



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

INPE-10068-PRE/5613

**REFINAMENTO DAS PREVISÕES DE TEMPERATURA DO
MODELO REGIONAL ETA**

Carlos Renato de Souza*
Chou Sin Chan
Jorge Luís Gomes

Seminário de Iniciação Científica do INPE – SICINPE 2003.

INPE
São José dos Campos
2003



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

REFINAMENTO DAS PREVISÕES DE TEMPERATURA DO MODELO REGIONAL ETA

RELATÓRIO FINAL DE PROJETO DE INICIAÇÃO CIENTÍFICA (INPE/PIBIC/CNPq)

Carlos Renato de Souza (UNITAU, Bolsista PIBIC/CNPq)

E-mail: carlosr@cptec.inpe.br

Dra. Chou Sin Chan (DOP/CPTEC/INPE, Orientadora)

E-mail: chou@cptec.inpe.br

Msc. Jorge Luís Gomes (DOP/CPTEC/INPE, Co-orientador)

E-mail: gomes@cptec.inpe.br

Maio de 2003

1. Introdução

O Modelo Regional Eta (Climanálise, 1996) de Previsão Numérica está implementado operacionalmente no Centro de Previsão de Tempo e Estudos Climáticos (CPTEC) desde 1996. Roda duas vezes ao dia (00UTC e 12UTC) e suas previsões se estendem até 72h de antecedência cobrindo grande parte da América do Sul. As previsões geradas pelo Eta têm representatividade sobre uma região de aproximadamente 40km x 40km e têm sido consideradas ferramentas úteis para a elaboração das previsões de tempo pelos meteorologistas.

Entretanto, em regiões de orografia e vegetação complexas, esta representatividade se torna deficiente e esses modelos apresentam erros que acabam por refletir na qualidade das previsões diminuindo assim a confiabilidade dos resultados.

Pretende-se aplicar uma correção nos resultados das previsões através de cálculos estatísticos e matemáticos na tentativa de aproximar ao máximo as previsões geradas pelo Eta aos valores observados. O MOC (Model Output Calibration) (Mao et al., 1999) apresenta um modelo matemático estatístico para ajustar, por exemplo, a temperatura do ar a 2m prevista pelo modelo determinístico em um ponto próximo à observação.

1.1. Objetivo

O objetivo deste trabalho é o refinamento das previsões de temperatura do ar a 2m e vento a 10m do Modelo Eta através de ajustes estatísticos utilizando o algoritmo MOC (Model Output Calibration) (Mao et al., 1999).

2. Metodologia

O algoritmo MOC consiste em formular equações de regressão linear multivariada com base nas séries de 2 a 4 semanas das previsões e observações, estimando e ajustando assim os erros da previsão. Define-se como “temperatura prevista” a temperatura não ajustada do modelo (T_{2m}), como “temperatura ajustada” ou “temperatura calibrada” a temperatura prevista pelo Eta com a calibração aplicada (T_{moc}), e como “temperatura observada” o valor real da temperatura observada pelas estações

meteorológicas (T_{obs}).

Considerando que $T_{2m}(t)$ é uma série temporal de previsões de temperatura do ar a 2m de uma estação localizada na latitude i e longitude j e t é o índice do tempo, considerando também que $T_{obs}(t)$ é a série temporal da temperatura observada na mesma posição e no mesmo período de tempo, o erro das previsões da temperatura do ar a 2m (T_{2m}) é definido como:

$$\begin{aligned}\Delta T_{i,j}(t) &= T_{2m_{i,j}}(t) - T_{obs_{i,j}}(t), \\ t &= 1, 2, \dots, N.\end{aligned}\tag{1}$$

Admitindo-se que $\Delta T_{i,j}(t)$ está intrinsicamente associado às outras variáveis e parâmetros de saída do modelo, então assumimos que $\Delta T_{i,j}(t)$ pode ser uma função de $X_{i,j,l}(t)$, onde l é o índice da variável associada:

$$\begin{aligned}\Delta T_{i,j}(t) &= \prod(X_{i,j,l}(t)), \\ t &= 1, 2, \dots, N; \\ l &= 1, 2, \dots, L.\end{aligned}\tag{2}$$

Forma-se assim o modelo de equação linear multivariada:

$$\begin{aligned}\Delta T_{i,j}(t) &= \Lambda_0 + \sum_{l=1}^L \Lambda_l X_{i,j,l}(t), \\ t &= 1, 2, \dots, N, \\ l &= 1, 2, \dots, L;\end{aligned}\tag{3}$$

onde os coeficientes Λ_0 e Λ_l podem ser determinados através da regressão linear multivariada. O erro da previsão, ΔT , é o preditando das equações, enquanto que as outras variáveis de saída do modelo (X_l) são os preditores. A calibração será aplicada no tempo seguinte, ou seja, em $N + 1$. As previsões refinadas são representadas por:

$$\begin{aligned}T_{moc_{i,j}}(t) &= T_{2m_{i,j}}(t) - \Delta T_{i,j}(t), \\ t &= N + 1, N + 2, \dots, N + M,\end{aligned}\tag{4}$$

onde M representa o tamanho da série de tempo dos experimentos usada para as equações. Foram feitos testes com $\Delta T_{i,j}$ e T_{obs} como preditando.

Foram utilizadas como variáveis predictoras somente aquelas que apresentaram correlação linear com o preditando superior a R_c . O número total de preditores

adotado nas equações não excedeu a cinco, quando foram qualificadas mais de cinco preditores somente as cinco variáveis com os maiores coeficientes de correlação entraram nas equações. A Tabela 1 mostra a lista das variáveis candidatas a preditores a serem qualificadas pelo MOC. Estas variáveis são fornecidas nas previsões operacionais do Eta.

Foram utilizadas previsões do Modelo Eta iniciadas às 0000UTC para 36h de antecedência. O esquema de calibração foi testado para o período de 02/03/2000 a 01/06/2000, e de 02/03/2003 a 01/05/2003. No primeiro período foram utilizados dados diários das 1200UTC da estação do Aeroporto de São Paulo (46,65°W, 23,61°S). No segundo período foram utilizados dados da estação de Cachoeira Paulista (45,00°W, 22,67°S). A fim de se obter um bom resultado de calibração, efetuou-se testes com diferentes limiares de correlação: $R_c = 0.6$, $R_c = 0.7$ e $R_c = 0.8$), e com diferentes tamanhos de série: $N = 14$, $N = 21$ e $N = 28$ dias.

2.1 Avaliação

Para uma avaliação das previsões de temperatura geradas pelo MOC, foram usadas três medidas estatísticas básicas: o erro médio (BIAS), o erro médio absoluto (MAE) e o erro quadrático médio (RMSE). Que são dados por:

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

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

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

onde M é o número total de previsões a serem analisadas, T_i é a i -ésima temperatura prevista com o ajuste do MOC ou sem o ajuste (Eta), e T_{obs_i} é a i -ésima temperatura observada. Também foi aplicado o “*Skill Score*” para medir em porcentagem o desempenho das previsões ajustadas sobre as não ajustadas. O “*Skill Score*” é dado por:

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

onde E_{Eta} ou E_{MOC} é uma das três medidas estatísticas BIAS, MAE ou RMSE para T_{2m} ou para T_{moc} .

Tabela 1: Lista dos candidatos a preditores qualificados pelo MOC.

Preditores	Descrição
T_{2m}	temperatura do ar a 2 metros [K]
T	temperatura do ar absoluta a 1000 ou 850 hPa [K]
P_{nmm}	pressão ao nível do mar (Método Mesinger) [hPa]
P_s	pressão da superfície [hPa]
Z	topografia [m]
LSM	máscara do oceano e continente [0,1]
Td	temperatura do ponto de orvalho [K]
U_{10}	componente zonal do vento a 10m [m/s]
V_{10}	componente meridional do vento a 10m [m/s]
P_t	precipitação total [kg/m ² /dia]
P_{cv}	precipitação convectiva [kg/m ² /dia]
Pr_{ge}	precipitação em grande escala [kg/m ² /dia]
$Neve$	neve [kg/m ² /dia]
L	fluxo de calor latente à superfície [W/m ²]
H	fluxo de calor sensível à superfície [W/m ²]
G	fluxo de calor no solo [W/m ²]
T_{sup}	temperatura da superfície [K]
q_{sup}	Umidade Específica na superfície [K]
T_s	temperatura da superfície do solo [K]
T_r	temperatura da zona das raízes [K]
q_s	umidade da superfície do solo [0-1]
q_r	umidade na zona das raízes [0-1]
q_{dis}	umidade disponível do solo [0-1]
R	runoff [kg/m ² /s]
Z_o	rugosidade [m]
Evp_s	evaporação potencial da superfície [m]
U_z	cisalhamento do vento zonal na superfície a 1000 ou 850 hPa [Pa]
V_z	cisalhamento do vento meridional na superfície a 1000 ou 850 hPa [Pa]
C_b	cobertura de nuvens baixas [0-1]
C_m	cobertura de nuvens médias [0-1]
C_a	cobertura de nuvens altas [0-1]

Preditores	Descrição
C	cobertura total de nuvens [0-1]
SW_{is}	radiação de onda curta incidente à superfície [W/M^2]
LW_{is}	radiação de onda longa incidente à superfície [W/M^2]
SW_{es}	radiação de onda curta emergente à superfície [W/m^2]
LW_{es}	radiação de onda longa emergente à superfície [W/m^2]
ROC	radiação de onda curta no topo da atmosfera [W/m^2]
ROL	radiação de onda longa no topo da atmosfera [W/m^2]
a	albedo [%]
$CAPE$	energia potencial disponível para convecção [m^2/s^2]
$CINI$	energia potencial disponível para inibição da convecção [m^2/s^2]
li	índice de levantamento (para 500 hPa) [K]
PW	água precipitável [kg/m^2]
P_{trop}	pressão da tropopausa [hPa]
Z_f	altura do nível de congelamento [m]
UR_f	umidade relativa do nível de congelamento [%]
P_{max}	pressão do nível do vento máximo [hPa]
U_{max}	vento zonal máximo [m/s]
V_{max}	vento meridional máximo [m/s]
P_b	pressão na base da nuvem [hPa]
P_t	pressão no topo da nuvem [hPa]
Geo	altura geopotencial [gpm]
U	vento zonal em 1000 ou 850 hPa [m/s]
V	vento meridional em 1000 ou 850 hPa [m/s]
UR	umidade relativa em 1000 ou 850 hPa [%]
W	movimento vertical do vento em 500 hPa [Pa/s]
q	umidade específica [kg/kg]
θ_e	temperatura potencial equivalente em 1000 ou 850 hPa [K]
l	água líquida da nuvem [kg/m^2]
$\frac{\partial T}{\partial t}$	tendência da temperatura ($T_t - T_{t-1}$) [K/dia]
$\Delta Geo _{500hPa}^{925hPa}$	espessura da camada de 925 a 500 hPa [m]
$\Delta Geo _{700hPa}^{850hPa}$	espessura da camada de 850 a 700 hPa [m]

Preditores	descrição
$\Delta Geo _{700hPa}^{925hPa}$	espessura da camada de 925 a 700 hPa [m]
$\frac{\partial T}{\partial z} _{500hPa}^{925hPa}$	gradiente vertical da temperatura entre 925 e 500 hPa [K/m]
$\frac{\partial T}{\partial z} _{700hPa}^{850hPa}$	gradiente vertical da temperatura entre 850 e 700 hPa [K/m]
$\frac{\partial T}{\partial z} _{700hPa}^{925hPa}$	gradiente vertical da temperatura entre 925 e 700 hPa [K/m]
$\frac{\partial U}{\partial z} _{300hPa}^{700hPa}$	cisalhamento do vento zonal [m/s]
$\frac{\partial V}{\partial z} _{300hPa}^{700hPa}$	cisalhamento do vento meridional [m/s]
ΔT_{solo}	$T_r - T_s$ [K]
Δq_{solo}	$q_r - q_s$ [Kg/Kg]

3. Resultados

O esquema foi testado para o período de abril a maio de 2000 para a estação do Aeroporto de São Paulo. O mês de março de 2000 foi utilizado para o cálculo dos coeficientes das equações. As temperaturas observadas estão em curvas azuis com círculos cheios. As temperaturas previstas pelo Eta sem a calibração estão em curvas verdes com círculos vazios e as temperaturas calibradas pelo MOC representadas pelas curvas vermelhas com quadrados vazios. À esquerda estão os gráficos dos erros estatísticos onde as barras cinza-claro representam os erros RMSE, MAE e BIAS do Eta em relação às temperaturas observadas, e as cinza-escuro da temperatura do MOC em relação às observadas. A seguir serão mostrados resultados testando o limiar de confiança R_c , o tamanho da série da regressão linear, N , a correlação com os preditandos ΔT ou T_{obs} .

A Figura 1 mostra os resultados do MOC para as previsões de São Paulo utilizando uma série de 28 dias, tomou-se como preditando ΔT , e restringiu-se as correlações aos limiares em 0.8, 0.7 e 0.6. Nota-se que a série das temperaturas previstas pelo Eta se apresentou bem próxima à série observada resultando assim em BIAS pequeno ($BIAS_{Eta} = 0,18$). A série temporal da T_{moc} ficou bem próxima à série da T_{2m} do Eta. Os erros não diminuíram significativamente, em alguns casos até aumentaram, como por exemplo RMSE e MAE com os limiares de 0.8 e 0.7. O BIAS diminuiu e em alguns casos mudou de sinal. Porém com o limiar 0.6 todos os erros diminuíram em relação aos erros do Eta.

Utilizou-se o nível de 850 hPa para as variáveis tridimensionais para a estação

Abr-Mai/2000 : SP,Capital - T2m : serie 28 dias (oC)

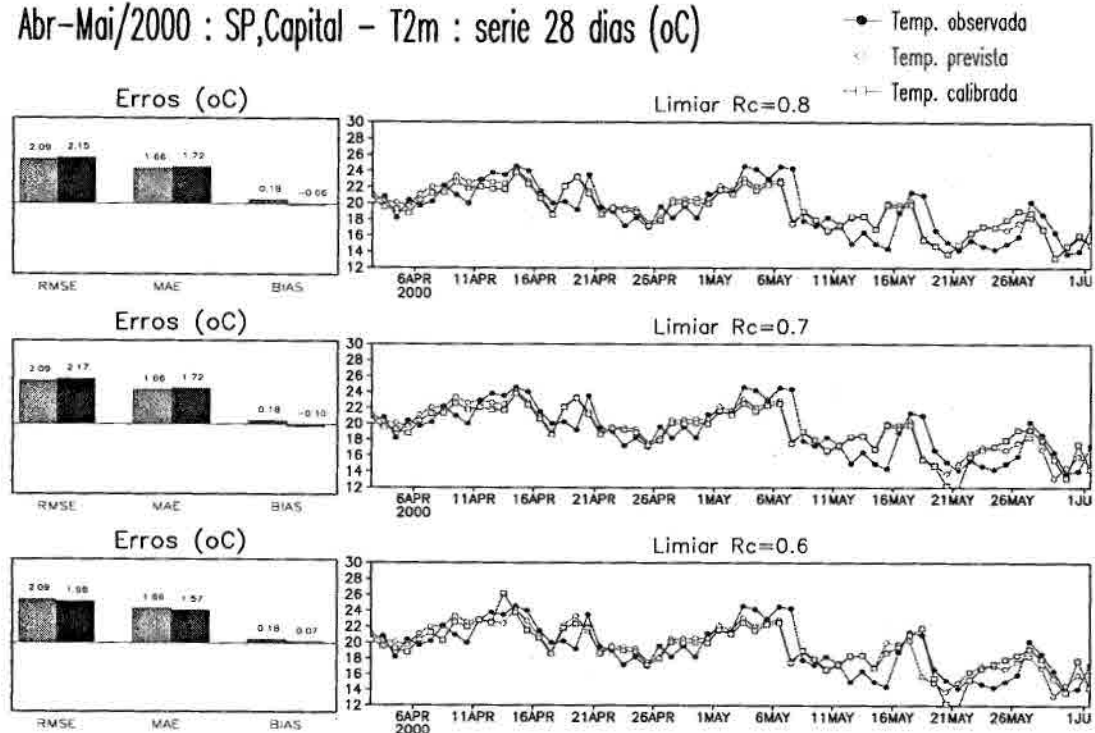


Figura 1: À direita: Temperatura do ar de 2m ($^{\circ}\text{C}$): observada (linha azul com círculo cheio), prevista pelo Eta (linha verde com círculo vazio) e calibrada pelo MOC (linha vermelha com quadrado vazio), com $N = 28$ dias, $R_c = 0.8, 0.7$ e 0.6 . À esquerda: Erros: RMSE, MAE e BIAS da temperatura prevista pelo Eta (barras cinza-claro) e temperatura calibrada pelo MOC (barras cinza-escuro).

de São Paulo. Testou-se também os níveis de 1000 e 925 hPa, porém não se observou melhoria significativa, pois as séries temporais calibradas não apresentaram diferenças significativas.

A Tabela 2 mostra o “Skill Score” do MOC utilizando diferentes limiares de correlação. Apesar de $R_c = 0.8$ produzir melhorias no BIAS, o $R_c = 0.6$ é em geral melhor.

Tabela 2: “Skill Score” de São Paulo, com preditando ΔT , baseado em RMSE, MAE e BIAS do MOC com $N = 28$ dias para São Paulo.

R_c	RMSE	MAE	BIAS
0.8	-3%	-4%	67%
0.7	-4%	-4%	44%
0.6	5%	5%	61%

A Figura 2 apresenta as séries temporais e os respectivos erros para a estação

da cidade de Cachoeira Paulista no período de abril de 2003. O mês de março foi usado para se obter as equações de regressão linear e o mês de abril foi usado para avaliar o esquema MOC. Foram utilizadas variáveis tridimensionais do modelo em 850 hPa e uma série de 28 dias.

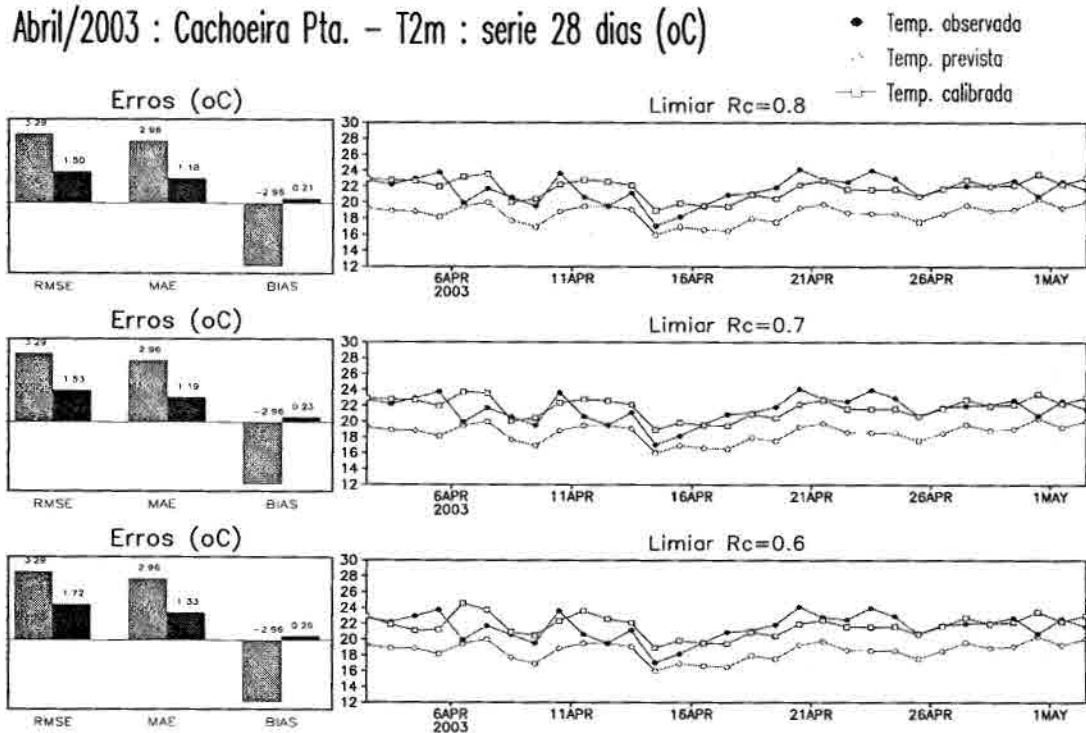


Figura 2: À direita: Temperatura do ar de 2m ($^{\circ}C$): observada (linha azul com círculo cheio), prevista pelo Eta (linha verde com círculo vazio) e calibrada pelo MOC (linha vermelha com quadrado vazio), com $N = 28$ dias, $R_c = 0.8, 0.7$ e 0.6 . À esquerda: Erros: RMSE, MAE e BIAS da temperatura prevista pelo Eta (barras cinza-claro) e temperatura calibrada pelo MOC (barras cinza-escuro).

Nota-se que a série da temperatura prevista pelo Eta se mantém afastada e negativa em relação à temperatura observada, o que resulta em um BIAS alto igual a $2,96^{\circ}C$. A série da temperatura calibrada se apresenta mais próxima à observada do que a prevista, o que gera menores erros. Apesar da T_{moc} se apresentar mais próxima à T_{obs} , ela não acompanha a variação diária da temperatura observada, mas acompanha a curva da T_{2m} . Nota-se o erro sistemático negativo das previsões do Eta. Comparando o ajuste por limiares diferentes, não se nota diferença significativa nas séries temporais, porém analisando os erros nota-se uma pequena melhoria com $R_c = 0.8$.

A Tabela 3 mostra o “Skill Score” do MOC para Cachoeira Paulista com $N = 28$

dias e com as variáveis tridimensionais em 850 hPa. Com $R_c = 0.8$ o MOC apresentou melhores resultados.

Tabela 3: “Skill Score” de Cachoeira Paulista, com preditando ΔT , baseado em RMSE, MAE e BIAS do MOC com $N = 28$ dias.

R_c	RMSE	MAE	BIAS
0.8	54%	60%	93%
0.7	53%	60%	92%
0.6	48%	55%	93%

A Figura 3 mostra o mesmo gráfico que a Figura 2 mas com N reduzido para 21 dias. Nota-se que não há diferenças significativas entre as séries temporais com 28 dias e 21 dias. Porém os erros mostraram diferenças, por exemplo, o RMSE aumentou com $R_c = 0.6$.

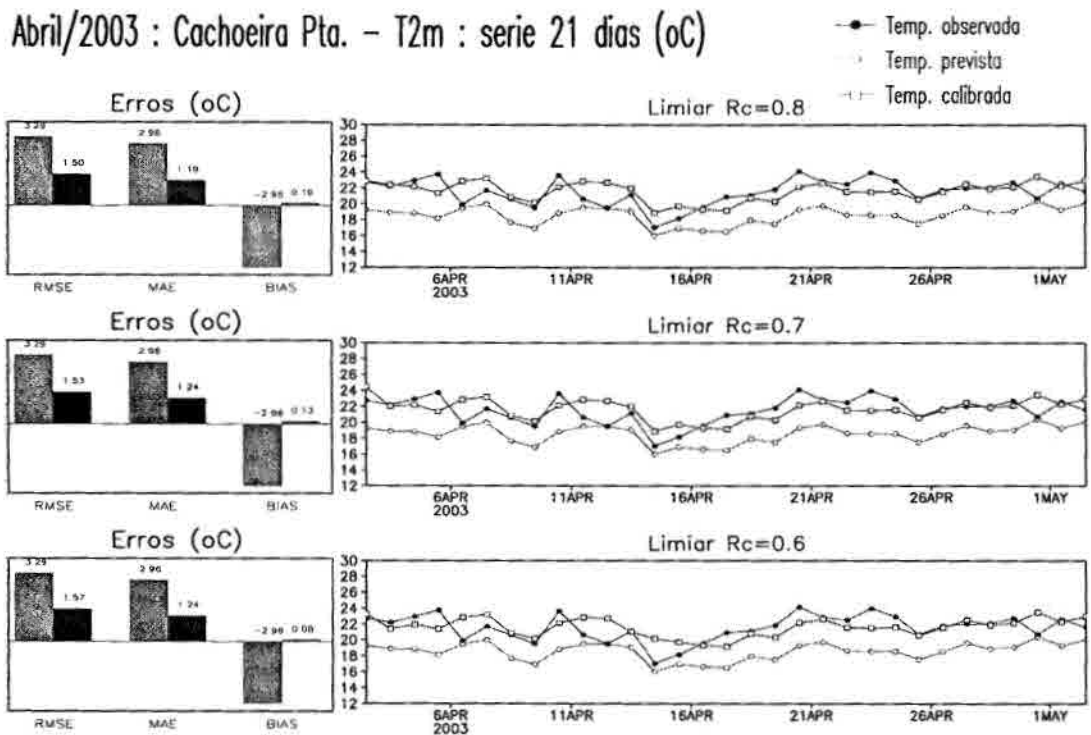


Figura 3: À direita: Temperatura do ar de 2m (°C): observada (linha azul com círculo cheio), prevista pelo Eta (linha verde com círculo vazio) e calibrada pelo MOC (linha vermelha com quadrado vazio), com $N = 28$ dias, $R_c = 0.8, 0.7$ e 0.6 . À esquerda: Erros: RMSE, MAE e BIAS da temperatura prevista pelo Eta (barras cinza-claro) e temperatura calibrada pelo MOC (barras cinza-escuro).

A Tabela 4 mostra o “Skill Score” do MOC para $N = 21$ dias e variáveis

Tabela 4: “Skill Score” de Cachoeira Paulista, com preditando ΔT e $N = 21$ dias.

R_c	RMSE	MAE	BIAS
0.8	54%	60%	97%
0.7	53%	58%	96%
0.6	52%	58%	97%

tridimensionais em 850 hPa. Nessa tabela nota-se as melhorias do MOC com $N = 21$ dias sobre o MOC com $N = 28$ dias.

A Figura 4 apresenta os gráficos semelhantes a Figura 3 mas com T_{obs} sendo a preditanda. Com $N = 28$ dias e com as variáveis tridimensionais do modelo em 850 hPa, nota-se que com $R_c = 0.8$, a série temporal da temperatura calibrada a partir do dia 8 de abril permaneceu quase invariável. Com $R_c = 0.7$ e $R_c = 0.6$ a calibração

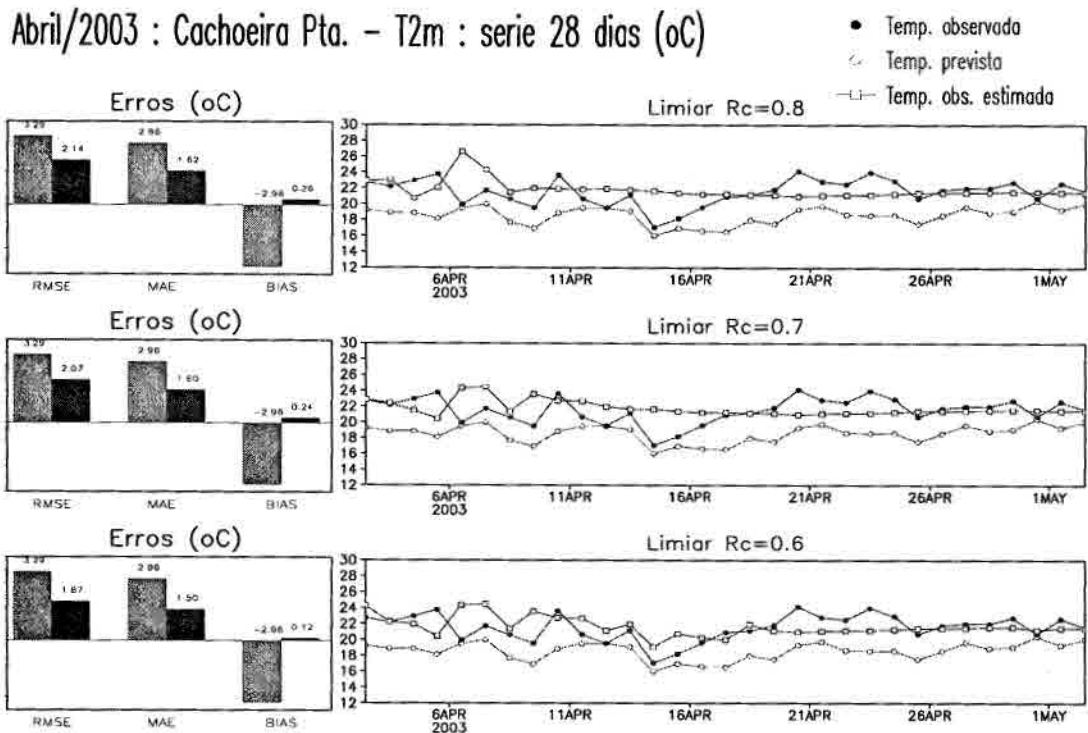


Figura 4: À direita: Temperatura do ar de 2m ($^{\circ}\text{C}$): observada (linha azul com círculo cheio), prevista pelo Eta (linha verde com círculo vazio) e calibrada pelo MOC (linha vermelha com quadrado vazio), com $N = 28$ dias, $R_c = 0.8, 0.7$ e 0.6 . À esquerda: Erros: RMSE, MAE e BIAS da temperatura prevista pelo Eta (barras cinza-claro) e temperatura calibrada pelo MOC (barras cinza-escuro).

não ocorreu a partir dos dias 10 e 19 de abril, respectivamente. Isso ocorre porque a correlação linear das variáveis do modelo com a T_{obs} é baixa, a temperatura prevista

pele Eta é calibrada por um valor quase constante. As correlações são maiores entre os preditores e o preditendo ΔT . No caso de T_{obs} , o valor de R_c deve ser menor para permitir a inclusão de mais preditores na equação de regressão.

Tabela 5: “Skill Score” de Cachoeira Paulista, com preditando T_{obs} , baseado em RMSE, MAE e BIAS do MOC com $N = 28$ dias.

R_c	RMSE	MAE	BIAS
0.8	35%	45%	91%
0.7	37%	46%	92%
0.6	43%	49%	96%

Apesar disso, os erros da T_{moc} se apresentaram menores do que os erros da previsão do Eta. O “Skill Score” do MOC com $N = 28$ dias e as variáveis tridimensionais em 850 hPa estão mostradas na Tabela 5. Com $R_c = 0.6$ o MOC apresentou melhores “Skill Scores”.

Abril/2003 : Cachoeira Pta. - T2m : serie 21 dias (oC)

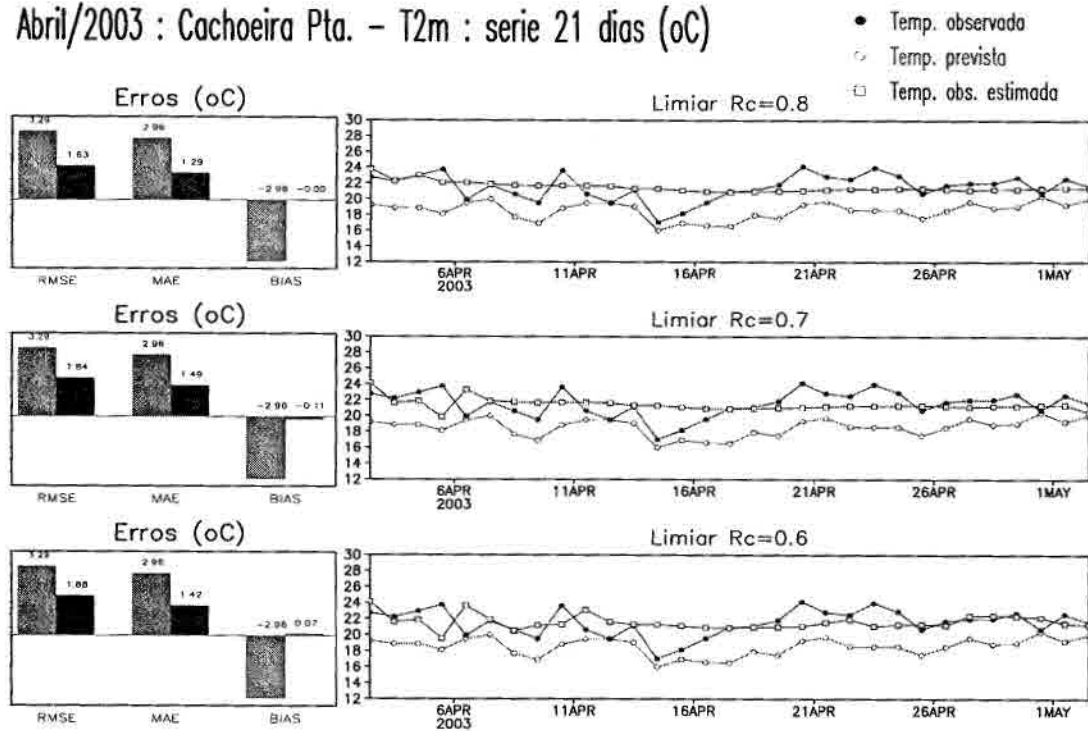


Figura 5: À direita: Temperatura do ar de 2m (°C): observada (linha azul com círculo cheio), prevista pelo Eta (linha verde com círculo vazio) e calibrada pelo MOC (linha vermelha com quadrado vazio), com $N = 28$ dias, $R_c = 0.8, 0.7$ e 0.6 . À esquerda: Erros: RMSE, MAE e BIAS da temperatura prevista pelo Eta (barras cinza-claro) e temperatura calibrada pelo MOC (barras cinza-escuro).

A Figura 5 mostra gráficos semelhantes à Figura 4, mas com $N = 21$ dias. Nota-se que a variação da curva estimada ocorreu até o dia 5 de abril com $R_c = 0.8$. Com os limiares $R_c = 0.7$ e $R_c = 0.6$, a variação ocorreu até os dias 6 e 12 de abril, respectivamente, permanecendo quase constante nos outros dias da série.

Os erros diminuíram em relação ao caso com a série de $N = 28$ dias, como mostra a comparação entre a Tabela 5 e a Tabela 6.

Tabela 6: “Skill Score” de Cachoeira Paulista, com preditando T_{obs} , baseado em RMSE, MAE e BIAS do MOC com $N = 21$ dias.

R_c	RMSE	MAE	BIAS
0.8	50%	56%	100%
0.7	44%	50%	96%
0.6	43%	52%	98%

3.1 Conclusão

O método MOC apresentou melhores refinamento das temperaturas de Cachoeira Paulista utilizando ΔT como preditando. Esta estação apresentou um erro sistemático que foi reduzido pelo método. Porém, a série calibrada não acompanhou o sinal da série observada, e sim o sinal da série prevista pelo Eta, a correção aplicada a T_{2m} foi praticamente constante.

Quando T_{obs} é utilizada como preditando de T_{moc} , os erros se apresentam ligeiramente superiores, apesar da série não não apresentar a variabilidade de T_{obs} ou de T_{2m} a partir de um certo dia.

As temperaturas de São Paulo são bastante semelhantes à série prevista pelo eta. O MOC pouco atuou. Gerou-se uma série T_{moc} muito próxima à série prevista pelo Eta e com erros ligeiramente menores. O método utilizando ΔT como preditando apresentou melhores resultados: com menores erros, maiores “Skill Scores” e com uma série calibrada mais próxima à série observada.

4. Próximos Passos

Os resultados não podem ser considerados conclusivos. Pretende-se ainda realizar mais experimentos com o método. Pretende-se acrescentar mais algumas variáveis derivadas e testar as variáveis tridimensionais em níveis mais elevados, observando os coeficientes de correlação linear dessas variáveis com ΔT e T_{obs} . Pretende-se também testar outros tamanhos de séries (N) relaxando a restrição dos limiares de correlação. Pretende-se aplicar o MOC em outras localidades onde se observa erro sistemático. Pretende-se também estender o método para T_{max} e T_{min} , vento a 10m e Nebulosidade.

5. Bibliografia

- Mao, Q., R. T. McNider, S. F. Muellere H. H. Juang, 1999: An Optimal Model Output Calibration Algorithm Suitable for objective Temperature Forecasting. *Wea. Forecasting*, **14**, 190-202.
- Chou, S. C., 1996: Modelo Regional Eta. In *Climanálise. Edição Comemorativa de 10 anos*. Instituto Nacional de Pesquisas Espaciais. Cachoeira Paulista, SP, Brasil.