

# Filtragem adaptativa para a redução de desvios em séries temporais de previsão numérica climática

Alexandre Ribeiro Freitas, Michel Pompeu Tcheou, Lisandro Lovisoló, Sin Chan Chou

**Resumo**—Este artigo 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. A variável climática estudada é o vento meridional na região coberta pelo modelo de previsão atmosférico Eta/CPTEC com resolução de 40km. Modelos numéricos apresentam imprecisão nas previsões climáticas. Utilizaram-se filtros adaptativos de forma a corrigir o erro de previsão. Resultados mostram que o filtro adaptativo reduz o erro quadrático médio das previsões do modelo Eta em relação às reanálises do NCEP, melhorando o desempenho da previsão climática.

**Palavras-Chave**—Filtro adaptativo, Previsão atmosférica, k-means, NLMS,RLS.

**Abstract**—This article presents results of the application of adaptive filters, using algorithms NLMS (*Normalized Least Mean Square*) and RLS (*Recursive Least Square*), for the reduction of deviations in weather forecasts. The climatic parameter that we consider is the meridional wind in the region covered by the atmospheric forecast model Eta/CPTEC with resolution of 40km. Numerical models provide atmospheric forecast that have errors. We used adaptive filters to correct the forecast error. Results show that the adaptive filter reduces the mean square error of the Eta model with respect to the re-analysis provided by NCEP, improving weather forecasting.

**Keywords**—Adapted filters, weather forecast, k-means, NLMS, RLS.

## I. INTRODUÇÃO

O modelo regional Eta [2], [3] é utilizado para previsão numérica pelo CPTEC/INPE (Centro de Previsão de Tempo e Estudos Climáticos/Instituto Nacional de Pesquisas Espaciais).

O modelo Eta considerado apresenta resolução espacial de 40km e utiliza como condição inicial os dados de análise fornecidos pelo NCEP (*National Centers for Environmental Prediction*) [1] e como condições de contorno, as previsões do modelo global de circulação geral atmosférico (MCGA) do CPTEC [3]. O NCEP fornece um 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*). A Figura 1 mostra em linha tracejada as grades do modelo de maior resolução espacial (Eta de 40 km), e em linhas contínuas as grades do modelo global MCGA de menor resolução e a região sombreada as condições de contorno laterais provenientes do NCEP. As condições de contorno lateral impactam fortemente

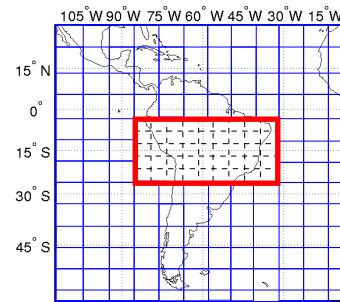


Fig. 1. Resolução espacial do modelo Eta, MCGA e das condições de contorno fornecidas pelo NCEP.

as integrações de prazos climáticos, de meses a décadas. Já as condições iniciais afetam as integrações a curto e médio prazo, de alguns dias a algumas semanas [4]. Para previsões sazonais (quatro meses e meio), os modelos numéricos apresentam erros na previsão atmosférica inerentes à modelagem [5]. O nível de qualidade da condição inicial utilizada compromete a qualidade das previsões obtidas pelo modelo e há erros ocasionados pelas condições de contorno lateral que podem propagar-se para o interior do domínio e comprometer as previsões [4].

Dados meteorológicos podem ser abordados como sinais discretos tridimensionais (em função da latitude, longitude e altitude) aleatórios no tempo. Além disso, para cada altitude, a área compreendida, pode ser entendida como imagem digital, onde os pixels contêm dados meteorológicos de uma determinada coordenada geográfica (Latitude/Longitude) [6]. Neste contexto, pretende-se utilizar filtros adaptativos para filtrar as previsões na área compreendida pelo modelo de modo a minimizar os erros entre as previsões do Eta de 40 km e os dados de reanálise do NCEP, e com isso minimizar os erros propagados para modelos de mais alta resolução cuja condição de contorno provem do Eta de 40km.

## II. ADAPTAÇÃO DE FILTROS

Neste trabalho adaptaram-se filtros com o objetivo de melhorar a previsão sazonal do Eta 40km relativamente às reanálises do NCEP. Esta abordagem foi escolhida pela sua aplicabilidade em diversas áreas envolvendo otimização e problemas de identificação de sistemas e equalização [7], [8]. Geralmente, a adaptação requer a existência de um sinal de referência que rege a aproximação do filtro. A Figura 2 mostra a configuração geral empregada para adaptar os filtros [8],  $x(n)$  é o sinal de entrada,  $y(n)$  é a saída do sistema adaptativo e  $d(n)$  é definido como o sinal 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 adaptável de modo que o sinal de erro  $e(n)$  seja minimizado [8].

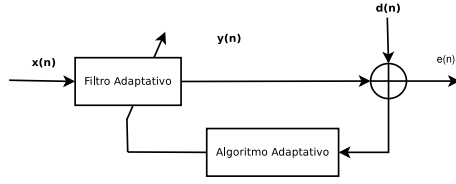


Fig. 2. Configuração geral de um sistema adaptativo.

A saída do sistema adaptativo é na forma

$$y(n) = \sum_{i=0}^N w_i(n)x(n-i) = \mathbf{w}^T(n)\mathbf{x}(n), \quad (1)$$

com  $\mathbf{x}(n) = [x(n) \ x(n-1) \ x(n-2) \ \dots \ x(n-N)]^T$  o vetor de amostras do sinal de entrada atrasado de  $N$  instantes de tempo, e  $\mathbf{w}(n) = [w_0(n) \ w_1(n) \ \dots \ w_N(n)]^T$  o vetor de coeficientes do filtro adaptável [8].

O algoritmo de adaptação empregado influencia uma série de aspectos do processo adaptativo, como por exemplo a complexidade computacional e a velocidade de convergência. Os algoritmos utilizados foram o NLMS (*Normalized Mean Square*) e o RLS (*Recursive Least Square*). Optou-se pelo algoritmo NLMS por possuir um fator de normalização proporcional às alterações de potência no sinal de entrada, possibilitando assim uma convergência mais rápida do algoritmo em comparação ao algoritmo LMS (*Least Mean Square*). A equação de atualização no LMS é dada por [8]

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

onde  $\mu$  é o passo (que é escolhido de forma a garantir a convergência). Já o algoritmo NLMS emprega a equação de atualização dos coeficientes,

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu_{norm}}{\gamma + \mathbf{x}^T(n)\mathbf{x}(n)} e(n)\mathbf{x}(n) \quad (3)$$

onde  $e(n) = d(n) - \mathbf{x}^T(n)\mathbf{w}(n)$ ,  $\mu_{norm}$  é o passo,  $\gamma$  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. De modo a garantir a estabilidade, o fator de convergência deve ser escolhido no intervalo  $0 < \mu_{norm} < 2$  [8].

O algoritmo RLS é um algoritmo de estimação paramétrica que procura minimizar os erros quadráticos da saída. O algoritmo RLS também apresenta melhores resultados para sistemas sujeitos a rápidas variações na amplitude do sinal [9]. A atualização dos coeficientes no RLS é dada por

$$\mathbf{w}(n) = \mathbf{R}_D^{-1}(n)\mathbf{p}_D(n) \quad (4)$$

onde

$$\mathbf{R}_D^{-1}(n) = \left[ \sum_{i=0}^n \lambda^{n-i} \mathbf{x}(i)\mathbf{x}^T(i) \right]^{-1}, \quad (5)$$

$$\mathbf{p}_D(n) = \sum_{i=0}^n \lambda^{n-i} \mathbf{x}(i)d(i), \quad (6)$$

sendo  $\mathbf{R}_D^{-1}(n)$  a inversa da matriz de correlação do sinal de entrada e  $\mathbf{p}_D(n)$  é a matriz de correlação cruzada entre o vetor

de entrada e o sinal desejado,  $\lambda$  é dita “memória” do algoritmo e o seu valor deve estar entre  $0 \leq \lambda \leq 1$  [7].

### III. METODOLOGIA

Devido a diferenças nas grades de resolução espacial do Eta e do NCEP, é necessário ajustá-las. Originalmente, a resolução espacial do Eta é uma matriz  $61 \times 126$  e a do NCEP uma matriz  $49 \times 101$ . Após o ajuste, as matrizes apresentam dimensões de  $244 \times 504$  com um total de 122976 elementos. Este ajuste é feito por meio da realocação dos índices da matriz do Eta. Os valores de cada linha e coluna do modelo Eta são replicados por quatro enquanto as reanálises do NCEP são replicados por cinco. Por fim aplica-se um filtro espacial passa baixas, com frequência de corte normalizada igual a 0,4, de modo a eliminar as possíveis variações bruscas dos valores entre as posições geográficas. A frequência de corte foi escolhida de modo a não causar perda de informação nas altas frequências do sinal.

O diagrama de blocos da Figura 3 mostra o método utilizado para a aplicação do filtro adaptável. Na etapa (i) da Figura 3 têm-se os dados volumétricos, com os eixos horizontal (Latitude) e vertical (Longitude) representando as posições geográficas das séries climáticas ao longo do tempo (profundidade). Em seguida, na etapa (ii) é selecionada uma região associada a um conjunto de séries climáticas. O tamanho destas séries depende do horizonte de previsão, podendo ser sazonal ou diário. Na etapa (iii) as séries climáticas são utilizadas para treinar o filtro adaptável. E por fim, na etapa (iv), o filtro encontrado é aplicado no mesmo horizonte de previsão, na mesma região do ano subsequente. Por exemplo, se o ano de treinamento foi 2002, o ano que será ajustado pelo filtro será o ano de 2003, e assim sucessivamente.

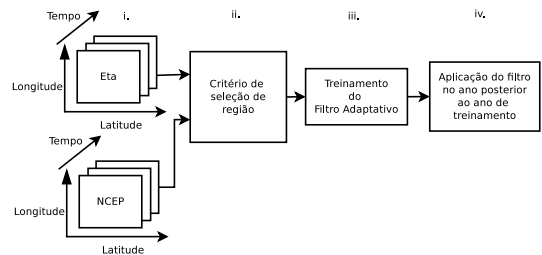


Fig. 3. Método utilizado para a implementação do filtro adaptável.

#### A. Critérios de seleção de regiões

A seleção de região mais trivial é utilizar toda a região para o treinamento do filtro, entretanto, testes mostraram que mesmo com um grande número de séries, os filtros assim obtidos não geraram redução no erro. Isso acontece pois realizações em pontos geográficos distantes uns dos outros que apresentam características estatísticas distintas, comprometem a qualidade do treinamento. Neste ponto, percebe-se a necessidade de agrupar as regiões que possuam alguma similaridade. Em (ii) foram utilizados três critérios de seleção de regiões de forma a agrupá-las em função de características estatísticas semelhantes.

##### A.1 - Primeiro Critério de Seleção (k-means)

O primeiro critério de seleção é a “clusterização” utilizando o algoritmo *k-means* [11], [10], que tem por objetivo classificar e separar regiões que apresentam comportamento semelhante conforme uma métrica pré estabelecida. Para realizar a classificação, uma função objetivo avalia a qualidade do particionamento de modo que os objetos dentro de um grupo sejam semelhantes um ao outro, mas diferentes de objetos de outros grupo [11]. Seja um conjunto de dados,  $D$ , contendo  $q$  objetos no espaço euclidiano. O *k-means* particiona os objetos em função de suas distâncias aos centróides dos clusters. O algoritmo *k-means* distribui os objetos de  $D$  em  $k$  clusters,  $C_j, \dots, C_k$ , com  $C_j \subset D$  e  $C_i \cap C_j = \emptyset$  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 representação da posição do cluster, é medido como  $dist(\mathbf{c}_i, \mathbf{v})$ , por exemplo, a distância euclidiana entre o objeto  $\mathbf{v}$ , e o centróide a ele mais próximo  $\mathbf{c}_i$ . Dessa forma a função objetivo é dada como a soma dos erros para todos os objetos no conjunto de dados [11]

$$E = \sum_{i=1}^k \sum_{\mathbf{v} \in C_i} dist(\mathbf{c}_i, \mathbf{v})^2 \quad (7)$$

o parâmetro  $k$  é o número de clusters,  $\mathbf{v}$  é o 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 euclidiana entre o objeto  $\mathbf{v}$ , e o centróide a ele mais próximo  $\mathbf{c}_i$ .

O funcionamento do *k-means* se dá primeiramente com a escolha aleatória de  $k$  distintos objetos em  $D$  que representam inicialmente os centros dos clusters. Depois associa-se cada objeto ao cluster mais semelhante a ele, com base na distância euclidiana entre o objeto e o centro do cluster. O algoritmo em seguida, de forma iterativa reduz a variação de  $E$  dentro do cluster. Para cada cluster, calcula-se o centróide usando os objetos atribuídos ao cluster na iteração anterior. Todos os objetos são, então, redistribuídos utilizando os centróides atualizados dos clusters. As iterações continuam até que a redistribuição fique estável, isto é, os clusters formados na rodada atual são os mesmos formados na rodada anterior [11].

O *k-means* é utilizado para agrupar regiões que apresentam médias semelhantes, de modo que estes dados sejam utilizados para o treinamento do filtro.

### A.2 - Segundo Critério de seleção (Bordas)

Este critério separa as regiões em função das bordas do modelo. As séries temporais de cada borda (superior, inferior e bordas laterais) são tratadas como conjuntos estatísticos distintos. Dentro deste critério teremos quatro possibilidades de seleção de regiões e cada uma será avaliada separadamente.

### A.3 - Terceiro Critério de seleção (região circular)

O terceiro critério de seleção consiste em separar regiões em círculos em torno de um determinado pixel específico (coordenada geográfica).

## B. Métricas de desempenho

As métricas de desempenho utilizadas para validar as séries temporais decorrentes da adaptação dos dados de previsão numérica do Eta em relação aos dados de reanálise do NCEP,

são o erro quadrático médio normalizado (EQMn) e sua variância.

O EQMn entre o sinal desejado (NCEP) e os sinais do modelo Eta (adaptado ou não) é usado para avaliar se o filtro encontrado realizou melhoras na previsão. O EQMn é dado pela seguinte equação,

$$EQMn = \frac{1}{T} \frac{\sum_{n=1}^T (\mathbf{d}(n) - \mathbf{x}'(n))^2}{\sum_{n=1}^T (\mathbf{d}(n))^2} \quad (8)$$

onde  $\mathbf{x}'(n)$  é uma série qualquer e  $\mathbf{d}(n)$  como a série de referência,  $T$  é a quantidade de pontos da série. A redução do EQMn denota que o filtro reduziu o erro existente entre as previsões fornecidas pelo Eta e as reanálises do NCEP, ou seja, houve melhora na qualidade dos dados de previsão numérica.

A variância  $var(X) = E((X - m_x)^2)$ , onde  $m_x = E(X)$  é o valor esperado (média) da variável aleatória  $X$ . Assim, estima-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.

## IV. RESULTADOS DE SIMULAÇÃO COMPUTACIONAL

### A. Configurações

O parâmetro climático utilizado para analisar nossa proposta de seleção de dados de previsão numérica foi o vento meridional, em  $m/s$  a pressão de 250 mbar. O período de previsão é de 2001 a 2010, compreendendo sempre o período de 13 de dezembro a 30 de abril do ano seguinte. As séries utilizadas, chamadas de previsão sazonal, apresentam um total de 133 dias de previsão [2], [3]. O vento meridional foi escolhido por ser uma das variáveis de previsão que mais influencia na modelagem atmosférica [3]. O modelo Eta e as reanálises do NCEP apresentam resolução temporal de seis horas, com isso, para cada dia, há quatro previsões. Com o objetivo de ajustar os dados provenientes do Eta, foi utilizado um filtro adaptativo. Foi utilizado um filtro com quatro coeficientes (um dia de previsão) e um filtro com trinta e dois coeficientes (oito dias de previsão). Para o algoritmo NLMS utilizou-se o fator de convergência  $\mu_{norm} = 0,1$  e para o RLS a constante de esquecimento foi  $\lambda = 0,999$ .

Para cada região, a estratégia da Figura 3 é aplicada dez vezes em cada ano, de modo a formar estatísticas confiáveis. A cada realização, a sequência de entrada das séries no filtro é aleatoriamente selecionada. No caso do critério de seleção utilizando uma região circular, os centros das circunferências que formam o conjunto de treinamento também são aleatoriamente sorteados. Um dos fatores importantes quando se trata de filtragem adaptativa é o número de séries usadas para treinar o filtro. As quantidades de séries utilizadas são definidas pelo critério de seleção das bordas, pois este critério tem limitação. As bordas superior e inferior possuem 1008 séries cada, enquanto as bordas laterais 504. Para se ter um mesmo critério de comparação, optou-se por utilizar aproximadamente 500 e 1000 séries em todos os outros critérios de seleção de regiões.

### B. Resultados

Pode-se utilizar a comparação da porcentagem de regiões nas quais se obtém redução do EQMn normalizado como

critério de avaliação global. As Tabelas I e II apresentam resultados referentes a aproximadamente 500 séries temporais no conjunto de treinamento para a aplicação de filtros adaptativos na previsão do Eta. A Tabela I é referente a ordem do filtro correspondente a um dia de previsão e a Tabela II a oito dias de previsão. As Tabelas III e IV apresentam resultados referentes a aproximadamente 1000 séries temporais no conjunto de treinamento para a aplicação de filtros adaptativos na previsão do Eta. A III é referente a ordem do filtro correspondente a um dia de previsão e a IV a oito dias de previsão.

TABELA I

PORCENTAGEM DE REGIÕES QUE SE OBTÉM REDUÇÃO NO EQMn APLICANDO FILTROS DE 1 DIA (COMPRIMENTO), COM CARDINALIDADE DO TREINAMENTO DETERMINADA PELAS FORÇANTES HORIZONTAIS

Ano de aplicação	k-means		Borda Esquerda		Borda direita		Região circular	
	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	100	100	81,7	93	82,4	92	100	100
2003	100	100	56	100	80	100	1	100
2004	94,7	100	95,9	100	47	100	20,5	35,1
2005	100	100	71,8	100	48	90,8	10	100
2006	100	100	71,8	100	28,1	87	30	100
2007	98,8	100	57,7	100	100	100	80	100
2008	87,3	100	61,6	100	5,4	89,8	82	100
2009	100	100	96,7	93	97,5	100	70	100
2010	100	100	40	100	10	70	0	100
Média	97,9	100	70,3	98,4	55,4	92,2	43,7	92,8

TABELA II

PORCENTAGEM DE REGIÕES QUE SE OBTÉM REDUÇÃO NO EQMn APLICANDO FILTROS DE 8 DIAS (COMPRIMENTO), COM CARDINALIDADE DO TREINAMENTO DETERMINADA PELAS FORÇANTES HORIZONTAIS.

Ano de aplicação	k-means		Borda Esquerda		Borda direita		Região circular	
	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	100	100	35,5	74,2	59	59	100	100
2003	100	100	80,8	100	46	100	100	100
2004	0	74	38,2	100	59	98,7	30	87,5
2005	39,6	96,7	84	100	28	90,4	12,1	49,8
2006	100	100	95,3	100	68	81,2	100	77,5
2007	100	99,8	93,7	100	62,6	99,2	34,8	83,5
2008	63,8	78	56,4	23,6	18,8	79,6	33	89
2009	100	100	77	100	52	96,5	100	100
2010	100	100	10	100	0	53	29	42
Média	77,8	94,3	63,4	88,6	43,7	84,2	59,9	81

As colunas das Tabelas I, II, III e IV apresentam resultados dos métodos de seleção da região do conjunto de treinamento na aplicação dos algoritmos adaptativos NLMS e RLS. Por estas tabelas sabe-se que o uso de filtros com comprimento de um dia de previsão gerou na média a redução do EQMn em uma porcentagem maior de regiões que os filtros de comprimento de oito dias de previsão. Percebe-se que os filtros adaptados por meio do algoritmo RLS reduziu o EQMn em uma porcentagem superior de regiões que os filtros adaptados pelo algoritmo NLMS. Observa-se ainda, que a separação de regiões utilizando k-means mostrou-se a melhor estratégia pois apresentou redução do EQMn em todas as partes das regiões nos ajustes de todos os anos. Por fim percebe-se que é melhor utilizar 500 amostras no treinamento do filtro, pois a quantidade de amostras não influenciou significativamente os resultados.

TABELA III

PORCENTAGEM DE REGIÕES QUE SE OBTÉM REDUÇÃO NO EQMn APLICANDO FILTROS DE 1 DIAS (COMPRIMENTO), COM CARDINALIDADE DO TREINAMENTO DETERMINADA PELAS FORÇANTES VERTICAIS.

Ano de aplicação	k-means		Borda Superior		Borda Inferior		Região circular	
	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	100	100	70	95,4	30	95,4	100	100
2003	100	100	79	100	89	100	0	100
2004	23	70	42	100	43,3	100	3,2	35
2005	100	100	52,5	98	32,7	97,6	20	100
2006	100	100	81,8	99	19	99,5	40	100
2007	100	100	41	94	88	94,1	85	100
2008	98	100	60	100	8,3	99,9	85,2	100
2009	100	100	65	100	9,1	99,9	100	100
2010	99	100	33	98	23	100	20	100
Média	91,1	96,7	58,1	98,4	38	98,5	50,4	92,8

TABELA IV

PORCENTAGEM DE REGIÕES QUE SE OBTÉM REDUÇÃO NO EQMn APLICANDO FILTROS DE 8 DIAS (COMPRIMENTO), COM CARDINALIDADE DO TREINAMENTO DETERMINADA PELAS FORÇANTES VERTICAIS.

Ano de aplicação	k-means		Borda Superior		Borda Inferior		Região circular	
	NLMS	RLS	NLMS	RLS	NLMS	RLS	NLMS	RLS
2002	100	100	68	97	49,1	90,5	98	100
2003	100	100	96,3	99	52	98	76,4	100
2004	72,4	83	30,1	91,5	92,5	100	40	67
2005	0	100	83	95	10,8	98	36,7	73
2006	0	100	96,7	100	35,3	93	98	9
2007	52,8	100	55	93	77,2	86,6	12	86
2008	57,1	82,4	40	75,5	40,4	92	15,4	65
2009	0	90,8	61	100	42,3	96,6	98	100
2010	0	61	50	77,4	32	93,3	32,8	65,4
Média	42,5	90,8	64,5	92	48	94,4	56,4	74

No ano de 2004, pelas Tabelas II e III percebe-se que nas regiões selecionadas, no uso do algoritmo NLMS, a porcentagem de regiões com EQMn reduzido é pequena. Em [12] mostra-se que o ano de 2004 foi um ano atípico, e com isso o ano de 2003 não forneceu bons parâmetros de modelagem para o ano seguinte. Assim o NLMS não conseguiu corrigir esses erros de previsão. Contudo, pela Tabela I, 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 95,8% e 94,7% das regiões. Na Tabela IV 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 92,5% e 72,4% 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.

C. Estudo de Caso

Com o objetivo de avaliar com mais detalhes o efeito do filtro sobre as séries ajustadas, selecionou-se uma série que obteve redução do EQMn em relação ao NCEP. O algoritmo utilizado foi o RLS sobre o ano de 2006. Na Figura 4 pode-se ver a série temporal do vento meridional observada (NCEP), a série prevista pelo Eta e a série obtida através do ajuste da

previsão do Eta pelo filtro adaptado.

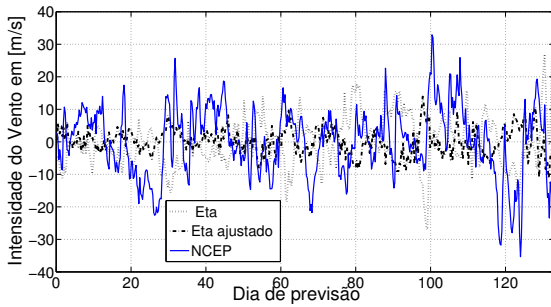


Fig. 4. Série do Vento meridional.

O EQMn entre as séries é calculado para estimar estatisticamente a melhora desta série ao longo do tempo. Para melhor identificar a distribuição do EQMn, utilizou-se uma janela deslizante com 32 coeficientes, ou seja, apenas 32 pontos da série em cada interação é utilizado para calcular o erro. A Figura 5 mostra o erro por janela e a Figura 6 mostra a distribuição estatística deste erro. Percebe-se que o EQMn entre o NCEP e a série ajustada apresenta redução em 84,86% do tempo de previsão em comparação com o EQMn entre o NCEP e a série original.

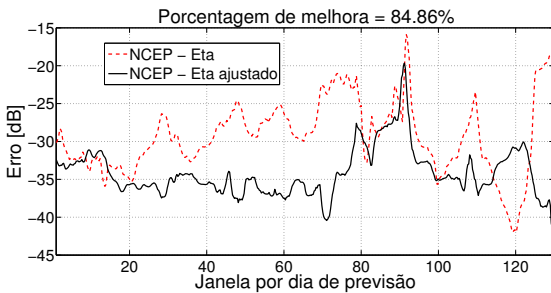


Fig. 5. Erro quadrático médio normalizado por bloco.

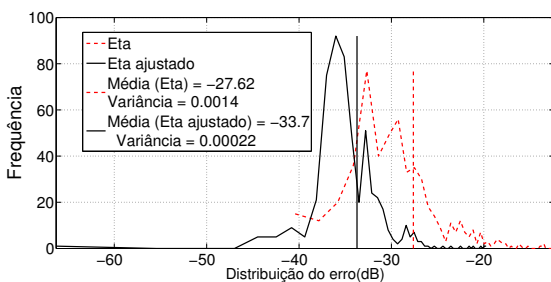


Fig. 6. Histograma do erro ao longo do bloco.

Pela Figura 6 percebe-se que a média do erro da série ajustada relativamente ao NCEP é reduzida de 6,1dB. A variância reduz aproximadamente a um sexto. A Figura 7 apresenta um retrato das séries na janela. Neste caso percebe-se claramente a melhora na previsão atmosférica.

## V. CONCLUSÕES

Este artigo apresenta resultados da aplicação de filtros adaptativos, utilizando os algoritmos NLMS e RLS, para a redução de desvios em previsões climáticas. Em geral, verificou-se que

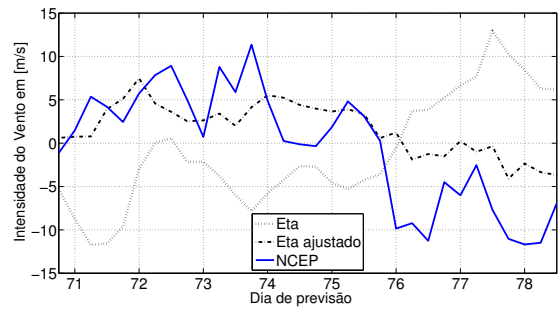


Fig. 7. Série da janela que apresentou maior redução do erro.

a aplicação do filtro adaptado leva à diminuição do EQMn entre a previsão do modelo Eta e os dados de reanálise do NCEP, produzindo uma previsão climática mais precisa. Os resultados obtidos mostram que o filtro adaptado pelo algoritmo RLS em conjunto com o critério de seleção por k-means apresenta o melhor desempenho ao reduzir o EQMn e a dispersão do erro.

Em trabalhos futuros pretende-se avaliar outros parâmetros climáticos como por exemplo a altura geopotencial e o vento zonal.

## REFERÊNCIAS

- [1] NCEP Climate Forecast System Reanalysis, "CISL Research Data Archive", <http://rda.ucar.edu/pub/cfsr.html>, acesso em 20 de Julho de 2015.
- [2] INPE/CPTEC, "Etamodel", <http://etamodel.cptec.inpe.br/>, acesso em 20 de Julho de 2015
- [3] J. R. Rozante, J. L. Gomes, J. F. Bustamante e C. S. Shan, "Uma Avaliação da previsibilidade de tempo do modelo Eta para a América do Sul" *Instituto Nacional de Pesquisas Espaciais - CPTEC/INPE, Revista Brasileira de Meteorologia*, v.20, n.1, 59-70, 2005
- [4] J. R. Rozante e I. F. A. Cavalcanti, "Influência da Condição Inicial na simulação do Evento de precipitação extrema na serra do Mar" , *Instituto Nacional de Pesquisas Espaciais - CPTEC/INPE, Congresso Brasileiro de Meteorologia*, 14. (CBMET), Florianópolis-SC, 20
- [5] J. M. Guimarães, C. P. Dereczynski, C. S. Shan, "Avaliação do modelo Eta durante um episódio de chuvas intensas na região da Serra do Mar" , *Congresso Brasileiro de Meteorologia, XIV - (CBMET), Anais...2006*.
- [6] R. C. Gonzalez, R. E. Woods, "Digital Image Processing", *Addison-Wesley Publishing Company*, 1993.
- [7] P. S. R. Diniz, E. A. B. da Silva, S. L. Netto, "Processamento Digital de Sinais, Projeto e Análise de Sistemas", *Bookman*, 2004.
- [8] P. S. R. Diniz, "Adaptive Filtering: Algorithms and Practical Implementation", *Springer*, 2008.
- [9] E. Eweda, "Comparison of RLS, LMS, and Sign Algorithms for Tracking Randomly Time-Varying Channels", *Vol. 42. NO. 11, 1994*.
- [10] G. S. Semaan, M D. Cruz, J. A. de M. Brito, L. S. Ochi, "Proposta de um método de classificação baseado em densidade para a determinação do número ideal de grupos em problemas de clusterização", *Learning and Nonlinear Models (L&NLM)- Journal of Brazilian computational Intelligence Society*, Vol. 10, Iss. 4, pp. 242-262, 2012.
- [11] Jiawei Han, Micheline Kamber, Jian Pei, "Data Mining: Concepts and Techniques", *Morgan Kaufmann*, 2012.
- [12] J. M. V. DE Almeida, W. de A. Bezerra, J. X. de Medeiros, "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", apresentado no XVI Congresso Brasileiro de Agrometeorologia de 22 a 25 de Setembro de 2009-Belo Horizonte-MG.