

POTENCIALIDADES DO SCAMTEC (SISTEMA DE AVALIAÇÃO DE MODELOS NUMÉRICO DE PREVISÃO DE TEMPO E CLIMA) NA AVALIAÇÃO DE PREVISÕES DE PRECIPITAÇÃO

LUIZ F. SAPUCCI¹, JOÃO G. Z. DE MATTOS², ARIANE F. DE MATTOS³,
LUIZ GUSTAVO G. DE GONÇALVES⁴, CARLOS BASTARZ⁵, DEREK V. SCHUBERT⁶.

¹ Matemático, Pesquisador, Centro de Previsão de Tempo e Estudos Climáticos – CPTEC - Instituto Nacional de Pesquisas Espaciais – INPE - Rodovia Presidente Dutra, km 40 - Cachoeira Paulista – SP. luiz.sapucci@cptec.inpe.br

² Meteorologista, Doutorando, CPTEC-INPE - Cachoeira Paulista – SP

³ Meteorologista, Doutorando, CPTEC-INPE - Cachoeira Paulista – SP

⁴ Meteorologista, Pesquisador, CPTEC-INPE - Cachoeira Paulista – SP

⁵ Matemático Mestre em Meteorologia, Bolsista, CPTEC-INPE - Cachoeira Paulista – SP.

⁶ Bel. em Tecnologia da Informação, Bolsista, CPTEC-INPE - Cachoeira Paulista – SP

Apresentado no XVII Congresso Brasileiro de Agrometeorologia – 18 a 21 de Julho de 2011
SESC, Centro de Turismo de Guarapari, Guarapari - ES.

RESUMO – A previsão numérica de precipitação é um importante produto para a agricultura, uma vez que o regime pluviométrico gera impactos significativos na produção agrícola de cada região, e previsões com alto índice de acerto podem contribuir com o planejamento das safras, minimizando possíveis prejuízos. Embora existam muitas fontes de incertezas no processo de previsão de precipitação, avanços significativos nessa área têm sido obtidos nos últimos anos no Brasil, como o emprego de melhores sistemas de assimilação de dados, a evolução dos modelos de previsão de tempo, bem como o aumento da resolução horizontal e vertical dos mesmos. Avaliar adequadamente o impacto de cada uma dessas melhorias na qualidade final das previsões de precipitação, requer uma metodologia apropriada. Visando contribuir com o tema, encontra-se em fase de implementação no Grupo de Desenvolvimento em Assimilação de Dados (GDAD) do CPTEC-INPE um sistema denominado SCAMTEC (Sistema de Avaliação de Modelos numéricos de Previsão de Tempo e Clima). O presente trabalho tem por objetivo apresentar as potencialidades desse sistema na avaliação da qualidade das previsões de precipitação geradas por modelos numéricos.

PALAVRAS-CHAVE: Previsão de Precipitação, Avaliação de Modelos, Previsão Numérica de Tempo e Clima.

POTENTIALITY OF THE SCAMTEC (SISTEMA DE AVALIAÇÃO DE MODELOS NUMÉRICOS DE PREVISÃO DE TEMPO E CLIMA) IN THE ASSESSMENT OF PRECIPITATION FORECAST

ABSTRACT - The numerical prediction of precipitation is an important product in agriculture since the pluviometric regime cause substantial impacts in the agricultural production of each region and correct predictions can contribute to the planning of the yield helping to minimize the losses. Despite of the many uncertainty sources in the process of precipitation forecast there were significant advances in Brazil due to the better data assimilation systems, improvements in the numerical prediction models, and the increase of temporal and spatial resolution. The correct assessment of impacts from each of the above mentioned improvements in the resulting quality of the precipitation predictions requires appropriated methodology. Aiming to contribute with this topic the Group for Data Assimilation Development (GDAD) at CPTEC-INPE has under development a system called SCAMTEC (Sistema de Avaliação de Modelos numéricos de previsão de Tempo e Clima). The goal of the present work is to present the potentialities of this system when used for assessment of the precipitation predictions quality produced by numerical weather and climate models.

KEYWORDS: Precipitation forecast, Model evaluation, Numerical Climate and Weather Prediction.

Introdução

A produção agrícola está ligada diretamente à disponibilidade espaço-temporal da precipitação, cuja escassez gera aumento dos custos da produção, quando se tem disponíveis sistemas de irrigação, ou em prejuízos na ausência desses. É consenso que a previsão de chuva é um dos produtos da Previsão Numérica de Tempo e Clima (PNTC) mais importantes para esse setor da atividade humana. Por outro lado, o emprego de modelos numéricos para caracterizar adequadamente as componentes principais do ciclo hidrológico, permanece como um grande desafio científico (ANTHES et al. 1989). Nesse desafio, incluem-se a Assimilação de Dados (AD), as parametrizações da temperatura e umidade do solo e a representação de processos físicos de pequena escala, como a convecção e a turbulência, além da microfísica das nuvens. Embora existam diversas fontes de incertezas no processo, significativos avanços nessa área têm ocorrido nos últimos anos, como o emprego de melhores sistemas de assimilação (SAPUCCI et al. 2010), evolução dos modelos de previsão de tempo e clima (NACHAMKIN et al. 2005), bem como o aumento da resolução horizontal e vertical de tais modelos. Avaliar adequadamente o impacto de cada uma dessas melhorias na qualidade final das previsões de precipitação requer uma metodologia apropriada. Visando contribuir com o tema, encontra-se em fase de implementação no Grupo de Desenvolvimento em Assimilação de Dados (GDAD) do CPTEC/INPE um sistema de avaliação de modelos numéricos baseado em softwares livres. Trata-se do SCAMTEC - Sistema de Avaliação de Modelos numéricos de Tempo e Clima (DE MATTOS et al. 2011). O presente trabalho tem por objetivo apresentar as potencialidades deste sistema na avaliação da qualidade das previsões geradas por modelos numéricos.

Sistema de Avaliação de Modelos numéricos de Previsão de Tempo e Clima

Para que toda a atividade ligada a PNTC obtenha sucesso e constantes aprimoramentos, é fundamental o uso de ferramentas adequadas para a avaliação estatística dos resultados das simulações numéricas produzidas. Quando se aplica uma melhoria ao sistema de PNTC, como por exemplo, a inclusão de uma nova fonte de dados no processo de AD, pretende-se mostrar que esta inclusão apresenta um impacto positivo na qualidade das previsões numéricas. Para isso a demanda por um sistema robusto de avaliação de produtos de PNTC é crescente. A atual implementação do SCAMTEC é proveniente de uma antiga demanda interna do GDAD em realizar avaliações robustas e unificadas entre diferentes modelos de PNTC que receberam o mesmo tipo de melhoria no processo de AD.

O SCAMTEC é composto por um *kernel* (um núcleo central), que corresponde a um conjunto de programas escritos em linguagem fortran 90 com a função de interpretar dados de diversos modelos de PNTC, fornecendo ao usuário resultados estatísticos em formatos pré-selecionados. Desenvolveu-se um módulo para o cálculo de métricas estatísticas básicas (Viés, Erro Quadrático Médio e Correlação de Anomalia), independente do kernel, sendo possível a implementação de novas e específicas metodologias de uma forma simples.

Um dos gargalos dos sistemas que integram resultados de diferentes fontes de dados e produtos é o acesso à informação no qual está presente uma grande quantidade de formatos de leitura e escrita de arquivos e conseqüentemente o uso de bibliotecas e rotinas computacionais. Em função disso, esse processo tem sido tratado adequadamente no SCAMTEC e uma metodologia bastante flexível está em implementação, a qual proporcionará que formatos diferenciados possam ser incluídos pelo usuário final. A versão inicial do sistema está preparada para ler dados binários, em suas diferentes versões, HDF4, NetCDF e GRIB1. Esse último é o formato padronizado pela OMM (Organização

Meteorológica Mundial) e é amplamente utilizado para o armazenamento de dados meteorológicos.

O SCAMTEC possui uma interface gráfica cujo objetivo é facilitar o acesso à ferramenta e suas configurações, a qual carrega em um arquivo, todas as decisões feitas pelo usuário com relação às opções disponíveis no sistema de avaliação, além de disparar o processo. Esta interface é escrita em linguagem wxPython (RAPPIN e DUNN, 2006) e permite ao usuário ajustar todas as opções possíveis para executar o sistema. Para a visualização dos resultados, além de tabelas convencionais, o sistema permite que os mesmos sejam apresentados em forma de gráficos. Para tal é utilizada a biblioteca Magics++, que é a versão mais recente do software de plotagem do ECMWF (*European Centre for Medium-Range Weather Forecasts*) - (<http://www.ecmwf.int/products/data/software/magics++.html>) (SIEMEN e LAMY-THEPAUT 2009). A Figura 1 apresenta: (a) a interface gráfica do SCAMTEC e (b) gráficos ilustrativos exemplificando algumas das formas disponíveis nesse sistema para a apresentação de resultados.

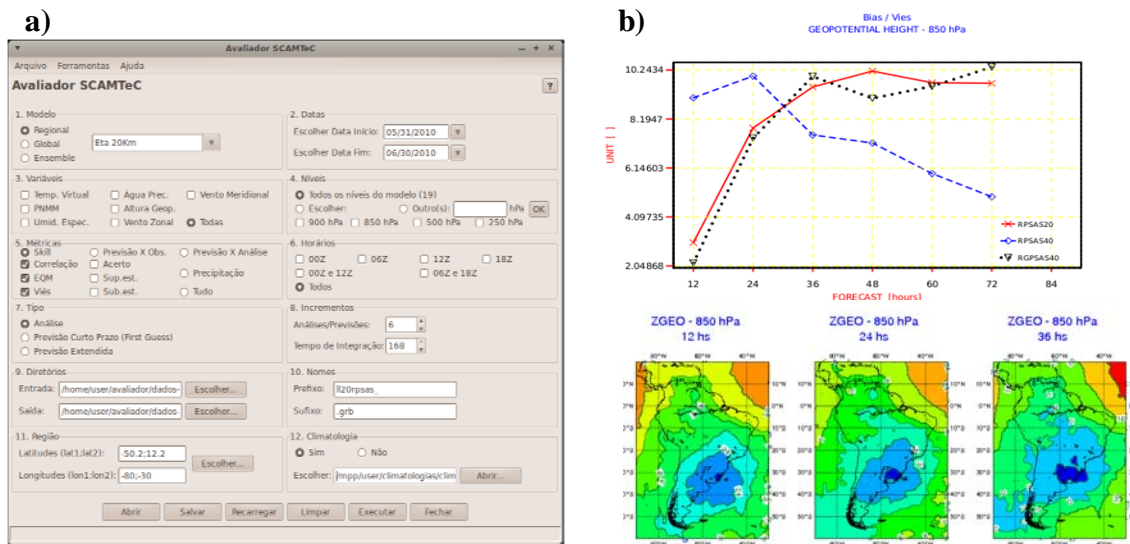


Figura 1: Exemplos das interfaces do SCAMTEC: (a) interface de configuração do sistema (modelos, datas, variáveis, níveis etc.) e (b) gráficos exemplificando algumas das formas disponíveis no sistema para a apresentação de resultados.

Métricas Estatísticas para avaliação da precipitação em implementação no SCAMTEC

A avaliação dos produtos de PNTC necessita de dois ingredientes fundamentais: dados observados e métricas estatísticas adequadas para cada tipo de variável analisada. Os produtos de previsão mais próximos dos valores observados, segundo tais métricas, podem ser considerados os melhores. Como a rede de observações meteorológicas não é adequadamente densa, essa comparação envolve interpolações espaciais que podem ser fontes adicionais de erro. No que diz respeito à precipitação, outras fontes de incertezas podem surgir no processo de comparação. Talvez a mais importante seja a qualidade e representabilidade espacial das próprias observações, em decorrência da distribuição inapropriada da rede de pluviômetros. Visando minimizar os possíveis ruídos que prejudiquem a avaliação, diversas metodologias de interpolação espacial e temporal (linear, polinomial, triangulação, *kriging*, entre outros) estão sendo implementadas no SCAMTEC. Sob o ponto de vista dos dados observacionais de precipitação, um avanço significativo foi a combinação de produtos de sensoriamento remoto com observações *in situ* (VILA et al. 2009). Essa metodologia realiza pequenos ajustes nos campos produzidos por satélites a fim de torná-los mais precisos e confiáveis, levando a possibilidade de utilização dos mesmos para a avaliação e validação de campos provenientes

de simulações numéricas.

A combinação de todos esses aspectos mostra que uma metodologia mais sofisticada deve estar disponível para a avaliação de precipitação. Além das métricas convencionais aplicáveis as demais variáveis, para a precipitação no SCAMTEC há diversas outras. Entende-se como métricas convencionais o Erro Médio para medida de tendência, Erro Médio Quadrático e Desvio Padrão para medidas de dispersão e o Coeficiente de Correlação de Anomalia, o qual é utilizado para avaliar o desempenho dos modelos de PNTC na representação das anomalias das análises baseadas em campos climatológicos. Para a precipitação as métricas em fase de estudos e implementação são brevemente descritas a seguir.

Tabela de Contingência: (WILKS, 1995) representação quantitativa e qualitativa de dados na qual os elementos da amostra ou da população são classificados de acordo com pelo menos dois fatores ou critérios, fornecendo vários parâmetros úteis para a avaliação de previsões, análises e observações. Várias avaliações ou *scores* podem ser feitas a partir de seus resultados, tais como: Porcentagem de Acerto (*Hit Rate*), *Critical Success Index (CSI)* ou *Thread Score (TS)*, Probabilidade de Detecção (POD) e Taxa de Alarme Falso (FAR), bem como o Viés.

Diagramas de interpretações de resultados envolvendo diversos modelos: TAYLOR (2001) e BOER e LAMBERT (2001) construíram diagramas (Diagrama de Taylor e Diagrama BLT, respectivamente) para sintetizar as várias informações obtidas por meio da tabela de contingência. Os índices utilizados neste diagrama são o POD, FAR, Viés e CSI. Tais diagramas são úteis para a visualização e a avaliação de múltiplos aspectos dos modelos (mudanças na resolução, nas parametrizações ou nas condições iniciais) ou mesmo medir a habilidade relativa entre diferentes modelos.

Receiver Operating Characteristic (ROC): este índice é representado em um gráfico que combina o *Hit-Rate* (eixo Y) e a FAR (eixo X) para diferentes limiares de decisão, a fim de fornecer uma visão clara do desempenho do classificador. Para um dado sistema classificador, uma curva característica da probabilidade de detecção p_d pode ser plotada como uma função da taxa de alarme falso p_f , dado um limiar de classificação. Um classificador produzirá resultados que têm altos valores de p_d e baixos valores de p_f , ou seja, eles se aproximam do topo esquerdo da curva do ROC. A localização da curva inteira em unidade quadrada é determinada pela capacidade de discriminação intrínseca do sistema de previsão, e a localização de pontos específicos em uma curva é fixada pelo limiar de decisão em que o sistema está operando.

Rank Correlation (RC): Associação monotônica linear e não linear entre previsões e observações podem ser medidas utilizando estatísticas baseadas em classificadores de dados (posição de valores quando arranjados em ordem crescente). Um *Rank Correlation* significa que os valores de previsão são funções crescentes das observações – há uma perfeita associação monotônica entre as previsões e as observações. O coeficiente de *Rank Correlation* é também mais resistente (menos sensível) a grandes valores discrepantes. Há diferentes métodos de *Rank Correlation*, e os mais conhecidos são os de *Spearman* e *Kendal*. Mais detalhes a respeito deste método pode ser encontrado em WILKS (1995), e JOLLIFFE e STEPHENSON (2003).

Considerações finais

O presente trabalho discute os problemas envolvidos para uma adequada avaliação das previsões de precipitação produzidas por modelos de PNTC e apresenta as potencialidades do sistema denominado SCAMTEC nessa tarefa. Nos últimos anos, ferramentas desenvolvidas pelo INPE inicialmente para atender a uma demanda interna foram posteriormente disponibilizadas para outras instituições interessadas, como é o caso do SPRING (Sistema de PROcessamento de INformações Georeferenciadas), o Terra LIB e

outros. O mesmo deve ocorrer com o SCAMTEC, com o objetivo de contribuir com os desenvolvedores de novos produtos envolvendo modelos de PNTC e os usuários desses produtos. As atividades ligadas à agrometeorologia são áreas bastante promissoras para a sua aplicação, em especial na avaliação e validação de produtos de precipitação, buscando uma melhoria dos mesmos, e acima de tudo possibilitando uma intercomparação mais justa entre sistema de diferentes centros de pesquisa. Esta ferramenta pode ainda ser útil no meio acadêmico, ao apresentar de forma fácil e prática o desempenho dos modelos atualmente disponíveis, bem como nas atividades produtivas do setor agrícola, avaliando adequadamente os riscos de decisões baseadas em produtos de modelos de PNTC.

Agradecimentos.

Os autores agradecem o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo apoio financeiro à pesquisa (Processo 481379/2010-0).

Referências bibliográficas

- ANTHES, R. A.; KUO, Y. H.; HSIE, E. H.; LOW-NAN, S.; BETTGE, T. W.; **Estimation of skill and uncertainty in regional numerical models.** Quarterly Journal of the Royal Meteorological Society, v. 115, n. 488A, p. 763–806, July 1989.
- BOER, G. J., S. J. LAMBERT, 2001: **Second order space-time climate difference statistics.** Climate Dyn., 17, 213–218.
- JOLLIFFE, I. T. e STEPHENSON, D. B.; **Forecast Verification: A Practitioner's Guide in Atmospheric Science.** John Wiley & Sons, Ltd. ISBN: 0-47-49759-2. 2003.
- DE MATTOS, J. G. Z, SAPUCCI L. F., BASTRAZ C., DE GONÇALVES G. G., SANTOS A. F. **SCAMTEC: Sistema de Avaliação de Modelos Numérico de Previsão de Tempo e Clima.** Boletim da Sociedade Brasileira de Meteorologia. Em submissão. 2011.
- NACHAMKIN, J. E., CHEN, S. SCHMIDT, J. **Evaluation of Heavy Precipitation Forecasts Using Composite-Based Methods: A Distributions-Oriented Approach.** Monthly Weather Review, 133, 2163-2177, 2005.
- RAPPIN, N. e DUNN, R.; **wxPython in Action.** Manning Publications Co., 620p. ISBN: 1-932394-62-1. 2006.
- SAPUCCI, L. F. ; HERDIES, D. L.; SOUZA, R. V. A.; MATTOS, J. G. F.; ARAVEQUIA, J. A. **Os últimos avanços na previsibilidade dos campos de umidade no sistema global de assimilação de dados e previsão numérica de tempo do CPTEC/INPE.** Revista Brasileira de Meteorologia v. 25, p. 100, 2010.
- SIEMEN, S. E; LAMY-THÉPAUT, S.; **Magics++ 2.8 - New developments in ECMWF's meteorological graphics library.** ECMWF Newsletter No. 122, 2009/10.
- TAYLOR, K.E.: **Summarizing multiple aspects of model performance in a single diagram.** J.Geophys. Res., 106, 7183-7192, 2001 (also see PCMDI Report 55, <http://www.pcmdi.llnl.gov/publications/ab55.html>)
- VILA, D. A.; DE GONCALVES, L. G. G.; TOLL, D. L.; ROZANTE, J. R. **Statistical Evaluation of Combined Daily Gauge Observations and Rainfall Satellite Estimates over Continental South America.** Journal of Hydrometeorology, v. 10, p. 533, 2009.
- WILKS, D. S.; **Statistical methods in the atmospheric sciences: an introduction.** First edition. San Diego: Academic Press, 465 p. (International Geophysics Series, v. 59), 1995.