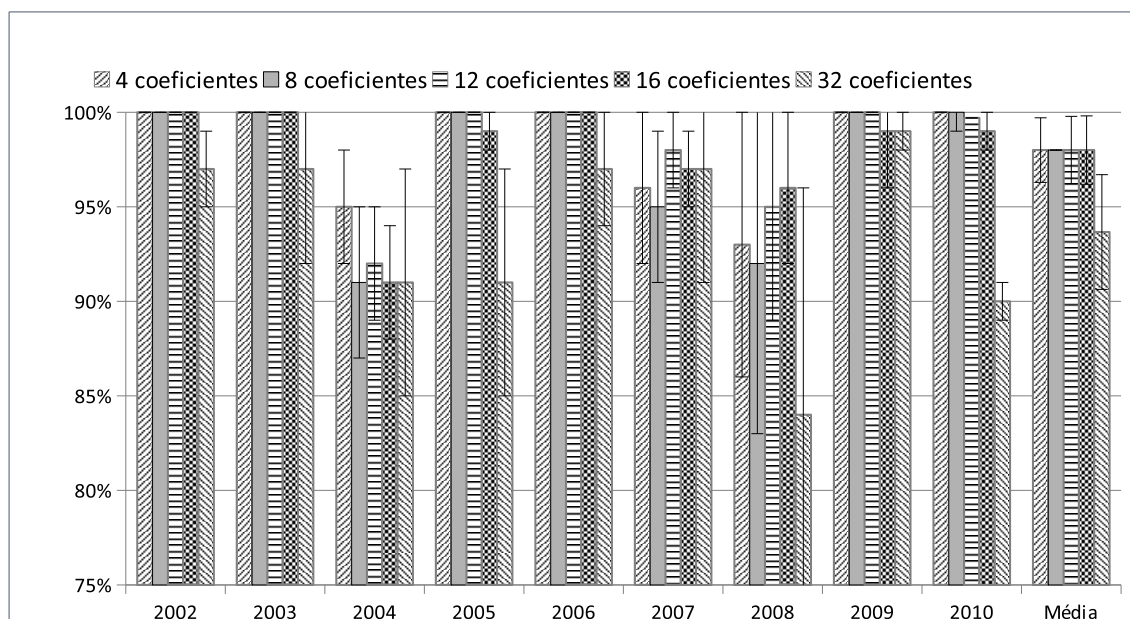


Pela Figura 86 percebe-se que entre a seleção de regiões retangulares na vertical o filtro com 4 coeficientes (2 dias de previsões) foi o que apresentou na média o melhor resultado, reduzindo na média o EQMn em 86% da região.

V. Filtro adaptado pelo algoritmo RLS e com conjunto de treinamento dado pela seleção de regiões em anéis retangulares

A Figura 87 mostra a porcentagem de regiões anulares com redução no EQMn da variável vento meridional, com filtros de 4, 8, 12, 16 e 32 coeficientes.

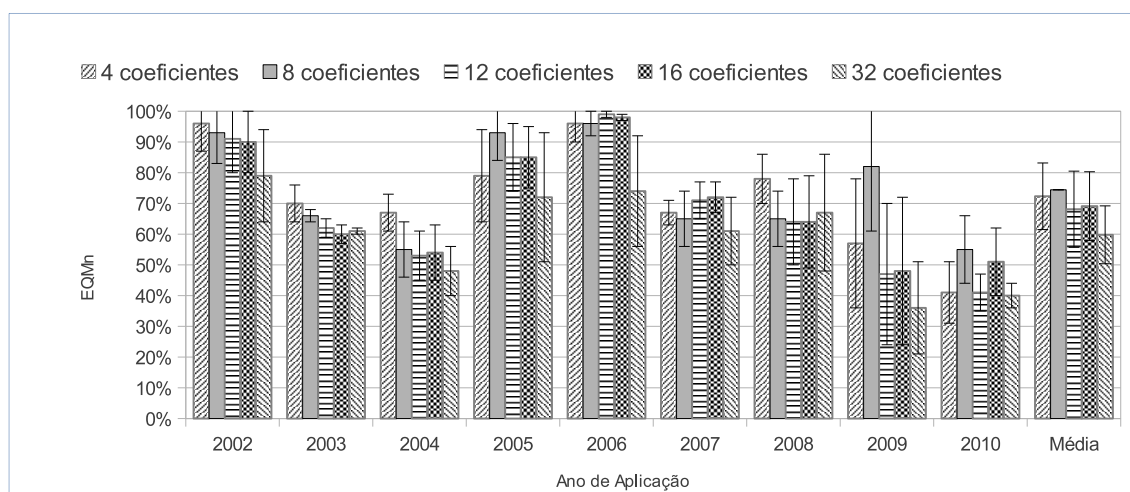
Figura 87 – Porcentagem de regiões com redução no EQMn no vento meridional, na aplicação do filtro adaptado pelo algoritmo RLS



Pelo gráfico da Figura 87 percebe-se que entre a seleção de regiões em anéis retangulares os filtro com 4, 8, 12 e 16 coeficientes apresentaram o melhor resultado, pois ambos reduziram na média o EQMn em 98% da região.

A Figura 88 mostra o gráfico dos resultados da porcentagem de regiões com redução no EQMn da variável altura geopotencial, com filtros de ordens iguais a 4, 8, 12, 16 e 32.

Figura 88 – Porcentagem de regiões com redução no EQMn da altura geopotencial, na aplicação do filtro adaptado pelo algoritmo RLS



Pela Figura 88 percebe-se que entre a seleção de regiões retangulares ao longo das ordenadas o filtro com 8 coeficientes (2 dias de previsões) foi o que apresentou na média o melhor resultado, reduzindo o EQMn em 74% da região.

4.2.2.1 Avaliação dos critérios propostos

Para verificarmos a intensidade da redução do EQMn para cada um dos métodos propostos, calculou-se a redução do EQMn para o ano de 2008 quando aplicado aos filtros resultantes dos métodos propostos.

A. Vento meridional

Como foi mostrado anteriormente, os filtros com 4 coeficientes adaptados através do algoritmo RLS e treinado com séries provenientes dos critérios de seleção de regiões *k-means*, ao longo das ordenadas, ao longo das abscissas e anel, reduziram na médio o EQMn em 100%, 99%, 97, 98% de região respectivamente obtendo resultados superiores aos filtros com 8, 12, 16 e 32. Já para o caso de regiões circulares o filtro com 8 coeficientes foi o que apresentou melhores resultados, reduzindo o EQMn em 97% da região.

Para identificar a distribuição do EQMn ao longo dos dias de previsões, utilizou-se uma janela deslizante com 4 coeficientes para os filtros de 4 coeficientes e uma janela de 8 coeficientes para filtros com 8 coeficientes. Assim, se calculou a média do EQMn de todas as 10 realizações.

A Figura 89 mostra a média do EQMn, considerando o critério de seleção por *k-means*, por uma janela de 4 coeficientes ao longo dos dias de previsões e a Figura 90 mostra a distribuição estatística deste erro. Neste caso houve redução do EQMn em 90,3% do tempo de previsão e o EQMn reduziu na média em 4,09dB.

Figura 89 – Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes do *k-means*, sobre a variável vento meridional

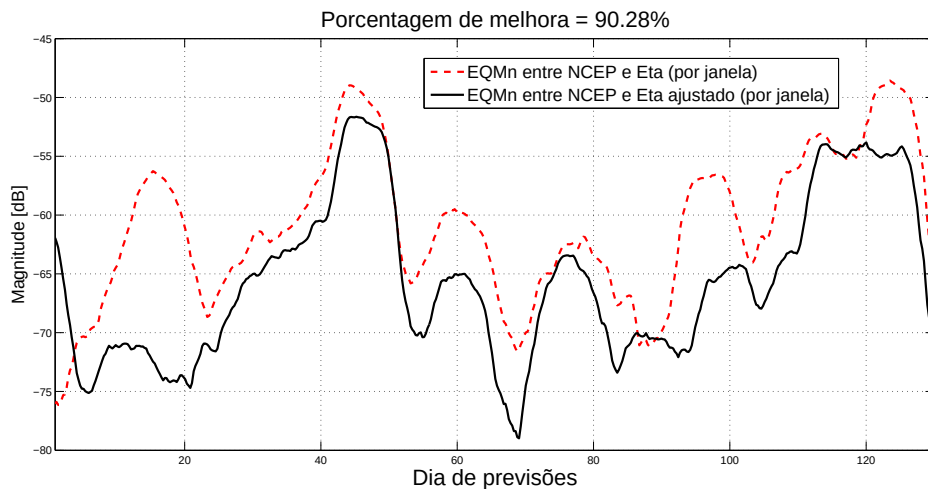
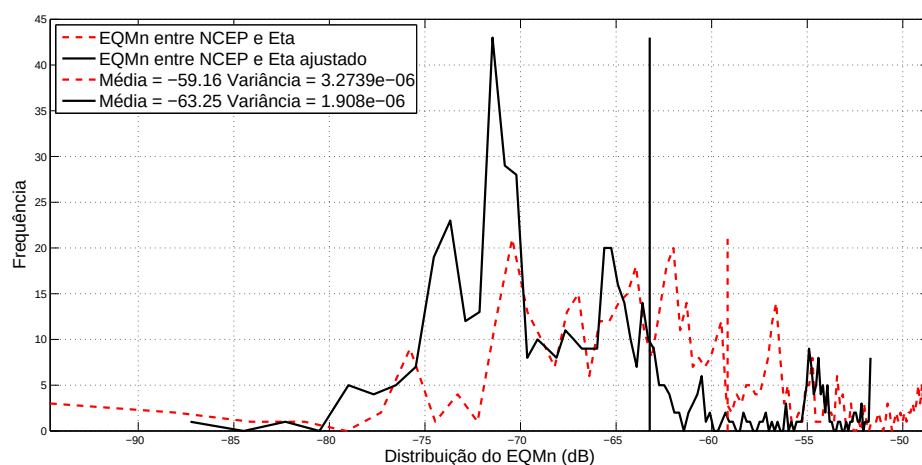


Figura 90 – Distribuição do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes do *k-means*, sobre a variável vento meridional



A Figura 91 mostra a média do EQMn, considerando o critério de seleção em retângulos ao longo das ordenadas, por uma janela de 4 coeficientes ao longo dos dias de previsões e a Figura 92 mostra a distribuição estatística deste erro. Neste caso houve redução do EQMn em 90,5% do tempo de previsão e o EQMn reduziu na média em 4,85 dB.

Figura 91 – Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção de regiões retangulares ao longo das ordenadas, sobre a variável vento meridional

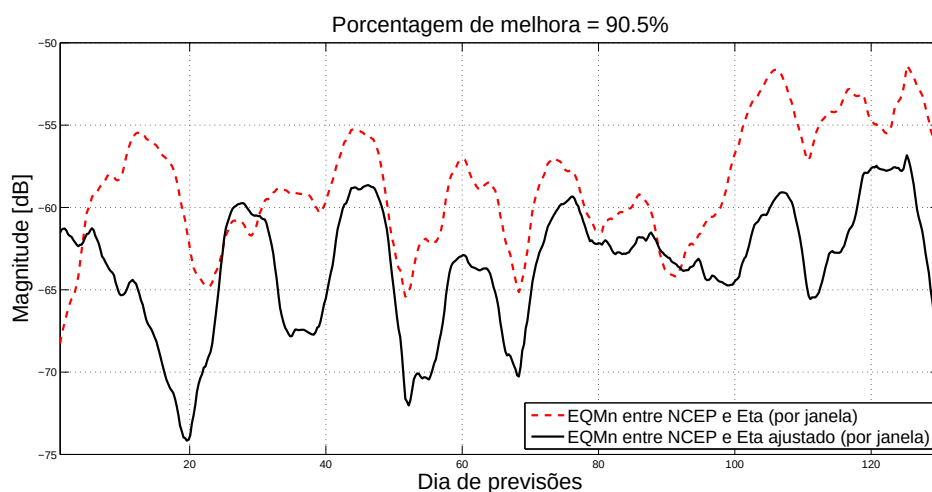
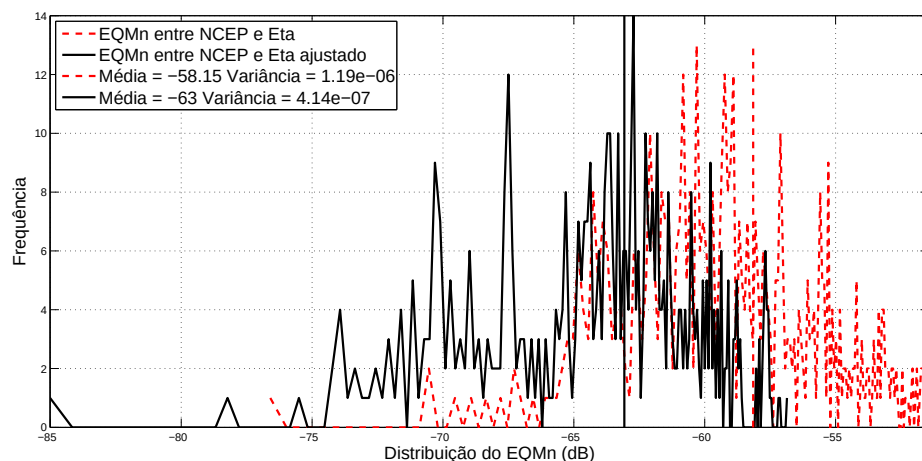


Figura 92 – Distribuição do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção em regiões retangulares ao longo das ordenadas, sobre a variável vento meridional



A Figura 93 mostra a média do EQMn considerando o critério de seleção em retângulos ao longo das abscissas, por uma janela de 4 coeficientes ao longo dos dias de previsões e a Figura 94 mostra a distribuição estatística deste erro. Neste caso houve redução do EQMn em 94,4% do tempo de previsão e o EQMn reduziu na média em 1,7 dB.

Figura 93 – Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção de regiões retangulares ao longo das abscissas, sobre a variável vento meridional

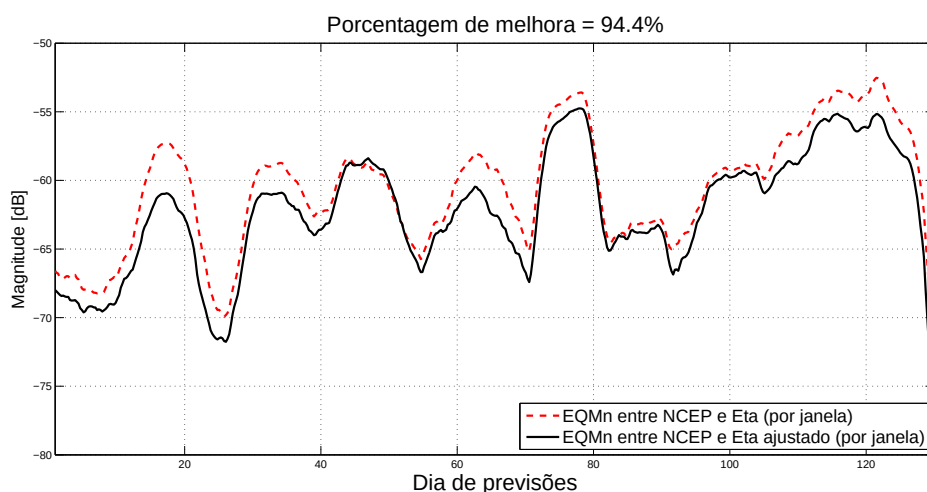
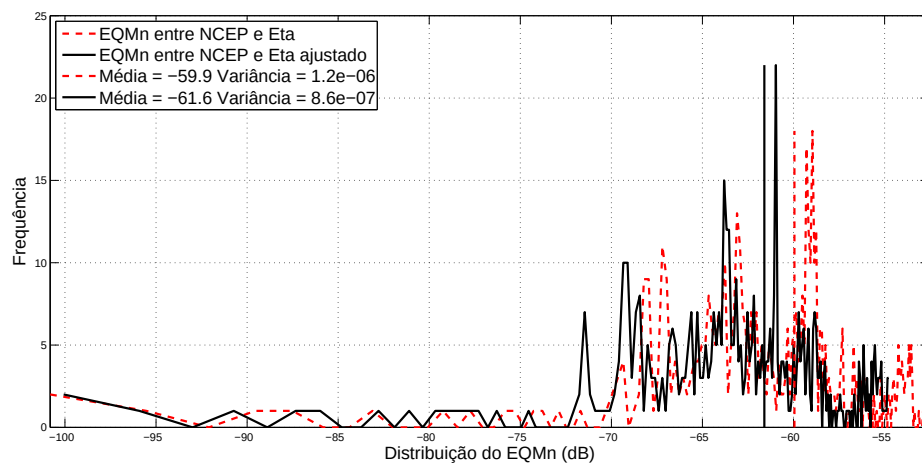


Figura 94 – Distribuição do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção em regiões retangulares ao longo das abscissas, sobre a variável vento meridional



A Figura 95 mostra a média do EQMn considerando o critério de seleção em regiões circulares, por uma janela de 8 coeficientes ao longo dos dias de previsões e a Figura 96 mostra a distribuição estatística deste erro. Neste caso houve redução do EQMn em 87,5% do tempo de previsão e o EQMn reduziu na média em 1,9 dB.

Figura 95 – Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões circulares, sobre a variável vento meridional

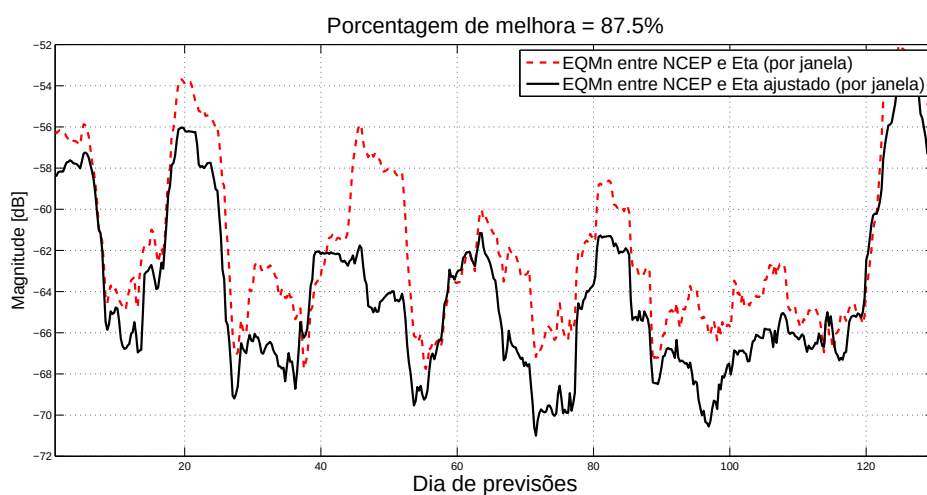
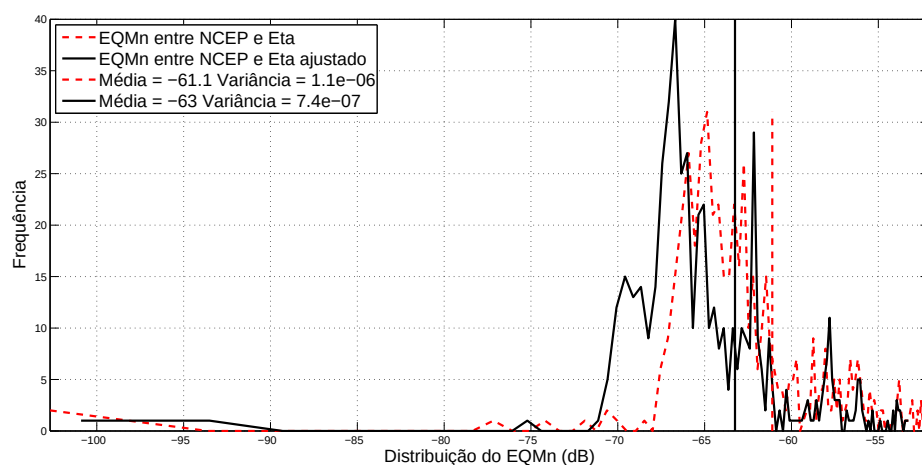


Figura 96 – Distribuição do EQMn por janela na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões circulares, sobre a variável vento meridional



A Figura 97 mostra a média do EQMn considerando o critério de seleção em regiões anulares, por uma janela de 4 coeficientes ao longo dos dias de previsões e a Figura 98 mostra a distribuição estatística deste erro. Neste caso houve redução do EQMn em 88,4% do tempo de previsão e o EQMn reduziu na média em 3,6 dB.

Figura 97 – Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões anulares, sobre a variável vento meridional

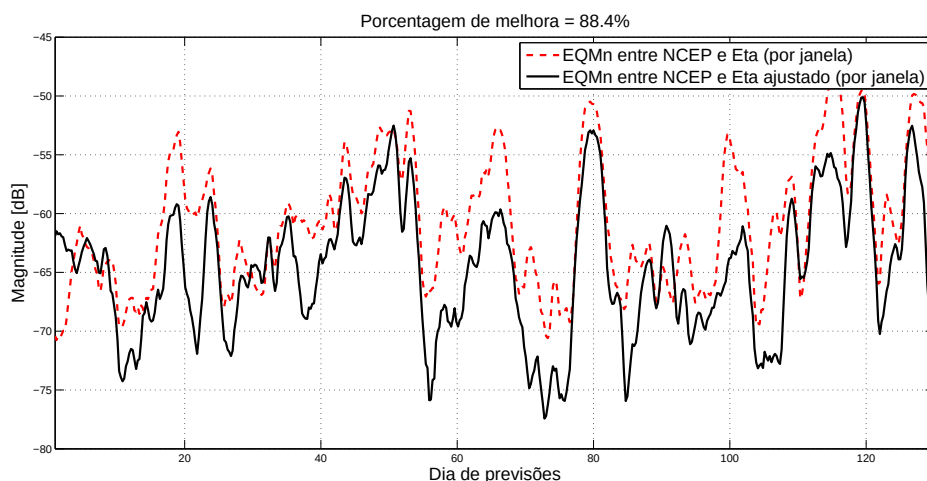
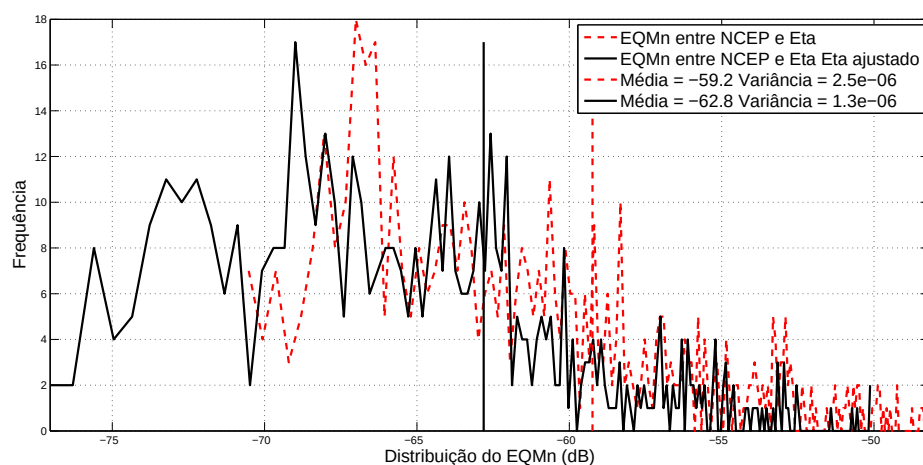


Figura 98 – Distribuição do EQMn por janela na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões anulares, sobre a variável vento meridional



Estes procedimentos foram feitos para os demais anos de previsão, e retirou-se a média dos resultados encontrados para o caso da Intensidade de redução do EQMn e a porcentagem do tempo de previsão que houve redução no EQMn. A Tabela 14 mostra um resumo dos resultados, sendo eles a média dos resultados de todos os anos, referente a variável vento meridional. A primeira coluna da tabela mostra o critério de seleção de região empregado. A segunda coluna mostra a porcentagem de regiões que tiveram redução no EQMn. Na terceira coluna tem-se a média da redução do EQMn em *dB*, ao longo das coordenadas geográficas. Na quarta coluna se tem a média da porcentagem do tempo de previsão que teve redução no EQMn.

Tabela 14 – Tabela comparativa dos critérios de seleção de regiões propostos sobre a variável vento meridional

Critério de seleção de regiões	EQMn ao longo das coordenadas geográficas	Intensidade de redução no EQMn	EQMn ao longo do tempo de previsão
<i>k-means</i>	100%	4,4 dB	91%
Ordenadas	99%	3,6 dB	79%
Abcissas	97%	3,9 dB	92%
Circular	97%	2,8 dB	74%
Anel	98%	3,8 dB	87%

Pela Tabela 14, na segunda coluna, percebe-se que o critério de seleção de regiões feito por meio do algoritmo *k-means* foi o que apresentou melhor resultado, reduzindo o EQMn em toda a região. Na terceira coluna, percebe-se que o critério de seleção de regiões por meio do algoritmo *k-means* também foi o que apresentou maior redução do EQMn, no caso, 4,4 dB. Na quarta coluna o critério de seleção de regiões retangulares ao

longo das abscissas foi o que apresentou o melhor resultado, reduziu o EQMn em 92% do tempo de previsões.

Assim, de acordo com os resultados obtidos para o vento meridional, para treinar filtros com 4 coeficientes e adaptados pelo algoritmo RLS, o melhor critério de seleção de regiões é o critério utilizando o *k-means*.

B. Altura geopotencial

Como visto anteriormente, para a variável altura geopotencial, os filtros com 4 coeficientes adaptados através do algoritmo RLS e treinados com séries provenientes dos critérios de seleção de regiões *k-means*, ao longo das ordenadas e circular tiveram resultados superiores aos filtros com 8, 12, 16 e 32. Já para o caso critério de seleção de regiões ao longo das abscissas e regiões anulares o filtro com 8 coeficientes foi o que apresentou melhores resultados.

A distribuição do EQMn foi identificada por meio de uma janela deslizante com 4 coeficientes para os filtros de 4 coeficientes e 8 coeficientes para filtros com 8 coeficientes. Assim, se calculou a média do EQMn de todas as 10 realizações.

A Figura 99 mostra o EQMn, considerando o critério de seleção: *k-means*, por uma janela de 4 coeficientes ao longo dos dias de previsões e a Figura 100 mostra a distribuição estatística deste erro. Neste caso houve redução do EQMn em 85% do tempo de previsão e na média o EQMn reduziu em 2,4 dB.

Figura 99 – Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes do *k-means*, sobre a variável altura geopotencial

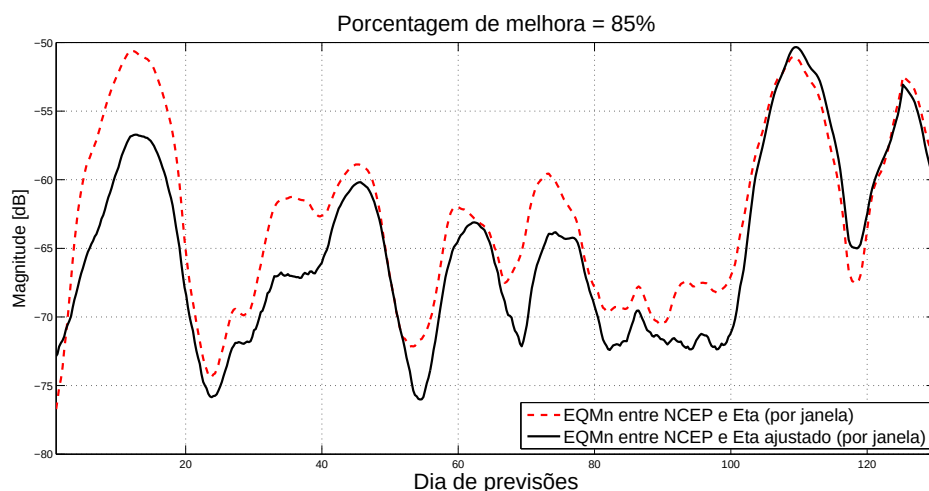
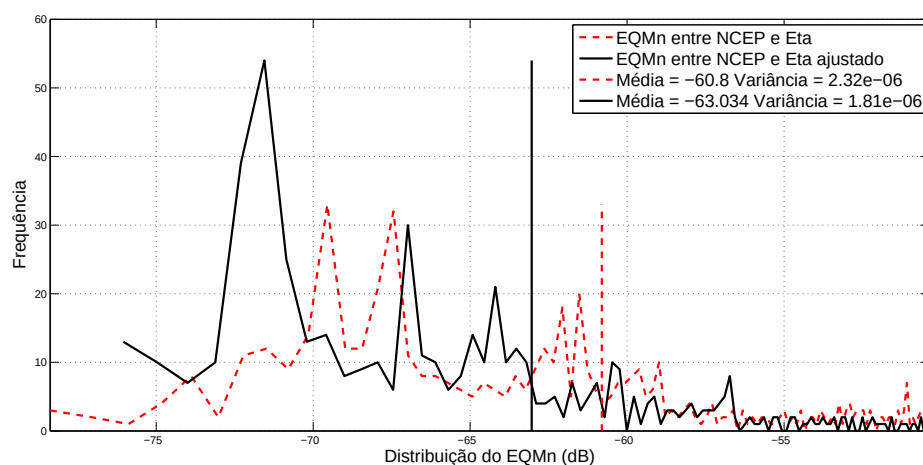


Figura 100 – Distribuição do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes do *k-means*, sobre a variável altura geopotencial



A Figura 101 mostra o EQMn, considerando o critério de seleção em retângulos ao longo das ordenadas, por uma janela de 4 coeficientes ao longo dos dias de previsões e a Figura 102 mostra a distribuição estatística deste erro. Neste caso houve redução do EQMn em 54% do tempo de previsão e na média o EQMn reduziu em 1,52 dB.

Figura 101 – Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção em regiões retangulares ao longo das ordenadas, sobre a variável altura geopotencial

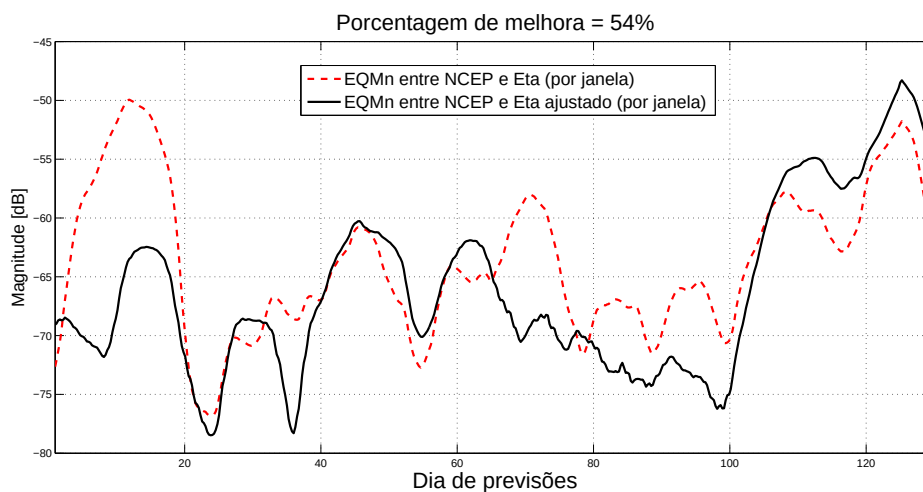
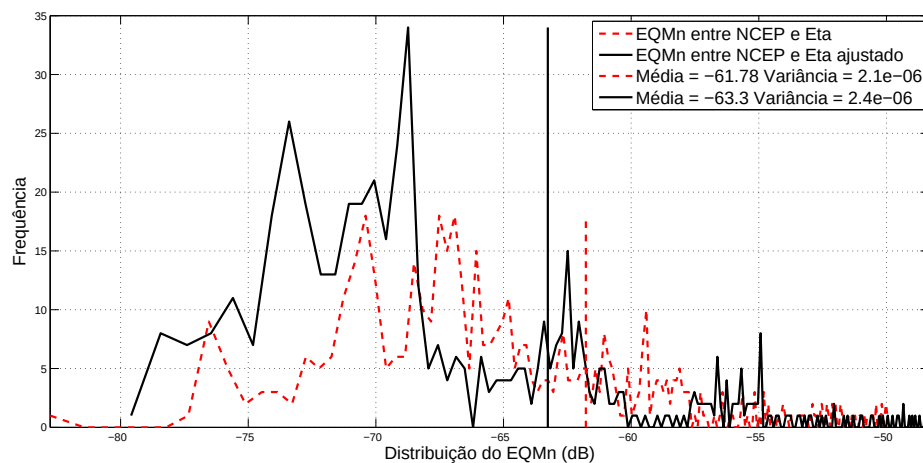


Figura 102 – Distribuição do EQMn por janela do vento meridional, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção em regiões retangulares ao longo das ordenadas, sobre a variável altura geopotencial



A Figura 103 mostra o EQMn, considerando o critério de seleção em retângulos ao longo das abscissas, por uma janela de 4 coeficientes ao longo dos dias de previsões e a Figura 104 mostra a distribuição estatística deste erro. Neste caso houve redução do EQMn em 60% do tempo de previsão e na média o EQMn reduziu em 0,5 dB.

Figura 103 – Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção em regiões retangulares ao longo das abscissas, sobre a variável altura geopotencial

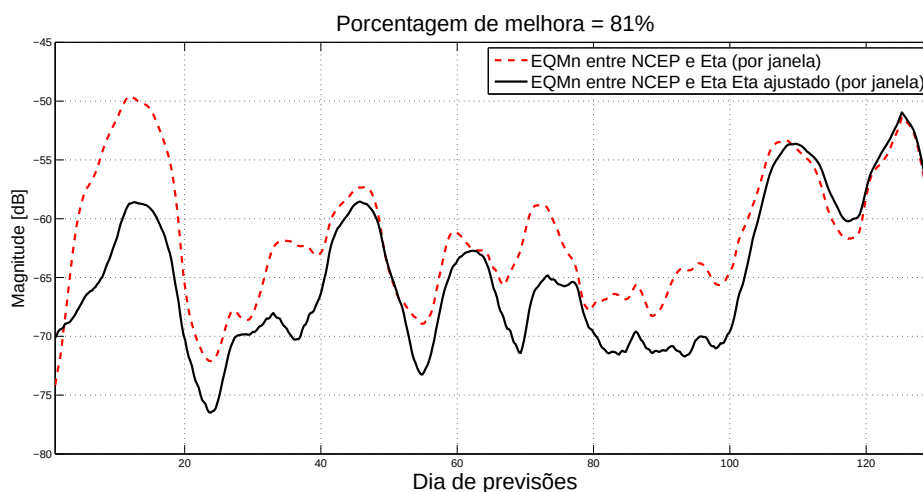
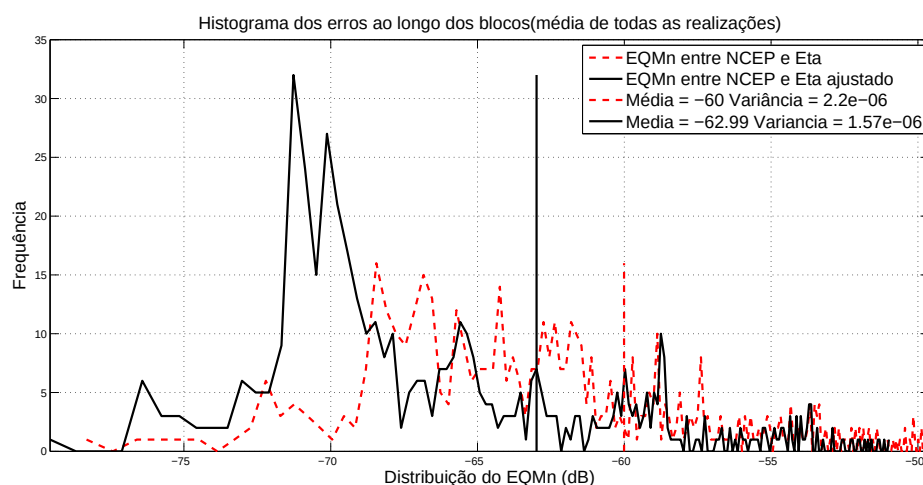


Figura 104 – Distribuição do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes da seleção em regiões retangulares ao longo das abscissas, sobre a variável altura geopotencial



A Figura 105 mostra o EQMn, considerando o critério de seleção em regiões circulares, por uma janela de 8 coeficientes ao longo dos dias de previsões e a Figura 106 mostra a distribuição estatística deste erro. Neste caso houve redução do EQMn em 71% do tempo de previsão e na média o EQMn reduziu em 3,8 dB.

Figura 105 – Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões circulares, sobre a variável altura geopotencial

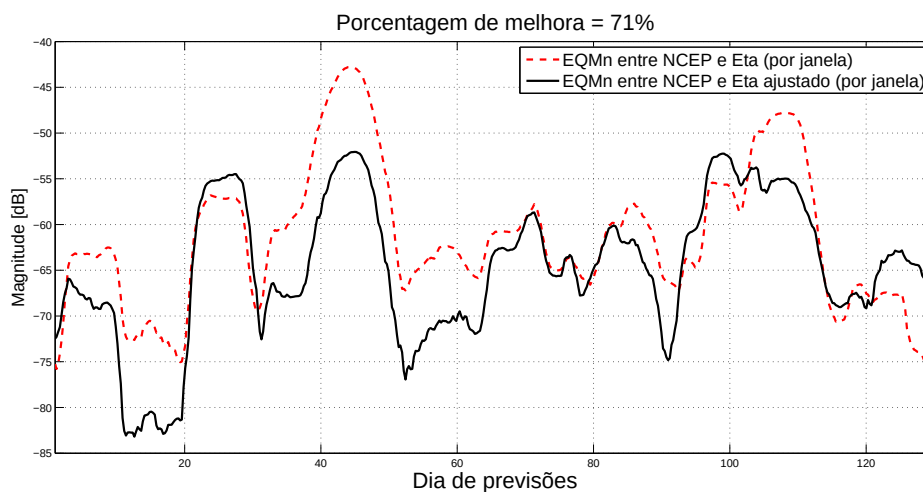
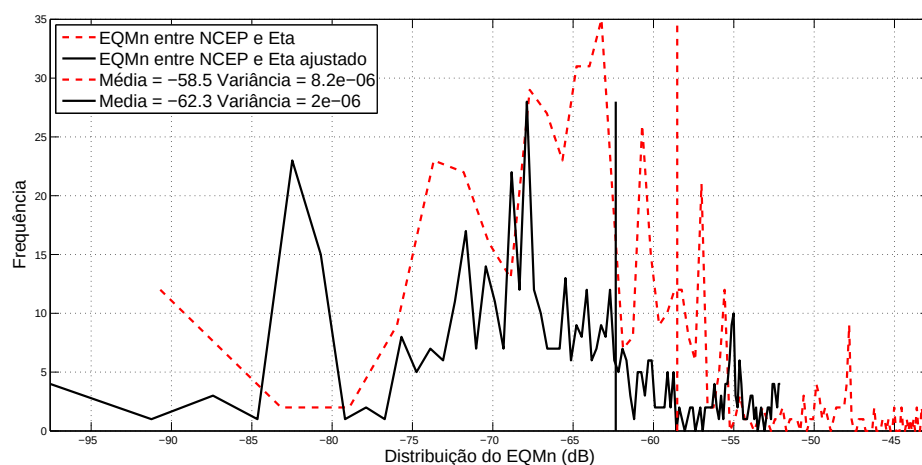


Figura 106 – Distribuição do EQMn por janela do vento meridional, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões circulares, sobre a variável altura geopotencial



A Figura 107 mostra o EQMn, considerando o critério de seleção em regiões anulares, por uma janela de 4 coeficientes ao longo dos dias de previsões e a Figura 108 mostra a distribuição estatística deste erro. Neste caso houve redução do EQMn em 54% do tempo de previsão e na média o EQMn reduziu em 0,7 dB.

Figura 107 – Média do EQMn por janela, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões anulares, sobre a variável altura geopotencial

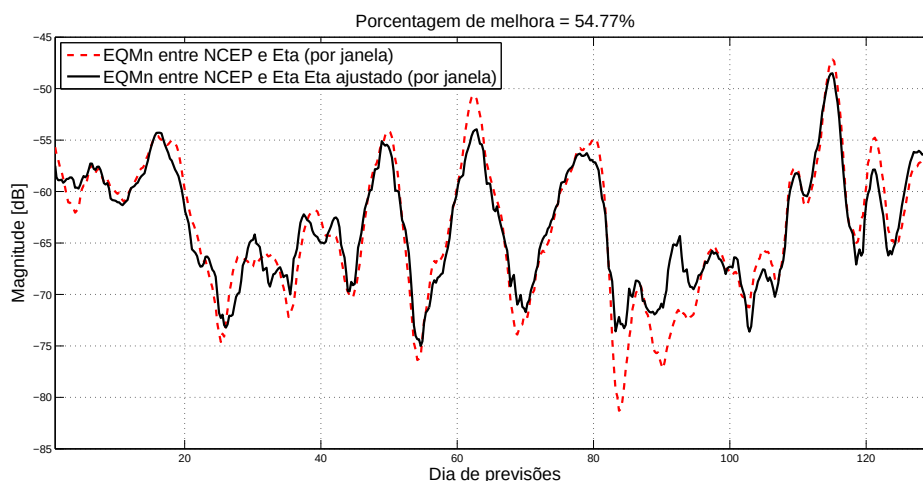
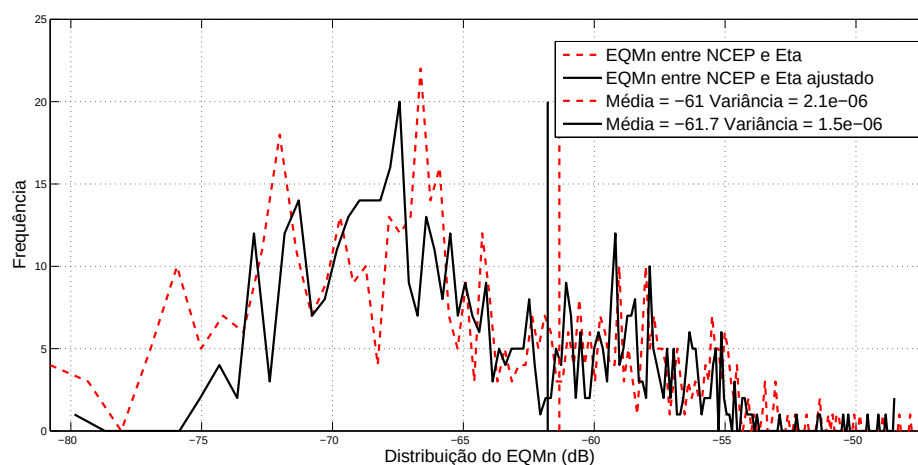


Figura 108 – Distribuição do EQMn por janela do vento meridional, na aplicação do filtro com 4 coeficientes, adaptado pelo algoritmo RLS e treinado pelas séries provenientes de regiões anulares, sobre a variável altura geopotencial



Assim como foi feito com a variável vento meridional, estes procedimentos foram repetidos para os demais anos de previsão, e calculou-se a média dos resultados encontrados para o caso da intensidade de redução do EQMn e a porcentagem do tempo de previsão que houve redução no EQMn. A Tabela 15 mostra um resumo destes resultados, sendo eles a média dos resultados de todos os anos, referente a variável altura geopotencial. A primeira coluna da tabela mostra o critério de seleção de região empregado. A segunda coluna mostra a porcentagem de regiões que tiveram redução no EQMn. Na terceira coluna tem-se a média da redução do EQMn em dB, ao longo das coordenadas geográficas. Na quarta coluna se tem a média da porcentagem do tempo de previsão que teve redução no EQMn.

Tabela 15 – Tabela comparativa dos critérios de seleção de regiões propostos sobre a variável altura geopotencial

Critério de seleção de regiões	EQMn ao longo das coordenadas geográficas	Intensidade de redução no EQMn	EQMn ao longo do tempo de previsão
<i>k-means</i>	100%	2,2 dB	76%
Ordenadas	76%	1,6 dB	73%
Abcissas	77%	2,0 dB	62%
Circular	84%	3,3 dB	63%
Anel	74%	2,0 dB	56%

Pela Tabela 15, na segunda coluna, percebe-se que o critério de seleção de regiões feito por meio do algoritmo *k-means* foi o que apresentou melhor resultado, reduzindo o EQMn em toda a região. Na terceira coluna, percebe-se que o critério de seleção de regiões circulares foi o que apresentou maior redução do EQMn, no caso, 3,3 dB. Na

quarta coluna o critério de seleção de regiões por meio do *k-means* foi o que apresentou o melhor resultado, reduziu na média o EQMn em 76% do tempo de previsões. Entretanto, percebe-se que os dois critérios de seleção de regiões mencionados, *k-means* e regiões circulares, apresentaram bons resultados nos três critérios estudados. Assim, de acordo com os resultados obtidos, para a altura geopotencial, para treinar filtros com 4 coeficientes e adaptados pelo algoritmo RLS, os melhores critérios de seleção de regiões são os critérios utilizando o *k-means* e a seleção de regiões circulares. O critério de seleção de regiões circulares é um bom critério para esta variável devido ao seu comportamento. Como pode ser visto pela Figura 22 a variável altura geopotencial, apresenta ao longo do espaço, isolinhas circulares, sendo assim, o critério de seleção em regiões circulares acaba selecionando series com características semelhantes entre si.

CONCLUSÃO

Este trabalho apresentou resultados da aplicação de filtros adaptativos, utilizando os algoritmos NLMS e RLS, para a redução de desvios em previsões climáticas. Verificou-se que a aplicação de filtros permite diminuir o EQMn entre a previsão do modelo Eta e os dados de reanálise do NCEP, produzindo uma previsão climática mais acurada.

Os resultados apresentados mostram que o filtro adaptado pelo algoritmo RLS apresenta melhores resultados por reduzir o EQMn e a dispersão do erro, do que filtros adaptados pelo algoritmo NLMS.

Percebe-se ainda que o número de séries utilizadas para realizar o treinamento do filtro pouco influencia nos resultados, para os casos de seleção de regiões utilizando o *k-means*, seleção de regiões retangulares e seleção de regiões circulares.

A estratégia de separação de regiões utilizando o *k-means* reduziu o EQMn em toda a região tanto para a variável vento meridional quanto para a variável altura geopotencial. Como mencionado, este critério é o único que utiliza o comportamento das variáveis para realizar a seleção de regiões, podendo agrupar as regiões de forma mais eficiente que os demais critérios, o que pode justificar seus resultados superiores perante os outros critérios.

Percebeu-se que para a variável vento meridional, o filtro de 4 coeficientes, adaptado pelo algoritmo RLS e séries climáticas escolhidas pelos critérios de seleção de regiões por meio do algoritmo *k-means* e seleção regiões retangulares ao longo das ordenadas e das abscissas, apresentaram melhores resultados do que os demais. O melhor conjunto de estratégia de separação em regiões retangulares ao longo das ordenadas reduziu o EQMn em 99% da região para o vento meridional. O vento meridional é uma componente do vento na direção do meridiano (Norte - Sul). Esta característica do vento meridional pode ser um fator que explica o bom resultado da aplicação do filtro com conjunto de treinamento definido por regiões retangulares ao longo do eixo das ordenadas.

No caso da variável altura geopotencial, de acordo com os testes feitos, percebeu-se que o filtro de 4 coeficientes, adaptado pelo algoritmo RLS e séries climáticas escolhidas pelos critérios de seleção de regiões por meio do algoritmo *k-means* e seleção regiões circulares, apresentaram melhores resultados do que os demais.

Os resultados obtidos mostram que o filtro com 4 coeficientes, adaptado pelo algoritmo RLS em conjunto com o critério de seleção de regiões por meio do algoritmo *k-means* se apresenta como a estratégia mais robusta e versátil na adaptação dos dados de previsão atmosférica. Isso é justificado pois ele reduziu o EQMn em toda a região selecionada tanto para a variável vento meridional quanto para a variável altura geopotencial.

Para trabalhos futuros pretende-se calcular o desempenho do filtro sobre outros níveis de pressão, e também sobre outras variáveis. Pretende-se ainda utilizar o *k-means* usando o vento meridional e a altura geopotencial conjuntamente na clusterização. Verificar entre outras coisas, o número de clusters necessários para se obter redução no EQMn em toda a região e se há mudança no comprimento do filtro que apresenta o melhor re-

sultado. Pretende-se ainda implementar no modelo Eta (em todos os níveis de pressão) a configuração que apresentou melhores resultados no treinamento do filtro adaptado, para com isso comprovar a aplicação prática do método proposto.

REFERÊNCIAS

- ALMEIDA, J. M. V. D.; BEZERRA, W. D. A.; MEDEIROS, J. X. D. Perdas de safra do programa de garantia da atividade agropecuária (proagro): Um estudo de caso para a cultura do milho no rio grande do sul no período de 1997-2007. *XVI Congresso Brasileiro de Agrometeorologia*, Setembro 2009.
- BLACK, T. L. The new nmc mesoscale eta model: Description and forecast examples. *Development Division, National Meteorological Center, NWS/NOAA, Whashington, D.C.*, June 1994.
- BONATTI, J. P. *Modelo de circulação geral atmosférico do CPTEC*. 2015. Disponível em: <http://climanalise.cptec.inpe.br/~rclimanl/boletim/cliesp10a/bonatti.html>.
- BRAVO, J. M.; COLLISCHONN, W.; TUCCI, C. E. M. Avaliação dos benefícios de previsões de vazão na operação de reservatórios. *RBRH - Revista Brasileira de Recursos Hídricos*, v. 1, n. 3, p. 197–213, Jan-Mar 2008.
- BUSTAMANTE, J. F. et al. Uma avaliação da previsibilidade de tempo do modelo eta para a américa do sul. *Revista Brasileira de Meteorologia*, v. 20, n. 1, p. 59–70, Abril 2005.
- CHEN, W.-K. *Passive, Active, and Digital Filters*. [S.l.]: CRC Press, 2009.
- CHOU, S. C.; BUSTAMANTE, J.; GOMES, J. L. Evaluation of eta model seasonal precipitation forecasts over south america. *Nonlinear Processes in Geophysics*, v. 12, p. 537–555, 2005.
- CHOU, S. C. et al. Refinamento estatístico das previsões de temperatura a 2 m do modelo eta em estações do nordeste do brasil. *Revista Brasileira de Meteorologia*, v. 22, n. 3, p. 287–296, Abril 2006.
- COHEN, L. *TimeFrequency Analysis*. [S.l.]: Prentice Hall PTR, 1995.
- COLA. *Grid Analysis and Display System (GrADS)*. 2015. Disponível em: <http://iges.org/grads/>.
- COLA. *Grid Analysis and Display System (GrADS)*. 2015. Disponível em: <http://www6.cptec.inpe.br/ManualGrADS/tutorial.html>.
- DINIZ, P. S. R. *Adaptative Filtering: Algorithms and Practical Implementation*. [S.l.]: Springer, 2008.
- DINIZ, P. S. R.; SILVA, E. A. B. da; NETTO, S. L. *Processamento Digital de Sinais, Projeto e Análise de Sistemas*. [S.l.]: Bookman, 2004.
- FELS, S. B.; SCHWARZKOPF, M. D. The simplified exchange approximation: A new method for radiative transfer calculations. *J. Atmos. Sci.*, p. 1475–1488, 1975.
- GUIMARÃES, J. M.; DERECHYNSKI, C. P.; C., C. S. Avaliação do modelo eta durante um episodio de chuvas intensas na regio da serra do mar. *Congresso Brasileiro de Meteorologia*, v. 4, 2006.

- HAN, J.; KAMBER, M.; PEI, J. *Data Mining: Concepts and Techniques*. 3. ed. [S.l.]: Elsevier, 2012.
- HAYKIN, S. *Adaptive Filter Theory*. 3. ed. [S.l.]: Prentice-Hall, 1995.
- HAYKIN, S.; VEEN, B. V. *Sinais e Sistemas*. [S.l.: s.n.], 2001.
- JAMES, L. K. I. et al. *The COLA Atmosphere-Biosphere General Circulation Model: Formulation*. [S.l.], 1997.
- JANJIC, Z. The step-mountain coordinate model: further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon Weather Rev*, p. 927–945, 1994.
- JONES, J. W. et al. *Potential benefits of climate forecasting to agriculture*. [S.l.]: ELSEVIER, 2000. 169–184 p.
- KUHN, E. V. *Modelagem estocástica do algoritmo NLMS: Revisão e aprimoramentos*. Dissertação (Mestrado) — Universidade Federal de Santa Catarina, 2012.
- LAPRISE, R. et al. Challenging some tenets of regional climate modelling. *Meteorology and Atmospheric Physics*, v. 100, p. 3–22, August 2008.
- MEIRES, L. S. *Reanálise CFSR - NCEP aplicada a modelo SiB2 na Simulação de Fluxos de CO₂ sobre o cultivo de Soja em Cruz Alta*. Dissertação (Mestrado) — Universidade Federal do Rio Grande do Sul, 2012.
- MESINGER. A blocking technique for representation of mountains in atmospheric models. *Rivista Meteor. Aeronautica*, v. 44, n. 1-4, p. 195–202, 1984.
- MESINGER, F.; ARAKAWA, A. Numerical methods used in atmospheric models. *GARP Publications Series*, v. 1, n. 17, p. 94, 1976.
- MESINGER, F. et al. The step-mountain coordinate: Model description, and performance for cases of alpine lee cyclogenesis and for a case of an appalachian redevelopment. *Mon. Wea.*, v. 116, p. 1493–1518, 1988.
- MOURA, R. G. de; HERDIES, D. L.; ROZANTE, J. R. Avaliação da pressão ao nível médio do mar no modelo eta utilizando as análises do cptec e ncep. In: *Instituto Nacional de Pesquisas Espaciais-CPTEC/INPE*. [S.l.: s.n.], 2010.
- NOAA. *Global Forecast System*. 2015. Disponível em: <<https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forecast-system-gfs>>.
- PINTO, L. I. C.; DINIZ, L. M. F.; COSTA, M. H. Comparação entre os campos de precipitação de reanálise do cptec e ncep/ncar para a américa do sul. *Revista Brasileira de Meteorologia*, 2008.
- ROZANTE, J. R. *O Impacto do aninhamento no modelo Eta nas previsões de Tempo*. Dissertação (Mestrado) — Instituto Nacional de Pesquisas Espaciais, 2004.
- ROZANTE, J. R.; CAVALCANTI, I. F. A. Influência da condição inicial na simulação do evento de precipitação extrema na serra do mar. *Congresso Brasileiro de Meteorologia*, v. 14, 2006.

SAHA, S. et al. The ncep climate forecast system reanalysis. *AMERICAN METEOROLOGICAL SOCIETY*, p. 1015–1057, August 2010.

SEMAAN, G. S. et al. Proposta de um método de classificação baseado em densidade para a determinação do numero ideal de grupos em problemas de clusterização. *Brazilian computational Intelligence Society*, v. 10, n. 4, p. 242–262, 2012.

SILVA, E. F. da et al. Desenvolvimento de um filtro de resposta ao impulso finito em vhdl. *Congresso Norte e Nordeste de Pesquisa e Inovação*, Outubro 2012.

WORLEY, S. *NCEP Climate Forecast System Reanalysis*. 2015. Disponível em: <http://rda.ucar.edu/pub/cfsr.html>.