

## REFINAMENTO ESTATÍSTICO DAS PREVISÕES HORÁRIAS DE TEMPERATURA A 2 M DO MODELO ETA EM ESTAÇÕES DO NORDESTE DO BRASIL

SIN CHAN CHOU<sup>1</sup>, CARLOS RENATO DE SOUZA<sup>1</sup>, JORGE L. GOMES<sup>1</sup>, EDGARD F.D. EVANGELISTA<sup>1</sup>, CHRISTIANE OSÓRIO<sup>2</sup> e MÁRCIO CATALDI<sup>2</sup>

<sup>1</sup>Instituto Nacional de Pesquisas Espaciais (INPE),  
Centro de Previsão de Tempo e Estudos Climáticos (CPTEC)  
Cachoeira Paulista, SP, CEP 12630-000

chou@cptec.inpe.br, carlosr@cptec.inpe.br, gomes@cptec.inpe.br, edgardf@hotmail.com

<sup>2</sup>Operador Nacional do Sistema Elétrico (ONS)  
Rio de Janeiro, RJ, CEP 20091-000  
cosorio@ons.org.br, cataldi@ons.org.br

Recebido Janeiro 2006 - Aceito Abril 2006

### RESUMO

Previsões de temperatura do Modelo Eta de 40 km de resolução são refinadas para produzir previsões em pontos localizados em sete cidades da região Nordeste: Salvador, Maceió, Recife, Natal, Fortaleza, São Luís e Teresina. O esquema de refinamento empregado é o Model Output Calibration (MOC) que se baseia em regressão linear do erro da previsão com as variáveis previstas pelo modelo. O esquema apresenta a vantagem de empregar período curto de treinamento, neste caso foi usado o período de somente 28 dias. A seleção dos preditores se baseou em correlação maior que 0,6. Os testes foram realizados com dados horários de temperatura do mês de setembro de 2003. Os erros sistemáticos das previsões foram substancialmente reduzidos para valores menores que 0,5 °C em média. Os erros quadráticos que eram da ordem de 2 °C passaram para valores em torno de 1 °C em média. As correlações entre as séries refinadas e as séries observadas foram geralmente mantidas altas e semelhantes às correlações entre as séries previstas pelo Modelo Eta e as séries observadas, indicando que a aplicação do esquema não prejudicou as tendências horárias da temperatura nas estações. As temperaturas de cidades onde as amplitudes dos ciclos diurnos não foram bem previstas pelo Eta, foram refinadas pelo esquema.

**Palavras-chave:** Refinamento estatístico, Modelo Eta, previsões de temperatura.

### ABSTRACT: STATISTICAL DOWNSCALING OF THE ETA MODEL 2-M TEMPERATURE HOURLY FORECASTS OVER NORTHEAST BRAZIL.

Statistical downscaling of 2 m air temperature forecasts from the Eta Model at 40 km resolution were carried out for 7 sites located in Northeast Brazil. The downscaling method was the Model Output Calibration (MOC), which is based on the linear regression between the forecast error and model output variables. One feature of the scheme is the use of short training period, in this work only 28 days was the chosen period. The model output variables with correlation higher than 0.6 were selected as predictors of the regression equation. The scheme was tested for one-month hourly forecasts in September 2003. Model systematic errors were strongly reduced to values below |0.5| °C on average. The root mean square errors, which initially were of the order of 2 °C, were reduced to about 1 °C. The linear correlation between the observations and Eta forecasts were generally kept the same after applying the downscaling scheme. Temperature forecasts with greatest errors in the amplitude of the diurnal cycle were corrected accordingly.

**Keywords:** Statistical downscaling, Eta Model, temperature forecasts.

## 1. INTRODUÇÃO

As previsões do Modelo Eta (Mesinger et al., 1988; Black, 1994) geradas pelo Centro de Previsão de Tempo Estudos Climáticos (CPTEC) do Instituto Nacional de Pesquisas Espaciais (INPE) têm sido consideradas ferramentas úteis para a elaboração das previsões de tempo pelos meteorologistas no Brasil. Estas previsões cobrem grande parte da América do Sul e são fornecidas duas vezes ao dia, às 0000 UTC e às 1200 UTC, para o prazo de antecedência de até 120 horas.

A resolução horizontal do modelo é de 40 km e a resolução vertical de 38 camadas. Isto significa que as previsões representam valores homogêneos em caixas formada pela grade horizontal de aproximadamente 40 km x 40 km de cada camada. Entretanto, em regiões de orografia e vegetação complexas esta representatividade pelo modelo se torna deficiente devido às diferenças nas características das superfícies que muitas vezes uma área de 40 km x 40 km pode apresentar. Por exemplo, duas cidades, que estão a menos de 40 km de distância uma da outra e que, entretanto, apresentam diferenças de altitude e vegetação, podem estar localizadas na mesma caixa de grade do modelo. O modelo numérico gera as previsões com base em dados médios, como altitude média e vegetação predominante da caixa de grade. Assim as previsões geradas pelo modelo numérico não se aplicam diretamente a uma cidade ou um ponto específico.

Para produzir previsões para um ponto específico é necessário aplicar uma correção estatística nas previsões numéricas. Estas correções geralmente se baseiam em Model Output Statistics – MOS (Glahn e Lowry, 1972) e redes neurais artificiais – RNA (Haykin, 1994). Estas técnicas geralmente requerem uma série longa de dados históricos e um modelo com suas características mantidas constantes, isto é, congeladas para permitir o treinamento de padrões. Sugahara (2000) empregou MOS para previsão das temperaturas mínimas na cidade de Bauru e obteve erros absolutos menores que 1,5 °C. A equação de regressão se baseou em variáveis do modelo global do CPTEC e observações. Ramirez (2003) utilizou RNA para correção das previsões espacial e pontual de precipitação do Modelo Eta para os períodos de verão e inverno. A técnica resultou em aumento do skill das previsões de precipitação sobre estados da região Sudeste, sendo maior skill no período de inverno. O período de treinamento utilizado foi maior que 2 anos. Guarnieri (2006) utilizou RNA para correção das previsões de radiação solar do Modelo Eta para as cidades de Florianópolis e São Martinho da Serra em Santa Catarina. A técnica produziu um ganho em aproximadamente 30% na redução dos erros, com aumento na correlação linear com as séries observadas. Neste trabalho foi empregado período de treinamento de pelo menos 1 ano. RNA tem sido aplicada para produtos específicos de precipitação, por exemplo o PERSIANN (Precipitation Estimation from

Remotely Sensed Information using Artificial Neural Network de Hsu et al. (1997).

Neste trabalho pretende-se aplicar uma correção simples sobre as previsões de temperatura do ar a 2 m da superfície do modelo Eta em 7 localidades do Nordeste do Brasil, onde o modelo apresenta discrepâncias na representatividade. A correção se baseia em regressão linear multivariada desenvolvida por Mao et al. (1999) e denominada por “Model Output Calibration” (MOC). O esquema é simples e adequado para variáveis bem comportadas como a temperatura do ar. Em geral, os esquemas mais complexos estimam diretamente a variável a ser corrigida ou prevista, o MOC, por sua vez, estima o erro da previsão da variável. Uma outra característica do MOC é o tamanho curto da série de treinamento, isso permite que o esquema facilmente se ajuste às mudanças e melhorias nas previsões do modelo. Previsões resultantes de modificações introduzidas ao modelo como, por exemplo, aumento de resolução ou mudança de parâmetros da física do modelo, podem rapidamente ser submetidas ao refinamento com a construção de um período de 28 dias de previsão com a nova configuração. Mudanças nos modelos podem ser introduzidas em períodos irregulares. Quando são identificados erros na integração, as correções destes erros são introduzidas imediatamente, entretanto, mudanças de parâmetros da física ou outras melhorias dependem dos resultados das avaliações e, finalmente, aumento de resolução geralmente depende de disponibilidade computacional do centro.

Portanto, este trabalho tem por objetivo aplicar o esquema de correção estatística, “Model Output Calibration”, nas previsões de temperatura horária do modelo Eta em estações localizadas na região Nordeste do Brasil. O trabalho está estruturado da seguinte forma: o esquema é apresentado na seção 2, os dados e a metodologia empregados são apresentados na Seção 3. As previsões corrigidas pelo esquema para cada cidade são mostradas e avaliadas na Seção 4. Finalmente, algumas discussões e conclusões são apresentadas na Seção 5.

## 2. O ESQUEMA

O esquema MOC consiste em formular equações de Regressão Linear Multivariada com base nos dados das séries de poucas semanas das previsões e das observações, estimando e ajustando assim os erros da previsão da variável a ser corrigida, neste caso a temperatura do ar a 2 metros da superfície.

Considerando que  $T_{2m}(t)$  é a série temporal de previsões de temperatura do ar a 2 metros de uma caixa de grade do modelo Eta, considerando também que  $T_{obs}(t)$  é a série temporal da temperatura observada de uma estação localizada dentro da caixa de grade, o erro da previsão de temperatura é definido como:

$$\Delta T(t) = T_{2m}(t) - T_{obs}(t) \quad (1)$$

Supõe-se que o erro da previsão de temperatura,  $\Delta T(t)$ , está intrinsicamente associado a diversas variáveis e parâmetros de saída do modelo Eta,  $X_i(t)$ , como, por exemplo, variáveis próximas ao solo (umidade do solo, temperatura da superfície, etc...). Constrói-se assim um modelo de Regressão Linear Multivariada na forma:

$$\Delta T(t) = \Lambda_0 + \sum_{l=1}^L \Lambda_l X_l(t) \quad (2)$$

$t = 1, 2, \dots, N,$   
 $l = 1, 2, \dots, L$

onde os coeficientes  $\Lambda_0$  e  $\Lambda_l$  são os coeficientes da equação de regressão. O erro da previsão,  $\Delta T$ , é o preditando das equações, a variável dependente, enquanto que as outras variáveis de saída do modelo,  $X_i(t)$ , são os candidatos a preditores. A calibração será aplicada no tempo seguinte, ou seja, em  $N+1$ , sendo  $N$  o tamanho da série tomada para cálculo dos coeficientes de regressão, isto é, o período de treinamento. As previsões refinadas são representadas por:

$$T_{\text{moc}}(t) = T_{2m}(t) - \Delta T(t), \quad (3)$$

$t = N+1$

Neste esquema os preditores não são definidos a priori, mas selecionados a partir da correlação entre o erro da previsão e variáveis previstas pelo modelo,  $R$ . Variáveis que apresentam correlação maior que um valor crítico,  $R_c$ , são selecionadas como preditoras. Os coeficientes da equação de regressão e as correlações são recalculados a cada nova previsão do modelo Eta.

### 3. DADOS E METODOLOGIA

#### 3.1. Dados

Foram utilizados dois conjuntos de dados do período de agosto a setembro de 2003:

1. Previsões: Valores de temperaturas horárias do Modelo Eta do CPTEC/INPE na resolução de 40 km, previstas da integração iniciada às 0000 UTC. Além da temperatura outras 72 variáveis são utilizadas como candidatas a preditores, sendo 46

delas fornecidas diretamente das previsões do Modelo Eta e 26 derivadas das 46 originais. As previsões são extraídas das caixas de grade que contém as cidades: São Luís, Teresina, Fortaleza, Natal, Recife, Maceió, e Salvador.

2. Observações: A correção horária utiliza observações de aeroportos codificadas em mensagem METAR, que são distribuídas de hora em hora. A frequência destes dados é relativamente satisfatória, por outro lado, a precisão é baixa, pois se apresentam somente como números inteiros. As estações utilizadas são identificadas pelo nome da cidade, código, número, latitude e longitude na Tabela 1. A maioria das estações, com exceção de Teresina, se localiza em área litorânea.

#### 3.2. Metodologia

##### a. Controle de qualidade das observações

A qualidade dos dados das estações meteorológicas brasileiras sofre variações no tempo e no espaço. Um esquema simples de controle de qualidade e preenchimento dos dados faltosos é aplicado para permitir a implementação do esquema. Observações que apresentam valores discrepantes são rejeitadas. Entende-se por valor discrepante quando a variação de temperatura entre dois intervalos de 1 hora é maior que 5 °C e há mudança de sinal da variação no intervalo, o dado no centro do intervalo é rejeitado. Dados ausentes ou rejeitados são preenchidos com a média entre as medidas do horário seguinte e do anterior da observação ausente. Se a ausência de dado ocorrer em intervalos maior ou igual que 2 medições consecutivas, o esquema preenche com o valor corrigido pelo MOC da previsão do dia anterior.

##### b. Construção do período de treinamento

O período de treinamento adotado foi de 28 dias. Períodos menores foram testados, como por exemplo de 21 dias, entretanto, a série mais longa apresentou resultados mais estáveis. Esquemas de refinamento como MOS ou redes neurais

**Tabela 1** – Identificação e localização das estações do Nordeste utilizadas.

|   | Cidade    | Código da estação | Número da estação | Latitude (graus) | Longitude (graus) |
|---|-----------|-------------------|-------------------|------------------|-------------------|
| 1 | Salvador  | SBSV              | 83248             | -12.90           | -38.33            |
| 2 | Maceió    | SBMO              | 82993             | -9.52            | -35.78            |
| 3 | Recife    | SBRF              | 82899             | -8.07            | -34.85            |
| 4 | Natal     | SBNT              | 82599             | -5.92            | -35.25            |
| 5 | Fortaleza | SBFZ              | 82398             | -3.78            | -38.53            |
| 6 | São Luís  | SBSL              | 82281             | -2.60            | -44.23            |
| 7 | Teresina  | SBTE              | 82579             | -5.05            | -42.82            |

utilizam geralmente longas séries, de pelo menos 2 anos, para construção do período de treinamento. Estes outros esquemas costumam apresentar resultados superiores ao MOC, entretanto, pela simplicidade e pela rápida adaptação a mudanças nos modelos de previsão de tempo, a relação custo-benefício torna-se favorável para o MOC.

#### c. Seleção dos preditores a partir das variáveis previstas pelo Modelo Eta

A seleção dos preditores se baseou na correlação de cada variável do modelo com o erro da previsão de temperatura. A correlação linear que apresentou em média maior redução dos erros da previsão foi geralmente 0.6. A cidade de São Luís utilizou  $R_c=0.8$ . O número de preditores foi limitado em 5; um número maior dificultou a construção da equação de regressão, em algumas vezes a solução não convergiu. Este número pode aumentar para 6 preditores quando as observações apresentarem maior precisão e qualidade. As variáveis escolhidas mais frequentemente sugerem quais são os processos do modelo diretamente relacionados com os erros da previsão de temperatura.

#### d. Correção das previsões

Após o cálculo dos coeficientes de correlação e dos coeficientes da equação de regressão multivariada, são construídas 24 equações para cada horário do dia, para cada cidade. O refinamento é aplicado para o ciclo de 24 a 48 horas de previsão.

## 4. RESULTADOS

### 4.1. Avaliação da redução dos erros

Para avaliar as previsões de temperatura geradas pelo MOC, foram usadas três medidas estatísticas básicas de erro: o erro médio (BIAS), o erro absoluto médio (MAE) e a raiz quadrada do erro quadrático médio (RMSE) definidos por:

$$\text{BIAS} = \frac{1}{M} \sum_{i=1}^M (T_i - T_{\text{obs}_i}), \quad (5)$$

$$\text{MAE} = \frac{1}{M} \sum_{i=1}^M |T_i - T_{\text{obs}_i}|, \quad (6)$$

$$\text{RMSE} = \sqrt{\frac{1}{M} \sum_{i=1}^M (T_i - T_{\text{obs}_i})^2}, \quad (7)$$

onde  $M$  é o número total de previsões corrigidas,  $T_i$  é a  $i$ -ésima temperatura prevista com o ajuste do MOC ou sem o ajuste, isto é obtidos diretamente do Modelo Eta, e  $T_{\text{obs}_i}$  é a  $i$ -ésima temperatura observada.

Espera-se que a correção aplicada mantenha a tendência da série observada, desta forma, a correlação entre a série observada e aquela prevista pelo Modelo Eta é comparada

com a correlação entre a série observada e a série corrigida pelo MOC. Pretende-se verificar se houve alguma mudança de fase, isto é, se a tendência da curva de temperatura se mantém após a correção.

Também foi calculado o índice SS (“Skill Score”), ou ganho, para medir em porcentagem o ganho das previsões corrigidas com relação às não-corrigidas. O índice é dado por:

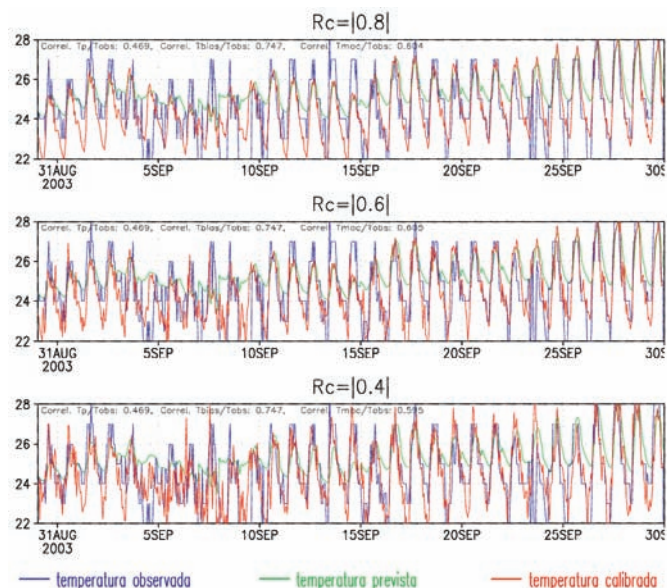
$$\text{SS} = \left( \frac{E_{\text{Eta}} - E_{\text{MOC}}}{E_{\text{Eta}}} \right) \times 100\% \quad (8)$$

onde  $E_{\text{Eta}}$  ou  $E_{\text{MOC}}$  é uma das três medidas estatísticas de erro: BIAS, MAE ou RMSE.

### 4.2. Séries temporais

Durante o período de setembro de 2003, os eventos de chuva ocorreram mais frequentemente no litoral leste do Nordeste do Brasil. Estes eventos de chuva estavam associados às perturbações geradas pela passagem de frentes frias em dissipação sobre o Oceano Atlântico, próximo ao litoral do Nordeste. Estas chuvas ocorreram nos primeiros 10 dias do mês. No restante deste mês, algumas chuvas fracas ocorreram no litoral, resultantes de perturbações de leste no escoamento, neste caso, as chuvas foram geradas por nuvens de pequena profundidade e de valores relativamente menores.

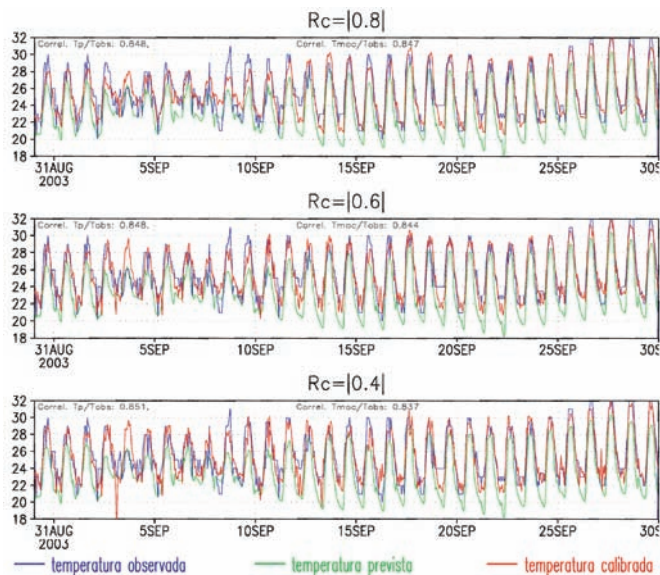
A Figura 1 mostra 3 séries horárias de temperatura do ar da estação do aeroporto de Salvador. As 3 séries se referem aos resultados do refinamento utilizando diferentes limiares de correlação linear,  $R_c$ , para seleção das variáveis do modelo Eta empregadas como preditores na equação de regressão. O período de treinamento adotado foi de 28 dias, enquanto que a correção foi aplicada no período de 30 dias. As curvas azuis se referem às medidas observadas no aeroporto, as curvas verdes se referem às temperaturas previstas pelo Modelo Eta para o local, no ciclo de 24 a 48 horas de previsão, enquanto que as curvas vermelhas se referem às temperaturas refinadas pelo MOC. Nesta cidade as previsões de temperatura do Eta apresentam amplitude térmica pequena demais. As temperaturas variaram somente entre 24 °C e 27 °C neste período. Apesar disso, o erro médio é menor que 1 °C. O esquema corrige as temperaturas máximas e mínimas reduzindo RMSE em 22% e o MAE em 21% (Tabela 3). O ganho no BIAS passou de 100% porque apesar de ter praticamente removido o erro sistemático, o erro restante mudou de sinal. Nesta série, destaca-se o aumento da correlação linear entre as séries observadas e corrigidas pelo esquema MOC. As séries produzidas por  $R_c=0.6$  ou  $R_c=0.8$  geraram resultados satisfatórios, sendo 0.6 ligeiramente superior. O ponto de grade onde se localiza a cidade de Salvador sofre grande influência do oceano, a grande proporção de superfície líquida neste ponto de grade acarreta em menor amplitude térmica.



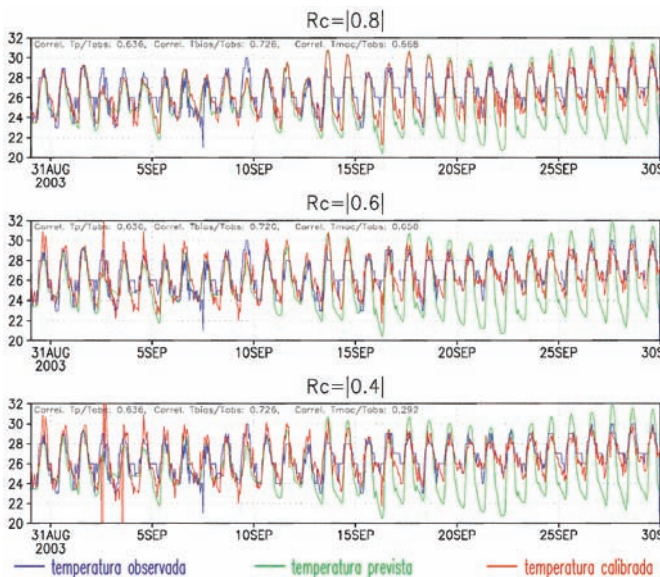
**Figura 1** – Temperatura do ar a 2m (°C) de Salvador em setembro de 2003, observada (azul), prevista pelo Eta (verde), corrigida pelo MOC (vermelha), para  $R_c=0.8; 0.6; 0.4$ .

A Figura 2 mostra as séries de Maceió para os 3 limiares de  $R_c$ . Nota-se que as previsões do Eta frequentemente subestimam as temperaturas máximas e as mínimas resultando em um erro sistemático (BIAS) de aproximadamente  $-2.1$  °C nesta cidade (Tabela 2). O esquema reduziu o erro sistemático para aproximadamente  $-0.1$  °C, o que corresponde a um ganho de aproximadamente 95%. A correção não alterou a fase da série de temperatura corrigida visto que a correlação entre a série observada e a prevista é semelhante à correlação entre a série observada e a série corrigida pelo MOC. Nota-se que a série corrigida acompanha a série calibrada, principalmente nos horários de máximas e mínimas. A redução no RMSE e no MAE foi da ordem de 47% e 52%, respectivamente. As chuvas nos primeiros 10 dias do mês se revelam na menor amplitude térmica, prevista pelo modelo Eta e refinada pelo MOC.

A Figura 3 mostra as séries de Recife construídas com os 3  $R_c$ 's. Nesta cidade predominam os erros negativos de previsão do Modelo Eta. Os RMSE do Eta são da ordem de 2 °C enquanto que após o refinamento os erros se reduzem a 1,2 °C (Tabela 2). O esquema MOC reduz os RMSE em cerca de 45% e o MAE em 47%. O BIAS restante é considerado pequeno, sendo 82% o ganho do esquema na redução deste erro (Tabela 3). As correlações entre as séries corrigidas e as séries observadas aumentaram em relação à série prevista pelo Eta. O emprego de  $R_c=0.8$  produz valores corrigidos mais estáveis para esta estação.

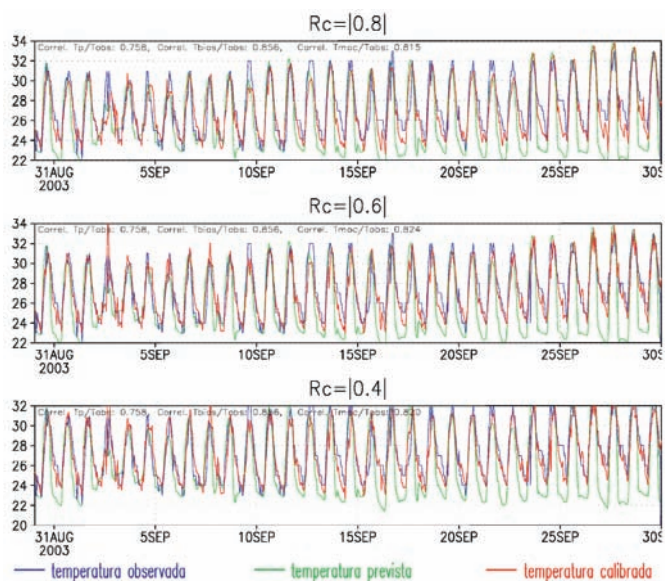


**Figura 2** – Temperatura do ar a 2m (°C) de Maceió em setembro de 2003, observada (azul), prevista pelo Eta (verde), corrigida pelo MOC (vermelha), para  $R_c=0.8; 0.6; 0.4$ .



**Figura 3** – Temperatura do ar a 2m (°C) de Recife em setembro de 2003, observada (azul), prevista pelo Eta (verde), corrigida pelo MOC (vermelha), para  $R_c=0.8; 0.6; 0.4$ .

A Figura 4 mostra as séries de Natal construídas com os 3  $R_c$ 's. Nesta estação as previsões do Eta subestimam as temperaturas em aproximadamente 1.3 °C, enquanto as máximas são ligeiramente superestimadas, as mínimas são fortemente subestimadas, principalmente nos últimos 20 dias do mês. O esquema apresenta redução dos três erros: sendo o ganho de 49% no RMSE, 51% no MAE e 79% no BIAS. As correlações entre as séries observadas e a corrigida aumentaram em relação à série prevista pelo Eta.

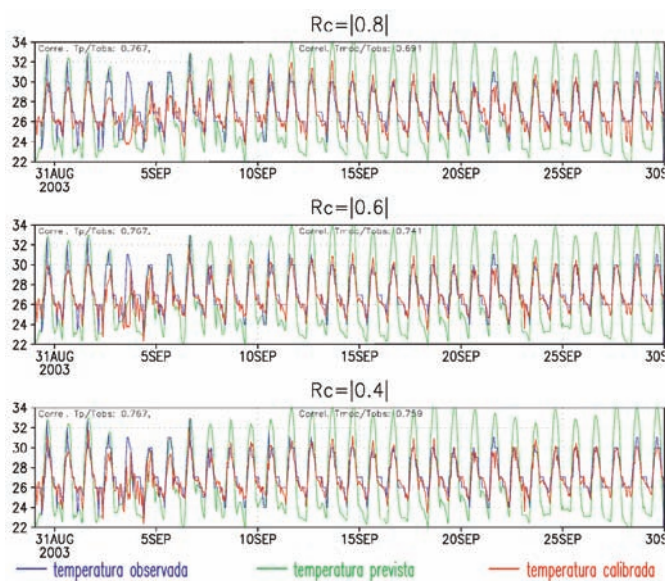


**Figura 4** – Temperatura do ar a 2m (°C) de Natal em setembro de 2003, observada (azul), prevista pelo Eta (verde), corrigida pelo MOC (vermelha), para  $R_c=0.8; 0.6; 0.4$ .

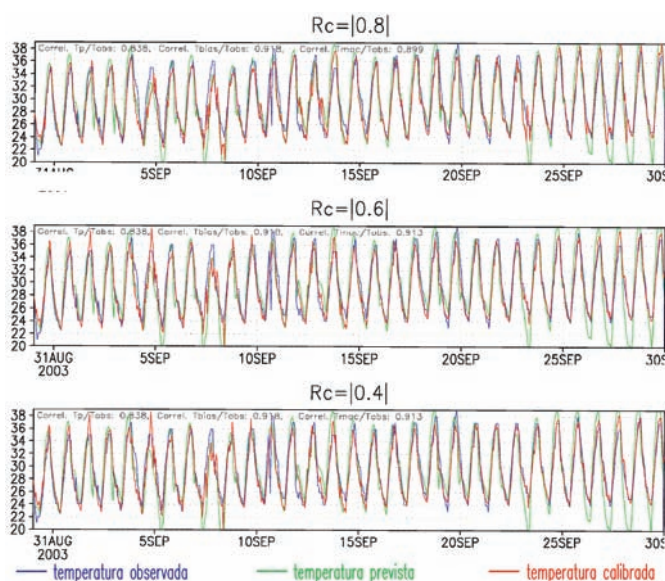
A Figura 5 mostra as séries de Fortaleza. Nesta estação as previsões do Eta superestimam as temperaturas máximas e subestimam as temperaturas mínimas, resultando em uma maior amplitude térmica. Os erros sistemáticos se cancelam e se apresentam pequenos, entretanto, o RMSE revela a presença de erros acumulados. O esquema MOC reduz o RMSE de 2,5 °C para 1 °C, o que representa um ganho de 59% para o RMSE. O MAE apresenta redução de 69%. As correlações das séries apesar de não apresentar aumento, também não apresentam diminuição significativa. Em Fortaleza, a série construída a partir de  $R_c=0.6$  apresentou-se ligeiramente superior à série de  $R_c=0.8$ .

A Figura 6 mostra as séries de Teresina construídas com os 3  $R_c$ 's. Nota-se que nesta estação as previsões praticamente não apresentam BIAS. O esquema acaba aumentando, em módulo, o BIAS e resultando em perda. Entretanto, os demais erros, RMSE e MAE são reduzidos em aproximadamente 1 °C, representando um ganho importante de 36% e 46%, respectivamente. As correlações das séries corrigidas apresentaram aumento em relação à série prevista pelo Eta. A  $R_c=0.6$  parece ser a melhor opção para esta cidade.

A Figura 7 mostra as séries de São Luís para os 3  $R_c$ 's. Nesta cidade as previsões do Eta subestimam frequentemente as temperaturas mínimas, resultando em BIAS de cerca de -1.2 °C. Nas horas matutinas, nota-se que o modelo tende a iniciar o aumento da temperatura algumas horas antes da observação. O esquema MOC acompanha corretamente a queda lenta de temperatura nos horários noturnos. Nesta cidade o uso de



**Figura 5** – Temperatura do ar a 2m (°C) de Fortaleza em setembro de 2003, observada (azul), prevista pelo Eta (verde), corrigida pelo MOC (vermelha), para  $R_c=0.8; 0.6; 0.4$ .



**Figura 6** – Temperatura do ar a 2m (°C) de Teresina em setembro de 2003, observada (azul), prevista pelo Eta (verde), corrigida pelo MOC (vermelha), para  $R_c=0.8; 0.6; 0.4$ .

$R_c=0.6$  resulta em grandes erros, sendo necessária a seleção dos preditores utilizando  $R_c=0.8$ . A redução do RMSE é de cerca de 42% e MAE de 50%. O skill score na redução do BIAS é de 91%. A opção  $R_c=0.8$  resulta em série de maior correlação entre a série corrigida e a série observada.

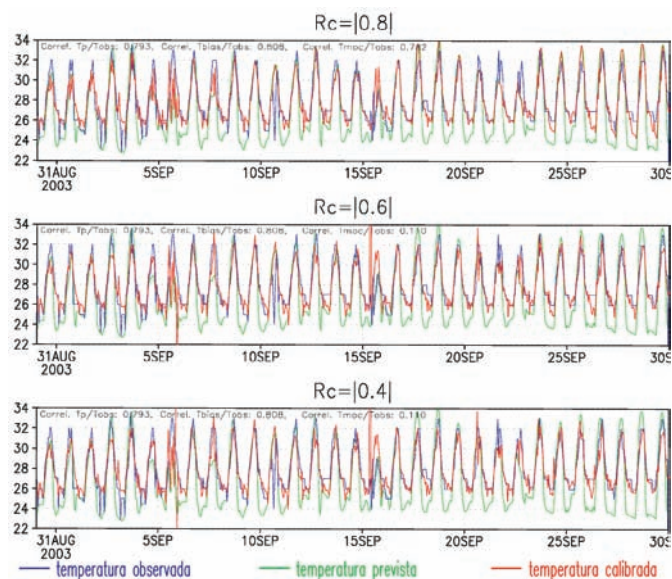


Figura 7 – Temperatura do ar a 2m (°C) de São Luís em setembro de 2003, observada (azul), prevista pelo Eta (verde), corrigida pelo MOC (vermelha), para  $R_c=0.8; 0.6; 0.4$ .

A Tabela 2 resume os erros em diferentes estações para o período de 30 dias de funcionamento do esquema MOC, para os diferentes  $R_c$ . Nota-se que o RMSE médio das estações após o refinamento foi de aproximadamente 1,2 °C. Esta foi a ordem de magnitude de erro das previsões de temperaturas mínimas obtidas por Sugahara (2000) empregando MOS.

Os dados observacionais apresentam erros freqüentes. A aplicação de um filtro simples nos dados produz maior estabilidade no esquema, enquanto que a ausência deste filtro produz variações na série e grandes erros na série refinada.

As variáveis selecionadas mais freqüentemente pelo esquema para construção da equação de regressão foram: temperatura do ar a 2m, temperatura potencial a 10m, temperatura máxima e mínima no intervalo de uma hora, temperatura e tendência temperatura à superfície, temperatura do solo, umidade específica e relativa a 2m, fluxos de calor sensível a superfície, evaporação potencial, radiação de onda longa emitida e incidente à superfície, radiação de onda curta incidente e refletida à superfície, radiação de onda curta incidente no topo da atmosfera.

Tabela 2 – Erros das previsões do Modelo Eta e após refinamento pelo esquema MOC em diferentes níveis de  $R_c$ : Raiz quadrada do erro quadrático médio (RMSE), erro absoluto médio (MAE) e erro médio (BIAS). A unidade é m/s. O valor de  $R_c$  empregado está indicado entre colchetes.

| Cidade    | RMSE |       |       |       | MAE  |       |       |       | BIAS  |       |       |       |
|-----------|------|-------|-------|-------|------|-------|-------|-------|-------|-------|-------|-------|
|           | Eta  | MOC   |       |       | Eta  | MOC   |       |       | Eta   | MOC   |       |       |
|           |      | [0.8] | [0.6] | [0.4] |      | [0.8] | [0.6] | [0.4] |       | [0.8] | [0.6] | [0.4] |
| Salvador  | 1,63 | 1,27  | 1,27  | 1,32  | 1,21 | 0,96  | 0,95  | 1,00  | 0,76  | -0,08 | -0,06 | -0,07 |
| Maceió    | 2,45 | 1,30  | 1,31  | 1,39  | 2,15 | 1,06  | 1,04  | 1,06  | -2,09 | -0,15 | -0,11 | -0,14 |
| Recife    | 2,16 | 1,27  | 1,18  | 3,34  | 1,72 | 1,00  | 0,92  | 1,11  | -1,14 | -0,26 | -0,20 | -0,32 |
| Natal     | 2,45 | 1,37  | 1,26  | 1,28  | 1,93 | 1,06  | 0,95  | 0,96  | -1,34 | -0,39 | -0,28 | -0,27 |
| Fortaleza | 2,51 | 1,18  | 1,03  | 0,96  | 2,22 | 0,80  | 0,68  | 0,68  | -0,31 | -0,19 | -0,27 | -0,25 |
| Teresina  | 2,58 | 1,65  | 1,44  | 1,44  | 2,01 | 1,21  | 1,09  | 1,08  | -0,14 | -0,40 | -0,38 | -0,35 |
| São Luís  | 2,08 | 1,21  | 12,96 | 12,96 | 1,71 | 0,85  | 1,27  | 1,28  | -1,19 | -0,11 | 0,29  | 0,34  |

Tabela 3 – Índice de desempenho (SS-Skill Score) (%) do refinamento pelo esquema MOC sobre as previsões do Modelo Eta. O valor de  $R_c$  empregado está indicado entre colchetes.

| Cidade    | SS MOC RMSE % |       |       | SS MOC MAE % |       |       | SS MOC BIAS % |       |       |
|-----------|---------------|-------|-------|--------------|-------|-------|---------------|-------|-------|
|           | [0.8]         | [0.6] | [0.4] | [0.8]        | [0.6] | [0.4] | [0.8]         | [0.6] | [0.4] |
| Salvador  | 22            | 22    | 19    | 21           | 21    | 17    | 111           | 108   | 109   |
| Maceió    | 47            | 47    | 43    | 51           | 52    | 51    | 93            | 95    | 93    |
| Recife    | 41            | 45    | -55   | 42           | 47    | 35    | 77            | 82    | 72    |
| Natal     | 44            | 49    | 48    | 45           | 51    | 50    | 71            | 79    | 80    |
| Fortaleza | 53            | 59    | 62    | 64           | 69    | 69    | 39            | 13    | 19    |
| Teresina  | 36            | 36    | 44    | 40           | 46    | 46    | -186          | -171  | -150  |
| São Luís  | 42            | -523  | -523  | 50           | 26    | 25    | 91            | 124   | 129   |

Esta lista mostra que, em grande parte, os erros destas previsões estavam relacionados com as variáveis de superfície e de balanço de energia. Algumas daquelas variáveis são geradas pelo esquema que trata das trocas turbulentas entre a superfície-biosfera-atmosfera, no caso do Modelo Eta, o esquema OSU (Chen et al., 1997) foi utilizado. Parte daquelas variáveis utilizadas para a construção da equação de regressão foram geradas pelo esquema de radiação. O Modelo Eta representa a radiação solar pelo esquema de Lacis e Hansen (1974), enquanto que a radiação de onda longa segue o esquema de Fels e Schwarzkopf (1975). O balanço de radiação depende fortemente do horário do dia, da cobertura de nebulosidade, da umidade do solo. Estes dois esquemas, de radiação e de superfície, interagem fortemente no modelo. A interação se dá através do controle do albedo da superfície, que por sua vez depende do tipo de cobertura vegetal, do tipo de solo, da umidade do solo. No trabalho de Hinkelman et al. (2002), detectou-se um excesso médio de  $50 \text{ Wm}^{-2}$  na radiação solar incidente prevista pelo Modelo Eta para uma localidade em Oklahoma; este erro foi atribuído em parte à baixa absorção do vapor d'água e aerossóis do modelo e em parte ao tratamento das nuvens do modelo. Este excesso de radiação de onda curta incidente também foi notado no trabalho de avaliação do Modelo Eta/SSiB para a América do Sul realizado por Chou et al (2002).

As chuvas ocorridas no litoral leste no período inicial de 10 dias do mês foram detectadas pelo Modelo Eta, como se pode notar pela menor amplitude das previsões de temperatura das estações, por exemplo, de Maceió e Recife, para este período. No período final do mês, quando as temperaturas voltam a apresentar menor amplitude devido à presença da nebulosidade baixa, o modelo não representa bem estas nuvens e produz previsão de temperatura com grande amplitude térmica gerando os erros. O esquema estatístico rapidamente se adapta aos novos erros do modelo e os corrige. Fenômenos localizados, como convecção isolada, são mais difíceis de serem detectados pelo modelo, enquanto que a convecção produzida por sistemas de grande escala, como passagens de frentes, é melhor prevista pelo modelo.

### 4.3. Ciclo diurno do erro da previsão

O ciclo diurno médio da temperatura e do RMSE das cidades durante o período de 30 dias de teste do esquema MOC pode indicar os períodos do dia em que os erros são frequentemente gerados.

Nota-se que em Salvador (Fig. 8a) as previsões do ciclo diurno apresentam amplitude térmica pequena demais e, em geral, superestimativa. O esquema corrige o ciclo diurno e a amplitude, mas deixa um erro médio residual. Os erros foram reduzidos principalmente nos horários noturnos. Estes erros da

previsão se devem à maior proporção de água nesta caixa de grade do modelo Eta.

Em Maceió (Fig. 8b), as previsões do ciclo diurno acompanham corretamente a variação da amplitude, entretanto, sistematicamente subestima as temperaturas. O esquema em média quase eliminou o erro sistemático, reduzindo o RMSE em todos os horários do dia.

Pode-se notar que, como comportamento médio, as previsões tendem a apresentar maior amplitude térmica do que o observado, isso ocorre em 4 localidades: Recife, Natal, Fortaleza e São Luís (Figs. 8 c, d, e, g). Ao contrário de Salvador, o modelo Eta nestes pontos apresentam menor proporção de superfície de água. Em Recife, Natal e São Luís, o esquema é bastante eficiente em reduzir o RMSE nos horários noturnos, entretanto, nos horários convectivos, o RMSE foi pouco alterado. Em geral, o esquema refina bastante as previsões do Modelo Eta, construindo nova curva do ciclo diurno para estas 4 cidades.

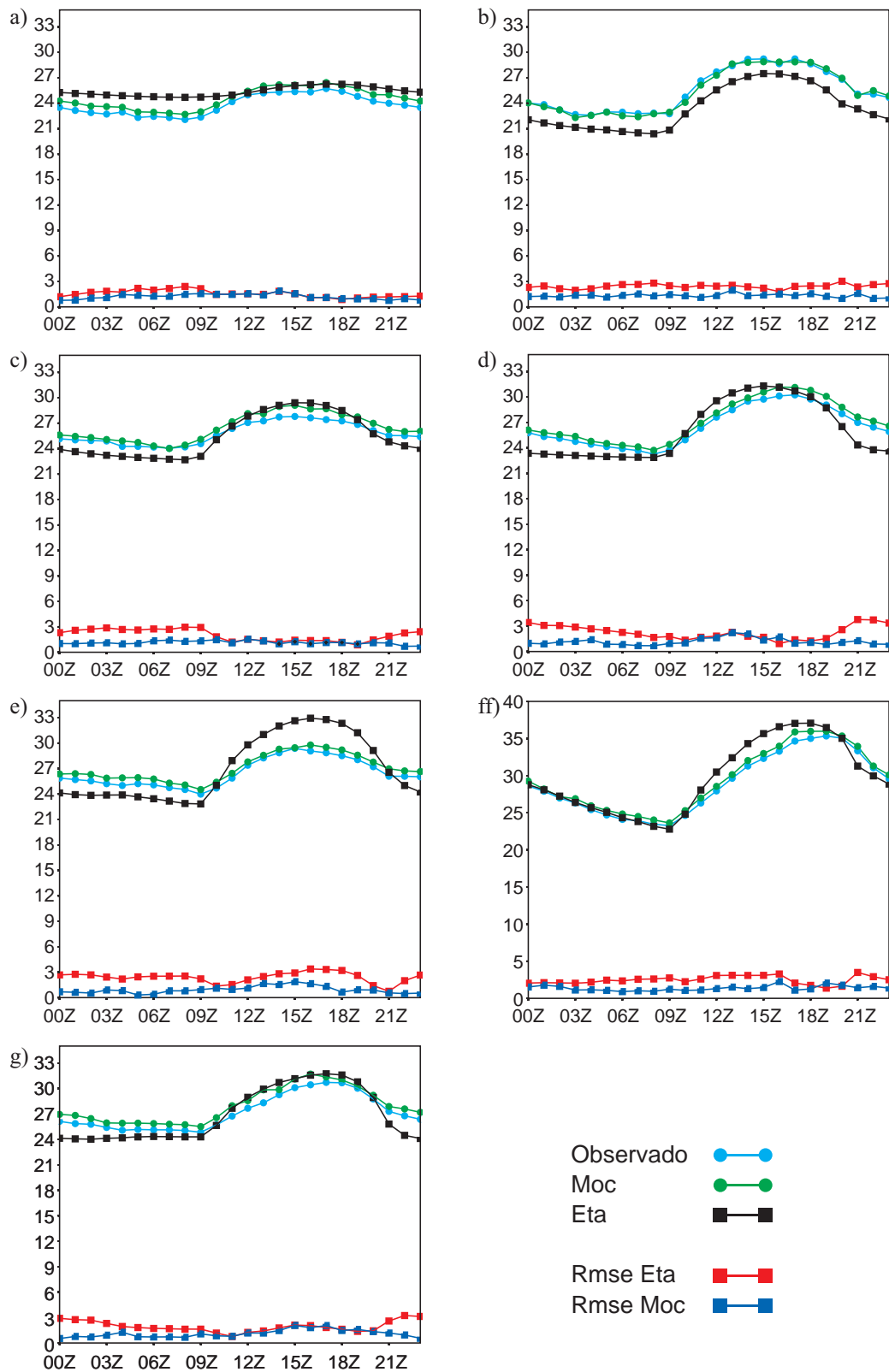
Em Teresina (Fig. 8f) e Natal (Fig. 8.), os erros ocorrem nos horários entre 12Z e 21Z, quando as temperaturas são superestimadas. Nota-se que estes erros são produzidos devido à curva de temperatura prevista estar adiantada em relação à observação.

## 5. CONCLUSÕES

O esquema de refinamento estatístico MOC foi aplicado nas previsões de temperatura do Modelo Eta em 7 estações do Nordeste do Brasil para o período de previsões de setembro de 2003 adotando um período de treinamento de 28 dias. A redução dos erros com a aplicação do MOC foi da ordem de 43% para o RMSE, enquanto que a redução dos erros sistemáticos (BIAS) foi freqüente e ficou abaixo de  $0.4^\circ\text{C}$ . O valor crítico da correlação linear entre o erro da previsão e os candidatos a preditores,  $R_c$ , utilizado para a escolha dos preditores mostrou gerar resultados superiores em 0.6. Mao et al. (1999) aplicaram MOC sobre as previsões de um modelo regional espectral para 6 estações nos Estados Unidos que apresentavam fortes erros sistemáticos. Eles obtiveram uma redução dos erros sistemáticos superior a 40%, o que é comparável com os resultados obtidos para a região Nordeste.

Falhas nas observações comprometeram a qualidade das estimativas das correções do esquema. Além da irregularidade das mensagens, algumas estações de aeroportos não costumam operar durante a madrugada agravando a qualidade das observações. As mensagens METAR não possuem precisão decimal, isto é, não reportam a primeira casa decimal, o que dificulta o esquema perceber as variações ou tendências de temperatura. Diante da baixa qualidade das mensagens METAR, recomenda-se testar o esquema utilizando as mensagens das estações automáticas de superfície que apresentam maior precisão.





**Figura 8** – Ciclo diurno médio das temperaturas: observada (azul claro), prevista pelo Eta (verde) e refinada pelo MOC (preto). Ciclo diurno médio da raiz dos erros quadráticos médios das previsões do Eta (vermelho) e do MOC (azul escuro), para as cidades: (a) Salvador, (b) Maceió, (c) Recife, (d) Natal, (e) Fortaleza, (f) Teresina, (g) São Luís. A unidade é °C.

O esquema pode ser facilmente estendido para outras cidades onde se identifica fortes erros sistemáticos na previsão do Modelo Eta. A escolha do  $R_c$  e o desempenho do esquema deve ser novamente avaliado. Da mesma forma, o esquema pode ser estendido para refinar outras variáveis bem comportadas desde que haja observações disponíveis. Em geral, o esquema MOC reduziu os erros de previsão e a curva produzida se aproxima mais da observação em todas as estações testadas para a variável utilizada.

A metodologia apresenta-se inadequada na aplicação em variáveis como precipitação ou nebulosidade. Por se basear em um modelo linear, variáveis com comportamento não linear e de ocorrência e não-ocorrência não são corretamente representados. Um outro aspecto que poderia ser uma limitação é a necessidade de atualização diária dos coeficientes da equação do MOC. Por outro lado, este cálculo diário permite que o esquema capture a sazonalidade, em regiões onde está mais marcada.

O esquema se apresenta vantajoso ao ser utilizado para refinar previsões de modelo que estão em constante aprimoramento visto que o período de treinamento é curto, além de ser de baixo custo computacional. Os erros de previsão detectados neste trabalho subsidiarão estudos para correção dos processos físicos representados pelo modelo numérico. Testes com o esquema em diferentes períodos do ano podem sugerir erros originados de diferentes processos físicos. O esquema pode ser testado em outras regiões sob diferente regime meteorológico.

## 6. AGRADECIMENTOS

Ao Instituto Nacional de Meteorologia – INMET e Departamento de Controle do Espaço Aéreo – DECEA pelos dados meteorológicos. Os autores Carlos Renato de Souza e Edgard Freitas Evangelista agradecem ao CNPq que financiou o desenvolvimento de parte deste trabalho através do programa CNPq/PIBIC/INPE.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

- BLACK, T. L. The new NMC mesoscale Eta model: Description and forecast examples. **Wea. Analysis and Forecasting**, v.9, p.265-278, 1994.
- CHEN, F.; JANJIC, Z. I.; MITCHELL, K. Impact of atmospheric surface-layer parameterization in the new land-surface scheme of the NCEP mesoscale Eta model. **Boundary Layer Meteorology**, v.85, p.391-421, 1997.
- CHOU, S. C.; TANAJURA, C. A. S.; XUE, Y-K.; NOBRE, C. A. Validation of the Eta/SSiB over South America. **J. Geophys. Res.**, v.107, D20, 2002.
- FELS, S. B.; SCHWARZKOPF, M. D. The simplified exchange approximation. A new method for radiative transfer calculations. **J. Atmos. Sci.**, v.32, p.1475-1488, 1975.
- GLAHN H. R; LOWRY, D. L. The use of Model Output Statistics in Objective Weather Forecasting. **Journal of Applied Meteorology**, v.11, p. 1203-1211, 1972
- GUARNIERI, R. A. **Emprego de redes neurais artificiais e regressão linear múltipla no refinamento das previsões de radiação solar do Modelo Eta**. 2006. 171f. Dissertação (Mestrado em Meteorologia) - Instituto Nacional de Pesquisas Espaciais. INPE-14172-TDI/1089, São José dos Campos, 2006.
- HAYKIN, S. **Neural Networks. A comprehensive foundation**. Macmillan College Publishing Company, 1994.
- HINKELMAN, L. M.; ACKERMAN, T. P.; MARCHAND, R. T. An evaluation of NCEP Eta Model predictions of surface energy budget and cloud properties by comparison with measured ARM data. **J. Geophys. Res.**, v.104, D16, p.19535-19549, 1999.
- HSU, K.; GAO, X.; SOROOSHIAN, S; GUPTA, H. Precipitation estimation from remotely sensed information using artificial neural networks. **J. Applied Meteorology**, v.36, n.9, p.1176-1190, 1997.
- LACIS, A. A.; HANSEN, J. E. A parameterization for the absorption of solar radiation in the Earth's atmosphere. **J. Atmos. Sci.**, v. 31, p.118-133, 1974.
- MESINGER, F.; JANJIC, Z. I.; NICKOVIC, S.; GAVRILOV, D.; DEAVEN, D. G.: The step-mountain coordinate: Model description and performance for cases of Alpine lee cyclogenesis and for a case of Appalachian redevelopment. **Mon. Wea. Rev.**, v.116, p.1493-1518, 1988.
- MAO, Q.; McNIDER, R. T.; MUELLERE, S. F.; JUANG, H. An Optimal Model Output Calibration Algorithm Suitable for objective temperature forecasting. **Wea. Analysis and Forecasting**, v.14, p.190-202, 1994.
- RAMIREZ, M. C. V. **Previsão e análise da precipitação sobre as regiões Sudeste e Sul do Brasil utilizando redes neurais artificiais**. Tese (Doutorado em meteorologia) - Instituto Nacional de Pesquisas Espaciais. INPE-13950-TDI/1061, São José dos Campos, 2003.
- SUGAHARA, S. Uma experiência com modelo estatístico (MOS) para a previsão da temperatura mínima diária do ar. **Revista Brasileira de Geofísica**, v.18, p.3-11, 2000.