

## METODOLOGIA PARA ANÁLISE DE DESEMPENHO DE SIMULAÇÕES DE SISTEMAS CONVECTIVOS NA REGIÃO METROPOLITANA DE SÃO PAULO COM O MODELO ARPS: SENSIBILIDADE A VARIAÇÕES COM OS ESQUEMAS DE ADVEÇÃO E ASSIMILAÇÃO DE DADOS

RICARDO HALLAK E AUGUSTO JOSÉ PEREIRA FILHO

Universidade de São Paulo, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Departamento de Ciências Atmosféricas (USP/IAG/DCA), São Paulo, SP, Brasil

hallak@model.iag.usp.br, apereira@model.iag.usp.br

Recebido Janeiro de 2010 – Aceito Fevereiro de 2011

### RESUMO

Tempestades seguidas de enchentes e alagamentos em pontos da Região Metropolitana de São Paulo são eventos recorrentes nas estações chuvosas. O desempenho do modelo ARPS nas simulações numéricas de alta resolução espacial, para o evento de 04 de fevereiro de 2004, é avaliado por meio de erros estatísticos e índices de acurácia, com base no confronto entre a distribuição espacial da precipitação acumulada simulada e a estimada com o radar meteorológico de São Paulo. Os resultados quantificam a influência dos termos de advecção das equações primitivas nos erros de fase dos sistemas convectivos simulados, para o tipo de configuração atmosférica dominante no caso estudado. A assimilação incremental de medidas obtidas em estações de superfície durante a execução do modelo, corrige significativamente a fase das células convectivas profundas na região de influência das medidas, a qual é definida durante o processo de análise objetiva. Um novo índice de destreza para modelos de mesoescala é proposto ( $D_{PIELKE}$ ). As análises preliminares para o desempenho deste índice indicam adequação de sua formulação para o tipo de simulação efetuada.

**Palavras-chave:** tempestades, simulações numéricas de mesoescala, erros de fase e amplitude, índices de acurácia.

**ABSTRACT:** METHODOLOGY FOR PERFORMANCE ANALYSIS OF SIMULATIONS OF CONVECTIVE SYSTEMS IN THE METROPOLITAN AREA OF SÃO PAULO WITH THE ARPS MODEL: SENSITIVITY TO VARIATIONS WITH THE ADVECTION AND THE DATA ASSIMILATION SCHEMES

Thunderstorms followed by flash flood events are recurrent in the Metropolitan Area of São Paulo during the wet seasons. The performance of several numerical simulations of the February 04, 2004 event carried out with the ARPS system in high horizontal resolution is evaluated through statistical errors and indices of accuracy. The evaluation is based on direct confrontation of the accumulated precipitation fields simulated by the model and measured by the São Paulo weather radar. The results quantify the influence of the advective terms of the primitive equations on phase errors of the simulated convective systems. The incremental assimilation of surface station data done during the model execution improved its performance in the area limited by the influence radius, which is defined during the objective analysis process. A new skill index formulation is presented ( $D_{PIELKE}$ ). The preliminary results indicate that the  $D_{PIELKE}$  index is adequate for assessing the simulation skill of small-scale convective systems.

**Keywords:** thunderstorms, mesoscale numerical simulations, phase and amplitude errors, indices of accuracy.

## 1. INTRODUÇÃO

A acurácia das previsões de fase e amplitude dos fenômenos meteorológicos é um dos importantes problemas científicos, que estimulam pesquisas ao redor do globo. A abrangência desta questão adquire dimensões sociais em áreas densamente povoadas do Brasil, como a Região Metropolitana de São Paulo (RMSP). A RMSP é periodicamente sujeita a eventos convectivos severos, que produzem intensa precipitação, cujos efeitos negativos para a população são agravados pelo uso e ocupação inadequados do solo (Silva et al., 2009). O monitoramento operacional da precipitação na RMSP é efetuado já há algum tempo com o uso do Radar Meteorológico de São Paulo (RADSP), operado pela Fundação Centro Tecnológico de Hidráulica (FCTH) da Universidade de São Paulo (USP) (Pereira Filho e Barros, 1998). Com resolução espaço-temporal equivalente a uma rede telemétrica com cerca de 1 pluviômetro a cada 2 km x 2 km, o RADSP proporciona, por exemplo, estimativas da precipitação para acompanhamento e previsões de chuva a curtíssimo prazo (15 minutos a 1 hora) pelo método de extrapolação (p.e., Farias, 2009). No entanto, a despeito do uso de radares meteorológicos, imagens de satélite mais frequentes e previsões numéricas globais com razoável resolução espacial entre outras medidas atmosféricas e seus subprodutos, a previsão quantitativa de precipitação é ainda bastante deficiente na escala de tempo de curto prazo (1 a 24 horas). De fato, as previsões de precipitação de curto prazo são fortemente dependentes da subjetividade do previsor de tempo. A modelagem numérica regional de alta resolução espacial é uma das opções para se obter avanços, tanto no aprimoramento da compreensão dos fenômenos atmosféricos severos, como no aumento da precisão de sua previsão.

Neste trabalho, a simulação explícita de sistemas convectivos profundos (isto é, aquela onde o ciclo de vida completo de nuvens cumulonimbus individuais se desenvolve de fato na solução numérica) é obtida em domínios numéricos de alta resolução espacial com espaçamentos de grade de 1 km. O objetivo das simulações nesta resolução de mesoescala  $\beta$  e  $\gamma$  (Orlanski, 1975) é reproduzir oscilações espaço-temporais de variáveis meteorológicas com comprimentos de onda de, no mínimo, uma dezena de quilômetros. No entanto, como o realismo das simulações é afetado pela componente estocástica das simulações de mesoescala, que aumenta conforme a escala espacial do domínio numérico decresce (Anthes, 1983), espera-se que erros de fase e amplitude cresçam e se propaguem pelo domínio de simulação. As estimativas de precipitação à superfície RADSP oferecem uma oportunidade de se caracterizar e compreender a origem dos erros de fase e amplitude nos campos de precipitação das previsões e simulações numéricas, contra as quais as estimativas RADSP podem ser comparadas.

São diversas as fontes de erros nos resultados finais de simulações numéricas de alta resolução espacial. Por exemplo, as parametrizações dos fenômenos de subgrade, que é a parte da física atmosférica não resolvida pela grade do modelo e que devem necessariamente estar contidas nas suas soluções numéricas quando se objetiva um maior realismo dos resultados, podem ser tratadas como fontes de perturbação das simulações. Geralmente, são geradas diferentes soluções numéricas para cada diferente opção de parametrização de um mesmo fenômeno físico, com consequências significativas para a previsão dos campos de precipitação. Por outro lado, a questão da esparsa distribuição espacial dos instrumentos meteorológicos no Brasil, bem como, a acurácia e precisão das suas medições, também afeta a qualidade das soluções numéricas finais na escala dos fenômenos convectivos. Observe-se que a atmosfera ambiente sofre oscilações dinâmicas e termodinâmicas em alta frequência no tempo e no espaço, quando se desenvolvem sistemas convectivos profundos. Como consequência, estas oscilações são raramente capturadas pela base atual de medições convencionais de superfície e aerológicas. Mesmo quando se tenta corrigir estas deficiências por meio de metodologias alternativas, como por exemplo, aquelas que envolvem a assimilação de dados observacionais convencionais ou remotos nos modelos de mesoescala, a acurácia limitada dos dados é fonte de erro nos resultados numéricos (p.e., Daley, 1991).

Este trabalho revisa e aplica metodologias de análises de erros aos resultados de simulações numéricas de alta resolução espacial, sob o ponto de vista da sua distribuição espacial. As metodologias propostas baseiam-se em análises objetivas de erros e índices de concordância e destreza sobre conjuntos de medidas atmosféricas comparadas a resultados numéricos de simulações explícitas da convecção atmosférica. As análises são aplicadas aos campos de precipitação simulados para um caso real de sistema convectivo profundo de mesoescala observado na RMSP. O modelo numérico adotado é o sistema ARPS (Advanced Regional Prediction System) (Xue et al., 2000; 2001; 2003). A análise objetiva de dados meteorológicos ADAS [ARPS Data Assimilation System (Xue et al., 1995)] é usada para a assimilação de conjuntos de dados de superfície e sondagens atmosféricas na elaboração de experimentos de sensibilidade do modelo. O ARPS é integrado no cluster de 16 processadores XEON-Pentium 4/3.0 GHz (Pereira Filho et al., 2008) com pico de processamento de 120 Gflops (média de 38 Gflops) do Laboratório de Hidrometeorologia (LABHIDRO) do Instituto de Astronomia, Geofísica e Ciências Atmosféricas (IAG) da USP.

## 2. DADOS E METODOLOGIA

De forma sumária, a caracterização da situação atmosférica, associada aos sistemas convectivos observados na Região RMSP,

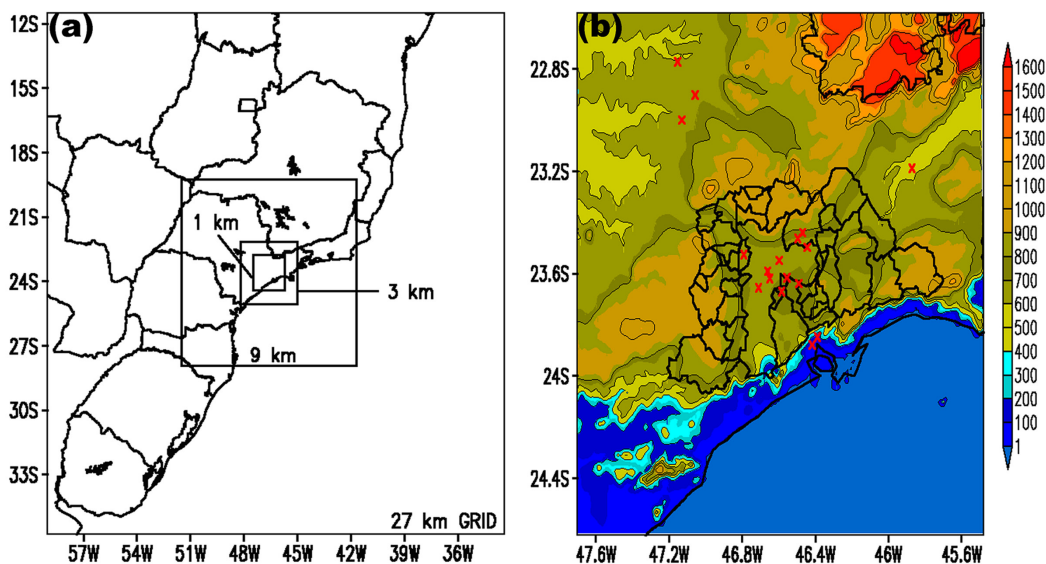
foi obtida pela análise diagnóstica de medidas meteorológicas convencionais e estimativas de variáveis obtidas por sensoriamento remoto. Neste sentido, estimativas de precipitação à superfície, obtidas pela conversão das refletividades RADSP por meio da relação Z-R de Marshall e Palmer (1948), foram usadas em conjunção às imagens dos satélites geoestacionários GOES-10 e GOES-12 para a identificação e seleção de casos de convecção profunda associados a enchentes repentinas em pontos da RMSP. As radiossondagens realizadas diariamente nos horários sinóticos das 0000 UTC e 1200 UTC no aeródromo de Campo de Marte (situado na Zona Norte da Capital Paulista), foram usadas para análise do perfil vertical das condições atmosféricas ambientes e, adicionalmente, para assimilação no modelo ARPS. Medidas de estações de superfície das redes SYNOP, METAR e CETESB foram também assimiladas no ARPS com o auxílio do sistema de análise objetiva ADAS. As análises GFS de 1° de espaçamento horizontal (cerca de 110 km na linha do equador), disponíveis a cada 6 horas, foram usadas para a caracterização sinótica da atmosfera nos casos selecionados. Além disso, estas análises globais alimentaram o modelo regional ARPS com condições iniciais e de fronteira na sua grade de menor resolução espacial. Os resultados gerais das análises diagnósticas efetuadas com o conjunto de dados descritos são encontrados em Hallak (2007).

## 2.1. Simulações numéricas

As simulações foram efetuadas em 4 domínios espaciais

com diferentes resoluções horizontais (27, 9, 3 e 1 km de espaçamento entre os pontos de grade) e com 72 níveis verticais, com alta resolução vertical no primeiro quilômetro acima da superfície. O aninhamento das 3 grades internas ao domínio de 27 km de espaçamento é mostrado na Figura 1a.

As características básicas da fisiografia do solo e superfície do modelo, a geometria dos domínios e as parametrizações de precipitação usadas em cada grade estão descritas na Tabela 1. Os campos de análises GFS são usados como condições iniciais e de fronteira do modelo regional ARPS em sua grade de menor resolução espacial (27 km). O aninhamento das grades é feito no modo unidirecional, de forma que não há influência das grades aninhadas sobre as grades exteriores, mas há transferência de informação no sentido inverso nas fronteiras de cada grade aninhada. O aninhamento unidirecional é o único modo atualmente disponível no ARPS. Neste esquema de aninhamento, cada grade é rodada independentemente das outras. Assim, uma grade de menor resolução alimenta com condições iniciais e de fronteira a grade imediatamente aninhada a ela. Na sequência, o processo se repete para a grade aninhada seguinte. Uma vantagem do aninhamento unidirecional é que diversos experimentos efetuados com as grades de maior resolução espacial podem ser obtidos sem que haja necessidade de se rodar as grades exteriores novamente. Por outro lado, qualquer alteração efetuada em uma grade externa implica na obrigatoriedade de se rodar novamente as grades interiores a ela.



**Figura 1** - (a) Domínios numéricos das 4 grades usadas no modelo ARPS no modo de aninhamento unidirecional. As resoluções horizontais das grades estão indicadas na figura. São também indicadas as fronteiras geopolíticas de vários estados brasileiros. As 4 grades estão centradas na Cidade de São Paulo. (b) Elevação topográfica (m) da grade de 1 km de espaçamento, sombreada de acordo com a barra de cores. Estão indicadas as fronteiras geopolíticas dos 39 municípios que compõe a RMSP. As marcas vermelhas em forma de X marcam a posição geográfica das estações meteorológicas de superfície das redes SYNOP, METAR e CETESB.

A Figura 1b mostra a elevação topográfica da grade de 1 km de espaçamento. Nesta figura estão também indicadas as fronteiras geopolíticas dos 39 municípios que compõem a RMSB. As marcas vermelhas em forma de X marcam as posições geográficas das estações meteorológicas de superfície das redes SYNOP, METAR e CETESB.

Neste trabalho, as grades com espaçamentos horizontais de 3 e de 1 km utilizam somente a parametrização de microfísica de nuvens de Lin et al. (1983) (Tabela 1). Esta parametrização permite o tratamento dos processos associados à distribuição espacial da água na atmosfera em suas três fases, incluindo a formação dos diversos tipos de nuvens. As grades de menor resolução espacial (27 km e 9 km) utilizam somente a parametrização de cúmulos de Kain e Fritsch (1993) nos experimentos de controle.

## 2.2. Análise dos resultados das simulações

Para a quantificação dos erros nas simulações numéricas e a possível determinação das suas fontes de erros, aplica-se uma série de índices e escores estatísticos, os quais são usados como ferramentas de avaliação da acurácia dos experimentos numéricos. A metodologia proposta oferece ferramentas adicionais às verificações diretas tradicionais das soluções numéricas, realizadas por meio da comparação visual dos campos simulados e derivados das quantidades básicas, procedimentos estes que se enquadram na categoria de métodos subjetivos. Além disso, as ferramentas aqui descritas oferecem uma alternativa às comparações estatísticas efetuadas

isoladamente em pontos de estações, que frequentemente penalizam exageradamente os resultados dos modelos numéricos (Anthes et al., 1989), em especial nos casos que envolvem o desenvolvimento de sistemas convectivos de mesoescala e de escala local. Dessa forma, são quantificados os erros das simulações em relação aos campos horizontais de distribuição de precipitação acumulada à superfície contra os campos de precipitação estimados pelo RADSP derivados dos CAPPs de 3 km de altitude nos horários correspondentes. Os erros são também calculados em função de variações na área de abrangência do domínio com o objetivo de se avaliar a sensibilidade das respostas do modelo para determinadas áreas do domínio numérico (Martin e Xue, 2006). Análises comparativas podem ser efetuadas sobre os campos de precipitação acumulada em variados períodos de tempo. Ressalta-se que qualquer outra variável meteorológica pode ser analisada pela mesma metodologia, desde que já esteja objetivamente analisada em pontos de grade, ou seja, distribuída apropriadamente numa grade regular de pontos.

### 2.2.1. Preliminares

Pereira Filho e Nakayama (2001) estabeleceram os limites espaciais onde as estimativas de precipitação obtidas pelo RADSP podem ser consideradas de boa qualidade, para além da qual a qualidade das estimativas cai com a distância ao radar. Segundo seus resultados, a área onde as estimativas são mais confiáveis corresponde a um círculo centrado no radar com raio de 120 km.

**Tabela 1** - Características fisiográficas gerais das simulações ARPS, configurações geométricas das 4 grades e parametrizações de precipitação utilizadas em cada grade.

Características Gerais	Espaçamento Horizontal Entre Pontos de Grade			
	27 km	9 km	3 km	1 km
Pontos na Horizontal	131 x 131	131 x 131	131 x 131	179 x 179
Níveis Verticais	72	72	72	72
Altura do 1º. Nível	40 m	20 m	10 m	10 m
Topo do Modelo	~ 30 km	~ 29 km	~ 28.5 km	~ 28.5 km
Tempo de Integração	24 h	24 h	12 h	12 h
Passo de Tempo	60 s	20 s	6 s	2 s
Condições Ini./Front.	GFS (1°)	ARPS 27 km	ARPS 9 km	ARPS 3 km
Topografia (USGS)	900 m	900 m	900 m	900 m
Parametr. Cúmulos	Sim	Sim	Não	Não
Parametr. Microf.	Não	Não	Sim	Sim
Modelo de Solo	2 Camadas	2 Camadas	2 Camadas	2 Camadas
Tipo de Solo (USGS)	Heterogêneo	Heterogêneo	Heterogêneo	Heterogêneo
Tipo de Veget. (USGS)	Heterogêneo	Heterogêneo	Heterogêneo	Heterogêneo



A metodologia de verificação estatística das simulações permite a comparação entre os diversos experimentos numéricos a partir da confrontação direta entre as métricas derivadas. Com base nestas comparações, os experimentos que minimizam os erros das simulações podem então ser objetivamente determinados. No entanto, a metodologia proposta somente pode ser aplicada para campos simulados e observados que possuam a mesma resolução espacial. Como o espaçamento horizontal original das estimativas RADSP se aproxima de 2 km, estas foram interpoladas para uma nova grade pelo método de interpolação bilinear com espaçamento  $dx = dy = 0.01^\circ$  (aproximadamente 1 km em longitude e latitude). Por sua vez, os campos de precipitação simulados com o ARPS na grade de 1 km foram degradados para exatamente  $dx = dy = 0.01^\circ$  pelo método “box averaging”, mais eficiente para a obtenção de degradações de resoluções em contraste aos métodos de interpolação de dados. Com ambos os campos de precipitação representados pelo mesmo conjunto de pontos de grade, procede-se então a avaliação objetiva das simulações, uma vez que ambos os processos de regragem preservam a distribuição espacial e a amplitude do campo original com desprezível perda de informação.

### 2.2.2. Volume de água precipitada à superfície

O valor total da precipitação acumulada à superfície é uma variável de interesse na comparação quantitativa entre observado e simulado. Neste trabalho, as comparações quanto à precipitação são obtidas em termos do volume total precipitado para cada um dos subdomínios selecionados para os cálculos estatísticos. Assim, para cada subdomínio de uma dada grade, o volume de água precipitada é calculado em função da área do pixel e da precipitação acumulada a cada hora durante um evento. Em seguida, o volume acumulado em cada pixel do subdomínio é somado para se achar o volume total acumulado naquela área durante um determinado período de tempo da duração do evento. O valor do volume de água precipitado contém informações sobre o desempenho do modelo numérico e, em adição, permite aplicação prática do ponto de vista hidrológico, como na previsão de enchentes repentinas e no manejo de recursos hídricos numa determinada bacia hidrográfica.

Um elemento de área da superfície terrestre  $S_j$  em coordenadas esféricas, definida para cada conjunto  $j$  de quatro pontos de grade nas direções  $x$  e  $y$ , é dado pela Equação 1. O quadrilátero assim definido, também chamado pixel de grade, é limitado pelos intervalos  $[\theta_i, \theta_f]$  em longitude e  $[\varphi_i, \varphi_f]$  em latitude, onde os índices  $i$  e  $f$  referem-se, respectivamente, aos valores “iniciais” e “finais” dos intervalos. Na Equação 1,  $r$  é o raio da terra:

$$S_j = \int_{\theta_i, \varphi_i}^{\theta_f, \varphi_f} r^2 \cos \varphi \partial \varphi \partial \theta. \quad (1)$$

A somatória do produto da precipitação  $P_j$  por cada elemento de área  $S_j$ , tomados no quadrilátero de área  $j$ , fornece a quantidade desejada do volume total precipitado ( $V_i$ ) no subdomínio analisado, desde que se considere que houve precipitação de forma homogênea em cada pixel da grade. Esta última condição penaliza os cálculos na medida inversamente proporcional à resolução do domínio de grade. Dessa forma, quanto mais alta a resolução espacial do domínio, menor é a limitação imposta pela condição de homogeneidade espacial da precipitação no pixel. O total do volume precipitado está expresso na Equação 2:

$$V_i = \sum_j P_j S_j. \quad (2)$$

### 2.2.3. Erros, métricas, escores e índices estatísticos

O cérebro humano, alimentado pela experiência pessoal e pelas informações sensoriais fornecidas pelos olhos, é uma poderosa ferramenta de análise e identificação de padrões geométricos. Este é o caso de comparações visuais entre campos simulados e medidos em meteorologia (método subjetivo). Por outro lado, a quantificação das comparações entre campos meteorológicos é muitas vezes mais precisa se obtida por intermédio do uso de computadores, que são calculadoras poderosas, e a ciência da computação e seus algoritmos, que ordenam os cálculos no computador. A análise objetiva baseia-se no segundo método.

Ressalta-se, inicialmente, que o tratamento dos erros inerentes às estimativas do RADSP não é aqui considerado. As métricas estatísticas calculadas neste trabalho para avaliação das simulações são baseadas nos escores estatísticos descritos, por exemplo, em Fox (1981), Willmot (1982), Anthes (1983), Anthes et al. (1989), Pielke (2002) e Wilks (2006). No que segue, as somatórias indicadas são aplicadas a campos horizontais com  $N$  pontos de grade, onde  $N = n_x \cdot n_y$ , com  $n_x$  representado o número de pontos da grade em longitude e  $n_y$ , em latitude. Contudo, observa-se que  $N$  poderia se referir a número de eventos numa série temporal, sem perda de generalidades.

Seja  $P$  uma variável qualquer discreta ou contínua no espaço. Segundo Pielke (2002), a correspondência ponto a ponto entre a solução numérica e as medidas observacionais de uma mesma variável, proporciona um teste quantitativo para medir a habilidade (ou destreza) do modelo em reproduzir ou prever dados observados. Sejam  $P_{is}$  e  $P_{io}$ , respectivamente, a precipitação acumulada simulada e observada num mesmo ponto de grade  $i$  num domínio numérico com  $N$  pontos de grade. A quantidade fundamental para o estudo de erros é a diferença ( $d_i$ ) entre o previsto ou simulado no ponto  $i$  do espaço

bidimensional ( $i = 1, 2, 3, \dots, N$ ) no tempo  $t$  [ $P_{is} = P(x, y, t)$ ] e o medido ou observado para uma mesma variável, no mesmo ponto no espaço e no tempo [ $P_{io} = O(x, y, t)$ ], o que é expresso simplesmente por:

$$d_i = P_{is} - P_{io}. \quad (3)$$

Basicamente,  $d_i = 0$  indica simulação perfeita naquele ponto  $i$ , enquanto  $d_i \gg 0$  ou  $d_i \ll 0$ , simulação imperfeita. Quanto mais distante de 0 o valor de  $d_i$ , mais imperfeita a simulação. No entanto, embora  $\sum d_i$  forneça uma idéia da qualidade de uma simulação para uma dada variável, o mesmo não explica as fontes ou características particulares das magnitudes dos erros. A partir da quantidade básica  $d_i$ , contudo, é possível derivar erros que reflitam diferentes componentes do erro total. No que segue, será dada preferência ao uso de siglas em língua inglesa para a identificação dos diversos erros, para uma melhor conformação à literatura científica disponível atualmente. É importante também notar que todas as operações matemáticas, que envolvam média no espaço, devem ser efetuadas considerando-se a esfericidade da superfície terrestre. Portanto, cálculos computacionais que envolvam área são sempre ponderados pelo cosseno da latitude, o que aumenta a precisão dos resultados. Além disso, os conceitos aqui desenvolvidos podem ser aplicados indiferentemente na verificação de simulações ou de previsões. No que segue, será usado sempre o termo “simulação”, em conformidade com o estudo de caso apresentado na seção 3.

A medida mais simples derivada de  $d_i$  é o Erro Médio (ME em Inglês) ou tendência ou, ainda, viés (BIAS em Inglês). O VIÉS mede a tendência de o modelo superestimar ou subestimar a precipitação em relação ao observado. Esta tendência, também chamada de erro sistemático, é definida como:

$$ME = BIAS = \bar{d} = \frac{1}{N} \sum_{i=1}^N d_i. \quad (4)$$

Em termos estatísticos, a média dos erros individuais ou viés definido pela Equação 4 é o primeiro momento da distribuição das diferenças. O viés não traz informação a respeito dos erros individuais e, portanto, não pode ser usado como medida de acurácia da simulação. Além disso, apesar de fornecer uma idéia da tendência ou erro sistemático, o viés é afetado pelo fato de que erros individuais positivos e negativos de mesma magnitude se cancelam na somatória, o que geralmente leva à subestimação do erro das simulações. A média dos erros absolutos (Erro Absoluto Médio  $|\bar{d}|$  ou MAE na sigla em Inglês) contorna este problema. Por ser menos afetado por pontos com valores anormalmente extremos, também chamados “pontos singulares” (ou “outliers” em Inglês), o MAE é considerado preciso e robusto como medida da habilidade de modelos numéricos em reproduzir a realidade (Fox, 1981). O MAE é definido por:

$$MAE = |\bar{d}| = \frac{1}{N} \sum_{i=1}^N |d_i|. \quad (5)$$

Uma medida frequentemente usada na verificação da acurácia de modelos numéricos é o Erro Quadrático Médio (MSE na sigla em Inglês) como descrito, por exemplo, em Wilks (2006). O MSE é similar ao MAE, porém mais sensível aos grandes erros, por elevar as diferenças individuais ao quadrado. MSE é sempre positivo. MSE = 0 indica simulação perfeita. MSE é definido por:

$$MSE = \frac{1}{N} \sum_{i=1}^N (P_{is} - P_{io})^2. \quad (6)$$

Em adição ao MSE, a raiz quadrada de MSE, ou Raiz do Erro Quadrático Médio (RMSE em Inglês), é comumente usada para expressar a acurácia dos resultados numéricos com a vantagem de que RMSE apresenta valores do erro nas mesmas dimensões da variável analisada. O RMSE é definido por:

$$RMSE = \left[ \frac{1}{N} \sum_{i=1}^N (P_{is} - P_{io})^2 \right]^{1/2}. \quad (7)$$

Por outro lado, segundo Takacs (1985), o Erro Quadrático Médio pode ser decomposto em um termo dissipativo (ou de amplitude,  $MSE_{diss}$ ) e outro, dispersivo (ou de fase,  $MSE_{disp}$ ). Para maior clareza, doravante o MSE como definido pela Equação 6 passa a ser chamado de Erro Médio Quadrático Total ( $MSE_{TOT}$ ),

$$MSE_{TOT} = MSE, \quad (8)$$

o qual pode ser escrito (Takacs, 1985):

$$MSE_{TOT} = MSE_{DISS} + MSE_{DISP}. \quad (9)$$

Na Equação 9, o componente dissipativo de  $MSE_{TOT}$  é dado por:

$$MSE_{DISS} = [\sigma(P_{io}) - \sigma(P_{is})]^2 + (\bar{P}_{io} - \bar{P}_{is})^2 \quad (10)$$

e o dispersivo por:

$$MSE_{DISP} = 2(1 - \rho)\sigma(P_{io})\sigma(P_{is}), \quad (11)$$

onde  $\sigma(x)$  é o operador desvio padrão da variável aleatória  $x$ ,  $\rho$  é o coeficiente de correlação espacial entre os campos horizontais de  $P_{io}$  e  $P_{is}$  e as barras sobrescritas indicam médias espaciais das variáveis. O desvio padrão para uma variável  $x$  qualquer,  $\sigma(x)$ , é definido como a raiz quadrada da variância de  $x$ :

$$\sigma(x) = \left[ \frac{1}{N} \sum (x - \bar{x})^2 \right]^{1/2}, \quad (12)$$

onde a variância de  $x$  é uma medida da dispersão das ocorrências de  $x$  em torno da sua própria média. Em outras palavras, a variância de uma variável  $x$  é o erro que seria cometido se todas as ocorrências de  $x$  fossem substituídas pelo valor esperado (ou

média). Note-se que o desvio padrão de uma variável também é afetado pelos pontos singulares.

O coeficiente de correlação espacial ( $\rho$ ) é definido da mesma forma que o coeficiente de correlação linear, com a diferença de ser aplicado às variáveis espacialmente distribuídas num mesmo domínio:

$$\rho = \frac{COV(P_{is}, P_{io})}{\sigma_s \sigma_o} \quad (13)$$

Na Equação 13, COV é a covariância entre o campo simulado e o observado de uma variável. Ainda,  $\sigma_s = \sigma(P_s)$  e  $\sigma_o = \sigma(P_o)$ . O exame conjunto dos três tipos de erros (MSE<sub>TOT</sub>, MSE<sub>DISS</sub> e MSE<sub>DISP</sub>), pode ser efetuado para indicar, por exemplo, os experimentos numéricos que minimizam os erros num conjunto de simulações de um mesmo caso.

Numa amostra obtida de um universo de medidas qualquer, a variância das diferenças ( $S_d^2$ ) é o segundo momento da distribuição das diferenças, definida por:

$$S_d^2 = \frac{1}{N-1} \sum_{i=1}^N (d_i - \bar{d})^2 = \frac{1}{N-1} \sum_{i=1}^N (d_i - \bar{d})(d_i - \bar{d}), \quad (14)$$

onde  $S_d^2$  fornece informação sobre a distribuição dos erros  $d_i$  em torno da média dos erros. Note-se que a solução computacional para a covariância (COV) é a mesma para o cálculo da variância, uma vez que a variância de uma variável  $x$  é um caso especial da covariância de  $x$  com ela mesma, como está explícito na segunda igualdade da Equação 14. A raiz quadrada da variância é o desvio padrão da distribuição das diferenças. Como mostrado em Fox (1981), o erro médio quadrático total (MSE<sub>TOT</sub>) está relacionado à tendência ou viés ( $\bar{d}$ ) e à variância ( $S_d^2$ ) por:

$$MSE_{TOT} = \frac{N-1}{N} (S_d^2) + \bar{d}^2. \quad (15)$$

Vê-se, portanto, que MSE<sub>TOT</sub> contém mais informação do que o viés ou a variância isoladamente. Note-se que, para variáveis com distribuição normal, o viés tem distribuição normal, enquanto a variância tem distribuição de chi-quadrado. O MSE<sub>TOT</sub>, de acordo com a Equação 15, tem então, distribuição composta (Fox, 1981). Por esta razão, e pelo fato de que o viés não representa adequadamente as características individuais dos erros, Willmott (1982) considera MAE e RMSE (ou MSE) as melhores medidas gerais de desempenho de modelos.

Por outro lado, Wilks (2006) ressalta que MSE<sub>TOT</sub> pode também ser escrito em função do quadrado do erro médio (viés) e do coeficiente de correlação linear de Pearson, entre os valores simulados e observados ( $r_{so}$ ):

$$MSE_{TOT} = (\bar{P}_{is} - \bar{P}_{io})^2 + \sigma_{is}^2 + \sigma_{io}^2 - 2\sigma_{is}^2 \sigma_{io}^2 r_{so}. \quad (16)$$

Neste caso, vê-se que o coeficiente de correlação linear  $r_{so}$  (ou, analogamente, o coeficiente de determinação  $r_v^2$ ) pode

levar a falsas conclusões, quando interpretado como índice de acurácia de previsões ou simulações. Basta, por exemplo, imaginar um caso em que, para um mesmo valor de MSE<sub>TOT</sub>, tenha-se alto valor de  $r_{so}$  e um igualmente alto valor de viés (primeiro termo do lado direito da Equação 16). Neste caso,  $r_{so}$  e viés podem mutuamente se cancelar na Equação 16. Portanto, para uma situação como a descrita, uma variável cuja simulação e observação sejam altamente correlacionadas, o alto valor de  $r_{so}$  (ou  $r_v^2$ ) ocultará a tendência das simulações superestimarem ou subestimarem o observado. Como notado por Willmott (1982), o parâmetro de intercepção e a declividade da reta de regressão são mais informativos do que o próprio coeficiente de correlação. Por analogia, o mesmo problema pode então ser atribuído ao coeficiente de correlação espacial ( $\rho$ ). Assim, quando analisado isoladamente,  $\rho$  não é adequado para indicar completamente a medida de acurácia de uma simulação.

Willmott (1982) ressalta que, apesar de MAE e RMSE proporcionarem estimativas do erro médio, nenhum dos dois proporciona informações sobre a natureza ou tipo das diferenças. Então, é importante determinar qual porção do erro descrito por RMSE é sistemático e qual porção do erro é não-sistemático. Em outras palavras, para uma simulação considerada boa, a diferença sistemática deve se aproximar a zero, enquanto a diferença não-sistemática deve se aproximar de RMSE. Para quantificar o erro sistemático em MSE (MSE<sub>sis</sub>) e o não-sistemático (MSE<sub>nsis</sub>), Willmott (1982) propõe os seguintes cálculos:

$$MSE_{sis} = \frac{1}{N} \sum_{i=1}^N (\hat{P}_i - P_{io})^2, \quad (17)$$

onde  $\hat{P}_i = a + bP_{io}$  é a equação da reta de regressão linear entre os campos simulado e observado (sendo  $\mathbf{a}$  o parâmetro de intercepção do eixo  $y$  e  $\mathbf{b}$  a declividade da reta) e para os erros não-sistemáticos.

$$MSE_{nsis} = \frac{1}{N} \sum_{i=1}^N (P_{is} - \hat{P}_i)^2. \quad (18)$$

No entanto, por serem não-limitados, nenhum dos erros definidos acima fornecem informação sobre o tamanho relativo das diferenças. O Índice de Concordância (IC), proposto por Willmott (1982), possui esta propriedade ( $0 \leq IC \leq 1$ ) e pode ser usado, por exemplo, para comparar diferentes simulações de um mesmo fenômeno. IC é definido por:

$$IC = 1 - \frac{\sum_{i=1}^N (P_{is} - P_{io})^2}{\sum_{i=1}^N (|P_{is} - \bar{P}_{io}| + |P_{io} - \bar{P}_{io}|)^2}, \quad (19)$$

com as médias (variáveis com barra sobrescrita) tomadas na área em análise. IC = 1 indica concordância perfeita entre o campo simulado e o observado.

Pielke (2002) sugere uma metodologia adicional para a determinação da qualidade de uma simulação. Dessa forma, se uma simulação preenche os três critérios sugeridos a seguir, a destreza da simulação fica demonstrada. Os critérios são:

- 1)  $\sigma_s \cong \sigma_o$  ou, equivalentemente,  $\frac{\sigma_s}{\sigma_o} \cong 1$ ;
- 2)  $RMSE < \sigma_o$ ;
- 3)  $RMSE_{bias} < \sigma_o$ .

onde  $RMSE_{bias}$  é o RMSE após a remoção de um bias constante associado à tendência do modelo, definido por:

$$RMSE_{bias} = \left\{ \frac{1}{N} \sum_{i=1}^N [(P_{is} - \bar{P}_{is}) - (P_{io} - \bar{P}_{io})]^2 \right\}^{1/2} \quad (21)$$

Um teste adicional, para se determinar a simulação mais realística dentro de um conjunto de simulações de um mesmo caso, foi elaborado com base nos três critérios propostos por Pielke (2002), conforme Equações 20. Note-se, primeiramente, que da Equação 20 (1)  $\frac{\sigma_s}{\sigma_o} \cong 1 \Rightarrow 1 - \frac{\sigma_s}{\sigma_o} \cong 0$  se os desvios padrão se equivalerem (ou seja, se o padrão de distribuição espacial da variável simulada é equivalente à observada). Caso contrário, a razão  $\sigma_s/\sigma_o$  pode assumir qualquer outro valor real e a diferença se distancia de zero. Considerando-se, então, um termo absoluto com a magnitude:

$$ABS\left(1 - \frac{\sigma_s}{\sigma_o}\right)$$

e que, da Equação 20 (2):

$$RMSE < \sigma_o \Rightarrow 0 \leq \frac{RMSE}{\sigma_o} < 1, \quad (22)$$

e da Equação 20 (3):

$$RMSE_{BIAS} < \sigma_o \Rightarrow 0 \leq \frac{RMSE_{BIAS}}{\sigma_o} < 1, \quad (23)$$

então uma possível formulação para um teste de destreza de modelo, segundo os critérios de Pielke (2002), seria:

$$D_{PIELKE} = ABS\left(1 - \frac{\sigma_s}{\sigma_o}\right) + \frac{RMSE}{\sigma_o} + \frac{RMSE_{BIAS}}{\sigma_o} \quad (24)$$

Na Equação 24,  $D_{PIELKE}$  é, então, um índice baseado nos critérios propostos em Pielke (2002), que fornece uma medida da destreza do modelo numérico regional em simular realisticamente um determinado campo meteorológico. Note-se que:

- 1) A destreza do modelo fica demonstrada se  $D_{PIELKE} < 2$ ;
- 2) A melhor simulação num conjunto de várias simulações de um mesmo caso é aquela que obtiver o menor valor de  $D_{PIELKE}$ ;
- 3) A simulação perfeita fornece  $D_{PIELKE} = 0$ .

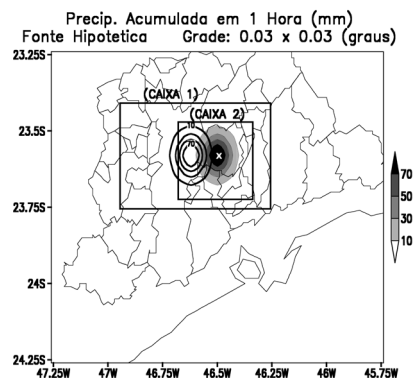
Apesar de sempre positivo,  $D_{PIELKE}$  não é um valor limitado pelo lado positivo do conjunto dos números reais, como é o IC da Equação 19. Portanto, uma desvantagem do índice  $D_{PIELKE}$  em relação ao índice de concordância (IC) da Equação 19, proposto por Willmot (1982), é que  $D_{PIELKE}$  não é limitado.

Willmot (1982) ressalta que índices que variam dentro de limites bem definidos são mais úteis para comparação entre diferentes modelos e para comparações entre diferentes simulações de um único caso, obtidas com um mesmo modelo. No entanto, como vantagem,  $D_{PIELKE}$  considera explicitamente a razão entre os desvios padrão [Equação 20(1)]. No presente trabalho, o índice  $D_{PIELKE}$  será testado e comparado ao índice de concordância (IC) da Equação 19 para o caso de 04 de fevereiro de 2004.

Um aspecto adicional na questão dos erros em simulações explícitas de convecção deve ainda ser abordado. Lembrando mais uma vez que o realismo de uma simulação é afetado pela componente estocástica das simulações de mesoescala, que aumenta conforme a escala espacial do domínio numérico decresce (Anthes, 1983), as estatísticas de verificação podem indicar um resultado pobre ainda que as simulações sejam bastante realísticas (Anthes et al., 1989). O exemplo hipotético a seguir ilustra este paradoxo.

### 2.2.4. Matriz de correlações

A Figura 2 mostra uma distribuição espacial hipotética de precipitação acumulada acima de 10 mm em 1 hora, em sombreado, a qual emula um campo observado de superfície. Nesta figura, o domínio mostrado representa uma grade com 3 km de espaçamento entre seus pontos. A previsão de precipitação acima de 10 mm, também hipotética e para o mesmo horário de acumulação, é mostrada na mesma figura (contornos contínuos). O campo de previsão foi obtido do campo de observação



**Figura 2** - Distribuição espacial hipotética de precipitação acumulada à superfície acima de 10 mm em 1 hora (sombreado conforme indicado pela escala de tons cinzas) na Zona Leste da Cidade de São Paulo e previsão hipotética da mesma variável (contornos) com erro apenas em fase. Grade de 3 km de espaçamento. As áreas escolhidas para cálculos de análises estatísticas neste caso hipotético são os retângulos internos, que indicam os limites das áreas nas quais o coeficiente de correlação espacial entre o campo hipotético (o qual emula um campo observado) e o previsto é positivo (CAIXA 1) e negativo (CAIXA 2). Mais detalhes são encontrados no texto. A letra X marca a posição, também hipotética, de um pluviômetro de superfície.





mostrada a comparação entre duas simulações, que diferem quanto ao grau do esquema advectivo de momento, em vista da maior sensibilidade do modelo a estas variações.

Inicialmente, deve-se notar que, no ARPS, as variáveis meteorológicas são decompostas em um termo relativo ao estado básico da atmosfera no instante inicial e um termo de perturbação deste estado básico, como mostrado nas Equações 25. Nestas equações,  $u$ ,  $v$  e  $w$  são, respectivamente, os componentes zonal, meridional e vertical do vento,  $\theta$  a temperatura potencial,  $p$  a pressão atmosférica,  $\rho$  a densidade do ar,  $q_v$  a razão de mistura do vapor d'água e  $q_{ii}$  a razão de mistura total de água nas fases líquida e sólida. Nestas equações, a barra sobrescrita indica o estado básico da variável, considerada horizontalmente homogênea e, portanto, dependente apenas da sua distribuição vertical no tempo inicial, enquanto o apóstrofo indica perturbação da variável em relação ao seu estado básico. Observe-se que os estados básicos de  $w$  e de  $q_{ii}$  são considerados nulos.

$$\begin{aligned} u(x, y, z, t) &= \bar{u}(z) + u'(x, y, z, t) \\ v(x, y, z, t) &= \bar{v}(z) + v'(x, y, z, t) \\ w(x, y, z, t) &= w'(x, y, z, t) \\ \theta(x, y, z, t) &= \bar{\theta}(z) + \theta'(x, y, z, t) \\ p(x, y, z, t) &= \bar{p}(z) + p'(x, y, z, t) \\ \rho(x, y, z, t) &= \bar{\rho}(z) + \rho'(x, y, z, t) \\ q_v(x, y, z, t) &= \bar{q}_v(z) + q'_v(x, y, z, t) \\ q_{ii}(x, y, z, t) &= q'_{ii}(x, y, z, t) \end{aligned} \quad (25)$$

As variáveis prognósticas no ARPS são,  $u$ ,  $v$ ,  $w$ ,  $\theta'$  e  $q_{\psi}$ , onde o índice  $\psi$  refere-se aos 6 tipos de espécies da substância água no modelo (vapor d'água, água de nuvem, água de chuva, cristal de gelo, neve e granizo). Além destas, o modelo resolve também uma equação para  $p'$ , sendo que os valores de  $\rho$  são posteriormente diagnosticados por meio da aplicação da equação de estado (Equação 26). Na Eq. 26,  $\varepsilon = R_d/R_v \approx 0.622$ , onde  $R_d$  e  $R_v$  são, respectivamente, as constantes dos gases para o ar seco e vapor d'água e  $T$ , a temperatura do ar.

$$\rho = \frac{p}{R_d T} \left( 1 - \frac{q_v}{\varepsilon + q_v} \right) (1 + q_v + q_{ii}) \quad (26)$$

O termo de advecção é importante, pois influencia diretamente a evolução temporal das variáveis prognosticadas pelo modelo em cada ponto de grade. Tomando-se como exemplo a equação prognóstica da variável escalar  $\theta'$  (Equação 27), vê-se que a mesma é composta pelo termo de advecção da perturbação (1º. termo do lado direito da equação), mais o termo de advecção vertical do seu estado básico (2º. termo do lado direito da equação), acrescida dos termos devidos à mistura turbulenta ( $D_\theta$ ) e fontes ou sorvedouros da variável ( $S_\theta$ ).

$$\frac{\partial(\rho\theta')}{\partial t} = - \left[ u \frac{\partial\theta'}{\partial x} + v \frac{\partial\theta'}{\partial y} + w \frac{\partial\theta'}{\partial z} \right] - \left[ \rho w \frac{\partial\bar{\theta}}{\partial z} \right] + D_\theta + S_\theta \quad (27)$$

Equações prognósticas semelhantes à Equação 27 são atribuídas às outras variáveis prognósticas escalares ( $q_{\psi}$  e  $p'$ ). Para as variáveis associadas a momento, tem-se:

$$\begin{aligned} \frac{\partial(\rho u)}{\partial t} &= - \left[ u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} \right] - \\ &- \left[ \frac{\partial}{\partial x}(\rho') + \frac{\partial}{\partial z}(\rho') \right] + [\rho f_x v - \rho f_z w] + D_u \end{aligned} \quad (28)$$

para o componente zonal do vento;

$$\begin{aligned} \frac{\partial(\rho v)}{\partial t} &= - \left[ u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} \right] - \\ &- \left[ \frac{\partial}{\partial y}(\rho') + \frac{\partial}{\partial z}(\rho') \right] - \rho f_x u + D_v \end{aligned} \quad (29)$$

para o componente meridional do vento;

$$\begin{aligned} \frac{\partial(\rho w)}{\partial t} &= - \left[ u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} \right] - \\ &- \frac{\partial}{\partial z}(\rho') + \rho B + \rho f_z u + D_w \end{aligned} \quad (30)$$

para o componente vertical do vento.

Nas Equações. 28 a 30, os primeiros termos do lado direito das equações referem-se à advecção da variável prognosticada. Os termos que envolvem  $p'$  referem-se à força do gradiente horizontal ou vertical de pressão. Há também termos associados à força de Coriolis, que são função de  $f$ , o parâmetro de Coriolis (com  $f_x = 2\Omega \sin \varphi$  e  $f_z = 2\Omega \cos \varphi$ , sendo  $\Omega$  a velocidade angular da Terra e  $\varphi$  a latitude do ponto considerado). Os termos de dissipação turbulenta e molecular de momento estão representados pela letra D, com subscrito indicando a variável a que se refere o termo. Na Equação 30, B é o termo de flutuação das parcelas de ar, mostrado na Equação 31, onde as variáveis são as mesmas já apresentadas. A variável  $c_s$ , dada pela Equação 32, é a velocidade das ondas de som, dependentes da temperatura do ar T e  $\gamma$ , dada pela Equação 33, a razão entre o calor específico do ar à pressão constante ( $c_p$ ) e a volume constante ( $c_v$ ).

$$B = -g \frac{\rho'}{\bar{\rho}} = g \left[ \frac{\theta'}{\bar{\theta}} - \frac{p'}{\bar{\rho} c_s^2} + \frac{q'_v}{\varepsilon + \bar{q}_v} - \frac{q'_v + q'_{ii}}{1 + \bar{q}_v} \right] \quad (31)$$

$$c_s = \sqrt{\gamma R_d \bar{T}} \quad (32)$$

$$\gamma \equiv \frac{c_p}{c_v} \quad (33)$$

As equações contínuas para o prognóstico de momento (Equações 28 a 30) são discretizadas no ARPS com o uso de técnicas de diferenças finitas centradas. O ARPS emprega a técnica de integração "time splitting" proposto por Klemp e Wilhelmson (1978). Segundo esta técnica, os termos das equações diferenciais discretas, que geram ondas de gravidade e de som (estas últimas são devidas ao fato de que as equações

descrevem uma atmosfera compressível) em suas soluções, são integrados em passos de tempo que são submúltiplos (“small time step” em Inglês) do passo de tempo empregado na integração dos termos numericamente estáveis. Os termos de advecção não geram ondas de som ou de gravidade de alta frequência e, por este motivo, podem ser integrados no passo de tempo maior (“big time step” em Inglês).

O esquema numérico empregado na integração dos termos de advecção é o explícito centrado no espaço. O truncamento de menor ordem num esquema numérico centrado é o de grau 2, o qual apresenta erro de truncamento proporcional ao quadrado do passo utilizado. No esquema centrado no espaço, o valor de uma variável qualquer  $\Phi$  no instante  $t$  num ponto de grade é obtido a partir das informações de pontos de grade vizinhos no mesmo instante  $t$ . Para simplificar a notação, define-se inicialmente a diferença no espaço ( $\delta$ ) de uma variável dependente  $\Phi = \Phi(x, y, z)$  para uma dada coordenada  $s$  ( $s = x, y$  ou  $z$  em coordenadas cartesianas) como:

$$\delta_{ns} \Phi = [\Phi(s + n\Delta s/2) - \Phi(s - n\Delta s/2)] / (n\Delta s) \quad (34)$$

onde  $n$  é um número inteiro conveniente associado ao grau do esquema numérico.

As variáveis do modelo ARPS são distribuídas segundo o definido para uma grade C de Arakawa (Xue et al., 1995). Neste tipo de grade, as variáveis de momento e de estado não são todas definidas num mesmo ponto, mas em pontos alternados da grade tridimensional ou a meio caminho entre dois pontos de grade. Mostra-se que este procedimento garante a conservação de energia no interior da grade numérica (Xue et al., 1995), um princípio básico fundamental em modelagem numérica. Ao se efetuarem cálculos com as variáveis prognósticas, são primeiramente obtidas médias espaciais destas variáveis sobre os eixos coordenados. Nas notações a seguir, serão omitidas as indicações de médias espaciais, sem perdas de generalidades. Pode-se discretizar o termo de advecção num esquema numérico de segunda ordem para uma variável qualquer  $\Phi$  como na Equação 35:

$$-\vec{V} \circ \vec{\nabla} \Phi = u\delta_{1x} \Phi + v\delta_{1y} \Phi + w\delta_{1z} \Phi \quad (35)$$

Por sua vez, os termos de advecção de 4ª. ordem no ARPS são escritos como a média ponderada de dois termos:

$$-\vec{V} \circ \vec{\nabla} \Phi = \frac{4}{3} [u\delta_{1x} \Phi + v\delta_{1y} \Phi + w\delta_{1z} \Phi] - \frac{1}{3} [u\delta_{2x} \Phi + v\delta_{2y} \Phi + w\delta_{2z} \Phi] \quad (36)$$

Ambas as formulações (Equações 35 e 36) são quadraticamente conservativas (Xue et al., 1995). Para fluxos não constantes, pode-se mostrar que o erro de truncamento no esquema de 4ª. ordem é proporcional ao gradiente do campo de velocidade (Xue et al., 1995). O esquema numérico de 4ª.

ordem é, em teoria, mais preciso que o de 2ª. ordem em relação ao erro de truncamento (Haltiner e Willians, 1980). Haltiner e Willians (1980) discutem os erros de truncamento envolvidos em esquemas numéricos comumente utilizados em modelagem numérica da atmosfera e mostram que o esquema advectivo de 2ª. ordem tem erro de truncamento da ordem de  $O[(\Delta s)^2]$ , enquanto que o esquema centrado de 4ª. ordem apresenta erro de truncamento da ordem de  $O[(\Delta s)^4]$ . No entanto, o menor erro de truncamento, no esquema centrado de 4ª. ordem, vem associado ao aumento do tempo computacional necessário para efetuar um mesmo período de integração numérica, em relação ao esquema de 2ª. ordem.

### 3. RESULTADOS

Embora sejam diversos os locais de acumulação de águas pluviais à superfície na RMSP, a Zona Leste da cidade de São Paulo é especialmente afetada pelos eventos convectivos localizados de longa duração. Pereira Filho et al. (2008) destacam a atuação de ilha de calor urbana naquela área como fator pré-condicionante favorável aos eventos de forte intensidade, como parece ter sido o caso de 04 de fevereiro de 2004, apresentado a seguir.

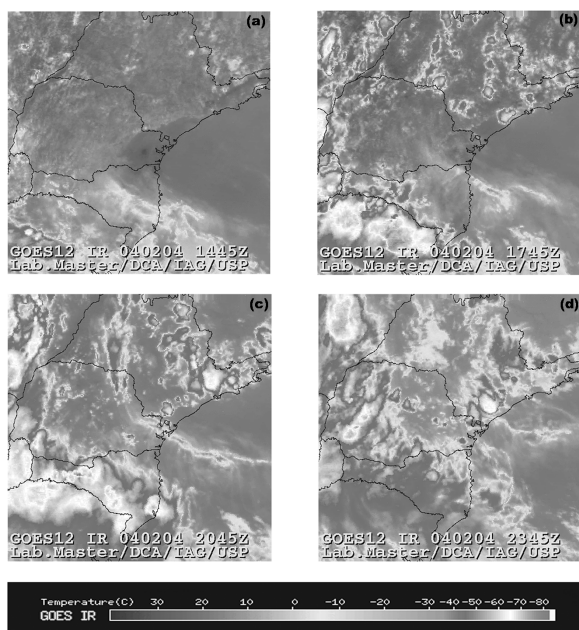
#### 3.1 Caracterização do caso de 04 de fevereiro de 2004

O evento de 04 de fevereiro de 2004 (de agora em diante referido como 04FEV2004) foi considerado bastante severo e está documentado em Barros et al. (2004). A área mais afetada pelas chuvas na RMSP naquele dia foi a Zona Leste da Cidade de São Paulo. Hallak (2007) verificou subjetivamente a existência de erros consideráveis de fase e amplitude nos campos de precipitação acumulada à superfície, simulados pelo modelo ARPS, quando comparados aos campos de precipitação estimados pelo RADSP, devido à alta complexidade das interações físicas que ocorreram durante a formação dos sistemas convectivos naquele evento. Basicamente, múltiplas interações entre diferentes escoamentos nos níveis mais baixos da atmosfera, como frentes de rajadas, brisa marítima e escoamento vale-montanha, influenciaram sobremaneira a posição das células convectivas naquele dia.

A Figura 3 mostra a seqüência de imagens do satélite GOES-12 no canal infravermelho para o dia 04FEV2004. Às 1445 UTC (Figura 3a), observa-se presença generalizada de nebulosidade baixa sobre o continente. Na região costeira do Estado de São Paulo, o aquecimento superficial já havia diluído a nebulosidade matutina naquela área por intermédio de movimentos térmicos turbulentos. Às 1745 UTC (Figura 3b), alguns sistemas convectivos em desenvolvimento destacavam-se em São Paulo, enquanto uma frente fria é observada no Rio

Grande do Sul, a qual avançou para Santa Catarina nas horas seguintes, com perda de intensidade.

O sistema convectivo que atingiu a Zona Leste da Capital é identificado em sua máxima atividade na imagem das 2045 UTC (Figura 3c). Este sistema foi responsável pelo evento de enchente discutido em Barros et al. (2004). Na mesma figura, identificam-se formações convectivas que se alinham à costa do Estado de São Paulo, aparentemente associadas à atuação da frente de brisa marítima no período da tarde e início da noite daquele dia. No entanto, o alinhamento da convecção paralelo à costa não foi predominante por todo o período e se desconfigurou nas horas seguintes, conforme se observa às 2345 UTC (Figura 3d). A sequência temporal das estimativas RADSP, para a tarde de 04FEV2004 (Figura 4), mostra que as primeiras células de precipitação surgiram pelo aquecimento superficial em associação ao levantamento forçado de parcelas de ar pela topografia (observem-se as Figuras 4a e 4b em comparação à Figura 1b). O movimento de propagação da frente de brisa marítima em direção ao interior do continente tem contribuição no desenvolvimento de sistemas de precipitação na Zona Leste da Capital a partir das 1930 UTC (Figura 4c), os quais se intensificam nos minutos seguintes (Figura 4d). O estacionamento do sistema convectivo na área da Zona Leste da Capital por quase 2 horas, como observado nas estimativas RADSP (Figuras 4c a 4f), pode ter sido motivado pelos efeitos da presença da ilha de calor urbana sobre a Zona Leste, fator que



**Figura 3** - Imagens do satélite GOES-12 no canal infravermelho em alta resolução espacial (4 km) em 04 de fevereiro de 2004 para (a) 1445 UTC, (b) 1745 UTC, (c) 2045 UTC e (d) 2345 UTC. Estão indicados contornos geopolíticos. Os tons de cinza estão associados com temperaturas do pixel segundo a barra de escalas. Fonte: Laboratório Master/DCA/IAG/USP.

também intensificaria o sistema convectivo. Na sequência de campos da Figura 4 observa-se que há alinhamento dos sistemas precipitantes na direção noroeste/sudeste do domínio a partir das 2000 UTC (Figura 4d), em concordância ao escoamento local predominante na baixa e média troposfera (campos não mostrados).

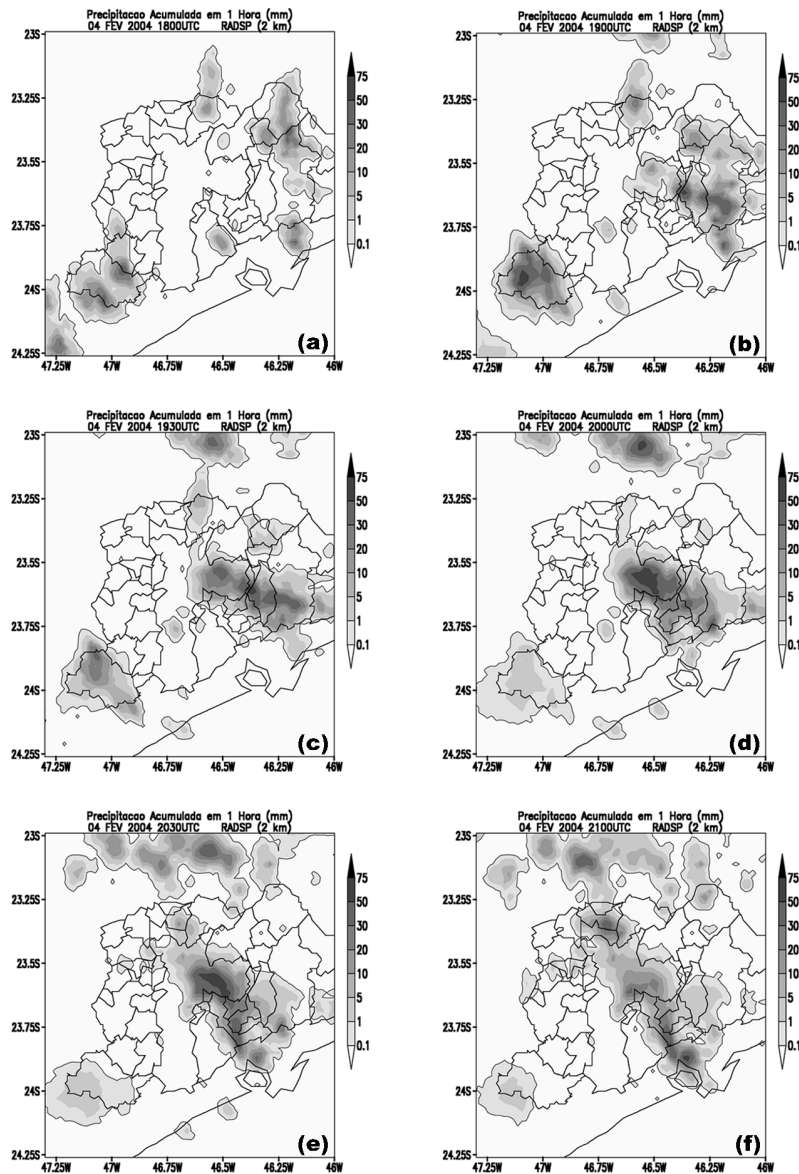
### 3.2. Erros e índices de destreza nas simulações de 1 km de espaçamento

Para os testes de sensibilidade na grade de 1 km, as simulações controle de menor resolução espacial (grades de 27, 9 e 3 km) foram mantidas constantes, aproveitando-se da vantagem proporcionada pelo aninhamento unidirecional de grades no ARPS. Este procedimento permite a execução de testes de sensibilidade, iniciando-se o modelo diretamente na grade desejada com o mesmo conjunto de condições iniciais e de fronteira, sem que se alterem os resultados das grades exteriores. Posteriormente, testes de sensibilidade poderiam ser efetuados com alterações específicas nas grades de menor resolução espacial avaliando-se o impacto consequente destas alterações sobre os resultados explícitos de convecção na grade de 1 km de espaçamento.

Desde que os testes de sensibilidade do modelo, quanto à precipitação acumulada à superfície, podem ser executados em variados períodos de acumulação, aqui serão mostradas apenas as avaliações correspondentes aos campos de precipitação acumulada entre 1700 UTC e 2100 UTC de 04FEV2004, que engloba o período de maior atividade convectiva naquele dia na RMS (Figura 4). A Figura 5a mostra o campo de precipitação acumulada obtido pelas estimativas RADSP, regrada para 1 km pelo método de interpolação bilinear a partir do campo original de 2 km para fins de comparação com as simulações. A Figura 5a mostra também os subdomínios SD 1, SD 2 e SD 3 sobre os quais foram computados os índices e erros para comparação objetiva com os experimentos numéricos de 1 km de espaçamento descritos na Tabela 3, sendo que SD 1 corresponde ao domínio total da grade numérica de 1 km.

A simulação controle (CTL) para a grade de 1 km de espaçamento horizontal foi configurada com 179 x 179 pontos em latitude e longitude e 72 níveis na vertical, com topo a aproximadamente 29 km acima da superfície e com camada de absorção de ondas de alta frequência nos níveis superiores a 20 km de altitude (Tabela 1). O experimento CTL foi também configurado com esquema de advecção de 4ª ordem para variáveis de momento e de 2ª ordem, para as escalares. Em adição, CTL foi definida com 15 pontos nas fronteiras laterais e formulação completa da física atmosférica [como radiação de ondas longas e curtas formulada pelo Goddard Space Flight Center, inclusão da força de Coriolis, microfísica de nuvens de

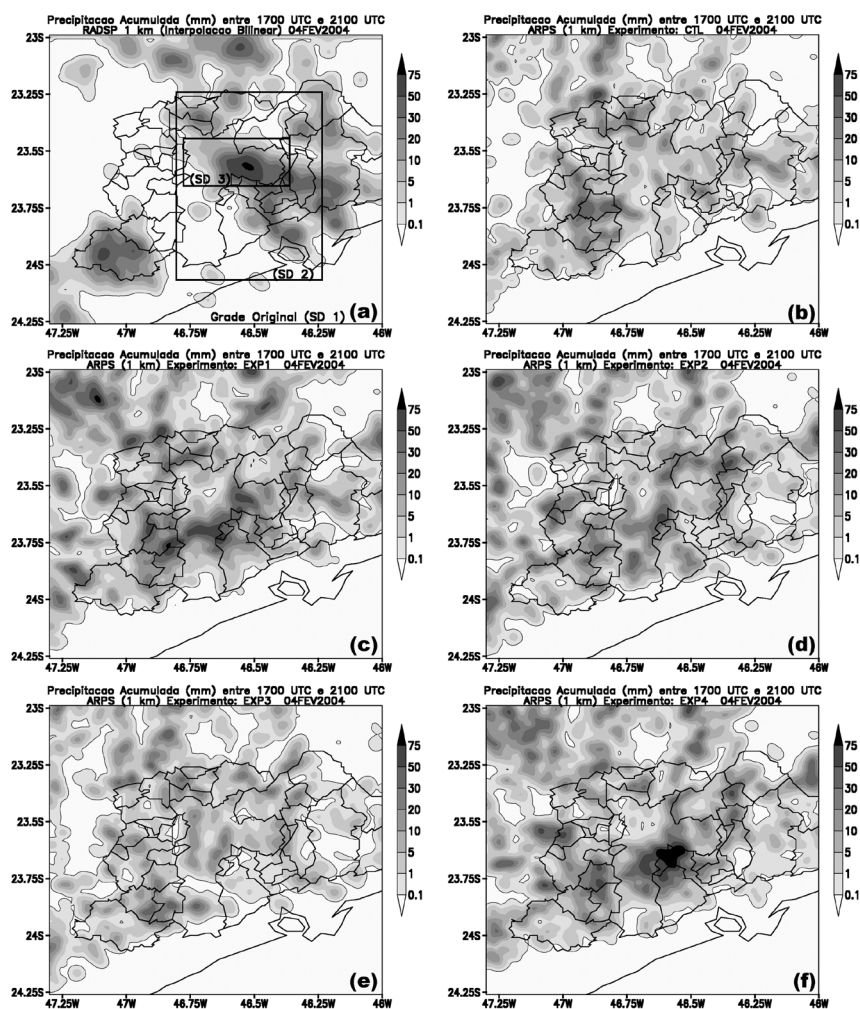




**Figura 4** - Série temporal dos campos de precipitação acumulada em 1 hora (mm) na RMSP e proximidades calculados a partir das estimativas RADSP em 04FEV2004 para (a) 1800 UTC, (b) 1900 UTC, (c) 1930 UTC, (d) 2000 UTC, (e) 2030 UTC e (f) 2100 UTC. Valores sombreados de acordo com a barra de escalas.

**Tabela 3** - Configurações dos experimentos numéricos de 1 km para testes de sensibilidade.

Código dos Experimentos	Experimentos Numéricos (GRADE: 1 km)			
	Ordem dos Esquemas de Advecção		ADAS 12 UTC Sondagem	ADAS 18 UTC Superfície
	Momento	Escalar		
CTL	4 <sup>a</sup>	2 <sup>a</sup>	Não	Não
EXP1	4 <sup>a</sup>	2 <sup>a</sup>	Não	Sim
EXP2	4 <sup>a</sup>	2 <sup>a</sup>	Sim	Sim
EXP3	4 <sup>a</sup>	2 <sup>a</sup>	Sim	Não
EXP4	2 <sup>a</sup>	2 <sup>a</sup>	Sim	Sim



**Figura 5** - Precipitação acumulada (mm) entre 1700 UTC e 2100 UTC de 04FEV2004. (a) Campo RADSP regradeado para 1 km pelo método de interpolação bilinear, derivado do campo original de 2 km de espaçamento. Experimentos ARPS para o mesmo período e variável que em (a) conforme descritos na Tabela 3, correspondentes a (b) experimento controle (CTL), (c) EXP1, (d) EXP2, (e) EXP3 e (f) EXP4. Os retângulos em destaque em (a) indicam os subdomínios nos quais foram calculados os erros e índices de acurácia (SD 1 corresponde ao domínio total da grade de 1 km de espaçamento).

Lin et al. (1983), modelo de solo de 2 camadas, parametrização de camada limite planetária com fechamento de turbulência de ordem 1.5 e características fisiográficas de 1 km de espaçamento horizontal]. Observe-se que não houve assimilação de dados adicionais por meio do pacote de análise objetiva ADAS na simulação CTL (Tabela 3).

A partir do experimento CTL, foi testada a sensibilidade do modelo a outras configurações, mantendo-se as características básicas mostradas na Tabela 1. A Tabela 3 mostra o código dos experimentos e as alterações de configuração efetuadas que os diferem do experimento CTL. Assim, o EXP1 é o mesmo que o CTL, exceto para a assimilação incremental de dados de estações de superfície (Figura 1b) durante a execução do

modelo, com maior peso central às 1800 UTC (Bloom et al., 1996), codificado como ADAS 18 UTC na Tabela 3. O EXP2 é o mesmo que o EXP1, acrescentando a assimilação de dados da sondagem de Campo de Marte às 1200 UTC na condição inicial, codificado como ADAS 12 UTC na Tabela 3. O EXP3 é o mesmo que CTL, exceto pelo uso isolado de ADAS 12 UTC na condição inicial. Finalmente, o EXP4 é o mesmo que EXP2, exceto pela configuração dos esquemas de advecção de momento e de escalares, ambos de 2ª ordem no EXP4. No esquema ADAS 18 UTC, a assimilação incremental de dados de estações de superfície se faz durante a execução do modelo, o que garante o constante ajuste das variáveis do modelo aos valores observados, com mínima geração de ondas espúrias de

alta frequência (Bloom et al., 1996). A janela de assimilação foi definida com 2 horas de integração incremental centrada no horário das 1800 UTC, com intervalos de incremento a cada 2 passos de tempo de integração.

As Figuras 5b a 5f mostram que todos os experimentos, incluindo o experimento controle (CTL), apresentam maior dispersão dos sistemas precipitantes no domínio numérico em relação ao observado (Figura 5a) e com menores quantidades acumuladas. Ainda, centros observados de altíssima acumulação, como na Zona Leste da Cidade de São Paulo, não foram adequadamente capturados nas simulações.

A Tabela 4 mostra os índices de acurácia e erros calculados para a totalidade da grade simulada (SD 1) e subdomínios SD 2 e SD 3 da Figura 5a. Os valores mostrados na Tabela 4 correspondem à diferença entre os volumes totais estimados e simulados ( $\Delta V$ ), viés, MAE,  $\sigma_S/\sigma_o$ ,  $RMSE_{TOT}$ , raízes quadradas de  $MSE_{DISS}$  e  $MSE_{DISP}$ , coeficiente de correlação espacial  $\rho$ , índice de concordância IC e índice de destreza  $D_{PIELKE}$ . Os resultados da Tabela 4 geram um grande número de análises, das quais serão aqui reportadas aquelas consideradas mais relevantes para este trabalho.

O maior índice de concordância e menor índice de destreza observados na Tabela 4 referem-se ao EXP1 (apenas acréscimo de ADAS 18 UTC em relação ao CTL), para os subdomínios SD 2 e 3. Em SD 1, o EXP1 obteve o maior IC, mas não o menor  $D_{PIELKE}$ , pois a acurácia de EXP1 foi penalizada pelo excessivo  $RMSE_{TOT}$ . Por outro lado, todos os experimentos obtiveram coeficiente de correlação espacial maior que zero, embora muitos deles bastante pequenos. Como consequência, o componente dispersivo dos erros é quase sempre maior do que o dissipativo, indicando que os maiores problemas na simulação de precipitação no caso de 04FEV2004 estão associados à fase dos sistemas. De fato, alguns experimentos chegam a mostrar

erros de fase 3 vezes superiores aos associados à amplitude. A Tabela 4 mostra também, que os erros associados à amplitude dos sistemas crescem quando se considera o subdomínio SD 3, onde se analisa a área em que os sistemas convectivos mais intensos foram observados por meio do RADSP.

Em termos de volume total de precipitação para toda a grade (SD 1), o EXP2, que assimilou a sondagem das 1200 UTC e também os dados de superfície das 1800 UTC, obteve um valor bastante próximo ao estimado pelo RADSP, o que pode ser considerado um resultado útil em termos de manejo hidrológico de bacias. O acréscimo da sondagem das 1200 UTC na inicialização do EXP1 (dando origem ao EXP2) contribuiu moderadamente para a melhora do desempenho do modelo para o caso de 04FEV2004 no maior subdomínio (SD 1), com tendência (viés) próxima de zero. Ao mesmo tempo, o EXP2 tem o menor IC entre os 5 experimentos, penalizado pelo baixo coeficiente de correlação espacial. Contudo,  $D_{PIELKE}$  parece ter capturado a maior acurácia geral de EXP2 em relação ao experimento EXP1.

Note-se que os subdomínios SD 2 e SD 3 desconsideram os pontos da grade que estão sobre o Oceano Atlântico, onde não se observou nem se simulou precipitação (Figura 5). Com isso, houve tendência do  $RMSE_{TOT}$  crescer nestes subdomínios em relação ao SD 1. É interessante notar que o EXP3, o qual difere de CTL apenas pela assimilação ADAS 12 UTC na condição inicial, obteve maiores valores de IC em SD 2 e SD 3, quando comparados ao IC obtido em SD 1. Isto ocorre apesar de EXP3 ter obtido os maiores valores nos erros (compare a coluna EXP3 nos subdomínios SD 2 e SD 3). Neste caso específico,  $D_{PIELKE}$  parece identificar com maior eficiência o desempenho inferior de EXP3 nas áreas analisadas.

No subdomínio SD 3, o EXP1 atingiu o máximo IC (IC = 0.717) e o índice de destreza muito próximo de 2.0 demonstra a qualidade dos resultados, apesar da razão relativamente baixa

**Tabela 4** - Erros e índices estatísticos da precipitação acumulada entre 1700 UTC e 2100 UTC de 04FEV2004 na grade de 1 km calculados para os subdomínios SB 1,2 e 3 definidos na Figura 5a.

ÍNDICES E ERROS	VALORES OBTIDOS PARA CADA EXPERIMENTO														
	SUBDOMÍNIO SD 1					SUBDOMÍNIO SD 2					SUBDOMÍNIO SD 3				
	CTL	EXP1	EXP2	EXP3	EXP4	CTL	EXP1	EXP2	EXP3	EXP4	CTL	EXP1	EXP2	EXP3	EXP4
$\Delta V$ ( $10^6$ m <sup>3</sup> )	0.415	0.253	0.020	0.468	0.237	0.280	0.029	0.107	0.270	0.045	0.173	0.081	0.124	0.154	0.050
BIAS (mm)	2.2	-1.3	-0.1	2.4	-1.2	5.3	0.7	2.1	5.1	-0.7	17.5	8.4	12.3	15.7	4.5
MAE (mm)	6.3	8.6	8.1	6.1	8.7	9.1	10.4	11.4	9.3	12.4	20.0	15.7	20.9	20.5	22.6
$\sigma_S/\sigma_o$ (adim)	0.60	1.10	0.86	0.53	1.13	0.43	0.89	0.68	0.43	1.15	0.21	0.68	0.52	0.30	1.13
$RMSE_{TOT}$ (mm)	13.0	15.9	14.7	12.5	16.2	17.0	17.7	18.6	16.8	21.6	29.5	21.4	28.9	28.9	32.7
$(MSE_{DISS})^{1/2}$ (mm)	5.0	1.7	1.6	5.9	1.9	10.1	1.8	5.3	10.2	2.5	26.3	11.5	17.5	23.3	5.6
$(MSE_{DISP})^{1/2}$ (mm)	12.0	15.8	14.7	11.1	16.1	13.6	17.6	17.8	13.8	21.4	14.5	17.9	23.5	18.3	32.5
$\rho$ (adim)	0.046	0.095	0.004	0.080	0.084	0.079	0.249	0.001	0.099	0.145	0.082	0.572	0.030	0.004	0.153
IC (adim)	0.322	0.371	0.294	0.348	0.348	0.377	0.530	0.335	0.382	0.417	0.479	0.717	0.448	0.463	0.480
$D_{PIELKE}$ (adim)	2.699	2.923	2.768	2.677	3.015	2.738	2.436	2.744	2.725	2.980	3.067	2.070	2.832	2.967	2.909

entre os desvios padrão simulado e observado (0.68). De acordo com a matriz de correlações, o coeficiente de correlação espacial de 0.572 atinge o máximo de 0.610 a apenas 2 km ao norte do simulado para este experimento. Para o EXP2, por exemplo, que obteve  $\rho = 0.030$  em SD 3, a matriz de correlações mostra um máximo de 0.450 com um deslocamento de 7 km para norte e 9 km para oeste. Neste caso, o componente dissipativo de RMSE passaria de 15.5 mm para 15.1 mm e o dispersivo, de 23.5 mm para 18.3 mm (uma diminuição relativa de 22%). Em geral, no entanto, as falhas do modelo ainda se devem à maior contribuição do erro associado ao termo dispersivo de MSE.

### 3.3. Comparação quanto ao esquema de advecção de momento

Os experimentos EXP2 e EXP4 são diretamente comparáveis, pois diferem somente quanto à especificação da ordem do esquema de advecção de momento (Tabela 3). A inspeção dos resultados da Tabela 4 indica que, com um esquema mais apurado para a advecção de momento, o EXP2 apresenta menores erros quadráticos médios nos três subdomínios, bem como menores erros absolutos. Os maiores efeitos do uso de um esquema de diferenças finitas de maior precisão quanto às variáveis de momento aparecem no componente dispersivo de MSE, nos 3 subdomínios. No entanto, o erro sistemático foi maior no esquema de 4ª. ordem em todos os subdomínios. Quanto aos índices de destreza, o melhor ajuste das fases dos sistemas precipitantes no EXP2 é refletido no menor índice  $D_{PIELKE}$  para este experimento. Novamente o índice  $D_{PIELKE}$  parece capturar com maior propriedade um aumento na acurácia da simulação EXP2 em relação ao EXP4 do que o índice de concordância IC.

## 4. CONCLUSÕES

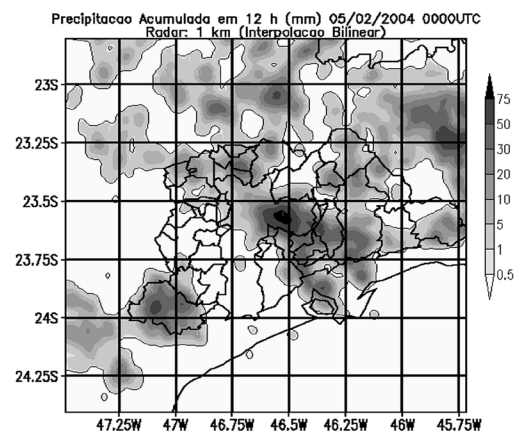
As soluções do modelo ARPS para o caso de 04FEV2004 tenderam ao desenvolvimento de precipitação acumulada em 4 horas com maior distribuição espacial do que o observado, obtendo-se indicações de que os sistemas convectivos desenvolvidos na grade de 1 km de espaçamento possuíam maior velocidade de deslocamento do que o observado. As análises dos erros e índices de acurácia indicaram significativa contribuição do termo de advecção das equações primitivas, no controle da distribuição espacial dos fenômenos observados nos baixos níveis da atmosfera.

A assimilação incremental dos dados observados de superfície às 1800 UTC (EXP1), efetuada durante a execução do modelo, mesmo obtida com dados de estações com má distribuição espacial (Figura 1b), atua na correção de boa parte da fase dos sistemas precipitantes. Os melhores resultados para destreza dentre os experimentos propostos foram obtidos

no subdomínio SD 3 do EXP1, que atingiu o máximo IC ( $IC = 0.717$ ) e o índice de destreza  $D_{PIELKE}$  muito próximo de 2.0, enquanto o coeficiente de correlação espacial foi de 0.610 a apenas 2 km a norte do simulado. Este resultado leva à questão da necessidade de se instalar uma rede de estações meteorológicas de superfície de mesoescala que cubra de forma adequada a RMSP e suas vizinhanças, e cujas medidas estejam disponíveis em tempo real para uma melhor iniciação dos modelos numéricos regionais.

Os resultados preliminares parecem também indicar que quanto menor a escala espacial dos fenômenos analisados, maior a restrição ao tamanho das áreas sobre as quais são calculados os erros espaciais. Por outro lado, é recomendável que os subdomínios de análise de erros espaciais estejam apropriadamente distantes das fronteiras do domínio numérico, que é uma região afetada pelas inconsistências inerentes às grades de diferentes resoluções espaciais. Estes resultados sugerem que testes de sensibilidade por intermédio de cálculos de índices de acurácia e erros sejam efetuados em vários subdomínios de áreas equivalentes, e que cubram, por exemplo, a maior parte da grade simulada, como apresentado na Figura 6, onde os subdomínios possuem 25 km<sup>2</sup> cada. Assim, a sensibilidade de cada porção do domínio numérico poderia ser determinada em função de variações em diversos parâmetros físicos do modelo.

De uma forma geral, os resultados iniciais da pesquisa indicam que a metodologia proposta é promissora para os objetivos estabelecidos. Note-se que o aumento da resolução espacial dos modelos numéricos, certamente acrescenta detalhes importantes às simulações, assim como, aumenta também o grau de complexidade dos fenômenos atmosféricos envolvidos. Quanto a esta questão, o índice de destreza proposto ( $D_{PIELKE}$ ) parece responder bem ao aumento das exigências ao se considerarem fenômenos de menor escala espacial, com a



**Figura 6** - Sugestão de subdomínios predefinidos para testes de sensibilidade do modelo ARPS em relação aos erros e índices estatísticos na grade de 1 km de espaçamento.



ressalva de que falta ainda uma comprovação mais rigorosa de sua utilidade.

Sugere-se a análise de um maior número de casos sobre a influência dos esquemas de advecção na distribuição horizontal das variáveis meteorológicas, para uma compreensão mais completa do problema da baixa reprodutibilidade da fase dos sistemas de precipitação, em simulações e previsões em alta resolução espacial.

Por fim, ressalta-se que a metodologia proposta pode ser usada para um número qualquer de membros de um sistema de previsão de tempo por conjuntos, bem como, na quantificação das diferenças entre um determinado experimento numérico controle e suas variações.

## 5. AGRADECIMENTOS

Os autores agradecem à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo apoio a esta pesquisa e aos dois revisores anônimos da Revista Brasileira de Meteorologia por suas contribuições à melhoria geral deste artigo. Os autores também agradecem à Universidade de Oklahoma pelo código do sistema ARPS. Agradece-se ainda ao Departamento de Águas e Energia Elétrica (DAEE) do Estado de São Paulo, à Companhia Ambiental do Estado de São Paulo (CETESB) e ao Laboratório MASTER (IAG/USP) pela cessão dos arquivos de dados meteorológicos empregados neste trabalho.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

- ANTHES, R. A. Regional Models of the Atmosphere in Middle Latitudes. **Monthly Weather Review**, v. 111, p. 1306-1335, 1983.
- ANTHES, R. A.; KUO, Y.-A.; HSIE, E.-Y.; LOW-NAM, S.; BETTGE, T.W. Estimation of Skill and Uncertainty in Regional Numerical Models. **Quarterly Journal of the Royal Meteorological Society**, v. 115, p. 763-806, 1989.
- BARROS, M. T. L.; PEREIRA FILHO, A. J.; GANDU, A. W.; HALLAK, R.; ANDRIOLI, C. P. Impacto Hidrológico das Precipitações Observadas na Região Metropolitana de São Paulo nos Dias 04 de Fevereiro de 2004 e 29 de Janeiro de 2004. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 13, 2004, Fortaleza, CE. Anais..., 2004. CD-ROM.
- BLOOM, S. C.; TAKACS, L. L.; DA SILVA, M.; LEDVINA, D. Data Assimilation Using Incremental Analysis Updates. **Monthly Weather Review**, v. 124, p. 1256-1271, 1996.
- DALEY, R. **Atmospheric Data Analysis**. 1a Edição, Canadá, Cambridge University Press, 1991. 420 p.
- FARIAS, J. F. S. **Previsão de Chuva a Curtíssimo Prazo na Abrangência do Radar Meteorológico de São Paulo**. 2009. 82 f. Dissertação (Mestrado em Meteorologia) -Departamento de Ciências Atmosféricas, Universidade de São Paulo, São Paulo, 2009.
- FOX, D. G. Judging Air Quality Model Performance. **Bulletin of the American Meteorological Society**, v. 62, p. 599-609, 1981.
- HALLAK, R. **Simulações Numéricas de Tempestades Severas na RMSP**. 2007. 219 f. Tese (Doutorado em Ciências) - Departamento de Ciências Atmosféricas, Universidade de São Paulo, São Paulo, 2007.
- HALTNER, G. J.; WILLIAMS, R. T. **Numerical Prediction and Dynamic Meteorology**. 2a Edição, Estados Unidos da América, John Wiley & Sons, Inc., 1980. 477 p.
- KAIN, J. S.; FRITSCH, J. M. **Convective Parameterization for Mesoscale Models: The Kain-Fritsch Scheme**. The Representation of Cumulus Convection in Numerical Models. In: AMS Monograph Series, American Meteorological Society, p. 165-170, 1993.
- KLEMP, J. B.; WILHELMSON, R. B. The Simulation of Three-Dimensional Convective Storm Dynamics. **Journal of the Atmospheric Sciences**, v. 35, p. 1070-1096, 1978.
- LIN, N.-T.; FARLEY, R. D.; ORVILLE, H. D. Bulk Parameterization of the Snow Field in a Cloud Model. **Journal of Climate and Applied Meteorology**, v. 22, p. 1065-1092, 1983.
- MARSHALL, J. S.; PALMER, W. M. K. The Distribution of Raindrops with Size. **Journal of Meteorology**, v. 5, p. 165-166, 1948.
- MARTIN, W. J.; XUE, M. Sensitivity Analysis of Convection of the 24 May 2002 IHOP Case Using Very Large Ensembles. **Monthly Weather Review**, v. 134, p. 192-207, 2006.
- ORLANSKI, I. A Rational Subdivision of Scales for Atmospheric Processes. **Bulletin of the American Meteorological Society**, v. 56, p. 527-530, 1975.
- PEREIRA FILHO, A. J., BARROS, M. T. L. Flood Warning System for Megacities: A Brazilian Perspective. In: INTERNATIONAL CONFERENCE ON HIDROLOGY IN A CHANGING ENVIRONMENT, British Hydrological Society, Exeter, 1998.
- PEREIRA FILHO, A. J.; NAKAYAMA, P. T. Intercomparison of Radar Rainfall Estimates and Rain Gage Measurements in Sao Paulo, Brazil. In: FIFTH INTERNATIONAL SYMPOSIUM ON HYDROLOGICAL APPLICATIONS OF WEATHER RADAR. Kyoto, Japan, 2001.
- PEREIRA FILHO, A. J.; MASSAMBANI, O.; KARAM, H. A.; HALLAK, R.; HAAS, R.; VEMADO, F. Sistema de Previsão Hidrometeorológica para a Bacia do Alto Tietê. **Boletim da Sociedade Brasileira de Meteorologia**, v. 32, p. 42-60, 2008.

- PIELKE SR., R. A. **Mesoscale Meteorological Modeling**. International Geophysics Series. 2ª. Edição, Estados Unidos da América, Academic Press, v. 78, 2002. 676 p.
- SILVA, F. D. S.; PEREIRA FILHO, A. J.; HALLAK, R. Classificação de Sistemas Meteorológicos e Comparação da Precipitação Estimada Pelo Radar e Medida Pela Rede Telemétrica na Bacia Hidrográfica do Alto Tietê. **Revista Brasileira de Meteorologia**, v. 24, p. 292-307, 2009.
- TAKACS, L. A Two-Step Scheme for the Advection Equation with Minimized Dissipation and Dispersion Errors. **Monthly Weather Review**, v. 113, p. 1050-1065, 1985.
- WILKS, D. S. **Statistical Methods in the Atmospheric Sciences**. International Geophysics Series. 2a. Edição, Estados Unidos da América, Academic Press, v. 91, 2006. 627 p.
- WILLMOT, C. J. Some Comments on the Evaluation of Model Performance. **Bulletin of the American Meteorological Society**, v. 63, p. 1309-1313, 1982.
- XUE, M.; DROEGEMEIER, K. K.; WONG, W.; SHAPIRO, A.; BREWSTER, K. **Advanced Regional Prediction System Users Guide**. 1a. Edição, Estados Unidos da América, Universidade de Oklahoma, 1995. 380 p.
- XUE, M.; DROEGEMEIER, K. K.; WONG, V. The Advanced Regional Prediction System (ARPS) - A Multi-Scale Nonhydrostatic Atmospheric Simulation and Prediction Model. Part I: Model dynamics and verification. **Meteorological Atmospheric Physics**, v. 75, p. 161-193, 2000.
- XUE, M.; DROEGEMEIER, K. K.; WONG, V.; SHAPIRO, A.; BREWSTER, K.; CARR, F.; WEBER, D.; LIU, Y.; WANG, D.-H. The Advanced Regional Prediction System (ARPS) - A Multi-Scale Nonhydrostatic Atmospheric Simulation and Prediction Model. Part II: Model Physics and Applications. **Meteorological Atmospheric Physics**, v. 76, p. 143-165, 2001.
- XUE, M.; WANG, D.-H.; GAO, J.; BREWSTER, K.; DROEGEMEIER, K. K. The Advanced Regional Prediction System (ARPS), Storm-Scale Numerical Weather Prediction and Data Assimilation. **Meteorological Atmospheric Physics**, v. 82, p. 139-170, 2003.