

## AVALIAÇÃO DE ESTIMATIVAS DE CAMPOS DE PRECIPITAÇÃO PARA MODELAGEM HIDROLÓGICA DISTRIBUÍDA

ADRIANO ROLIM DA PAZ<sup>1</sup> E WALTER COLLISCHONN<sup>2</sup>

<sup>1</sup>Depto de Engenharia Civil e Ambiental, Universidade Federal da Paraíba, João Pessoa, PB, Brasil

<sup>2</sup>Instituto de Pesquisas Hidráulicas, Universidade Federal do Rio Grande do Sul,  
Porto Alegre, RS, Brasil

adrianorpaz@yahoo.com.br, collischonn@iph.ufrgs.br

Recebido Novembro 2008 – Aceito Julho 2010

### RESUMO

É crescente a disponibilidade e utilização de campos de chuva estimados por sensoriamento remoto ou calculados por modelos de circulação da atmosfera, os quais são frequentemente utilizados como entrada para modelos hidrológicos distribuídos. A distribuição espacial dos campos de chuva estimados é altamente relevante e deve ser avaliada frente aos campos de chuva observados. Este artigo propõe um método de comparação espaço-temporal entre campos de chuva observados e estimados baseado na comparação pixel a pixel e na construção de tabelas de contingência. Duas abordagens são utilizadas: (i) a análise integrada no espaço gera índices de performance que retratam a qualidade do campo de chuva estimada em reproduzir a ocorrência de chuva observada ao longo do tempo; (ii) a análise integrada no tempo produz mapas dos índices de performance que resumem a destreza das estimativas de ocorrência de chuva em cada pixel. Como exemplo de aplicação, é analisada a chuva estimada na climatologia do modelo global de circulação da atmosfera CPTEC/COLA sobre a bacia do Rio Grande. Utilizando-se cinco índices de performance, o método proposto permitiu identificar variações sazonais e padrões espaciais na performance das estimativas de chuva em relação a campos de chuva derivados de observações em pluviômetros.

**Palavras-Chave:** Índice de Performance, Estimativa de Precipitação, Tabela de Contingência, Modelo Hidrológico

### ABSTRACT: ASSESSMENT OF ESTIMATED PRECIPITATION FIELDS FOR DISTRIBUTED HYDROLOGIC MODELING

There is an increasing availability and application of precipitation fields estimated by remote sensing or calculated by atmospheric circulation models, which are frequently used as input for distributed hydrological models. The spatial distribution of the estimated precipitation fields is extremely important and must be verified against observed precipitation fields. This paper proposes a method for spatiotemporal comparison between observed and estimated precipitation fields based on a pixel by pixel comparison and on contingency tables. Two distinct approaches are carried out: (i) the spatial integrated analysis produces skill scores denoting the ability of the estimated precipitation field in reproducing the occurrence of observed precipitation along the time; (ii) the time integrated analysis generates maps of skill scores showing the reliability of the precipitation estimates in each pixel. As an example of application, the estimated precipitation climatology of the CPTEC/COLA global circulation model over the Rio Grande basin is assessed. Using five different skill scores, the proposed method identified seasonal variations and spatial patterns in the performance of the estimated precipitation fields in relation to precipitation fields derived from pluviometers measurement.

**Keywords:** skill score, rainfall estimate, contingency table, hydrologic model

## 1. INTRODUÇÃO

Atualmente, há uma tendência crescente do emprego de sistemas sensores remotos, como radares e satélites para estimativa de precipitação (por exemplo: Scofield, 1987; Vicente et al., 1998; Kummerow et al., 2005; Zinevich et al., 2008). Embora os dados provenientes de sensores remotos sejam estimativas e estejam sujeitos a erros de magnitude variável conforme o sensor, a região e diversos outros fatores, o uso dessas estimativas tem duas grandes vantagens em relação ao uso de medição via redes de pluviômetros e pluviógrafos: (i) maior velocidade de obtenção e facilidade de disponibilização da informação, principalmente visando o uso em sistemas de alerta e controle de enchentes; (ii) produção da informação na forma de um campo espacial de precipitação que abrange grandes áreas.

Paralelamente, também é crescente o uso de modelos numéricos de circulação da atmosfera para simular o clima em escalas global e regional (Shaffrey et al., 2009; Hasler et al., 2009; Lowrey e Yang, 2008; Sampaio et al., 2007; Cavalcanti et al., 2002; Avissar e Verstraete, 1990; Xu, 1999; dentre centenas de referências). Estudos dos fenômenos meteorológicos em geral, variabilidade climática, previsão de tempo e clima, previsão de vazões, projeções de mudanças climáticas e de cenários de intervenções antrópicas são algumas das finalidades do uso de modelos atmosféricos, em alguns casos acoplados com modelos hidrológicos. Um dos principais produtos de tais modelos são os campos de chuva calculados para uma grade de pontos ao nível da superfície. São realizadas previsões de chuva com horizontes de alguns dias até meses, de grande utilidade em áreas como previsão de safras agrícolas, previsão e controle de enchentes e planejamento da operação de reservatórios hidroelétricos, entre outras.

Dentre as variáveis meteorológicas de interesse para estudos climáticos e hidrológicos, a precipitação é a que apresenta a maior variabilidade espaço-temporal. Tanto as estimativas de chuva por sensoriamento remoto, quanto aquelas calculadas pelos modelos de circulação atmosférica precisam ser avaliadas quanto à destreza em corretamente reproduzir a chuva observada. Erros sistemáticos detectados ou falhas grosseiras, por exemplo, podem ser corrigidos antes de usar esses campos de chuva calculados para alimentar um modelo hidrológico (Hay e Clark, 2003).

Basicamente, estimativas espaciais de chuva realizadas ao longo do tempo podem apresentar três tipos de erros (Habets et al., 2004): (i) localização espacial do evento chuvoso; (ii) intensidade da chuva; (iii) instante de tempo da ocorrência do evento. Os dois primeiros tipos de erros podem ser avaliados por comparação entre campos de precipitação observada e estimada referentes a um mesmo instante de tempo. O terceiro tipo de erro só pode ser inferido diretamente analisando os campos de

precipitação em uma janela de tempo, o que adiciona mais um grau de complexidade à análise e não é tratado neste artigo.

É fundamental avaliar os campos de chuva estimados levando em consideração sua distribuição espacial, principalmente no caso de serem usados como forçantes de entrada em modelos hidrológicos distribuídos (Benoit et al., 2000; Collischonn et al., 2005; Habets et al., 2004; Kite e Haberlandt, 1999). Nesse tipo de modelos hidrológicos, a bacia hidrográfica é discretizada em uma grade regular interconectada por uma rede de drenagem (Singh e Frevert, 2002), podendo levar em conta a variabilidade espacial da precipitação. Uma das entradas desse tipo de modelo é, portanto, o campo de chuva sobre a bacia, a cada passo de tempo de cálculo. Em determinados casos, como em um sistema de alerta em tempo real, a verificação dos campos de precipitação idealmente deve ser feita de forma automática e operacional previamente à entrada no modelo hidrológico (Feidas et al., 2007; Moore et al., 2005).

Uma forma usual de quantificar o grau de acerto ou concordância entre os valores estimados ou previstos de chuva e os valores observados é o cálculo de índices de performance (*skill scores*). Tratando de forma binária as estimativas e observações de chuva, isto é, transformando os dados em ocorrência ou não de chuva, pode-se construir uma tabela de contingência e calcular índices como Probabilidade de Detecção (*hit rate*), Taxa de Alarme Falso (*false alarm ratio*), Índice de Sucesso Crítico (*critical success index* ou *threat score*), etc (Wilks, 2006; Kuligowski, 2002). Comumente, a análise é feita considerando valores médios (estimados e observados) da precipitação sobre a área de interesse (Ebert et al., 2003), como uma bacia hidrográfica por exemplo. Esse procedimento simplificado não avalia a estimativa do padrão espacial da chuva, e é suficiente apenas quando o interesse é na chuva média sobre pequenas regiões. Outra forma de análise consiste em comparar direta e pontualmente a chuva estimada com a observada em pluviômetros. Contudo, como os campos de chuva estimados por modelos numéricos ou por sensores remotos são constituídos por valores médios de precipitação sobre a área de cada elemento da grade, a comparação pontual com dados de pluviômetros é prejudicada pela própria variabilidade espacial da precipitação sobre essa área (Barrera, 2005).

Este artigo apresenta um método prático de análise de séries temporais de estimativas de campos de chuva, em relação a séries temporais de campos de chuva tomados como referência. A abordagem empregada visa quantificar a qualidade das estimativas de chuva utilizadas como entrada em modelos hidrológicos distribuídos, enfocando tanto de forma espacial sobre uma determinada área, quanto ao longo do tempo. Como exemplo de aplicação, são apresentados resultados da avaliação dos campos de chuva referentes à Climatologia do Modelo de

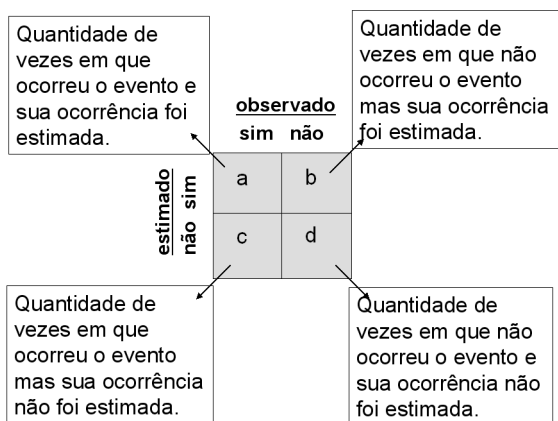
Circulação Global da Atmosfera CPTEC/COLA, operado pelo Centro de Previsão de Tempo e Estudos Climáticos (CPTEC) do Instituto Nacional de Pesquisas Espaciais (INPE). O estudo de caso é a bacia do Rio Grande, que tem uma área de drenagem de 145.000 km<sup>2</sup> e está localizada entre os Estados de São Paulo e Minas Gerais.

## 2. METODOLOGIA

Seja uma série temporal de campos de chuva observada e uma série temporal de campos de chuva estimada por um método qualquer, como sensoriamento remoto ou modelagem atmosférica. Descreve-se a seguir o procedimento de comparação pixel a pixel entre pares de campos de chuva observada e estimada referentes a um mesmo instante de tempo, e como a performance das estimativas podem ser avaliadas a partir de tais comparações.

### 2.1 Índices de performance

Tratando de forma binária a ocorrência/não ocorrência de um determinado evento, bem como a estimativa de sua ocorrência ou não ocorrência, pode-se construir uma tabela de contingência da forma ilustrada na Figura 1. As grandezas “a” e “d” são a quantidade de acertos na estimativa de ocorrência e de não ocorrência do evento, respectivamente. A quantidade “b” denota o número de vezes em que foi estimada a ocorrência do evento, mas ele não aconteceu. Analogamente, o valor “c” é a quantidade de vezes em que a ocorrência do evento foi observada, mas sua ocorrência não foi estimada. A partir dos valores da tabela de contingência, diversos índices podem ser deduzidos com enfoques diferentes quanto à avaliação da performance das estimativas realizadas de ocorrência/não ocorrência do evento (Wilks, 2006; Stephenson, 2000).



**Figura 1** – Esquema da tabela de contingência com análise comparativa dos acertos e erros nas estimativas da ocorrência (“sim”) e não ocorrência (“não”) de um determinado evento.

Um aspecto interessante do uso da tabela de contingência é permitir analisar a performance da estimativa sob diferentes aspectos, conforme o tipo de evento estudado, o tipo de estimativa realizada e o propósito do uso dessa informação. Por exemplo, considerando a previsão de ocorrência de eventos extremos como tornados ou chuvas muito intensas, a previsão da não ocorrência do evento é muito mais fácil do que a previsão da ocorrência do evento. Nesse caso, assumir um mesmo peso para os acertos do tipo “a” e do tipo “d” não seria adequado para uma verificação da performance do estimador. Nesse sentido, alguns índices permitem um enfoque mais restrito, como o percentual de acerto em conseguir detectar a ocorrência do evento dado que ele ocorreu (calculado pela expressão  $a/(a+b+c)$ ), desprezando acertos da não ocorrência (“d”). Por isso é importante empregar vários índices de performance para garantir uma análise integral das estimativas sob diferentes aspectos (Cherubini et al., 2002).

Neste estudo, foram utilizados os índices de performance denominados proporção correta (PC), probabilidade de detecção (POD), taxa de alarme falso (FAR), índice de sucesso crítico (CSI) e o índice BIAS, cujas formulações, significados e variações de valores esperados são apresentados na Tabela 1. Diversos outros índices são propostos e utilizados na literatura (Wilks, 2006; Kuligowski, 2002; Stephenson, 2000; Mesinger, 2008) e poderiam também ser utilizados dentro da metodologia apresentada neste artigo. Independente do índice de performance utilizado, ressalta-se a importância de entender sua formulação para poder avaliar mais coerentemente os resultados (Hamill e Juras, 2006).

### 2.2 Tipos de evento para análise

A avaliação da performance pode ser feita sob dois aspectos distintos: (i) a ocorrência em si da chuva e (ii) a

**Tabela 1** – Relação dos índices de performance derivados da tabela de contingência utilizados neste estudo.

Índice	Formulação	Significado	Valor
PC (proporção correta)	$PC = \frac{a+d}{n}$ , onde $n = a+b+c+d$	Percentual de acertos geral, sem distinção entre acertos da ocorrência ou não ocorrência do evento.	Varia de 0 a 1; Quanto maior o valor, melhor a performance.
POD (probabilidade de detecção)	$POD = \frac{a}{a+c}$	Dado que o evento ocorreu, percentual de acertos em estimar sua ocorrência.	Varia de 0 a 1; quanto maior o valor, melhor a performance.
FAR (taxa de alarme falso)	$FAR = \frac{b}{a+b}$	Dentre as vezes em que foi estimada a ocorrência do evento, percentual em que o evento não ocorreu.	Varia de 0 a 1; quanto menor o valor, melhor a performance.
CSI (índice de sucesso crítico)	$CSI = \frac{a}{a+b+c}$	Percentual de acertos nas estimativas, descontando as vezes em que a não ocorrência do evento foi corretamente prevista.	Varia de 0 a 1; quanto maior o valor, melhor a performance.
BIAS (taxa de tendência)	$BIAS = \frac{a+b}{a+c}$	Relação entre o número de estimativas de ocorrência do evento e o número de eventos ocorridos.	Assume qualquer valor > 0; Quanto mais próximo de 1 melhor a performance; se > 1 indica superestimativa da ocorrência do evento; se < 1 indica subestimativa.

quantidade de chuva (Ebert et al., 2003; Mearns et al., 1995; McBride e Ebert, 2000). A motivação de trabalhar com as duas abordagens é resultante da forma como as estimativas numéricas de chuva são produzidas. Em algoritmos usados para estimativa de chuva por sensoriamento remoto, por exemplo, em geral parte do algoritmo decide sobre a ocorrência ou não de chuva e outra parte estima quantitativamente a chuva.

No primeiro caso (i), tem-se uma análise do tipo chuva/não chuva, usada para verificar a habilidade do estimador quanto à distinção da ocorrência ou não de chuva. O evento a ser estimado é a ocorrência de chuva no intervalo de tempo considerado, não importando em qual intensidade. Comumente, por consequência da formulação dos algoritmos de interpolação e de estimativa da chuva, há dificuldade das estimativas numéricas em gerar valores exatamente nulos de chuva. Esses valores são gerados frequentemente muito próximos de zero, mas não nulos. Para contornar isso, considera-se um limiar ( $P_{lim}$ ) não nulo para diferenciar as classes chuva e não chuva, ou seja: se  $P \geq P_{lim}$ , ocorreu o evento; e se  $P < P_{lim}$ , não ocorreu o evento.

A segunda abordagem (ii) visa inferir sobre a destreza em estimar a quantidade de chuva ocorrida. São tomados apenas os casos em que houve acerto na estimativa da ocorrência da chuva (quando há um acerto tipo “a” na tabela de contingência), sendo então avaliado se houve acerto ou não em estimar a magnitude da chuva. Nesse tipo de análise, o evento é caracterizado pela ocorrência de chuva acima de um determinado limiar ( $P_{max}$ ). Por exemplo, tomando um limiar de 10 mm, o evento

é considerado caso a chuva tenha sido igual ou superior a 10 mm ( $P \geq P_{max}$ ), caso contrário considera-se a não ocorrência do evento ( $P < P_{max}$ ). Variando-se o valor do limiar (de 5 em 5 mm, por exemplo), obtém-se uma indicação da destreza em estimar quantitativamente a chuva (Feidas et al, 2007; Kuligowski, 2002; Cherubini et al., 2002), dado que já aconteceu o acerto em estimar a ocorrência da chuva.

No exemplo de aplicação apresentado neste trabalho, apenas a abordagem de distinguir eventos chuvosos/não chuvosos foi empregada, tomando o valor de  $P_{lim} = 1,0$  mm como ponto de corte (McBride e Ebert, 2000). Entretanto, a inferência da estimativa quantitativa de chuva poderia ser também utilizada com o mesmo procedimento de comparação pixel a pixel entre campos de chuva observados e estimados.

### 2.3 Comparação entre campos de chuva

A comparação entre as séries temporais dos campos de chuva observada e estimada foi realizada tomando pixel a pixel as duas imagens referentes a cada instante de tempo, em abordagem semelhante à empregada por Feidas et al. (2007). Em um determinado instante  $t$ , a ocorrência ou não do evento em um pixel  $i$  é verificada pela análise do valor de chuva observado nesse pixel. Analogamente, toma-se a chuva estimada em tal pixel para definir se o evento foi estimado ou não para ocorrer nesse pixel. Comparando as duas verificações, tem-se que ocorreu um acerto (tipo “a” ou “d”) ou um erro (tipo “b” ou “c”) para o pixel  $i$  no instante  $t$ . Esse procedimento de comparação

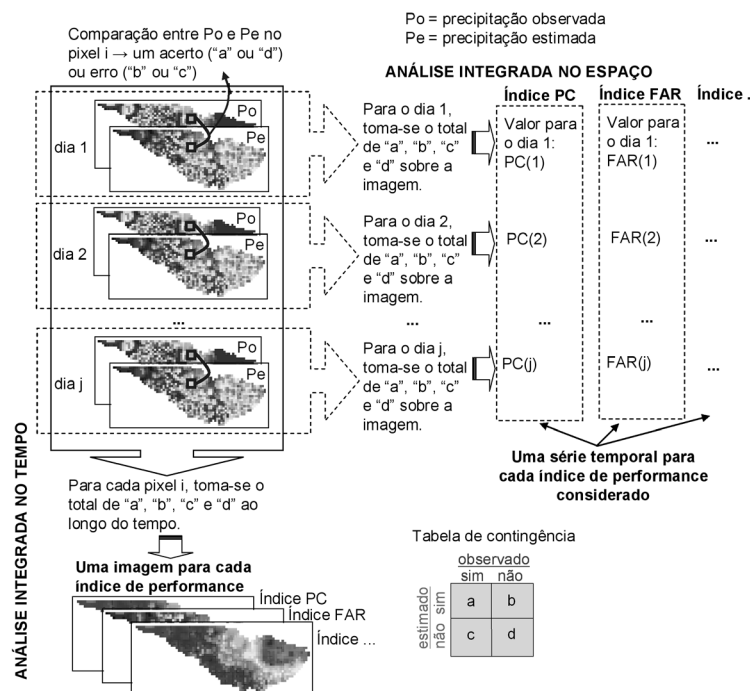


Figura 2 – Esquema da metodologia de comparação pixel a pixel entre campos de chuva observada e estimada.

é repetido para todos os pixels da imagem nesse instante de tempo  $t$ , e em seguida para todas as imagens ao longo do tempo. Para montar a tabela de contingência e determinar os índices de performance, duas abordagens são adotadas (Figura 2): (1) integração no espaço e (2) integração no tempo. Os itens a seguir detalham as duas formas de abordagem.

### 2.3.1 Análise integrada no espaço

Nessa abordagem, para cada instante de tempo  $t$  da série, o número de acertos dos tipos “a” ou “d” e o número de erros dos tipos “b” ou “c” são contabilizados considerando todos os pixels da área considerada. Obtém-se uma tabela de contingência específica desse instante  $t$  (ou seja, uma determinada quantidade de “a”, “b”, “c” e “d”, sendo que  $a + b + c + d =$  número de pixels da imagem).

A partir dessa tabela, são derivados os índices de performance correspondentes que resumem a performance geral da estimativa do campo de precipitação naquele instante de tempo  $t$ .

Repetindo-se o procedimento para todos os intervalos de tempo, tem-se uma série temporal de valores de cada índice de performance. Tais séries podem ser usadas para avaliar a performance das estimativas de campos de chuva quanto à existência de sazonalidade ou de mudança de comportamento ao longo do tempo.

### 2.3.2 Análise integrada no tempo

Na análise denominada integrada no tempo, o número de acertos (tipo “a” ou “d”) e erros (tipo “b” ou “c”) ao longo do tempo ocorrido para um pixel  $i$  é tomado para construção de uma tabela de contingência específica desse pixel. Nesse caso,  $a + b + c + d =$  quantidade de intervalos de tempo da série. Repetindo-se a análise para todos os pixels da imagem, gera-se uma tabela de contingência para cada pixel e obtém-se os índices de performance correspondentes. Dispondo espacialmente os

valores dos índices de performance obtidos para cada pixel, são geradas imagens que refletem o padrão espacial de desempenho do estimador.

## 3. APLICAÇÃO

### 3.1 Bacia do Rio Grande

O Rio Grande é o principal afluente da parte alta do Rio Paraná e apresenta uma área de drenagem de aproximadamente 145.000 km<sup>2</sup>, a qual se estende pelos Estados de Minas Gerais e São Paulo (Figura 3). A precipitação anual média sobre a bacia é de cerca de 1.400 mm, concentrada nos meses de Novembro a Abril. A bacia em questão tem grande importância no contexto da geração hidroeétrica brasileira, sendo responsável por aproximadamente 11,7% da produção nacional, com uma capacidade instalada em torno de 7.722 MW (ANEEL, 2005). Previsões de chuva e de vazão para a bacia do Rio Grande trazem, portanto, benefício em potencial para a operação dos reservatórios hidroeétricos e incremento da eficiência da geração de energia. Com essa motivação, foi realizado recentemente um estudo sobre previsão de vazões usando um modelo hidrológico distribuído com base em previsão quantitativa de chuva gerada por modelos atmosféricos global e regional (Paz et al., 2007; Tucci et al., 2008).

Neste estudo, dados diários de 273 postos pluviométricos distribuídos por toda a bacia do Rio Grande foram utilizados para caracterizar a chuva observada. Os dados são provenientes da base de dados HidroWeb/ANA e do SIGRHSP/DAEE, disponíveis na Internet, e os postos selecionados são os mesmos utilizados para compor a chuva observada no estudo de modelagem hidrológica anteriormente citado.

### 3.2 Climatologia do modelo global CPTEC/COLA

No Brasil, o modelo CPTEC/COLA tem sido usado operacionalmente na previsão de tempo e simulações climáticas

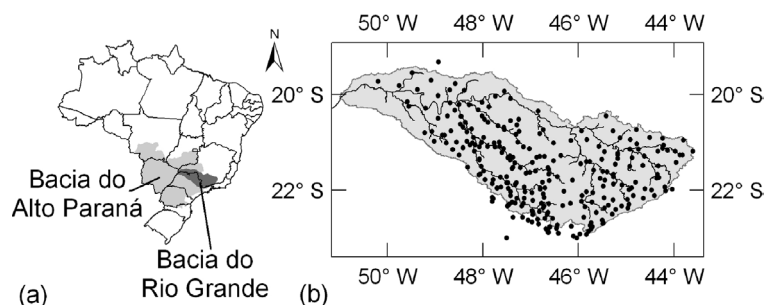


Figura 3 – Localização da bacia do Rio Grande e dos postos pluviométricos utilizados neste estudo.



(Cavalcanti et al., 2001; Cavalcanti et al., 2002; Marengo et al., 2003). Rodadas de modelos atmosféricos para longos períodos de tempo passado são usadas para produzir a climatologia do modelo, na qual se espera que o modelo represente satisfatoriamente o clima observado e sua variabilidade ao longo do tempo (Cavalcanti et al., 2002). Um importante uso da climatologia de modelos atmosféricos é a identificação de erros sistemáticos na estimativa de precipitação sobre determinadas áreas, os quais podem então ser removidos das previsões de precipitação produzidas pelo mesmo modelo, antes de servir como entrada para realizar previsões de vazão (Hay e Clark, 2003; Tucci et al., 2004; Tucci et al., 2008).

Neste trabalho, foi utilizada a precipitação da climatologia do modelo de circulação atmosférica global CPTEC/COLA, referente a um período de 51 anos (1jan1951 a 31dez2001), disponibilizada em uma grade de resolução  $1,875^\circ$  (Figura 4), e com um conjunto (*ensemble*) de 9 membros. O uso da metodologia de previsão por conjunto em simulações atmosféricas torna-se necessário devido ao comportamento caótico da atmosfera, ou seja, da sensibilidade do sistema às condições iniciais e devido aos erros do próprio modelo. Cada membro do conjunto nada mais é do que uma rodada distinta do modelo com as condições iniciais alteradas. Para compor o conjunto, as condições iniciais de simulação são alteradas e o modelo é rodado, fornecendo ao final um conjunto de resultados. Idealmente, espera-se que as várias respostas do modelo tenham um grau de dispersão adequado, nem pequeno o suficiente para não englobar a realidade (dados observados), nem grande o bastante para caracterizar uma alta dispersão e diminuir a confiabilidade da resposta do modelo (Wilks, 2006).

Obviamente, uma avaliação completa da qualidade das estimativas de precipitação da climatologia do modelo CPTEC/COLA, deveria levar em conta o conjunto de rodadas e o grau de dispersão dos resultados entre elas. Isso se torna ainda mais importante considerando-se que a área em estudo está localizada em uma região onde o *skill* do modelo é menor

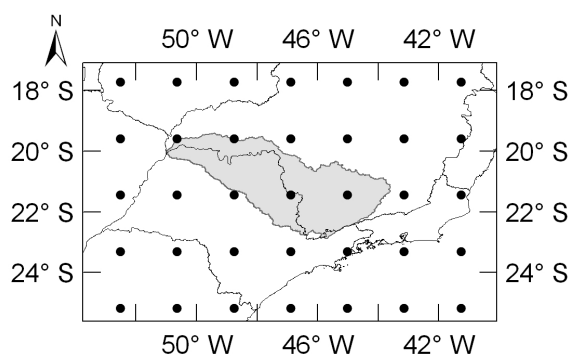
para previsão sazonal (Cavalcanti et al., 2002). Entretanto, o enfoque deste estudo não é traçar um diagnóstico da climatologia do modelo CPTEC/COLA, mas sim o de mostrar a viabilidade e o funcionamento do método proposto através da comparação de campos de chuva estimados e observados. A climatologia do modelo CPTEC/COLA é utilizada apenas como fonte de estimativas de campos de chuva para ilustrar o método e os tipos de análises que podem ser derivadas. Logo, simplificada mente foi tomado apenas um dos membros do conjunto, isto é, as estimativas de chuva geradas apenas em uma das rodadas do conjunto.

### 3.3 Campos de chuva observada e estimada

Para proceder à análise comparativa pixel a pixel, foram gerados campos de chuva observada e estimada em uma grade de  $0,1^\circ \times 0,1^\circ$  e com passo de tempo diário. A discretização da grade segue a discretização adotada na modelagem hidrológica da bacia do Rio Grande com o modelo MGB-IPH, desenvolvida para o estudo de previsão de vazões a partir de previsão quantitativa de chuva (Paz et al., 2007; Tucci et al., 2008). Apenas as células da grade que definem a bacia foram consideradas na análise dos campos de chuva (total de 1.248 células).

O esquema normalmente adotado para interpolação e preparação dos dados de chuva para entrada no modelo MGB-IPH foi utilizado para gerar os campos de chuva observados a partir dos dados de pluviômetros (Collischonn et al., 2007). Em tal esquema, a chuva em cada ponto da grade é interpolada pelo método do inverso do quadrado da distância, tomando os postos de chuva situados em um raio de até cinco vezes a distância entre o ponto e o posto mais próximo com dado sem falha naquele instante de tempo. Procedimento análogo foi utilizado para converter as estimativas diárias de precipitação da climatologia do modelo global da grade original desse modelo para a grade de  $0,1^\circ \times 0,1^\circ$ . Foram gerados campos de chuva observada e estimada com passo de tempo diário para o período de 1jan1951 a 31dez2001, ou seja, duas séries temporais com 18.628 campos de precipitação cada uma.

Sabe-se que as medições observadas em pluviômetros têm forte influência local e não capturam completamente a variabilidade espacial da chuva (Barrera, 2005). Logo, a geração de campos de chuva a partir de uma rede de pluviômetros não retrata exatamente o campo de precipitação ocorrido, e tais campos tendem a se tornar mais distintos com a diminuição da densidade de pluviômetros. Sistemas sensores remotos como satélites geoestacionários produzem campos de precipitação observada que idealmente devem ser utilizados, sozinhos ou em conjunto com dados de pluviômetros, para avaliar estimativas espaciais de chuva geradas por modelos numéricos de circulação



**Figura 4** – Grade do modelo global CPTEC/COLA sobre a bacia do Rio Grande.

da atmosfera (Feidas et al., 2007). Entretanto, como neste artigo o foco é o método de comparação entre pares de campos de precipitação, a simplificação de considerar os campos de chuva gerados por interpolação da rede de pluviômetros como os campos de chuva observada é aceitável e não prejudica o exemplo de aplicação do método.

#### 4. RESULTADOS E DISCUSSÃO

##### 4.1 Análise integrada no espaço

A análise integrada no espaço resulta em séries temporais dos índices de performance, onde cada valor da série de um

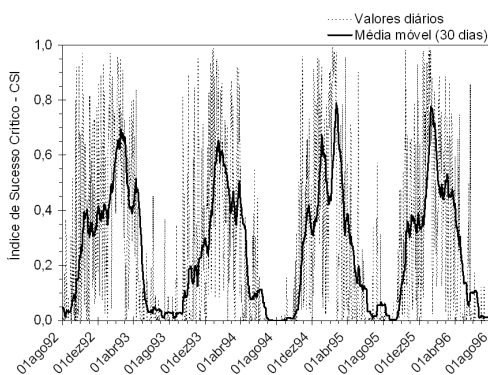


Figura 5 – Parte da série temporal do índice CSI obtida pela análise integrada no espaço.

determinado índice traduz de forma global o grau de acertos/erros do campo de chuva estimado para um instante de tempo. Por exemplo, na Figura 5 é apresentada parte da série temporal do índice CSI. Observa-se que a variação dos valores obtidos para esse índice ao longo do tempo segue um padrão sazonal. Há tendência de valores de CSI entre 0,4 e 0,8 para os dias do período chuvoso e de valores inferiores a 0,2 para o período mais seco. Esse índice representa o acerto de estimativas da ocorrência de chuva, desprezando o acerto de estimativas da não ocorrência. É mais fácil acertar a ocorrência de eventos chuvosos durante a estação úmida, do que os eventos isolados de chuva que ocorrem na estação seca.

O comportamento sazonal da performance das estimativas de chuva também fica claro nos resultados obtidos para os demais índices. Tomando os valores médios dos índices referentes a cada mês, nitidamente o padrão de desempenho do estimador é função do próprio padrão da ocorrência de chuvas (Figura 6). O índice PC representa a relação entre os acertos de ocorrência e não ocorrência do evento em relação ao número total de eventos ou pontos analisados, considerando os dois tipos de acertos sem ponderação diferenciada. A curva média obtida ao longo do ano para esse índice indica dois valores máximos, um no meio da estação seca ( $PC = 0,73$ ) e outro no meio da estação úmida ( $PC = 0,60$ ) (Figura 6a). Tendo em vista os resultados dos demais índices, pode-se concluir que ocorre o máximo de acertos da não ocorrência de chuva no meio da estação seca e o máximo de acertos da ocorrência de chuva no meio da estação úmida.

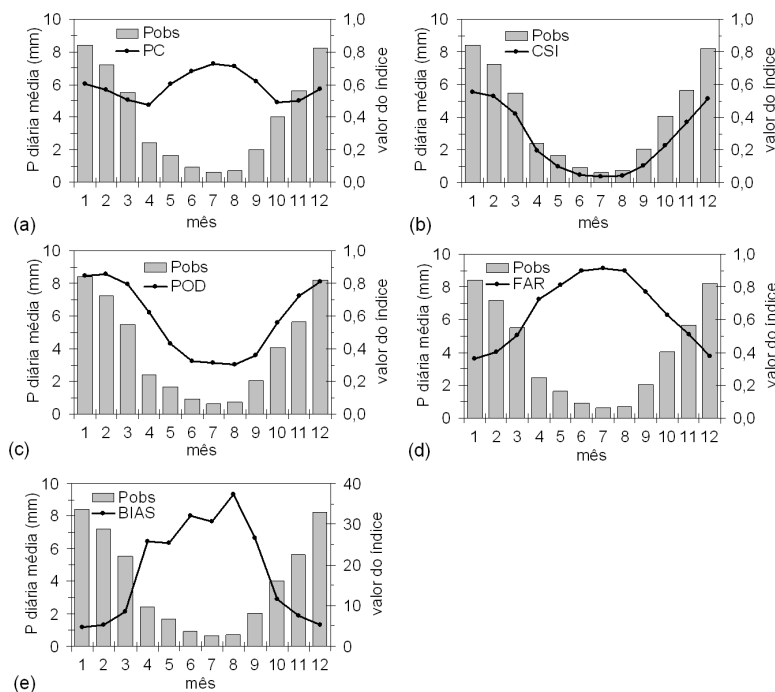
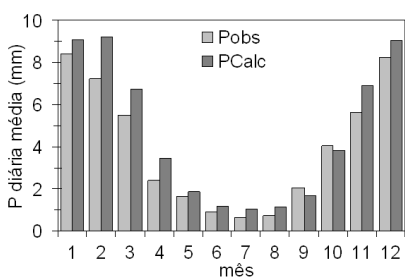


Figura 6 – Valores diários médios por mês da chuva observada e dos índices PC (a), CSI (b), POD (c), FAR (d) e BIAS (e), para a bacia do Rio Grande – índices calculados referentes à chuva estimada por uma das rodadas da climatologia do modelo CPTEC/COLA.

No período de transição entre as estações ocorrem os mínimos valores de PC (0,47), ou seja, menor quantidade de acertos. Durante o período chuvoso, o padrão segue aquele mencionado para o índice CSI em função da maior facilidade de acertar a ocorrência de chuva nesse período. Por outro lado, é na estação seca, e mais intensamente no meio dela, que existe maior facilidade em estimar a não ocorrência de chuvas.

O índice de performance POD (probabilidade de detecção) expressa a quantidade de acertos dos eventos que efetivamente ocorreram. Para o caso em análise, foi constatada uma maior facilidade de detecção de eventos no período chuvoso e menor no período seco (Figura 6c). Em fevereiro, por exemplo, houve em média 86% de acertos dos eventos diários ocorridos de chuva, enquanto apenas 30% dos eventos de chuva ocorridos em agosto, na média, foram estimados (ou detectados).

A maior facilidade de detecção da ocorrência de chuva no período chuvoso também induz a uma menor taxa de alarmes falsos (índice FAR), isto é, o número de vezes em que foi feita a estimativa de ocorrência de chuva e esta não ocorreu. A curva média sazonal dos valores de FAR variou de um mínimo de 0,36 em janeiro a um máximo de 0,92 em julho (Figura 6d). Em outras palavras, na média de janeiro em 36% das vezes foi estimada a ocorrência de chuva e ela não ocorreu, enquanto que esse percentual é de 92% em julho. Esse resultado, em conjunto com os valores obtidos para o índice BIAS, indica uma tendência de superestimativa da ocorrência de chuva para o membro considerado da climatologia do modelo global CPTEC/COLA. O índice BIAS varia de um valor mínimo 5 no meio do período chuvoso a um máximo de quase 40 em agosto (Figura 6e), indicando que essa tendência de superestimativa torna-se bastante elevada no período seco. Como consequência da superestimativa, são obtidos valores altos do acerto da ocorrência de chuva como explicitado pelo índice POD, mas ao custo de valores bastante elevados também para a taxa de alarmes falsos (FAR). Esse tipo de análise não poderia ser feito observando apenas, por exemplo, os valores médios de precipitação observada e estimada em cada mês (Figura 7), o que ressalta a importância de usar diversos índices que abrangem aspectos distintos de performance.



**Figura 7** – Chuva diária média ao longo do ano observada (Pobs) e estimada pela climatologia do modelo global CPTEC/COLA (Pcalc).

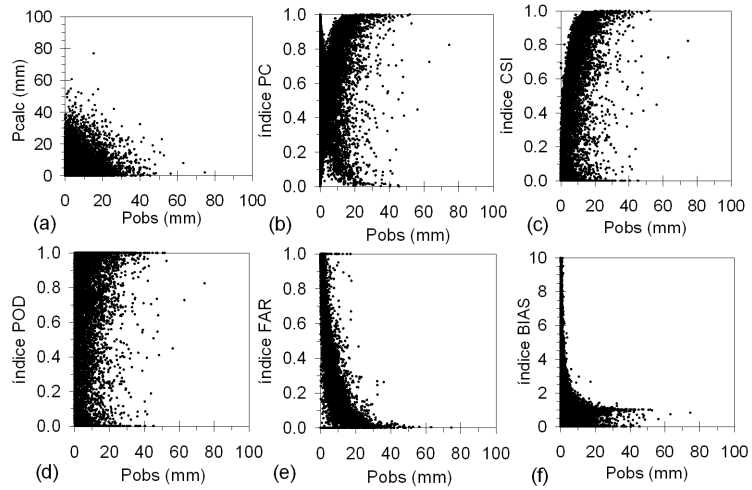
Ainda no escopo da análise integrada no espaço, é interessante questionar a performance das estimativas de ocorrência de chuva em função da intensidade da chuva observada. Traçando um gráfico dos valores estimados contra os valores observados de chuva diária, não se torna aparente nenhuma relação ou padrão (Figura 8a). Entretanto, os valores dos índices de performance obtidos indicam padrões de variação em função da intensidade da chuva diária observada, como ilustrado nas Figuras 8b a 8f. Nos gráficos dessas figuras, a chuva observada corresponde à média da chuva de todos os pixels da área de estudo em cada dia. Embora o valor do índice seja dado pela integração de acertos e erros das estimativas sobre todos os pixels (análise integrada no espaço), onde cada pixel tem valores específicos observado e previsto de chuva a cada dia, a consideração do valor médio da chuva sobre a bacia permite associar a performance espacial das estimativas em função da intensidade média do evento chuvoso observado.

Os resultados do índice PC, que mede o grau de acertos tanto da ocorrência quanto da não ocorrência de chuva, indicam duas tendências opostas (Figura 8b): uma de aumentar o valor do índice com o aumento da chuva observada (pontos alinhados segundo uma curva ascendente) e outra de diminuir o valor do índice com o aumento da chuva observada (pontos alinhados segundo uma curva descendente). Considerando os valores de PC alcançados para cada mês separadamente (Figura 9), constata-se que a tendência de aumentar o valor do índice com o aumento da chuva observada corresponde principalmente ao período chuvoso. Por outro lado, na estação seca, os valores elevados de PC ocorrem predominantemente pelo acerto da não ocorrência de chuva, sendo difícil a estimativa das chuvas esporádicas que acontecem no período. Dessa forma, nesse período há a tendência de diminuir o valor de PC com o aumento da intensidade da chuva observada – esse padrão pode ser observado para os meses de junho a agosto. Nos meses de transição entre estações seca e chuvosa, como setembro ou maio, as duas tendências estão presentes em menor intensidade.

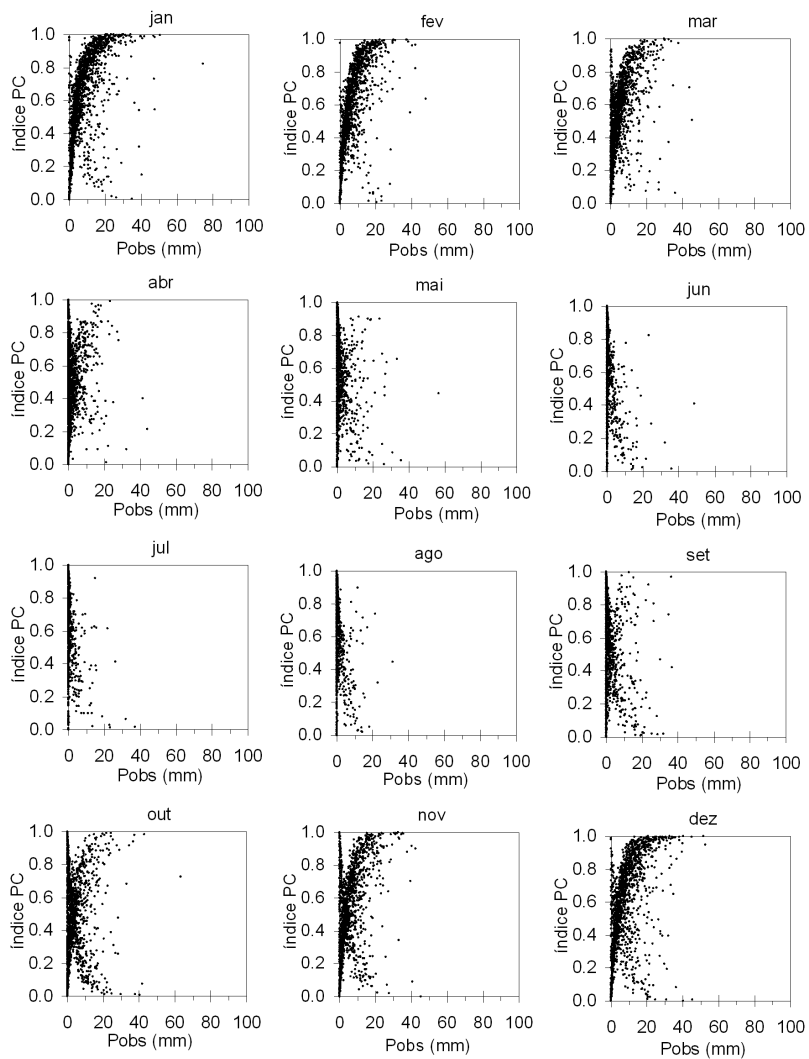
Ao contrário do índice PC, o índice de sucesso crítico (CSI) avalia apenas os acertos da ocorrência de chuva e, logo, foi obtido um padrão de performance único que é o de aumentar o valor desse índice para as maiores chuvas (Figura 8c). Padrão de desempenho análogo ao do índice CSI foi obtido para o índice POD (Figura 8d). Contudo, enquanto o índice POD avalia apenas os acertos de ocorrência de chuva nos dias em que efetivamente ocorreu chuva, o índice CSI leva em conta também as estimativas de chuva que não se concretizaram. Logo, o acréscimo da performance com o aumento da intensidade da chuva é mais forte para o índice POD do que para o índice CSI.

Quanto à taxa de alarme falsos (FAR), há a tendência de diminuir com o aumento da chuva ocorrida, já que pode ser considerada mais difícil a estimativa das chuvas de menor





**Figura 8** – Chuva média diária observada sobre a bacia (Pobs) versus: (a) chuva média diária calculada (Pcalc); (b),(c),(d),(e),(f) índices de performance diários PC, CSI, POD, FAR e BIAS respectivamente.



**Figura 9** – Chuva média diária observada sobre a bacia (Pobs) versus índice de performance diário PC, para cada mês.

intensidade (Figura 8e). O índice BIAS, por sua vez, aumenta consideravelmente com a diminuição da intensidade do evento de chuva observada, indicando a maior ocorrência de superestimativa para as menores chuvas (Figura 8f). Os valores de BIAS atingem até o valor de 1.248, que equivale à situação em que ocorre uma chuva observada sobre um único elemento da grade, mas há a estimativa de chuva sobre todos os 1.248 elementos da bacia. Na Figura 8f, o eixo y foi cortado no valor 10 para melhor visualização da nuvem principal de pontos.

#### 4.2 Análise integrada no tempo

Nesta abordagem de análise, os acertos e erros de estimativa de ocorrência de chuva em cada pixel são somados isoladamente ao longo do tempo, gerando índices de performance exclusivos do pixel, sem qualquer influência dos acertos/erros nos demais. Foram geradas quatro imagens que representam a distribuição espacial de cada índice sobre a discretização da bacia do Rio Grande na malha de  $0,1^\circ \times 0,1^\circ$  (Figura 10).

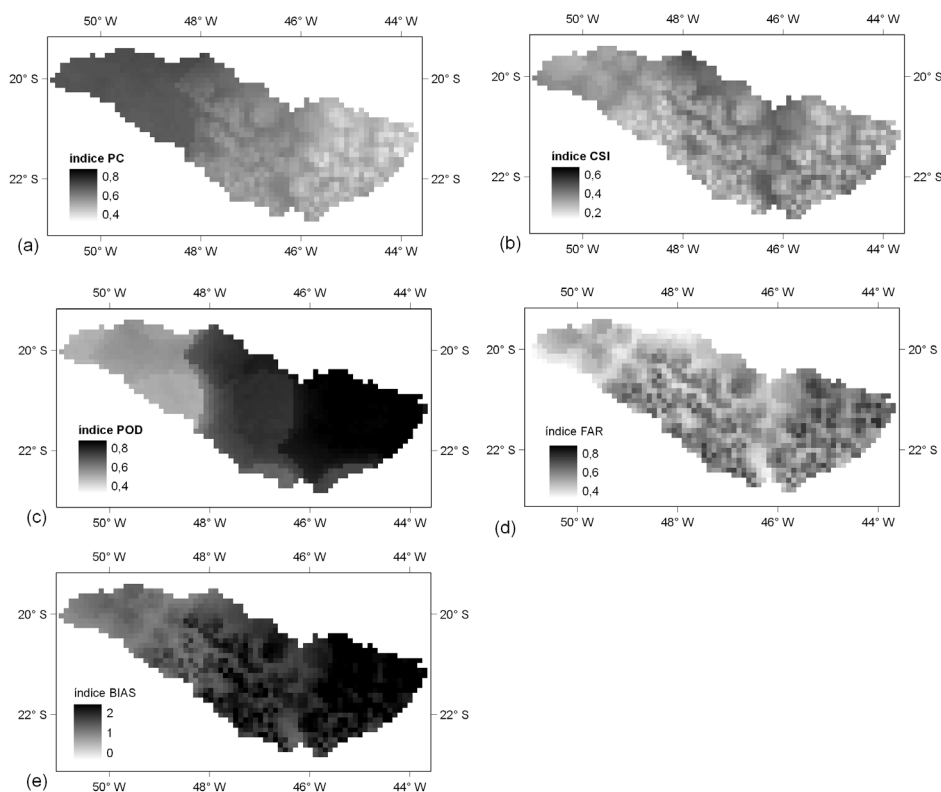
Para o índice PC, observa-se um gradiente espacial no sentido leste-oeste, com valores de até 40% na cabeceira do Rio Grande aumentando até 80% próximo ao exutório da bacia (Figura 10a). O índice POD apresenta distribuição espacial com gradiente inverso, aumentando de oeste para leste (Figura 10c).

Esses resultados indicam uma tendência de superestimativa da precipitação na porção leste da bacia e de subestimativa na porção oeste, o que é corroborado pelos valores obtidos para o índice BIAS: valores superiores a 1 na porção leste e inferiores ou mais próximos a 1 na porção oeste (Figura 10e).

Embora neste estudo apenas um dos membros da climatologia do modelo global CPTEC/COLA esteja sendo analisado, Paz et al. (2008) constataram a tendência de superestimativa de precipitação na região leste da bacia e de subestimativa na região oeste, mas de menor intensidade do que a primeira, para todos os membros do conjunto.

Por superestimar a ocorrência de chuva na porção leste, as chuvas que efetivamente ocorreram nessa região têm maior chance de detecção e por isso foram obtidos valores de POD em torno de 80%. Mas tal superestimação também conduz a errar com maior frequência a não ocorrência da chuva, reduzindo o índice PC para até 40% em alguns pixels. Por outro lado, a tendência de subestimativa de chuva na região oeste da bacia leva à menor probabilidade de detecção ( $POD < 60\%$ ), mas aumenta a quantidade total de acertos (tanto de ocorrência quanto de não ocorrência de chuva) e o índice PC alcança até 80%.

A distribuição espacial do índice FAR não apresenta um padrão de variação tão forte quanto dos índices PC e POD, mas



**Figura 10** – Imagens com distribuição espacial dos índices PC (a), CSI (b), POD (c), FAR (d), BIAS (e) sobre a bacia do Rio Grande, resultantes da análise integrada no tempo.

ainda assim é possível identificar uma semelhança. A tendência de superestimativa da ocorrência de chuva na porção leste da bacia conduz a maiores taxas de alarme falso, enquanto valores mais reduzidos foram obtidos para a região bem no extremo oeste onde há subestimativa da chuva.

Por sua vez, o índice CSI também não apresenta um padrão espacial bem definido, mas observa-se que maiores valores foram obtidos na porção central da bacia. Por não levar em conta os acertos da não estimativa de chuva, mas considerar ambos os tipos de erros (deixar de estimar uma chuva que ocorreu ou estimar uma chuva que não ocorreu), as tendências de superestimativa e de subestimativa conduzem a valores mais reduzidos desse índice. Em outras palavras, esse índice “pune” tanto a superestimativa quanto a subestimativa da ocorrência do evento chuvoso, de forma que valores maiores foram obtidos para a região onde há menor intensidade de ambas as tendências.

## 5. CONCLUSÕES

O método proposto foi aplicado para a avaliação de um dos membros do conjunto de realizações da climatologia do modelo global CPTEC/COLA sobre a bacia do Rio Grande. Esse estudo de caso ressaltou como o método proposto constitui uma alternativa interessante de avaliação de estimativas de campos de precipitação utilizadas como entrada para modelos hidrológicos distribuídos. A análise das séries temporais dos índices de performance possibilitou avaliar a existência de sazonalidade na destreza em estimar ocorrência/não ocorrência de chuva. Além disso, a geração de mapas que retratam a variação espacial da performance evidenciou tendências espaciais em superestimar/subestimar a ocorrência de precipitação em porções distintas da bacia.

O método apresentado pode ser aplicado para avaliar séries temporais de campos de precipitação gerados por modelos de circulação da atmosfera ou estimados a partir de imagens de satélite ou radar. Pelo seu caráter de comparação pixel a pixel, o referido método se apresenta bastante útil, principalmente como forma de verificar a qualidade de estimativas de chuva usadas como entrada em modelos hidrológicos distribuídos. Os campos de chuva podem ser analisados na mesma discretização espacial do modelo hidrológico, o que facilita inclusive a correção de eventuais erros sistemáticos previamente à utilização das estimativas de chuva como entrada para tais modelos.

Um incremento que pode ser aplicado à metodologia descrita neste artigo é a avaliação da significância estatística (Wigley e Santer, 1990; Mearns et al., 1995), bem como a consideração das estimativas por conjunto (Venugopal et al., 2005). Além disso, a própria comparação pixel a pixel poderia ser alterada para levar em conta uma vizinhança do pixel. A comparação pixel a pixel utilizada ignora o que ocorre nos pixels

vizinhos ao avaliar a estimativa de chuva em um determinado pixel. Esse procedimento tende a ser muito rígido na avaliação e subestimar a performance das estimativas, ao contrário de métodos que trabalham com a identificação de formas ou feições presentes nos campos de precipitação (Venugopal et al., 2005).

## 6. AGRADECIMENTOS

Ao CNPq, pela concessão de bolsa de Doutorado ao primeiro autor; parte desta pesquisa foi financiada pela FINEP-CT-Hidro.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

- ANEEL, Agência Nacional de Energia Elétrica. **Atlas de Energia Elétrica do Brasil**. Brasília: ANEEL, 2005. 2ª ed, 243 p.
- AVISSAR, R.; VERSTRAETE, M. M. The representation of continental processes in atmospheric models surface. **Reviews of Geophysics**, v. 28, n. 1, p. 35-52, 1990.
- BARRERA, D. F. Análisis comparativo de los mapas de precipitación obtenidos a partir de datos pluviométricos y de estimaciones satelitales. In: CONGRESO NACIONAL DE METEOROLOGÍA, IX, 2005, Buenos Aires. **Actas...** Buenos Aires: 2005, CD-ROM.
- BENOIT, R. et al. Toward the use of coupled atmospheric and hydrologic models at regional scale. **Monthly Weather Review**, v. 128, p. 1681-1706, 2000.
- CAVALCANTI, I. F. A. et al. **Climate characteristics in an ensemble simulation using the CPTEC/COLA atmospheric global circulation model**. São José dos Campos: INPE, 2001, (INPE-8150-RPQ/717). 71 p.
- CAVALCANTI, I. F. A. et al. Global Climatological Features in a Simulation Using the CPTEC-COLA AGCM. **Journal of Climate**, v. 15, n. 21, p. 2965-2988, 2002.
- CHERUBINI, T.; GHELLI, A.; LALAURETTE, F. Verification of precipitation forecasts over the Alpine region using a high-density observing network. **Weather and Forecasting**, v. 17, p. 238-249, 2002.
- COLLISCHONN, W. et al. Forecasting River Uruguay flow using rainfall forecasts from a regional weather-prediction model. **Journal of Hydrology**, n. 205, p. 87-98, 2005.
- COLLISCHONN, W. et al. The MGB-IPH model for large scale rainfall runoff modeling. **Hydrological Sciences Journal**, v. 52, n. 5, p. 878-895, 2007.
- EBERT, E. E. et al. The WGNE assessment of short-term quantitative precipitation forecasts. **Bulletin of the American Meteorological Society**, v. 84, n. 4, p. 481-492, 2003.
- FEIDAS, H. et al. A GIS tool for the evaluation of the precipitation forecasts of a numerical weather prediction

- model using satellite data. **Computers & Geosciences**, n. 33, p. 989-1007, 2007.
- HABETS, F.; LEMOIGNE, P.; NOILHAN, J. On the utility of operational precipitation forecasts to serve as input for streamflow forecasting. **Journal of Hydrology**, n. 293, p. 270-288, 2004.
- HAMILL, T. M.; JURAS, J. Measuring forecast skill: is it real skill or is it the varying climatology? **Quarterly Journal of the Royal Meteorological Society**, n. 132, p. 2905-2923, 2006.
- HASLER, N.; WERTH, D.; AVISSAR, R. Effects of tropical deforestation on global hydroclimate: a multimodel ensemble analysis. **Journal of Climate**, v. 22, n. 5, p. 1124-1141, 2009.
- HAY, L. E.; CLARK, M. P. Use of statistically and dynamically downscaled atmospheric model output for hydrologic simulations in three mountainous basins in the western United States. **Journal of Hydrology**, n. 282, p. 56-75, 2003.
- KITE, G. W.; HABERLANDT, U. Atmospheric model data for macroscale hydrology. **Journal of Hydrology**, n. 217, p. 303-313, 1999.
- KULIGOWSKI, R. A self-calibrating real-time GOES rainfall algorithm for short-term rainfall estimates. **Journal of Hydrometeorology**, n. 3, p. 112-130, 2002.
- KUMMEROW, C. et al. The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. **Journal of Applied Meteorology**, n. 39, p. 1965-1982, 2005.
- LOWREY, M. R. K.; YANG, Z. -L. Assessing the capability of a regional-scale weather model to simulate extreme precipitation patterns and flooding in central Texas. **Weather and Forecasting**, v. 23, n. 6, p. 1102-1126, 2008.
- MARENCO, J. et al. Ensemble simulation of regional rainfall features in the CPTEC/COLA atmospheric GCM. **Climate Dynamics**, v. 21, n. 5-6, p. 459-475, 2003.
- MCBRIDE, J. L.; EBERT, E. E. Verification of quantitative precipitation forecasts from operational numerical weather prediction models over Australia. **Weather and Forecasting**, v. 15, p. 103-121, 2000.
- MEARNS, L. O. et al. Analysis of the variability of daily precipitation in a nested modeling experiment: comparison with observations and 2xCO results. **Global and Planetary Change**, n. 10, p. 55-78, 1995.
- MESINGER, F. Bias Adjusted Precipitation Threat Scores. **Advances in Geosciences**, n. 16, p. 137-142, 2008.
- MOORE, R. J.; BELL, V. A.; JONES, D. A. Forecasting for flood warning. **Comptes Rendus Geoscience**, n. 337, p. 203-217, 2005.
- PAZ, A. R. et al. **Data assimilation in a large-scale distributed hydrological model for medium range flow forecasts**. IAHS Publ. 313. Wallingford: IAHS Press, 2007, p. 471-478.
- PAZ, A. R.; SAMPAIO, G.; COLLISCHONN, W. Análise da climatologia do modelo global CPTEC/COLA com ênfase aos eventos extremos de precipitação na bacia hidrográfica do Rio Grande. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, XV, São Paulo, **Anais...** São Paulo: SBMET, 2008, CD-ROM.
- SAMPAIO, G. et al. Regional climate change over eastern Amazonia caused by pasture and soybean cropland expansion. **Geophysical Research Letters**, v. 34, n. L17709, 2007.
- SCOFIELD, R. A. The NESDIS operational convective precipitation estimation technique. **Monthly Weather Review**, n. 115, p. 1773-1792, 1987.
- SHAFFREY, L. C. et al. The new U.K. High-Resolution Global Environment Model - model description and basic evaluation. **Journal of Climate**, v. 22, n. 8, p. 1861-1896, 2009.
- SINGH, V. P.; FREVERT, D. K. **Mathematical models of large watershed hydrology**. Chelsea: Water Resources Publications, 2002. 891 p.
- STEPHENSON, D. B. Use of the "odds ratio" for diagnosing forecast skill. **Weather and Forecasting**, v. 15, p. 221-232, 2000.
- TUCCI, C. E. M. et al. **Previsão de vazões na bacia do Rio São Francisco com base na previsão climática**. Brasília: ANEEL, Projeto OMM/98/00 - Relatório Técnico Final, 2004. 368 p.
- TUCCI, C. E. M. et al. Short- and long-term flow forecasting in Rio Grande watershed. **Atmospheric Science Letters**, v. 9, p. 1-4, 2008.
- VENUGOPAL, V.; BASU, S.; FOUFOULA-GEORGIOU, E. A new metric for comparing precipitation patterns with an application to ensemble forecasts. **Journal of Geophysical Research**, v. 110, n. D08111, 2005.
- VICENTE, G.; SCOFIELD, R.; MENZEL, P. The operational GOES infrared rainfall estimation technique. **Bulletin of the American Meteorological Society**, n. 79, p. 1883-1898, 1998.
- WIGLEY, T. M. L.; SANTER, B. D. Statistical comparison of spatial fields in model validation, perturbation, and predictability experiments. **Journal of Geophysical Research**, v. 95, n. D1, p. 851-865, 1990.
- WILKS, D. S. **Statistical methods in the atmospheric sciences**. San Diego: Academic Press, 2006. 630 p.
- XU, C. -Y. Climate change and hydrologic models: a review of existing gaps and recent research developments. **Water Resources Management**, n. 13, p. 369-382, 1999.
- ZINEVICH, A.; ALPERT, P.; MESSER, H. Estimation of rainfall fields using commercial microwave communication networks of variable density. **Advances in Water Resources**, 2008, n. 31, p. 1470-1480.