



ENSEMBLE DE MODELOS PROBABILÍSTICOS DE PREVISÃO HIDROLÓGICA

Alana Renata Ribeiro

Maurício Felga Gobbi

alanar89@gmail.com

mfgobbi@gmail.com.com

PPGMNE - UFPR

Rua Francisco H. dos Santos, 81530-900, Paraná, Curitiba, Brasil

Eduardo Alvim Leite

alvim@simepar.br

SIMEPAR

Rua Francisco H. dos Santos, 81530-900, Paraná, Curitiba, Brasil

Diandra Akemi Alves Kubo

diandraakemi@gmail.com

PPGInf - UFPR

Rua Francisco H. dos Santos, 81530-900, Paraná, Curitiba, Brasil

Resumo. Neste artigo um ensemble de modelos de previsão hidrológicos probabilísticos é desenvolvido, com o objetivo de representar as incertezas inerentes ao processo determinístico de previsão de vazão. A inferência bayesiana é utilizada, por meio dos algoritmos de amostragem do tipo MCMC (Markov Chain Monte Carlo), para a construção dos modelos, pois permite a combinação entre informações prévias, sobre parâmetros de distribuições probabilísticas, com os dados de erros de diferentes modelos determinísticos de previsão. Os dados de vazão e precipitação utilizados pertencem à sub-bacia de São Bento, na bacia do rio Iguaçu, no estado do Paraná, Brasil. Um sistema de visualização dos resultados demonstra que em casos de difícil previsão, tais como cheias e estiagens, o ensemble dos modelos probabilísticos superou as abordagens determinísticas, e probabilísticas individuais. Este sistema torna-se uma ferramenta útil para o monitoramento do comportamento da vazão hidrológica.

Palavras-Chave: Inferência bayesiana, Monte Carlo via Cadeias de Markov, Previsão de vazão

1 INTRODUÇÃO

Eventos hidrológicos extremos como cheias, estiagens e seus impactos, são apenas alguns dos motivos que levam pesquisadores e empresas de monitoramento do mundo inteiro a se preocuparem com a validade dos mais diversos tipos de dados hidrometeorológicos.

A operação de reservatórios também conta com dados hidrometeorológicos coletados, como dados de nível e precipitação, e com dados calculados a partir deles, tal como os dados de vazão. Todavia, este monitoramento está sujeito à ocorrência de falhas. Um problema bastante característico é a incerteza desses valores calculados.

A análise probabilística desses dados, estimativas e previsões, é uma abordagem que traz aos pesquisadores e usuários de sistemas de informações dados secundários, relevantes principalmente no processo de tomada de decisões.

Leite (2008) avaliou o valor da informação hidrometeorológica para o alerta de cheias no município paranaense de União da Vitória. Suas análises podem ser utilizadas para estabelecer níveis de atenção, alerta e urgência sobre cheias do Rio Iguaçu que venham a causar prejuízos no município.

Em um estudo posterior Leite, Breda e Saboia (2011) montaram um sistema de previsão probabilística para União da Vitória mostrando como, de fato, a abordagem probabilística é importante na eficácia da tomada de decisão sobre a ocorrência de cheias e remoção de bens e pessoas dos locais de risco.

As previsões probabilísticas possuem a vantagem de ser mais completas, pois expressam as incertezas e permitem a derivação da melhor estimativa sob o ponto de vista do usuário, ao contrário das previsões determinísticas que são incompletas e só expressam a melhor estimativa sob o ponto de vista do previsor.

Além disso, as previsões probabilísticas são apontadas como cientificamente mais honestas, pois admitem e expressam as incertezas, permitem a emissão de avisos e alertas baseados em riscos, favorecem a tomada de decisão racional do usuário, agregam potencial para benefícios econômicos adicionais para toda a sociedade e evitam mal entendidos nas atribuições de responsabilidade entre previsor e tomador de decisão, desacoplando os processos de previsão e uso das informações (Leite, 2008).

Além disso, um meio muito utilizado nos últimos anos para se considerar os diversos tipos de incerteza no processo de previsão hidrometeorológica, de uma maneira geral, é a técnica de previsão por conjunto, ou ainda, *ensemble* de previsões.

Segundo Fan (2014) a comunidade meteorológica começou a reconhecer a incerteza existente nas previsões de modelos de previsão do tempo, e desenvolveu maneiras de lidar e expressar essa incerteza para uso em modelos de previsão de vazão (Buizza et al., 1999; Krzysztofowicz, 2001; Bowler et al., 2008; Hamill et al., 2008).

Este fato culminou na adoção das previsões por conjunto como forma de considerar a incerteza na previsão hidrológica (Chen, Yu, 2007; Demeritt et al., 2007; Davolio et al., 2008; Pappenberger et al., 2008; Reggiani, Weerts, 2008; Cloke, Pappenberger, 2009; Thiemig et al., 2010; Bao et al., 2011; Bogner, Pappenberger, 2011; Calvetti, 2011; Cuo et al., 2011; Schellekens et al., 2011; Alfieri et al., 2012).

De maneira geral, previsões por *ensemble* podem ser desenvolvidas, por exemplo, com

perturbações nas condições iniciais de modelos determinísticos, gerando uma saída do modelo para cada uma dessas condições. Ou, com perturbações físicas dos modelos, e diferentes parametrizações, mantendo-se as condições iniciais. Ou ainda, com a combinação de resultados de diferentes modelos. Resultando em previsões hidrológicas probabilísticas.

Neste trabalho, um *ensemble* de modelos probabilísticos desenvolvidos por meio de inferência bayesiana e algoritmos de amostragem MCMC (*Markov Chain Monte Carlo*) será desenvolvido com o intuito de construir apenas uma previsão probabilística de vazão para a sub-bacia de São Bento.

Os modelos probabilísticos combinados no *ensemble* consideram informações obtidas com quatro modelos determinísticos diferentes, do tipo chuva-vazão, e com a climatologia da região, obtida por meio de uma longa série de dados históricos de vazão. Dessa maneira, os resultados da previsão do modelo probabilístico por *ensemble* tornam-se completos pois incorporam todas estas informações, bem como todas as incertezas inerentes aos modelos determinísticos.

2 Estudo de Caso

Os dados de precipitação e vazão utilizados neste trabalho são provenientes de uma estação hidrometeorológica monitorada pelo SIMEPAR (Sistema Meteorológico do Paraná). Esta estação pertence à exutória da sub-bacia de São-Bento, localizada na cabeceira da bacia do Rio Iguaçu, no estado do Paraná, na região sul do Brasil (Fig. 1), com área de drenagem de 2012km² e vazão média de 34,9 m³/s.

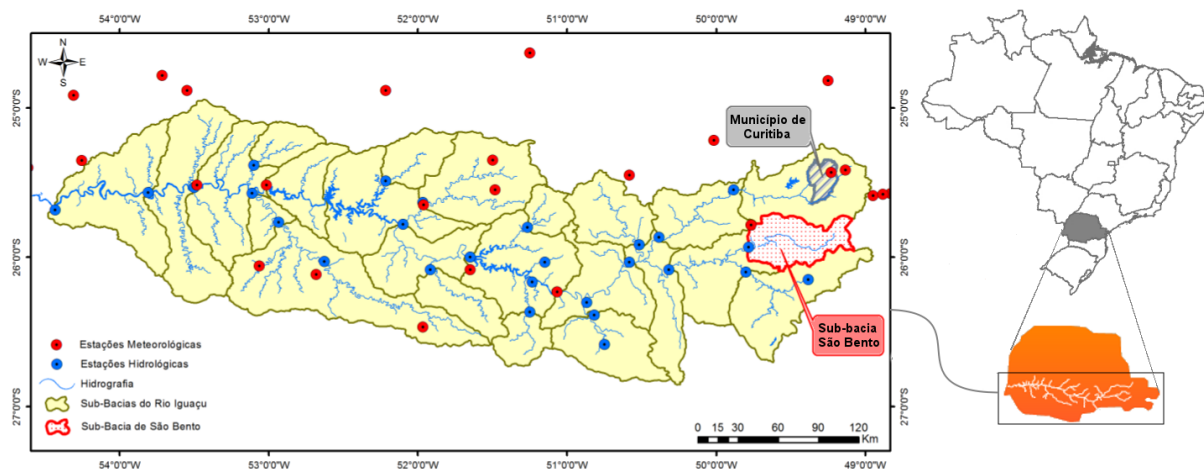


Figura 1: Localização da Sub-bacia de São Bento pertencente à bacia do rio Iguaçu, no estado do Paraná, região sul do Brasil.

Segundo Sema (2013), a bacia hidrográfica do Iguaçu possui uma área total, dentro do Estado do Paraná, de 54.820,4 Km², cerca de 28% da área total do estado, e uma população de 4.405.882 habitantes, em torno de 43% do total do estado, dos quais cerca de 2,5 milhões vivem na área metropolitana.

É nas cabeceiras da bacia hidrográfica, na região metropolitana de Curitiba, que há uma grande concentração populacional e as atividades industriais, comerciais e de serviços são as mais importantes.

3 MODELOS HIDROLÓGICOS DETERMINÍSTICOS

Modelos hidrológicos são ferramentas extremamente úteis que permitem, através da equacionalização dos processos, representar, entender e simular de forma aproximada o comportamento de bacias hidrográficas (Tucci, 1998).

A construção dos modelos probabilísticos que farão parte do *ensemble*, neste trabalho, faz uso dos resultados de quatro modelos hidrológicos determinísticos, e de um modelo construído especialmente para incluir informações climatológicas ao *ensemble* (descrito na subseção Climatologia).

Os modelos determinísticos estão subdivididos entre modelos conceituais e modelos empíricos.

Os modelos conceituais são baseados na representação dos processos físicos que envolvem o fenômeno estudado. Os modelos Sacramento (Cunha, 2004) e 3R- *rainfall-runoff-routing* (Neto et al., 2007) são exemplos de modelos conceituais e têm sido utilizados pelo SIMEPAR para a obtenção da previsão determinística de vazão.

Modelos empíricos, por sua vez, são basicamente definidos por equações matemáticas que estabelecem uma relação estável entre variáveis de entrada e saída, sem considerar neste processo quais leis físicas governam o processo natural de transformação da chuva em vazão. Dessa forma, um modelo empírico Regressivo Linear Múltiplo (Davison, 2003), e uma rede neural do tipo MLP - Multilayer perceptron (Haykin, 2001), também foram desenvolvidos para realizar previsões determinísticas de vazão.

Os modelos determinísticos chuva-vazão, Sacramento, 3R, Regressivo Linear Múltiplo e MLP, foram calibrados com dados de 05/01/1998 até 25/12/2010, e aplicados, a fim de realizar previsões de 6h em 6h, para um horizonte total de 72h, em dados de 24/05/2013 até 25/06/2016, chamado de período de verificação, totalizando 4515 casos de previsão.

Como entrada para os modelos probabilísticos, descritos na seção 4, utilizou-se o erro cometido pelos modelos determinísticos no período de verificação. Este erro é um vetor que contém a diferença entre valor observado de vazão e valor previsto pelo modelo determinístico (Eq. (1)).

$$\epsilon = q_{obs} - q_{prev} \quad (1)$$

3.1 Climatologia

O novo modelo determinístico desenvolvido, chamado de *Naif*, visa promover a incorporação da climatologia da sub-bacia em estudo ao *ensemble*, tornando a previsão de vazão muito mais completa.

A previsão *Naif* é realizada com base em uma longa série com 30 anos de dados de vazão observados, considerando a vazão atual e o último degrau de vazão (diferença entre o valor de vazão atual e o valor anterior). O último degrau de vazão é reproduzido de acordo com o horizonte que se deseja prever. Se a previsão for realizada para 12 horizontes à frente, então o degrau de vazão será reproduzido 12 vezes. Dessa forma o comportamento da vazão deve-se manter ao longo do tempo.

A Figura 2 apresenta a comparação entre os valores de REQM (Eq. (2)) calculados para os quatro modelos determinísticos (Sacramento, 3R, Regressivo Linear Múltiplo e MLP), e

também para o novo modelo *Naif*. Na Eq. (2), $n = 4515$ representa a quantidade de valores de vazão observados e previstos pelos modelos determinísticos.

$$REQM = \sqrt{\frac{1}{n} \sum_{i=1}^n (q_{obsi} - q_{previ})^2} \quad (2)$$

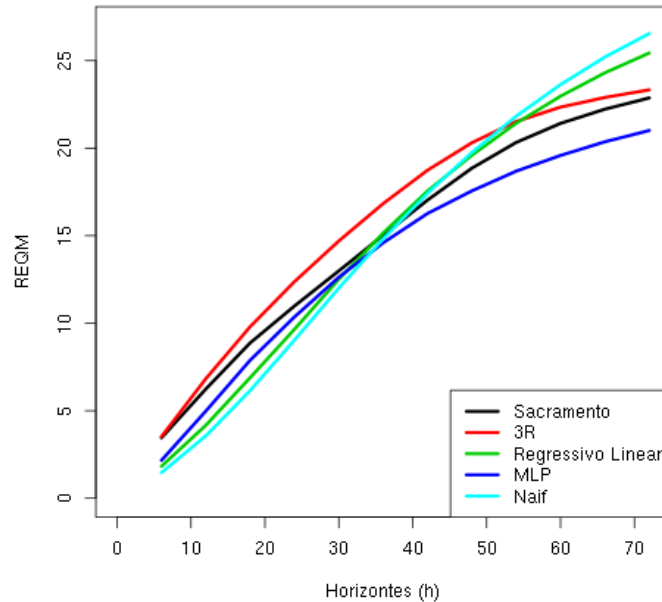


Figura 2: Valores de REQM para previsões dos quatro modelos determinísticos e para o novo modelo climatológico *Naif*

Os resultados do *Naif* (Fig. 2) demonstraram melhores previsões para horizontes de até 36h, que dos demais modelos determinísticos. Ressaltando a importância da inclusão da climatologia ao *ensemble*. Além disso, as curvas de desempenho demonstram que as incertezas com o passar dos horizontes contribuem para o aumento do erro de previsão de vazão.

O coeficiente de eficiência de *Nash-Sutcliffe* (Nash, 1970) também é muito utilizado para avaliar o poder preditivo de modelos hidrológicos. Definido pela Eq. (3), em que $\bar{q}_{obs}$ é a média dos valores de vazão observados, a eficiência pode variar de $-\infty$ a 1. Um coeficiente igual a um corresponde a um modelo de previsão perfeito, igual a zero indica que as previsões do modelo são tão precisas quanto a média dos dados observados, e menores do que zero que a média observada é melhor preditor do que o modelo.

$$NS = 1 - \frac{\sum_{i=1}^n (q_{obsi} - q_{previ})^2}{\sum_{i=1}^n (q_{obsi} - \bar{q}_{obs})^2} \quad (3)$$

A Figura 3 apresenta a comparação dos valores dos coeficientes de *Nash-Sutcliffe* para os quatro modelos determinísticos, e para o modelo *Naif*, no período de verificação para os 12 horizontes de previsão.

Em geral, as Figuras 2 e 3 evidenciam as diferenças dos resultados dos modelos determinísticos para diferentes horizontes de previsão, indicando a necessidade de um método robusto que seja capaz de considerar essas variações. Assim, o objetivo da construção dos modelos

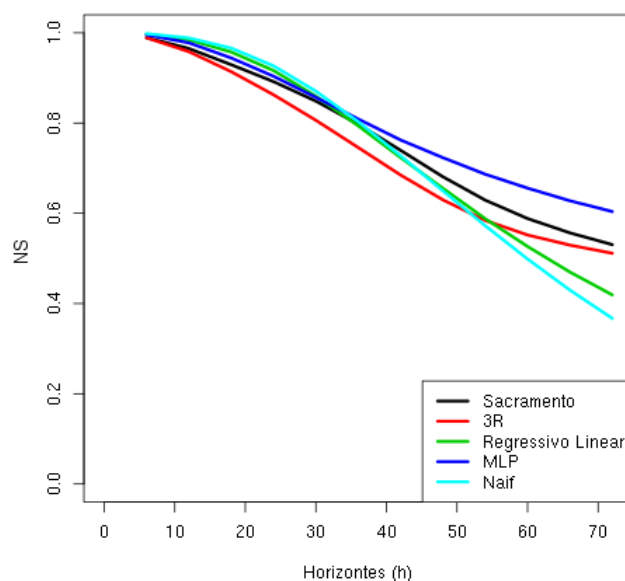


Figura 3: Valores do coeficiente de Nash-Sutcliffe para os quatro modelos determinísticos e para o *Naif*

probabilísticos de erros, apresentados na seção 4, é justamente agrupar as diferenças existentes entre os diferentes modelos determinísticos a fim de construir uma previsão probabilística que suporte todas essas incertezas.

Neste sentido, um estudo aprofundado da distribuição de probabilidade dos erros dos modelos determinísticos torna-se relevante.

Devido ao comportamento assimétrico das densidades de probabilidade dos erros, e das correlações existentes entre os erros nos diferentes horizontes de previsão dos modelos determinísticos, a calibração do modelo probabilístico de previsão de vazão também necessita de um estudo prévio sobre o tipo de distribuição de probabilidade ideal que deve ser utilizada.

4 MODELOS HIDROLÓGICOS PROBABILÍSTICOS

A construção dos modelos probabilísticos baseados nos erros dos modelos hidrológicos determinísticos é estruturada de acordo com a variação de seus parâmetros. Dessa forma, dois tipos de modelos probabilísticos são desenvolvidos.

No modelo probabilístico do tipo estático os parâmetros são considerados fixos para qualquer momento de previsão, ou seja, não variam ao longo do tempo.

Para o modelo probabilístico do tipo dinâmico um conjunto de preditores faz com que os parâmetros se tornem variáveis ao longo do tempo, e com que o modelo consiga incorporar informações hidrometeorológicas da sub-bacia em estudo.

O modelo dinâmico é uma melhoria do modelo estático, pois o modelo estático é utilizado como informação à priori no desenvolvimento do modelo dinâmico.

Os modelos probabilísticos, estáticos ou dinâmicos, são estruturados de forma que 12 horizontes de previsões de 6h em 6h (totalizando 72h) sejam correlacionados por meio de uma

análise estatística multivariada.

Dessa forma, fez-se uso da inferência bayesiana, com os algoritmos de amostragem MCMC (*Markov Chain Monte Carlo*), desenvolvidos em linguagem Stan (Gelman et al., 2013), para a construção dos modelos probabilísticos.

A linguagem Stan, em conjunto com o software estatístico R (R Core Team, 2012), permite que o usuário especifique um modelo, forneça valores iniciais, e implemente simulações da cadeia de Markov automaticamente para as distribuições posteriores resultantes.

Os modelos probabilísticos multivariados, estático e dinâmico, possuem como posteriori a distribuição de probabilidade multivariada *skew-t* (Arellano, 2010).

4.1 Modelos Probabilísticos Estáticos

O modelo probabilístico estático multivariado, que possui como posteriori a distribuição *skew-t* multivariada $ST_p(\xi, \Omega, \lambda, \nu)$, é composto de quatro parâmetros. O parâmetro de localização ξ (vetor de 12 dimensões), escala Ω (matriz de covariância de dimensão 12×12), inclinação λ (vetor de 12 dimensões) e curtose ν (número real). Esta distribuição pode ser aplicada à variável aleatória que corresponde ao vetor de erros de cada um dos modelos determinísticos.

Otimizações do Stan, de quatro cadeias de Markov em paralelo, são aplicadas para os 12 horizontes de previsão simultaneamente com o intuito de determinar representantes da distribuição posteriori dos parâmetros.

A avaliação dos modelos probabilísticos estáticos aplicados a cada um dos quatro modelos determinísticos e ao modelo *Naif* pode ser realizada pela comparação entre os valores dos logaritmos das probabilidades posteriores totais (L_{pt}), somados em todos os momentos e horizontes de previsão. Ou por meio dos logaritmos das probabilidades posteriores totais somadas por horizonte (L_{pth}).

4.2 Modelos Probabilísticos Dinâmicos

Os modelos dinâmicos multivariados são um aprimoramento dos modelos estáticos. Variáveis preditoras condicionam os parâmetros do modelo de acordo com o estado hidrológico do momento da previsão e os resultados dos modelos estáticos tornam-se prioris dos dinâmicos.

A quantidade de parâmetros do modelo dinâmico é igual à quantidade de parâmetros do modelo estático acrescida de oito parâmetros, chamados de parâmetros de translação ($\xi_{t1}, \xi_{t2}, \lambda_{t1}, \lambda_{t2}$) e parâmetros de rotação ($\omega_{r1}, \omega_{r2}, \nu_{r1}, \nu_{r2}$) do modelo dinâmico.

A distribuição multivariada *skew-t* $ST_p(\xi_{new}, \Omega_{new}, \lambda_{new}, \nu_{new})$ é composta por quatro parâmetros: localização (Eq. (4)), escala (Eq. (5)), inclinação (Eq. (6)) e curtose ou graus de liberdade (Eq. (7)).

$$\xi_{new} = \xi + (\xi_{t1} + \xi_{t2}p_\xi), \quad \xi \in [\xi_{L0}, \xi_{L1}] \quad (4)$$

$$\Omega_{new} = \Omega(\omega_{r1} + \omega_{r2}p_\omega), \quad (\omega_{r1} + \omega_{r2}p_\omega) \in [\omega_{L0}, \omega_{L1}] \quad (5)$$

$$\lambda_{new} = \lambda + (\lambda_{t1} + \lambda_{t2}p_\lambda), \quad \lambda \in [\lambda_{L0}, \lambda_{L1}] \quad (6)$$

$$\nu_{new} = \nu(\nu_{r1} + \nu_{r2}p_\nu), \quad (\nu_{r1} + \nu_{r2}p_\nu) \in [\nu_{L0}, \nu_{L1}] \quad (7)$$

Esta distribuição combina os parâmetros de translação e de rotação com informações de preditores do estado hidrológico ($p_\xi, p_\omega, p_\lambda, p_\nu$) às informações a priori dos modelos estáticos

$(\xi, \Omega, \lambda, \nu)$, ocasionando uma perturbação no modelo estático controlada pelos limites $\xi_{L0}, \xi_{L1}, \omega_{L0}, \omega_{L1}, \lambda_{L0}, \lambda_{L1}, \nu_{L0}, \nu_{L1}$.

Os limites da perturbação podem ser escolhidos empiricamente, sempre observando o comportamento da distribuição posteriori obtida.

Para a seleção dos preditores mais relevantes à formação de cada uma das perturbações $(p_\xi, p_\omega, p_\lambda, p_\nu)$, uma otimização individual de cada um dos parâmetros é realizada a fim de encontrar valores ótimos $(\xi^*, \omega^*, \lambda^*, \nu^*)$ para os quais o valor da soma do logaritmo da probabilidade posterior (L_{pt}) no período de treinamento seja maximizado.

Modelos lineares de regressão foram construídos individualmente. Nos quais os valores ótimos $(\xi^*, \omega^*, \lambda^*, \nu^*)$ são tratados como variáveis dependentes, e 32 preditores do estado hidrometeorológico da sub-bacia em estudo, como variáveis regressoras. Os preditores são construídos de acordo com 3 conjuntos chamados de: estado hidrometeorológico da bacia estudada no momento da previsão ou em um passado recente; informações das previsões hidrometeorológicas determinísticas; erros recentes cometidos pelo sistema previsor.

O método *stepwise* (Chambers, 1992) é aplicado a esses modelos para que se faça uma seleção dos preditores mais significativos para cada perturbação dos parâmetros. Por fim, previsões das perturbações $(p_\xi, p_\omega, p_\lambda, p_\nu)$ para o período de verificação são realizadas. E com estas previsões os parâmetros de translação e rotação são combinados para formarem as perturbações dos parâmetros da distribuição *skew-t* tornando o modelo dinâmico.

Otimizações do Stan, de quatro cadeias em paralelo, são aplicadas com o intuito de determinar amostras representativas da distribuição posteriori dos parâmetros do modelo. Os modelos probabilísticos dinâmicos são aplicados aos determinísticos e ao *Naif*, e comparados e avaliados segundo os valores de L_{pt} e L_{pth} .

5 ENSEMBLE

Após a aplicação dos 5 modelos probabilísticos dinâmicos, que utilizaram os dados dos modelos determinísticos Sacramento, 3R, Regressivo Linear, MLP e *Naif*, estimativas dos parâmetros das distribuições posteriores são obtidas.

Com os parâmetros estimados torna-se possível a criação de distribuições de probabilidade *skew-t* multivariadas para cada momento de previsão (4515 momentos, incluindo o período de treinamento e de verificação dos modelos probabilísticos), de todos os modelos (5 modelos). Totalizando 22575 distribuições *skew-t* multivariadas diferentes.

Estas distribuições de probabilidade são combinadas de acordo com a multiplicação delas por uma matriz β formada pelas verossimilhanças apresentadas por horizonte de cada modelo, segundo a Eq. (8). Em Eq. (8), m representa cada um dos cinco modelos, h cada um dos doze horizontes, L_{pth} a verossimilhança para cada modelo em cada um dos horizontes e $k = 1, \dots, n$ a quantidade de dados, os momentos de previsão.

$$\beta_{m,h} = \frac{e^{L_{pth}(m,h)}}{\sum_{k=1}^n e^{L_{pth}(k,h)}}, \quad \sum_{m=1}^n \beta_{m,h} = 1 \quad (8)$$

Da combinação das distribuições pode-se obter a mediana resultante da distribuição combinada. Esta mediana representa o resultado determinístico da previsão de vazão por *ensemble*.

Com o novo resultado determinístico do *ensemble* os modelos probabilísticos, estático e dinâmico, são aplicados resultando em previsões probabilísticas de vazão por *ensemble*.

A avaliação do desempenho da aplicação dos modelos probabilísticos dinâmicos ao *ensemble*, pode ser realizada por meio do cálculo do erro quadrático médio Eq. (9), em que n é o tamanho da amostra, q_{obsi} os valores de vazão observados para o momento de previsão, e q_{previ} os valores pontuais de vazão previstos pelos modelos determinísticos e para o *ensemble*. Para os modelos probabilísticos dinâmicos aplicados aos determinísticos e ao *ensemble* q_{previ} equivale à moda da distribuição de probabilidade obtida a posteriori.

$$EQM = \frac{1}{n} \sum_{i=1}^n (q_{obsi} - q_{previ})^2 \quad (9)$$

Os resultados dos cálculos dos EQM em $(m^3/s)^2$ são apresentados na Fig. 4.

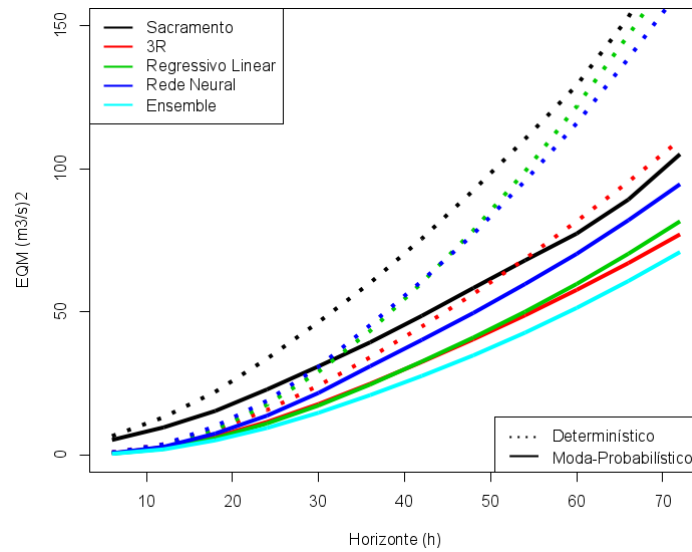


Figura 4: EQM dos modelos determinísticos, e das modas dos modelos probabilísticos dinâmicos

O EQM da moda do modelo probabilístico dinâmico aplicado ao *ensemble* é inferior a todos os EQM das modas dos modelos probabilísticos dinâmicos aplicados aos demais modelos determinísticos. As melhorias médias obtidas pela comparação entre o resultado do EQM do modelo probabilístico dinâmico aplicado ao *ensemble* e aos demais modelos determinísticos, são de 51% para o modelo Sacramento, de 16% para o 3R, 13% para o Regressivo Linear e 30% para a MLP.

O coeficiente de Nash-Sutcliffe (Eq. (3)) também é utilizado para avaliação dos modelos probabilísticos dinâmicos, e apresentado na Fig 5. O coeficiente do *ensemble* variou apenas de 0,99 a 0,88, ao longo de todos os horizontes de previsão.

A tabela 1 apresenta a comparação entre os desempenhos do modelo probabilístico dinâmico aplicado ao *ensemble*, e aos demais modelos determinísticos individuais. O modelo combinado pela mediana apresenta maior valor de L_{pth} em relação a todos os modelos dinâmicos, com uma melhoria média de 38%.

O modelo probabilístico dinâmico aplicado ao *ensemble* também apresenta maior verossimilhança (L_{pth}) em todos os horizontes de previsão, com uma melhoria média de 10%, estes resultados estão apresentados na figura 6.

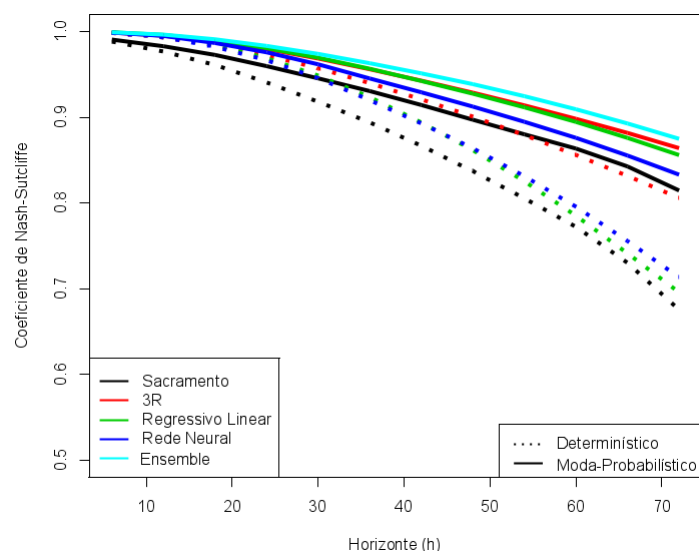


Figura 5: Valores dos coeficientes de Nash-Sutcliffe dos modelos determinísticos, e das modas dos modelos probabilísticos dinâmicos

Tabela 1: Avaliação de desempenho do modelo probabilístico baseado na combinação (*ensemble*) dos modelos hidrológicos probabilísticos dinâmicos

	Sacramento	3R	Regressivo	MLP	<i>Ensemble</i>
<i>Lpt</i>	-10145,28	-7419,88	-6531,30	-7192,65	-4830
Melhoria	52%	35%	26%	33%	

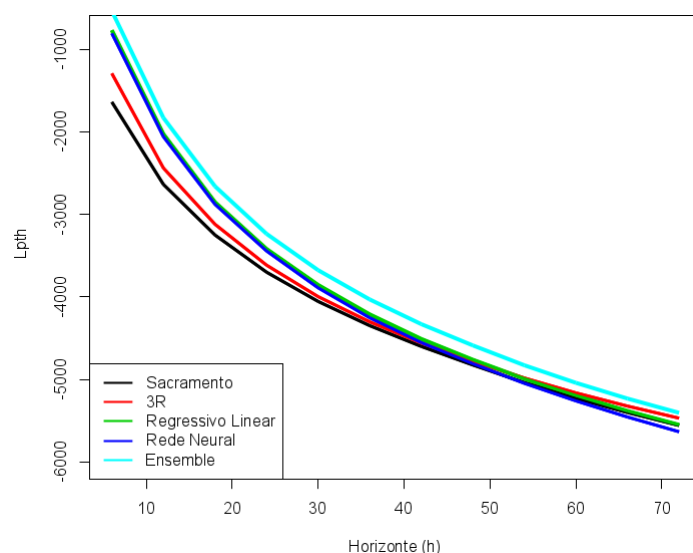


Figura 6: Valores de *Lpth* dos modelos probabilísticos dinâmicos aplicado ao *ensemble* e aos modelos determinísticos individuais

A melhoria obtida com a aplicação do *ensemble* permite a construção de um sistema de visualização dos resultados da previsão de vazão.

5.1 Visualização dos Resultados

A coletânea de figuras a seguir apresenta exemplos do sistema de visualização com os resultados dos modelos probabilísticos estáticos e dinâmicos aplicados ao *ensemble*. As quatro linhas, de cor magenta, contínuas representam os resultados dos quatro modelos determinísticos, e a linha pontilhada, de cor magenta, representa o resultado do novo determinístico *Naif*. A linha contínua de cor vermelha representa o resultado determinístico da previsão por *ensemble* dado pela mediana da distribuição probabilística combinada.

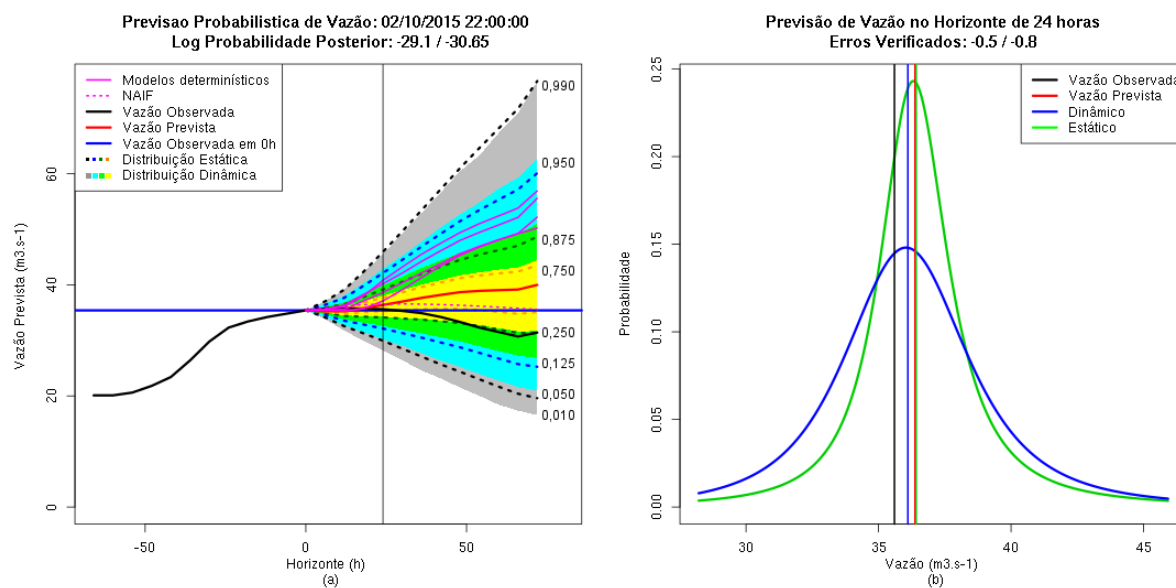


Figura 7: Resultados dos modelos probabilísticos aplicados ao *ensemble* na data 02/10/2015 às 22:00h (a) Previsões do modelo estático com as distribuições de probabilidade entre as linhas pontilhadas coloridas e previsões do modelo dinâmico com distribuições de probabilidade entre as regiões coloridas. A disposição das cores do modelo estático se referem às cores do modelo dinâmico cujas probabilidades são apresentadas (b) Distribuições de probabilidade dos modelos estático e dinâmico para o horizonte de 24h

A Figura 7 apresenta de maneira completa a visualização dos resultados dos modelos probabilísticos estático e dinâmico aplicados ao *ensemble* na data de 02/10/2015 às 22:00h, destacando o horizonte de 24h. Neste momento de previsão (Fig. 7 (a)) os quatro modelos determinísticos indicam grande aumento nos valores de vazão. O modelo *Naif* indica estabilidade nos valores ao longo dos horizontes. O *ensemble* determinístico resulta em um pequeno aumento nos valores de vazão, e os probabilísticos estático e dinâmico englobam os valores de vazão observados que, ao contrário dos resultados das previsões determinísticas, diminuem ao longo dos horizontes. Para o horizonte escolhido, de 24h a diferença entre o valor observado e o valor previsto pelo modelo probabilístico dinâmico é de apenas $-0,5 \text{ m}^3/\text{s}$, e entre o valor observado e o previsto pelo estático de $-0,8 \text{ m}^3/\text{s}$, as distribuições obtidas para este horizonte de previsão podem ser observadas na Fig. 7 (b).

Para simplificar a visualização dos resultados das previsões para o dia 20/09/2015 às 22:00h, na Fig. 8, apenas o modelo probabilístico dinâmico é apresentado. A variação no comportamento da vazão observada ao longo dos 12 horizontes de 6h, dificulta a realização da previsão. No entanto, o modelo probabilístico dinâmico construído para o *ensemble* foi capaz de englobar

em sua distribuição de probabilidade toda essa variação, demonstrando a vantagem de sua utilização e da utilização da abordagem multivariada que considera a correlação entre os horizontes de previsão.

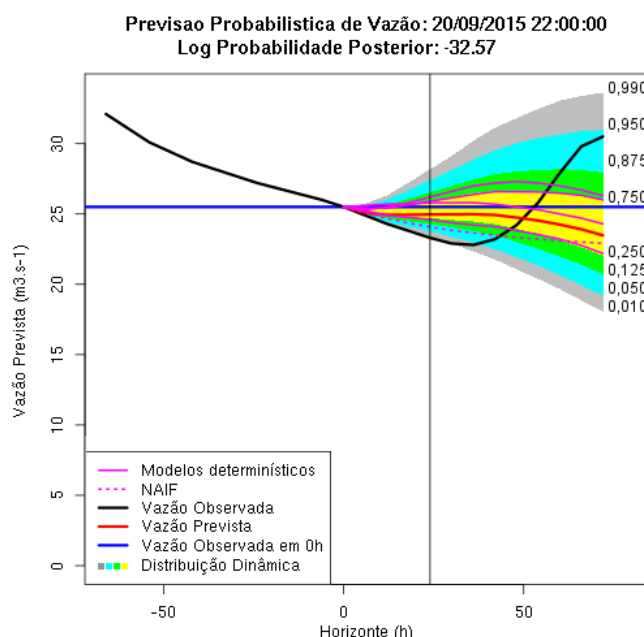


Figura 8: Resultado do modelo probabilístico dinâmico, com seus valores de probabilidade definidos entre as regiões coloridas, aplicados ao *ensemble* na data de 20/09/2015 às 22:00h

A Figura 9 apresenta uma sequência de momentos de previsão dos modelos probabilísticos, estático e dinâmico, na data de 18/05/2016 às 04h (figura 9 (a)), 10h (Fig. 9 (b)), 16h (Fig 9 (c)) e 22h (Fig 9 (d)), que apresentam casos de cheia.

No primeiro momento às 04h (Fig 9 (a)) o aumento dos valores de vazão foram previstos pelos modelos determinísticos, pelo modelo *Naif*, pelo *ensemble* e pelos probabilísticos estático e dinâmico. Porém, às 10h (Fig 9 (b)) e 16h (Fig 9 (c)) os valores continuaram aumentando e as previsões não acompanharam este comportamento. Às 22h (Fig 9 (d)) os modelos probabilísticos voltaram a englobar os valores de vazão observados. Este comportamento dos modelos de previsão se repete em muitos casos, demonstrando a importância de se acompanhar o desenvolvimento das previsões ao longo dos horizontes.

A Figura 10 apresenta uma sequência de momentos de previsão dos modelos probabilísticos, estático e dinâmico, na data de 23/09/2015 às 10h (Fig 10 (a)), 16h (Fig 10 (b)), 22h (Fig 10 (c)) e 04h do dia 24/09/2015 (Fig 9 (d)), que apresentam casos de estiagem. No início desta sequência apenas o modelo probabilístico estático é capaz de englobar na previsão os valores de vazão, porém ao longo dos horizontes a previsão se torna mais adequada.

As Figuras 11 e 12 apresentam exemplos dos bons resultados do modelo probabilístico dinâmico em casos de cheias observadas no período de verificação dos modelos probabilísticos, que não foram corretamente previstas pelos modelos determinísticos.

A cheia do dia 22/05/2016 (Fig 11 (a)) foi subestimada pelos modelos determinísticos, não foi prevista pelo modelo *Naif*, porém ficou dentro das faixas de probabilidade dos modelos probabilísticos estático e dinâmico. Às 04h deste dia os erros do modelo dinâmico foram de

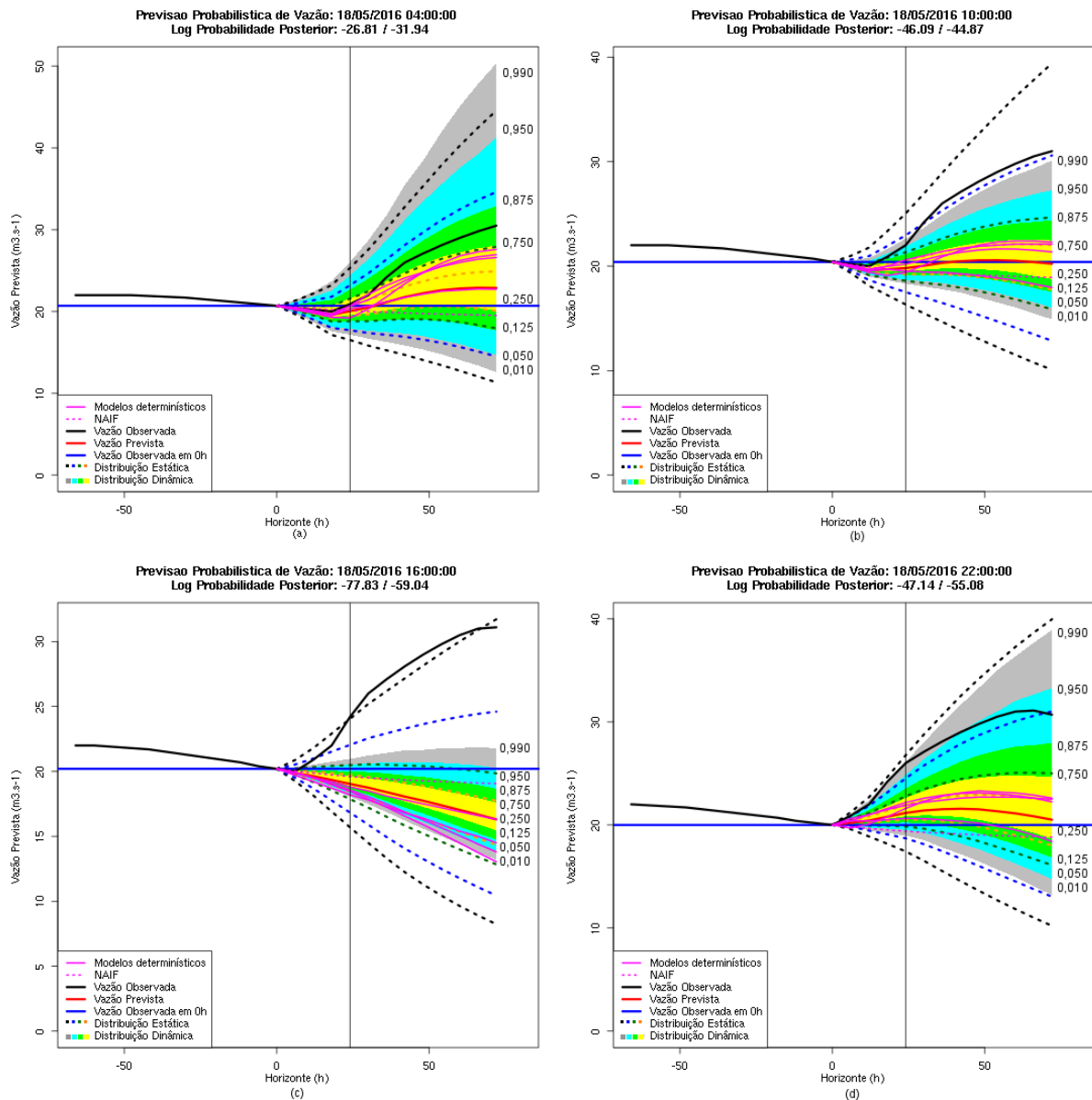


Figura 9: Resultados dos modelos probabilísticos aplicados ao *ensemble*. Previsões do modelo estático com as distribuições de probabilidade entre as linhas pontilhadas coloridas e previsões do modelo dinâmico com distribuições de probabilidade entre as regiões coloridas. A disposição das cores do modelo estático se refere às cores do modelo dinâmico cujas probabilidades são apresentadas (a) 04h do dia 18/05/2016 (b) 10h do dia 18/05/2016 (c) 16h do dia 18/05/2016 (d) 22h do dia 18/05/2016

apenas $0,6 \text{ m}^3/\text{s}$ e aproximadamente $0 \text{ m}^3/\text{s}$ para o modelo estático, no horizonte de 24h (Fig 11 (b)).

A cheia do dia 25/05/2016 (Fig 12 (a)) foi superestimada pelos modelos determinísticos, não foi prevista pelo modelo *Naif*, porém ficou dentro das faixas de probabilidade dos modelos probabilísticos estático e dinâmico. Às 04h deste dia os erros do modelo dinâmico foram de apenas $0,8 \text{ m}^3/\text{s}$ e aproximadamente $-2,2 \text{ m}^3/\text{s}$ para o modelo estático, no horizonte de 24h (Fig 12 (b)).

Este sistema de visualização, dos resultados das previsões dos modelos, pode se tornar

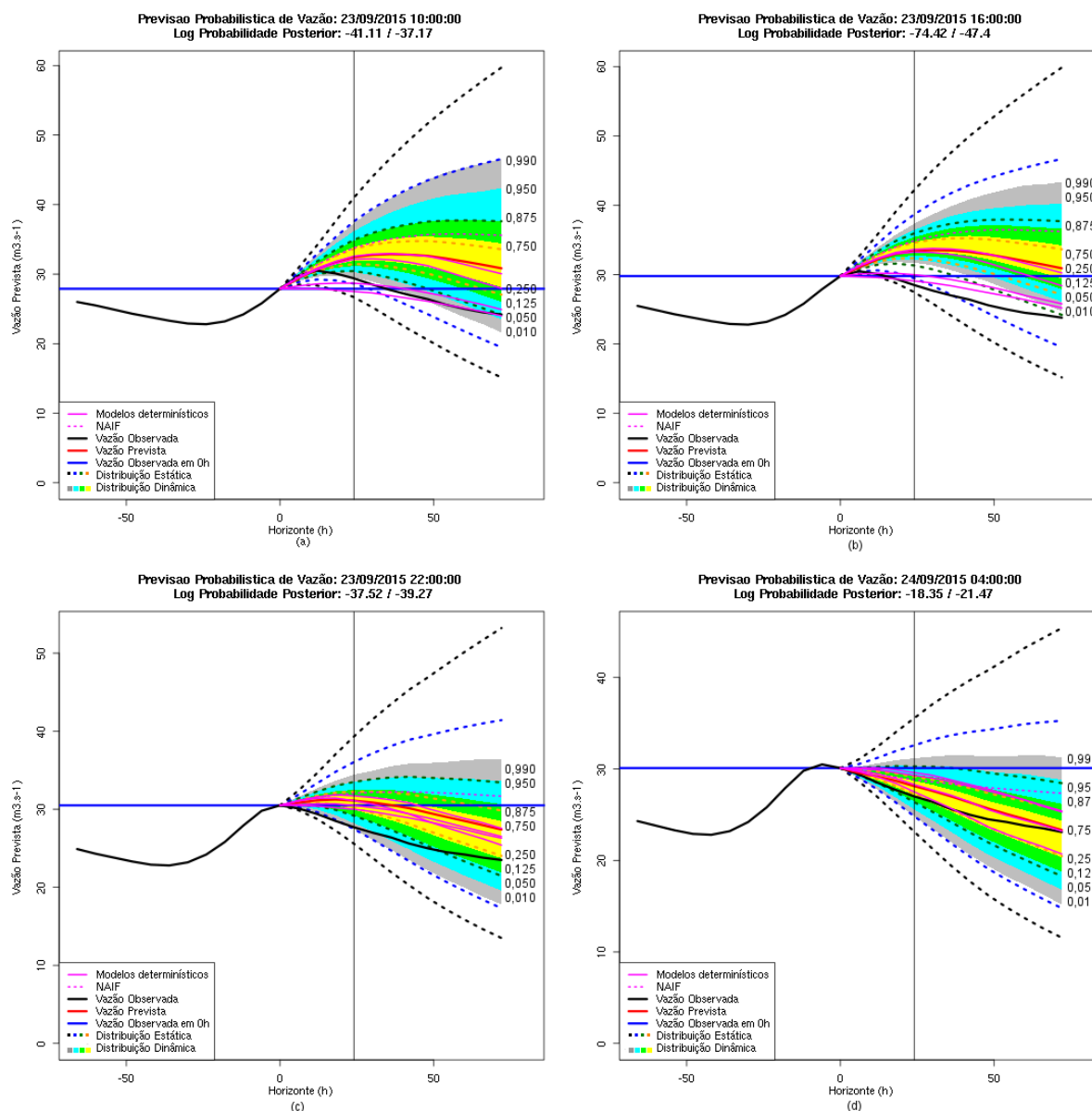


Figura 10: Resultados dos modelos probabilísticos aplicados ao *ensemble*. Previsões do modelo estático com as distribuições de probabilidade entre as linhas pontilhadas coloridas e previsões do modelo dinâmico com distribuições de probabilidade entre as regiões coloridas. A disposição das cores do modelo estático se refere às cores do modelo dinâmico cujas probabilidades são apresentadas (a) 10h do dia 23/09/2015 (b) 16h do dia 23/09/2015 (c) 22h do dia 23/09/2015 (d) 04h do dia 24/09/2015

extremamente útil para o acompanhamento, em tempo real, do desenvolvimento de cheias ou estiagens realizado por diversos órgãos de monitoramento. Além disso, ele permite o acompanhamento, por parte desses órgãos, do desempenho dos modelos e da necessidade de recalibrá-los com um período maior de dados que venham a ser disponibilizados.

6 CONCLUSÕES

A aplicação dos modelos probabilísticos multivariados (estáticos e dinâmicos) considerou os erros resultantes de quatro modelos determinísticos do tipo chuva-vazão, subdivididos entre

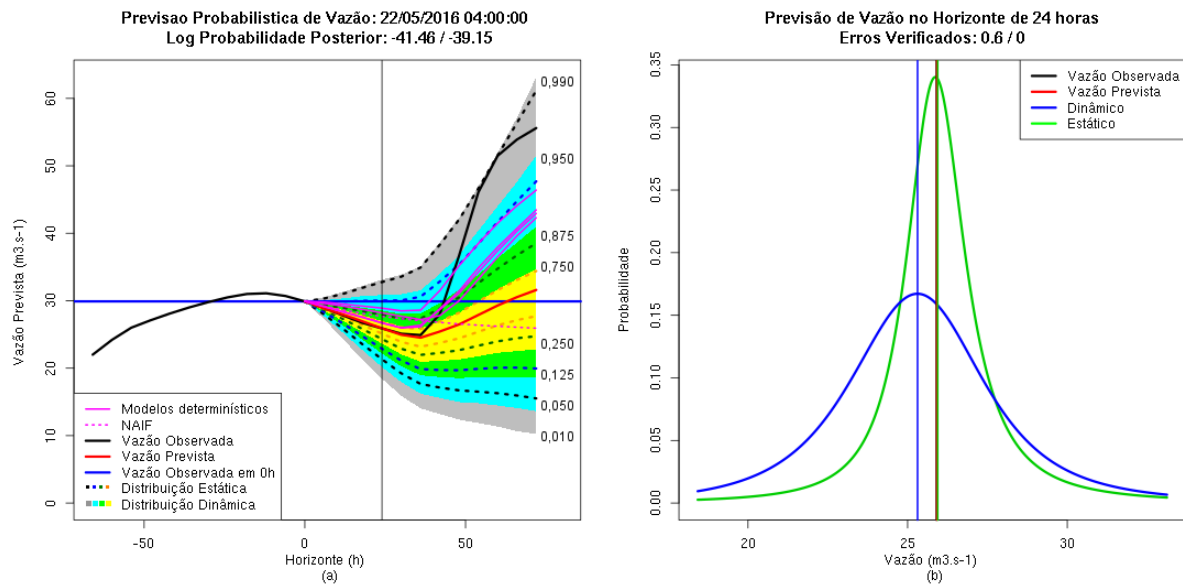


Figura 11: Resultados dos modelos probabilísticos aplicados ao *ensemble* na data 22/05/2016 às 04:00h (a) Previsões do modelo estático com as distribuições de probabilidade entre as linhas pontilhadas coloridas e previsões do modelo dinâmico com distribuições de probabilidade entre as regiões coloridas. A disposição das cores do modelo estático se referem às cores do modelo dinâmico cujas probabilidades são apresentadas (b) Distribuições de probabilidade dos modelos estático e dinâmico para o horizonte de 24h

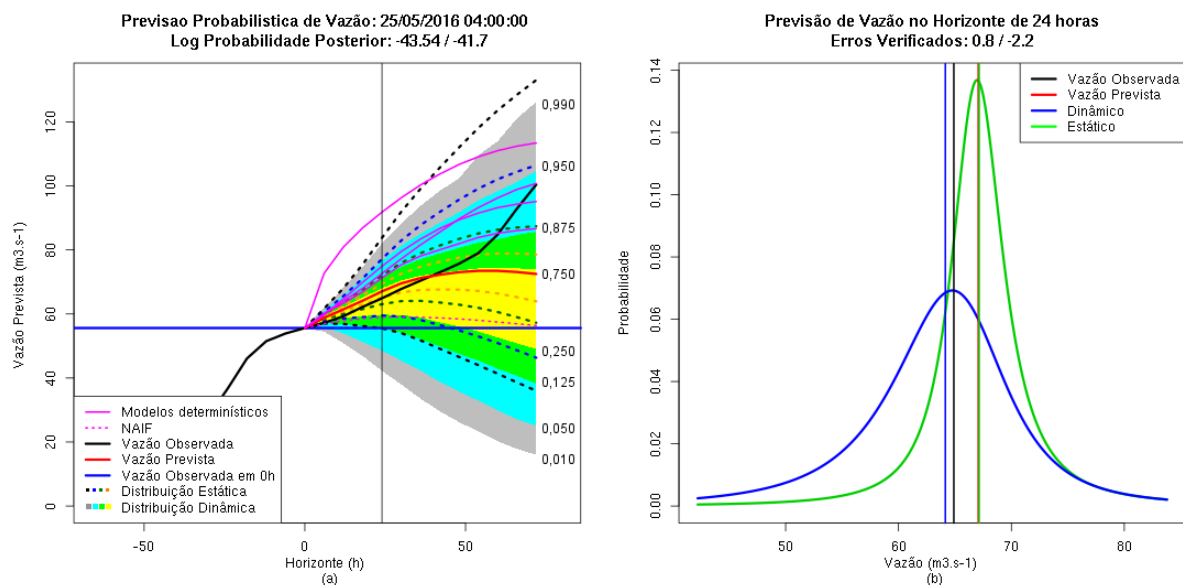


Figura 12: Resultados dos modelos probabilísticos aplicados ao *ensemble* na data 25/05/2016 às 04:00h (a) Previsões do modelo estático com as distribuições de probabilidade entre as linhas pontilhadas coloridas e previsões do modelo dinâmico com distribuições de probabilidade entre as regiões coloridas. A disposição das cores do modelo estático se referem às cores do modelo dinâmico cujas probabilidades são apresentadas (b) Distribuições de probabilidade dos modelos estático e dinâmico para o horizonte de 24h

modelos conceituais (Sacramento e 3R) e modelos empíricos (Regressivo Linear e rede neural MLP). Estes modelos foram escolhidos pois possuem características muito distintas.

A fim de completar a combinação dos modelos desenvolveu-se um novo modelo determi-

nístico, de maneira muito simplificada, que considera a climatologia da região em estudo. Este modelo utilizou 30 anos de dados de vazão observados e apenas com uma combinação desses dados criou uma nova previsão determinística chamada de *Naif*.

Os modelos probabilísticos foram aplicado aos erros do novo modelo determinístico *Naif* e incorporado ao *ensemble*, resultando em uma única previsão determinística que combinou a climatologia e os resultados dos modelos probabilísticos dinâmicos aplicados aos quatro modelos determinísticos do tipo chuva-vazão.

Os modelos probabilísticos, estático e dinâmico, foram aplicados novamente à previsão por *ensemble*, resultando em valores ainda maiores de *Lpt* e *Lpth*, do que os obtidos anteriormente. Seus resultados se demonstraram superiores aos testes que já haviam sido realizados.

A importância de se utilizar modelos multivariados que considerem todos os horizontes de previsão ao mesmo tempo e a correlação existente entre horizontes vizinhos, deve ser enfatizada.

Além disso, a dinâmica obtida com o uso de preditores nos parâmetros dos modelos desenvolvidos certamente contribuiu para a melhoria do desempenho dos modelos probabilísticos.

Por meio do desenvolvimento de um novo sistema de visualização destacou-se o uso do modelo probabilístico dinâmico multivariado aplicado ao *ensemble*. Demonstrando que mesmo em casos de difícil previsão tais como: cheias, estiagens, ou outros eventos com variações dos valores de vazão inesperados, o modelo dinâmico superou as abordagens determinísticas e estáticas. Mesmo em casos que os valores previstos pelos modelos determinísticos eram muito diferentes dos valores de vazão observados, a abordagem proposta foi capaz de construir regiões que representam a probabilidade de ocorrência das vazões.

O sistema de visualização dos resultados dos modelos probabilísticos estático e dinâmico aplicados à previsão por *ensemble* demonstrou que pode se tornar muito útil para o acompanhamento, realizado por diversos órgãos de monitoramento, do desenvolvimento de cheias ou estiagens em tempo real.

AGRADECIMENTOS

Os autores gostariam de agradecer à CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - pelo apoio financeiro dado a essa pesquisa.

REFERÊNCIAS

- Alfieri, L., Salamon, P., Pappenberger, F., Wetterhall, F., & Thielen, J., 2012. Operational early warning systems for water-related hazards in Europe. *Environmental Science Policy*, vol. 21, pp. 35–49.
- Arellano, R. B., & Genton, M. G., 2010. Multivariate extended skew-t distributions and related families. *METRON – International Journal of Statistics*, vol. 68, pp. 201–234.
- Bao, H-J., Zhao, L-N., He, Y., Li, Z-J., Wetterhall, F., Cloke, H. L., Pappenberger, F., & Manful, D., 2011. Coupling ensemble weather predictions based on TIGGE database with Grid-Xinjiang model for flood forecast. *Advances in Geosciences*, vol. 29, pp. 61–67.
- Bogner, K., & Pappenberger, F., 2011. Multiscale error analysis, correction, and predictive uncertainty estimation in a flood forecasting system. *Water Resources Research*, vol. 47.

- Bowler, N. E., Arribas, A., Mylne, K. R., Robertson, B., & Beare, S. E., 2008. The MOGREPS short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, vol. 134, pp. 703–722.
- Buizza, R., Hollingsworth, A., Lalaurette, F., & Ghelli, A., 1999. Probabilistic Predictions of Precipitation Using the ECMWF Ensemble Prediction System. *Weather and Forecasting*, vol. 14, pp. 168–189.
- Calvetti, L., 2011. *Previsão hidrometeorológica probabilística na Bacia do Alto Iguaçu-PR com os modelos WRF e TopModel*. PhD thesis, Universidade de São Paulo, São Paulo, Brasil.
- Chambers, J. M., & Hastie, T., 1992. *Statistical Models in S*. Wadsworth & Brooks/Cole Advanced Books & Software. Michigan.
- Chen, S-T., & Yu, P-S., 2007. Real-time probabilistic forecasting of flood stages. *Journal of Hydrology*, vol. 340, pp. 63–77.
- Cloke, H. L., & Pappenberger, F., 2009. Ensemble flood forecasting: A review. *Journal of Hydrology*, vol. 975, pp. 613–626.
- Cunha, L., 2004. *Sistemas avançados de previsão hidrológica e operação hidráulica em tempo real*. Msc thesis, Universidade Federal do Paraná, Curitiba, Brasil.
- Cuo, L., Pagano, T. C. & Wang, Q. J., 2011. A Review of Quantitative Precipitation Forecasts and Their Use in Short- to Medium-Range Streamflow Forecasting. *Journal of Hydrometeorology*, vol. 12, pp. 713–728.
- Davison, A. C., 2003. *Statistical models. Volume 11 de Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge. New York
- Davolio, S., Miglietta, M. M., Diomede, T., Marsigli, C., Morgillo, A., & Moscatello, A., 2008. A meteo-hydrological prediction system based on a multi-model approach for precipitation forecasting. *Natural Hazards and Earth System Sciences*, vol. 8, pp. 143–159.
- Demeritt, D., Cloke, H., Pappenberger, F., Thielen, J., Bartholmes, J., & Ramos, M-H., 2007. Ensemble predictions and perceptions of risk, uncertainty, and error in flood forecasting. *Environmental Hazards*, vol. 7, pp. 115–127.
- Fan, F. M., Collischonn, W., Meller, A., & Botelho, L. C. M., 2014. Ensemble streamflow forecasting experiments in a tropical basin: The São Francisco river case study. *Journal of Hydrology*, vol. 519, pp. 2906–2919.
- Gelman, A., Carlin, J. B., Stern, H. S., Rubin, D. B., Dunson, D., & Vehtari, A., 2013. *Bayesian Data Analysis. Thrid Edition*. Chapman and Hall. New York.
- Hamill, T. M., Hagedorn, R., Hitaker, & J. S. W., 2008. Probabilistic Forecast Calibration Using ECMWF and GFS Ensemble Reforecasts. Part II: Precipitation. *Monthly Weather Review*, vol. 136, pp. 2620–2632.
- Haykin, S., 2001. *Redes Neurais: princípios e prática*. Bookman. Porto Alegre. Brasil.
- Krzysztofowicz, R., 2001. The case for probabilistic forecasting in hydrology. *Journal of Hydrology*, vol. 249, pp. 2–9.
- Leite, E. A., 2008. *Gestão do valor da informação hidrometeorológica: O caso dos alertas de Inundação para proteção de bens móveis em edificações residenciais de União da Vitória*. PhD

thesis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.

Leite, E. A., Breda, A., & Saboia, J. P., 2011. Sistema de previsão probabilística de variáveis hidrológicas e sua aplicação no alerta de inundações, *XXI Seminário Nacional de Produção e Transmissão de Energia Elétrica*, Florianópolis, Brasil.

Nash, J. E., & Sutcliffe J. V., 1970. River flow forecasting through conceptual models part I — A discussion of principles*. *Journal of Hydrology*, vol. 10, pp. 282–290.

Neto, R. K., Guetter A. K., & Mine, M. R. M., 2007. Modelo Hidrológico com Atualização de Estado Parte I: Modelagem Hidrológica. *Revista Brasileira de Recursos Hídricos*, vol. 12, pp. 93–103.

Pappenberger, F., Bartholmes, J., Thielen, J., Cloke, H. L., Buizza, R., & Roo, A., 2008. New dimensions in early flood warning across the globe using grand-ensemble weather predictions. *Geophysical Research Letters*, vol. 35, pp. 7p.

R Core Team, 2012. *R: A Language and Environment for Statistical Computing*. Vienna, Austria. <http://www.R-project.org/>

Reggiani, P., & Weerts, A. H., 2008. Probabilistic Quantitative Precipitation Forecast for Flood Prediction: An Application. *Journal of Hydrometeorology*, vol. 9, pp. 76–95.

Schellekens, J., Weerts A. H., Moore, R. J., Pierce, C. E., & Hildon, S., 1980. The use of MOGREPS ensemble rainfall forecasts in operational flood forecasting systems across England and Wales. *Advances in Geosciences*, vol. 29, n. 1, pp. 77–84.

Sema, 2013. Bacias Hidrográficas do Paraná. Série Histórica. Segunda Edição *Governo do Estado do Paraná*.

Thiemig, V., Pappenberger, F., Thielen, J., Gadain, H., Roo, A., Bodis, K., Medico M. D., & Muthusi, F., 2010. Ensemble flood forecasting in Africa: a feasibility study in the Juba–Shabelle river basin. *Atmospheric Science Letters*, pp. 123–131.

Tucci, C. E. M., 1998. *Modelos Hidrológicos*. Editora da UFRGS. Porto Alegre.