

Refinamento das Previsões do Modelo Eta

Edgard de Freitas Diniz Evangelista edgard@cptec.inpe.br
Chou Sin Chan chou@cptec.inpe.br

Resumo

Modelos numéricos de previsão do tempo possuem simplificações para a solução das equações que regem o comportamento da atmosfera. As equações são discretizadas e mesmo que haja aumento de resolução dos modelos, as previsões apresentarão limitações e deficiências. Utiliza-se métodos estatísticos e dados observacionais para refinar as previsões geradas pelos modelos numéricos. Este trabalho tem por objetivo utilizar o algoritmo “Model Output Calibration” (MOC) para o refinamento das previsões do vento a 10m geradas pelo modelo Eta para a estação de Ilha Rasa, no Rio de Janeiro.

O MOC é baseado em equações de regressão linear multivariada, cujos preditores são selecionados a partir de correlações lineares. Preditores que apresentam coeficiente de correlação acima de valores limites são selecionados. A correlação é entre os preditores, que são as variáveis previstas pelo modelo Eta, e o preditando, o erro da previsão da magnitude do vento de 10 metros.

Foram utilizadas duas séries de treinamento para a execução do MOC. A primeira teve um período de 21 dias, a segunda, um período de 28 dias, sendo o mês de março utilizado como base para treinamento destas séries. Os meses de abril e maio foram utilizados para se realizar as calibrações, comparação com as observações e avaliação dos resultados. Testou-se as duas séries de treinamento para três limiares de correlação, sendo: 0,8, 0,6 e 0,4. Foram comparados entre si os resultados do MOC, os resultados do Eta e a previsão por persistência do erro da previsão. Em ambas séries de treinamento, o MOC obteve o melhor desempenho para o limiar de 0,8, com redução do RMSE sobre a previsão do Eta. O erro MAE do MOC também apresentou maior redução em relação ao Eta para o limiar de 0,8. Os RMSE e MAE utilizando a correção da previsão por persistência do erro foram maiores do que utilizando o MOC. Para o erro BIAS, notou-se que a aplicação do MOC reduziu drasticamente os erros sistemáticos do modelo Eta, porém as medidas de BIAS para a magnitude persistida foram as menores para as duas séries de treinamento e os três limiares de correlação.

O algoritmo MOC atingiu o objetivo de reduzir as medidas de erro do modelo Eta (medidas RMSE, MAE E BIAS), porém essa redução ainda pode ser otimizada. Para a série de treinamento de 28 dias, o MOC teve melhor desempenho, pois a redução dos erros da previsão do modelo Eta foi mais significativa. Os resultados, apesar de indicarem que o MOC é eficaz, precisam ser aperfeiçoados, pois as correlações estão baixas e espera-se obter uma maior redução das medidas de erros.

Os próximos passos do trabalho terão como objetivo a redução dos erros de calibração do MOC e o aumento da correlação linear entre o MOC utilizando novas variáveis como preditores e séries de menor intervalo horário.



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

REFINAMENTO DAS PREVISÕES DE INTENSIDADE DO VENTO DO MODELO ETA

**RELATÓRIO PARCIAL DE PROJETO DE INICIAÇÃO CIENTÍFICA
(PIBIC/CNPq/INPE)**

**Edgard de Freitas Evangelista (UNESP, Bolsista PIBIC/CNPq)
E-mail: edgard@cptec.inpe.br**

**Dra. Chou Sin Chan (DMD/CPTEC/INPE, Orientadora)
E-mail: chou@cptec.inpe.br**

COLABORADOR

Carlos Renato (CPTEC/INPE)

Mai de 2004

SUMÁRIO

CAPÍTULO 1 – INTRODUÇÃO	1
CAPÍTULO 2 – METODOLOGIA	2
2.1 Regressão linear multivariada	2
2.2 Escolha dos preditores	3
2.3 Calibração das previsões do Eta	4
2.4 Correção por persistência de erro	5
2.5 Avaliação dos erros	5
CAPÍTULO 3 - MATERIAIS E MÉTODOS	6
3.1 Dados observacionais	6
3.2 Correção dos dados	7
3.3 Utilização dos dados	9
3.4 implementação computacional	10
CAPÍTULO 4 - RESULTADOS E ANÁLISES	11
4.1 período de treinamento de 21 dias	11
4.1.1 Análise dos gráficos e índice de desempenho	11
4.1.2 Preditores melhor correlacionados	12
4.2 Período de treinamento de 28 dias	12
4.2.1 análise dos gráficos e índice de desempenho	12
4.2.2 Preditores melhor correlacionados	13
CAPÍTULO 5 - CONCLUSÕES E TRABALHOS FUTUROS	15
5.1 Conclusões	15
5.2 Trabalhos futuros	15

1 INTRODUÇÃO

O modelo regional Eta é um modelo numérico de previsão do tempo. Está em operação desde 1996 no CPTEC (Centro de Previsão de Tempo e Estudos Climáticos) do INPE (Instituto Nacional de Pesquisas Espaciais) em Cachoeira Paulista. Este modelo é executado duas vezes por dia (0000 UTC e 1200 UTC) e as previsões geradas têm alcance atualmente de 120 horas, a intervalos de 6 horas, tendo uma resolução de representatividade de 40kmx40Km, com 38 níveis verticais.

Apesar da eficiência deste modelo, há condições em que ele não consegue realizar as previsões com boa qualidade: por exemplo, em regiões onde há grandes variações na vegetação ou no relevo, o Eta utiliza valores médios de altitude na grade de resolução de 40 Km e considera a vegetação que predomina no local, o que acarreta geralmente um aumento nos erros das previsões: por exemplo, numa mesma célula de representação podemos ter duas cidades, sendo que uma delas possui vegetação abundante, havendo inclusive árvores de médio e grande porte, enquanto outra cidade situa-se em terreno quase totalmente desmatado. Caso o predominate da célula sejam árvores, a previsão da cidade que tem este tipo de vegetação apresentará melhor qualidade. Situações como a descrita acima são muito comuns em todos os modelos numéricos que descrevem a atmosfera de forma discretizada.

É importante, então, que haja uma ferramenta que refine as previsões do modelos numéricos, minimizando os erros; a ferramenta utilizada neste trabalho é o algoritmo MOC (Model Output Calibration) que consiste em cálculos estatísticos que buscam ajustar as previsões do modelo numérico, neste caso o modelo Eta.

Objetivo: o objetivo deste trabalho é refinar as previsões de intensidade do vento a 10 metros da superfície produzidas pelo Eta.

2 METODOLOGIA

Existem basicamente três métodos estatísticos utilizados em refinamento de previsão numérica. O MOS (Model Output Statistics), as redes normais e o MOC (Model Output Calibration). Os dois primeiros métodos precisam de um período de treinamento de pelo menos dois anos. Neste período o modelo numérico precisa estar “congelado” para não alterar as estatísticas. Apesar dos bons resultados destes métodos, o congelamento do modelo é uma desvantagem. Neste trabalho será testado o MOC para as previsões do Eta.

O MOC durante um período, que pode variar de duas a quatro semanas, coleta os dados de que necessitará para os cálculos: os erros de previsão da variável a ser prevista, os dados observacionais e todas as medidas das variáveis atmosféricas previstas pelo Eta.

O MOC se baseia em cálculos de regressão linear multivariada para realizar a calibração das previsões do Eta, pois o MOC faz a calibração subtraindo-se da magnitude prevista pelo Eta, o erro da previsão. Para calcular esse erro, é necessário que se conheçam os coeficientes da equação de regressão linear. Os coeficientes das equações são calculados para uma série inicial de 21 ou 28 dias. Estes períodos são definidos como período de treinamento. Assim, foi utilizado o mês de março para a construção da equação de regressão linear multivariada e os meses de abril e maio para se refinar as previsões.

Foram empregadas as seguintes definições: a magnitude prevista pelo Eta sem a calibração V_{ETA} ; magnitude do vento prevista pelo Eta e calibrada pelo MOC V_{MOC} e a magnitude do vento observada V_{OBS} . Há também outra grandeza que será útil para a avaliação do método e que chamaremos de magnitude do vento corrigida pela persistência do erro do vento previsto V_{PERS} .

A seguir será explicado com mais detalhes o cálculo de regressão linear multivariada realizado pelo MOC e a escolha dos preditores.

2.1 Regressão Linear Multivariada:

Consideremos a seguinte equação

$$Y = \Lambda_0 + \sum_{l=1}^L \Lambda_l X_l$$

onde a variável dependente Y (preeditando) é função linear das L variáveis X_l (preditores).

O cálculo da regressão linear multivariada consiste em, determinar os coeficientes Λ conhecendo-se os valores de Y e X_l . A intenção deste cálculo é avaliar a dependência que há entre preeditando e preditores, pois os Λ são as constantes de proporcionalidade.

Neste trabalho, o MOC toma como preeditando a diferença entre a magnitude do vento prevista pelo Eta e a observada, isto é, o preeditando é o erro da previsão da intensidade do vento a 10 metros.

$$\Delta V_{i,j}(t) = V_{ETAi,j}(t) - V_{OBSi,j}(t) \quad \text{com } t=1,2,\dots,N$$

onde i,j são as coordenadas geográficas do local para onde estão sendo realizadas as observações e previsões e N é o tamanho da série temporal dos dados observados.

Portanto, $\Delta V_{i,j}(t)$ está associado às 70 variáveis consideradas na previsão do Eta, pois espera-se que a magnitude do vento varia de acordo com a variação da temperatura, gradiente de pressão, etc. Pensando assim, pode-se escrever:

$$\Delta V_{i,j}(t) = \Lambda_0 + \sum_{k=1}^K \Lambda_k X_k(t)$$

onde K é a quantidade total de variáveis necessárias para a equação.

Para sabermos o significado físico dos coeficientes, raciocina-se da seguinte forma: a cada variável $X_{i,jk}$ corresponde um coeficiente da equação. Pois se este coeficiente for igual a zero, isso significa que por mais que $X_{i,jk}$ varie, não influenciará o valor de $\Delta V_{i,j}$. Isso indica que os coeficientes são uma medida da relação linear de cada variável para o cálculo das previsões. Em outras palavras, são as constantes de proporcionalidade para cada variável.

2.2 Escolha dos preditores

Para selecionar as variáveis com maior relação linear, utiliza-se o cálculo da Correlação Linear Múltipla. Por exemplo, quando se estuda a relação entre duas

variáveis, e uma tende a aumentar enquanto a outra aumenta, diz-se que há correlação linear positiva entre elas: ao contrário, se o aumento de uma variável ocorre enquanto há diminuição da outra, a correlação é negativa.

Quantitativamente, este coeficiente está em um intervalo entre -1 e 1 .

Quanto mais o módulo do coeficiente de correlação estiver próximo de 1 , mais forte e a correlação existente entre duas variáveis.

A expressão para obter os coeficientes de correlação linear é:

$$r_l = \frac{\sum_{t=1}^N [(\Delta V_t - \Delta \bar{V})(X_{l,t} - \bar{X})]}{\sqrt{\sum_{t=1}^N (\Delta V_t - \Delta \bar{V})^2 \sum_{t=1}^N (X_{l,t} - \bar{X})^2}}, \text{ com } l = 1, 2, \dots, L$$

onde L é o número total de variáveis consideradas, ΔV_t é o preditando da equação de regressão linear, $X_{l,t}$ é a l -ésima variável a qual desejamos saber o índice de correlação linear, N é o tamanho da série temporal dos dados observados e $\Delta \bar{V}$ é a média dos valores da variável considerada.

Cada preditor terá um valor próprio para este coeficiente. Estipula-se um valor de correlação tal que só serão levadas em consideração preditores que tiverem coeficiente igual ou maior este valor limite. Há liberdade para se escolher este limite. Recomenda-se, porém, ter seis variáveis para se obter uma solução estável da equação de regressão.

Uma vez que os coeficientes da regressão estão calculados, o MOC calcula o preditando, que é o erro da previsão a ser aplicado na próxima previsão.

$$\Delta V_{i,j}(t) = \Lambda_0 + \sum_{k=1}^K \Lambda_k X_{i,j,k}(t) \quad \text{com } t=N+1, N+2, \dots, N+M$$

onde M é o tamanho da serie temporal das previsões.

2.3 Calibração das previsões do Eta

Finalmente, esse valor calculado para o preditando é utilizado para corrigir V_{ETA} , tem-se para V_{MOC} :

$$V_{MOC}(t) = V_{ETA}(t) - \Delta V_{i,j}(t) \quad \text{com } t=N+1, N+2, \dots, N+M$$

2.4 Correção por persistência do Erro

Um método simples de se fazer um refinamento sobre a magnitude de vento previsto é a manutenção do erro da previsão do dia anterior, isto é, pela persistência do erro da previsão.

$$V_{PERSi,j}(t) = V_{ETAi,j}(t) - \Delta V_{i,j}(t-1)$$

Esta estimativa será útil para comparação com a previsões do MOC e avaliarmos o grau de ganho do MOC sobre a metodologia mais simples.

2.5 Avaliação Dos Erros

A estimativa dos erros das previsões se baseia em três medidas estatísticas:

O erro médio:
$$BIAS = \frac{1}{M} \sum_{i=1}^M (V_i - V_{OBS})$$

O erro absoluto médio:
$$MAE = \frac{1}{M} \sum_{i=1}^M |V_i - V_{OBS}|$$

A raiz quadrada do erro quadrático médio:
$$RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^M (V_i - V_{OBS})^2}$$

Onde: M é o o número total de previsões produzidas, V_i é a i-ésima magnitude prevista pelo Eta, MOC ou persistência de erro; V_{OBS} é a magnitude observada.

Estas três medidas foram utilizadas para comparar os erros das previsões do Eta, do MOC e a correção por persistência de erro.

Para comparação das medidas de erros entre si, utiliza-se o SKILL SCORE (SS) . Neste trabalho, será utilizado para compararmos os erros das previsões do MOC em relação ao Eta.

$$SS = \left(\frac{E_{ETA} - E_{MOC}}{E_{ETA}} \right) \times 100$$

3 MATERIAIS E MÉTODOS

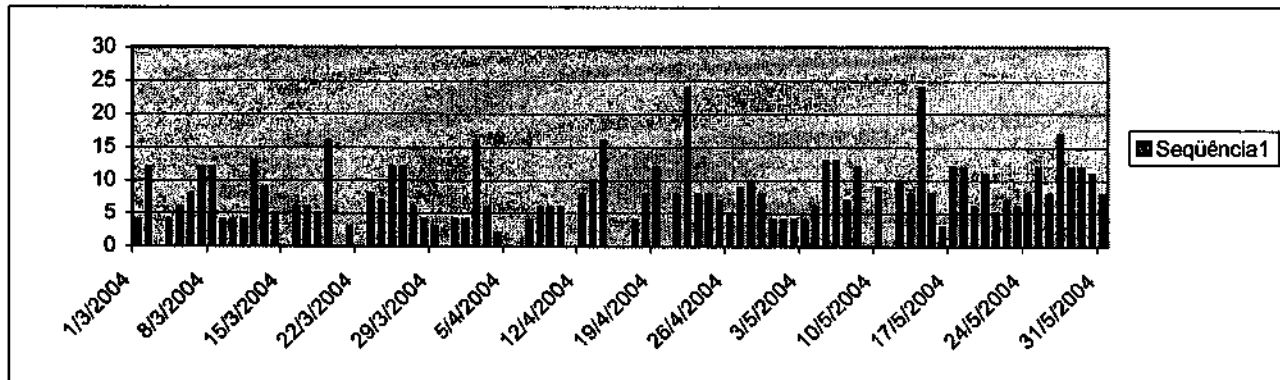
3.1 Dados observacionais

Os dados observacionais disponíveis que foram utilizados neste trabalho são da estação de Ilha Rasa-RJ. Estes dados referem-se ao ano de 2000 e foram coletados a cada seis horas (00:00z, 06:00z, 12:00z e 18:00z). Porém foram utilizados somente os dados para o horário de 12:00z para aplicar a calibração, durante os meses de março, abril e maio. Portanto, os dados são diários. A tabela para estes dados está abaixo.

1/mar	4.0	1/abr	4.0	1/mai	4.0
2/mar	12.0	2/abr	16.0	2/mai	4.0
3/mar	0.0	3/abr	6.0	3/mai	4.0
4/mar	4.0	4/abr	2.0	4/mai	6.0
5/mar	6.0	5/abr	0.0	5/mai	13.0
6/mar	8.0	6/abr	0.0	6/mai	13.0
7/mar	12.0	7/abr	4.0	7/mai	7.0
8/mar	12.0	8/abr	6.0	8/mai	12.0
9/mar	4.0	9/abr	6.0	9/mai	0.0
10/mar	4.0	10/abr	6.0	10/mai	9.0
11/mar	4.0	11/abr	0.0	11/mai	0.0
12/mar	13.0	12/abr	8.0	12/mai	10.0
13/mar	9.0	13/abr	10.0	13/mai	8.0
14/mar	5.0	14/abr	16.0	14/mai	24.0
15/mar	0.0	15/abr	0.0	15/mai	8.0
16/mar	6.0	16/abr	0.0	16/mai	3.0
17/mar	6.0	17/abr	4.0	17/mai	12.0
18/mar	5.0	18/abr	8.0	18/mai	12.0
19/mar	16.0	19/abr	12.0	19/mai	6.0
20/mar	0.0	20/abr	0.0	20/mai	11.0
21/mar	3.0	21/abr	8.0	21/mai	5.0
22/mar	0.0	22/abr	24.0	22/mai	7.0
23/mar	8.0	23/abr	8.0	23/mai	6.0
24/mar	7.0	24/abr	8.0	24/mai	8.0
25/mar	12.0	25/abr	7.0	25/mai	12.0
26/mar	12.0	26/abr	5.0	26/mai	8.0
27/mar	6.0	27/abr	9.0	27/mai	17.0
28/mar	4.0	28/abr	10.0	28/mai	12.0
29/mar	3.0	29/abr	8.0	29/mai	12.0
30/mar	0.0	30/abr	4.0	30/mai	11.0
31/mar	4.0			31/mai	8.0

Tabela A- tabela da magnitude do vento observado (mês/Valor)

Para melhor visualização dos dados, eles estão exibidos no gráfico a seguir:



3.2 Correção dos dados

Na tabela acima, podemos observar que há alguns horários em que estão registrados zeros para a magnitude do vento. Porém não pode haver dados indefinidos para a execução do MOC. Deve-se fazer a interpolação, utilizando dados vizinhos aos que estão anulados. Como os dados foram observados a cada seis horas e na tabela acima estão descritos somente os dados de 12:00z, foram utilizados os valores de 6:00z e 18:00z para a interpolação no horário das 12:00z. Consultando a tabela dos dados, anotamos os seguintes valores vizinhos aos horários de valores indefinidos (somente os valores relevantes aos cálculos da interpolação foram assinalados).

	00:00	06:00	12:00	18:00	00:00	06:00
03/mar		4.000	0.000	10.000		
15/mar	4.000	0.000	0.000	0.000	8.000	
20/mar		5.000	0.000	6.000		
22/mar	4.000	0.000	0.000	10.000		
30/mar	6.000	0.000	0.000	16.000		
05/abr		4.000	0.000	10.000		
06/abr		6.000	0.000	14.000		
11/abr		3.000	0.000	12.000		
15/abr		4.000	0.000	0.000	0.000	3.000
16/abr		3.000	0.000	4.000		
20/abr	18.000	0.000	0.000	5.000		
09/mai	18.000	68.000	0.000	18.000		
11/mai		12.000	0.000	9.000		

Quando há valores diferentes de zero em horários equidistantes do valor indefinido, faz-se uma média simples; se os valores não são equidistantes, faz-se uma média ponderada dando-se peso proporcional à distancia entre os dados existentes e o horário indefinido.

Outro procedimento a ser feito é observar se há valores muito diferentes dos demais, pois é necessário uma consistência temporal da série. O dado assinalado às 06:00z do dia 9 de maio será desconsiderado.

Os resultados das interpolações estão abaixo:

3 de março = 7

15 de março = 6

20 de março = 5.5

22 de março = 8.2 (atribuí peso 0.7 ao horário das 00:00z e peso 0.3 para as 18:00z)

30 de março = 13 (atribuí peso 0.7 ao horário das 00:00z e peso 0.3 para as 18:00z)

5 de abril = 7

6 de abril = 10

11 de abril = 7.5

15 de abril = 3.9 (atribui peso 0.9 para o horário das 06:00z e peso 0.1 para as 18:00z)

16 de abril = 3.5

20 de abril = 8.9 (peso de 0.7 para o horário das 00:00z e peso 0.3 para 18:00z)

9 de maio = 18 (já que o valor 68 foi desconsiderado)

11 de maio = 10.5

Estes dados interpolados foram colocados no horário dos dados indefinidos ou rejeitados da tabela A.

Quanto aos valores muito altos registrados na tabela dos dados das 12:00z, pode-se observar os valores em horários vizinhos, recorrendo-se novamente à tabela completa, como nos cálculos da interpolação (dados de 6:00z e 18:00z): se estes vizinhos também tiverem valores altos, é um bom indício de que estes valores estão certos. A tabela corrigida e pronta para ser utilizada está a seguir.

1/mar	4.0	1/abr	4.0	1/mai	4.0
2/mar	12.0	2/abr	16.0	2/mai	4.0
3/mar	7.0	3/abr	6.0	3/mai	4.0
4/mar	4.0	4/abr	2.0	4/mai	6.0
5/mar	6.0	5/abr	7.0	5/mai	13.0
6/mar	8.0	6/abr	10.0	6/mai	13.0
7/mar	12.0	7/abr	4.0	7/mai	7.0
8/mar	12.0	8/abr	6.0	8/mai	12.0
9/mar	4.0	9/abr	6.0	9/mai	18.0
10/mar	4.0	10/abr	6.0	10/mai	9.0
11/mar	4.0	11/abr	7.5	11/mai	10.5
12/mar	13.0	12/abr	8.0	12/mai	10.0
13/mar	9.0	13/abr	10.0	13/mai	8.0
14/mar	5.0	14/abr	16.0	14/mai	24.0
15/mar	6.0	15/abr	3.9	15/mai	8.0
16/mar	6.0	16/abr	3.5	16/mai	3.0
17/mar	6.0	17/abr	4.0	17/mai	12.0
18/mar	5.0	18/abr	8.0	18/mai	12.0
19/mar	16.0	19/abr	12.0	19/mai	6.0
20/mar	5.5	20/abr	8.9	20/mai	11.0
21/mar	3.0	21/abr	8.0	21/mai	5.0
22/mar	8.2	22/abr	24.0	22/mai	7.0
23/mar	8.0	23/abr	8.0	23/mai	6.0
24/mar	7.0	24/abr	8.0	24/mai	8.0
25/mar	12.0	25/abr	7.0	25/mai	12.0
26/mar	12.0	26/abr	5.0	26/mai	8.0
27/mar	6.0	27/abr	9.0	27/mai	17.0
28/mar	4.0	28/abr	10.0	28/mai	12.0
29/mar	3.0	29/abr	8.0	29/mai	12.0
30/mar	13.0	30/abr	4.0	30/mai	11.0
31/mar	4.0			31/mai	8.0

Tabela B- série dos dados de magnitude do vento observado corrigido

3.3 Utilização dos dados

O MOC utilizará os dados do mês de março da tabela B para, juntamente com as magnitudes do vento previstas pelo Eta, selecionar os preditandos para a equação de regressão linear multivariada. Os dados dos meses de abril e maio serão utilizados para comparação e cálculo dos erros e correlações após a calibrado pelo MOC.

Os dados da tabela B também serão utilizados para o cálculo da previsão por persistência de erro.

3.4 Implementação computacional

Todos os cálculos realizados pelo MOC descritos até aqui são muito complexos, e a única maneira de realizá-los é através de computação. Para isso, foi utilizada a linguagem FORTRAN-90, cujo compilador está instalado no sistema operacional UNIX TRU 64, que é utilizado nas estações de trabalho do CPTEC.

Para a visualização dos resultados, utilizou-se *software* GrADS, que é um visualizador em duas dimensões, que possui recursos de gráficos, tabelas e animações.

Neste trabalho, como os dados observacionais já estavam disponíveis, não foi necessário utilizar programas para extração do banco de dados: apenas foi necessário extrair os dados previstos pelo Eta para onde pudessem ser utilizados pelo MOC. O programa que realizou estas extrações está implementado na linguagem de programação do GrADS.

Após essa etapa de extração, tendo os dados observacionais e previstos em arquivos, executou-se o programa do MOC propriamente dito (escrito em FORTRAN-90), que calculou os coeficientes das equações de regressão linear, gerou arquivos binários para a posterior utilização pelo programa que gera os gráficos exibidos neste trabalho e produziu relatórios em formato de arquivos tipo texto onde estão descritas todas as variáveis utilizadas para os cálculos, bem como a frequência com que foram selecionadas.

Por último, executou-se o programa em GrADS para produzir a visualização dos resultados. Estes programas já foram implementados anteriormente a este trabalho pelo bolsista PIBIC Carlos Renato Sousa, e para a utilização com dados de vento, eles tiveram apenas de ser modificados.

4 Resultados e análise

Para a realização das experiências foram utilizadas séries de treinamento de 21 e 28 dias. Serão apresentadas as séries com o refinamento e os índices de acerto e desempenho do esquema MOC. Os gráficos fornecem a visualização das séries, as barras de erros apresentam o desempenho das três estimativas utilizadas.

Adiante serão exibidas também as listas das variáveis mais utilizadas nas equações de regressão linear.

4.1 Período treinamento de 21 dias

O algoritmo MOC foi testado para a estação de Ilha Rasa –RJ, durante o período de 1º abril a 31 de maio, sendo o mês de março utilizado como período de treinamento.

Primeiramente, o teste foi feito para um período de treinamento de 21 dias.

4.1.1 Análise dos gráficos e índice de desempenho

Pela análise dos gráficos, percebe-se que a correlação entre a magnitude avaliada pela persistência de erro e a magnitude observada é muito baixa, além de negativa; a correlação entre a magnitude prevista pelo Eta e magnitude observada também é baixa e negativa; a correlação entre a magnitude calibrada pelo MOC e observada só é negativa no limiar de correlação $R_C = 0,8$, porém esta correlação é bem pequena para os três limiares.

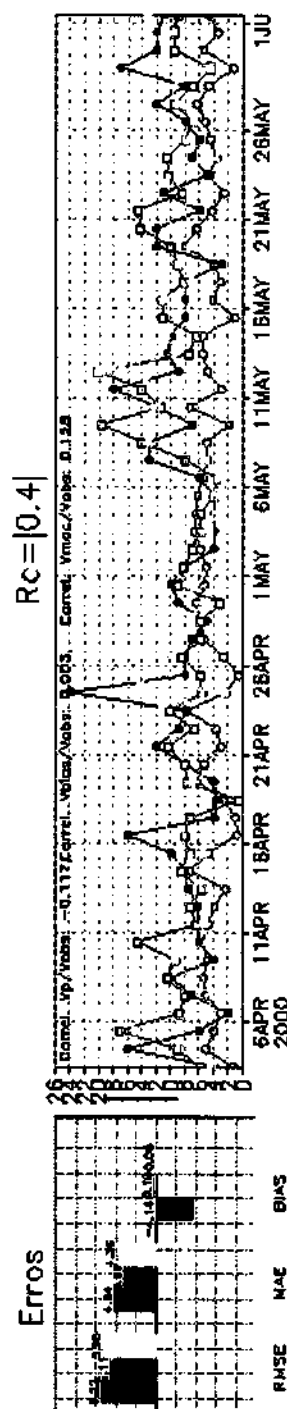
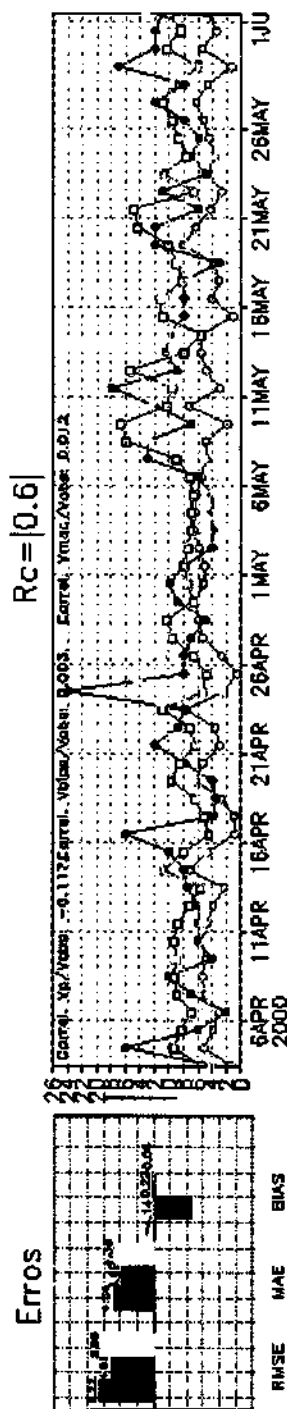
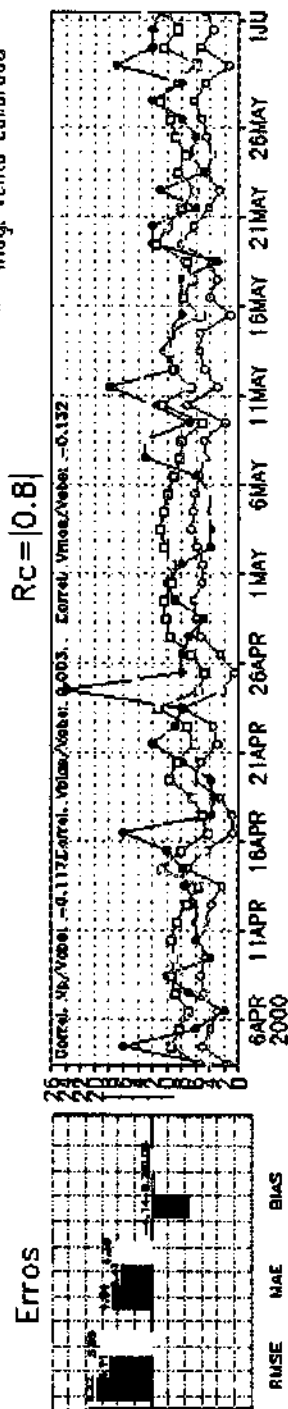
As medidas de erro RMSE, MAE e BIAS mostram que em geral o MOC teve melhor desempenho que o Eta. O MOC reduziu o RMSE em relação a previsão do Eta para os três limiares de correlação; o MOC também reduziu o MAE em relação ao Eta nos três limiares. Já para a medida BIAS, o V_{PERS} , apresentou os menores erros. Ainda assim, o MOC apresentou erros bem menores que o ETA para o erro BIAS nos três limiares de correlação.

Porém, mesmo diante da imprecisão das observações, percebe-se que o MOC teve melhor desempenho que o modelo ETA para os três limiares e em geral saiu-se melhor que a magnitude do vento avaliada por persistência de erro, baseado na redução do RMSE.

Gráfico para o período de treinamento de 21 dias. São exibidas as séries temporais de intensidade de vento e os gráficos dos erros.

Abr-Mai/2000 : Ilha Rasa-RJ - Delta V (nos) : serie 21 dias

- mag. vento observada
- mag. vento prevista
- mag. vento calibrada



Para uma melhor comparação entre os erros do ETA e do MOC, utilizou-se o índice de desempenho para cada um dos limiares e cada uma das medidas de erro:

Rc	RMSE	MAE	BIAS
0,8	24,30%	26,50%	95,20%
0,6	22,00%	21,30%	94,70%
0,4	17,80%	22,60%	95,40%

Tabela C-coeficientes de correlação entre o MOC eo Eta

Pela tabela acima, vê-se que em geral o MOC apresentou desempenho melhor que o Eta. Percebe-se isso principalmente pelo erro BIAS. Porcentagens negativas indicariam que o erro do MOC foi maior que o do Eta.

Portanto, o MOC teve melhor desempenho para o limiar de correlação de 0,8, indicando ter refinado as previsões em quase 30%.

4.1.2 Preditores melhor correlacionados

A seguir vê-se a tabela das dez variáveis mais utilizadas como preditores das equações de regressão linear utilizadas pelo MOC.

1	Magnitude do vento a 10 metros do solo	24/72
2	Umidade especifica da superficie	24/72
3	Comprimento da rugosidade da superficie	21/72
4	Temp. 925 – Temp. 500	20/72
5	Componente meridional do vento	18/72
6	Máximo do componente zonal do vento	17/72
7	U700 – U300	17/72
8	Radiação de onda longa emergente da superficie	13/72
9	Temperatura da superficie – temperatura na zona das raízes	12/72
10	Disponibilidade de umidade do solo	11/72

Tabela D - variáveis com maior índice de correlação linear
(classificação/variável/fração do número de vezes em que é utilizada)

4.2 PERÍODO DE COLETA DE DADOS DE 28 DIAS

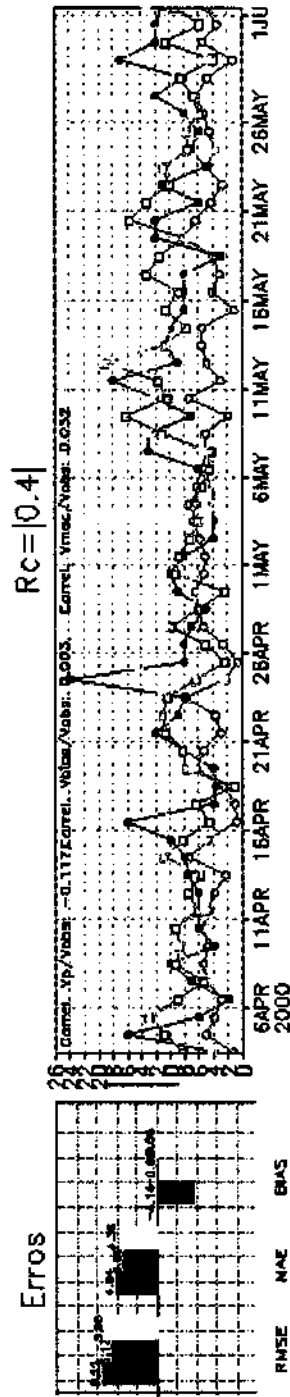
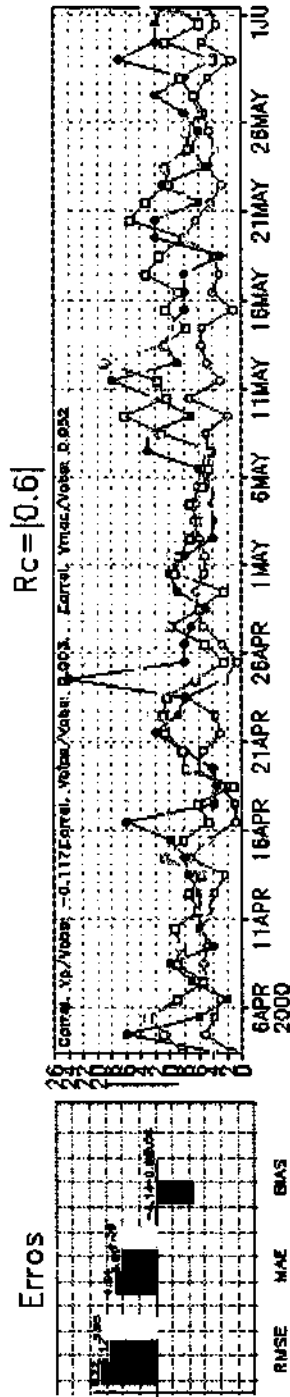
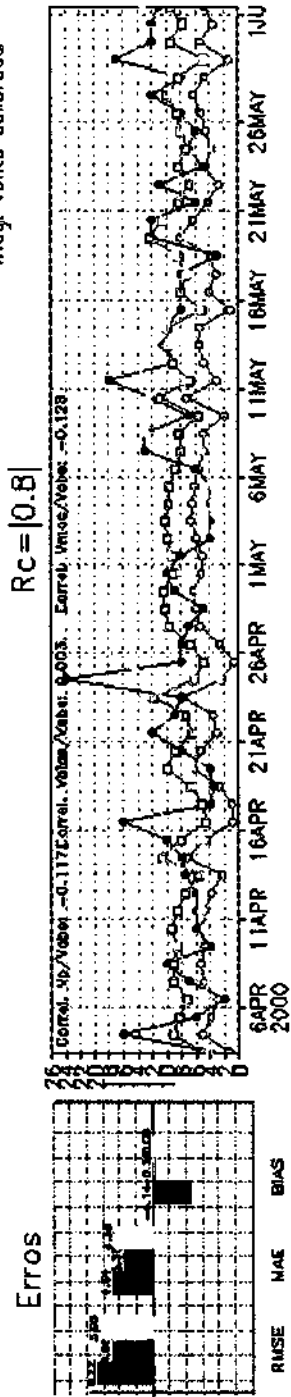
A experiência foi repetida para um período de treinamento de 28 dias, utilizando-se os mesmos limiares de correlação que anteriormente e a mesma série de dados observacionais. A intenção é realizar uma comparação entre as duas experiências.

4.2.1 Análise dos gráficos e índice de desempenho

Gráfico para a série de treinamento de 28 dias. São exibidas as séries temporais de intensidade do vento e os gráficos de barras.

Abr-Mai/2000 : Ilha Rasa-RJ - Delta V (nos) : serie 28 dias

- mag. vento observado
- mag. vento prevista
- mag. vento calibrada



A correlação entre as magnitudes prevista pelo MOC a observada teve ligeira redução para os limiares de correlação de 0.4 e 0.6, e teve um ligeiro aumento para o limiar de 0.8, em relação ao período de treinamento de 21 dias.

Pelos gráficos de barras, vemos que todos os erros do MOC diminuíram, com exceção do erro BIAS no limiar de 0.8, que teve pequeno aumento em relação às séries com limiar de 0,4 e 0,6.

O índice de desempenho do MOC em relação ao Eta para todas as medidas de erro está na tabela abaixo:

Rc	RMSE	MAE	BIAS
0,8	24,80%	27,80%	93,20%
0,6	17,50%	18,10%	97,80%
0,4	17,50%	18,10%	97,80%

Tabela E - coeficientes de correlação entre o Eta e o MOC

Para o limiar de correlação de 0,8 o MOC apresentou o melhor desempenho, tendo para a medida de erro RMSE um índice de aproximadamente 25%. Utilizando os limites de correlação de 0,4 e 0,6, o MOC teve índices idênticos em relação ao Eta.

4.2.2 Preditores melhor correlacionados

As dez variáveis mais utilizadas para as equações de regressão linear estão apresentadas a seguir.

1	Magnitude do vento	31/65
2	Componente meridional do vento em 850hPa	26/65
3	Comprimento da rugosidade da superfície	26/65
4	Umidade específica da superfície	23/65
5	disponibilidade de umidade do solo	21/65
6	deltaU300.700	17/65
7	deltaT500.925	15/65
8	Componente meridional do vento em 850hPa	14/65
9	máxima componente zonal do vento	14/65
10	deltaT700.850	07/65

Tabela F - variáveis com maior índice de correlação linear (classificação/variável/fração do numero de vezes em que é utilizada)

As variáveis que figuram entre as mais utilizadas são as mesmas para os dois outros limiares de correlação. Para a melhoria dos resultados, é desejável que se tente adicionar algumas variáveis, como por exemplo a temperatura potencial.

Finalmente, percebe-se pelas experiências anteriores que o MOC atuou de forma eficaz ao reduzir os erros do ETA, porém estes resultados devem ser melhorados.

5 CONCLUSÃO E TRABALHOS FUTUROS

5.1 Conclusão

Foi utilizada uma série de magnitude de vento observada na estação de Ilha Rasa- RJ, pertencente à marinha do Brasil. As calibrações das previsões de magnitude do vento para esta estação foram feitas utilizando-se o sistema MOC (Model Output Calibration). Foram utilizadas séries de treinamento de 21 e 28 dias.

Através das experiências feitas, percebemos que o algoritmo MOC desempenhou bem sua tarefa, pois teve medidas de erros inferiores aos erros do Eta: o MOC atuou de forma eficaz na redução dos erros RMSE e MAE; o erro BIAS foi diminuído de forma notável, pois o Eta o tinha alto e negativo, e o reduziu em magnitude e em alguns casos tornou-o positivo. Porém as correlações em geral foram baixas.

Apesar dos resultados positivos da aplicação do MOC, seria interessante além de experimentar outras séries, adequar a metodologia para melhorar ainda mais o desempenho do sistema MOC. Nesta experiência utilizaram-se dados diários das 12:00z. Em experiências com previsões de temperatura (Souza, 2003) foram utilizados dados diários de observação e geraram-se resultados em geral satisfatórios. Este método precisa ser ainda melhor adaptado para os próximos passos.

As experiências realizadas neste trabalho indicam que o MOC é uma boa ferramenta para o refinamento das previsões da intensidade do vento, e novas metodologias devem otimizar esse sistema.

Entretanto, apesar de se perceber que o MOC é eficaz, os resultados ainda não são satisfatórios ou conclusivos. O trabalho está em fase inicial e ainda há muito a ser feito para o aperfeiçoamento dos resultados.

5.2 Futuros trabalhos

Serão utilizadas séries de dados tomados em intervalos mais curtos, como por exemplo, tomados de hora em hora.

Serão introduzidas novas variáveis previstas pelo Eta como preditores para serem incluídos na equação de regressão linear multivariada. Exemplo de variáveis seriam a temperatura potencial equivalente, cisalhamento do vento na baixa troposfera, etc.

Entre os locais mais prováveis para a coleta de dados observados estão as estações em Florianópolis e na usina nuclear de Angra dos Reis cujas séries são horárias. Com isso, espera-se um desempenho melhor do algoritmo MOC.