

Previsão de precipitações utilizando a técnica de previsão por conjuntos e um algoritmo baseado no vizinho mais próximo.

Márcio Francisco Dutra e Campos¹, Maury Meirelles Gouvêa Júnior²

¹Departamento de Engenharia de Computação e Automação Industrial
FEEC – Universidade Estadual de Campinas (UNICAMP)
Caixa-postal: 6101 – CEP: 13083-852 – Campinas, SP – Brazil.

²Programa de Pós-Graduação em Engenharia Elétrica
Pontifícia Universidade Católica de Minas Gerais (PUC MINAS) – BH, MG – Brasil
marciocampos@pucminas.br, maury@pucminas.br

Abstract. *This work presents a precipitation forecasting method that uses the stochastic approach of ensemble forecasting and the nearest neighbor technique (KNN) to select, among several candidate probability distributions, those that best adhere to observed data. The research uses the historical series of hydrological data made available by the National Water Agency (ANA). The preliminary results show that the proposed method generates predictions that are, at least, comparable to several regression methods.*

Resumo. *Este trabalho apresenta um método de previsão de precipitações que utiliza a abordagem estocástica de previsão por conjuntos (ensembles) e a técnica do vizinho mais próximo (KNN) para selecionar, entre várias distribuições de probabilidades candidatas, aquelas que melhor aderem aos dados observados. A pesquisa utiliza as séries históricas de dados hidrológicos disponibilizados pela Agência Nacional de Águas (ANA). Os resultados preliminares mostram que o método proposto gera previsões que são, pelo menos, comparáveis a diversos métodos de regressão.*

1. Introdução

Realizar previsões a respeito do comportamento do clima (possibilidade de chuva, temperaturas mínimas e máximas, índice de radiação ultravioleta, umidade relativa do ar, etc) e das bacias hidrográficas (vazão, secas, inundações, entre outras) é de suma importância para o planejamento urbano e para várias atividades econômicas.

Os impactos das intempéries da natureza precisam ser medidos e, se possível, previstos, para evitar danos irreparáveis às pessoas, ao patrimônio e às finanças públicas. Por outro lado, é difícil fazer um planejamento, ou prognóstico, de médio ou longo prazo sem considerar os impactos das mudanças climáticas (em sentido amplo) localizadas.

Realizar previsões hidrológicas confiáveis é de extrema relevância, na medida em que a antecipação de eventos extremos permite, entre outras coisas, a emissão de alertas, a tomada de decisão na operação de obras hidráulicas, a execução de planos de reposta a emergências, a otimização do uso de recursos hídricos [Fan et al., 2015].

A Região Metropolitana de Belo Horizonte (RMBH) possui inúmeros cursos de água. A distribuição destes cursos de água e o relevo da região favorece, no período das chuvas, a formação de enchentes; por outro lado, no período de seca, o que se tem observado nos últimos anos é o decréscimo dos níveis dos reservatórios a valores críticos, exigindo, inclusive, estudos para implantação de políticas de racionamento. Neste contexto, a previsão de eventos severos, inundações ou secas, é fundamental para qualquer planejamento de médio ou longo prazo na região.

Este trabalho propõe um modelo/método de previsão de precipitações que utiliza a abordagem estocástica de previsão por conjuntos e a técnica do vizinho mais próximo (KNN) para selecionar, entre diversas distribuições de probabilidade candidatas, aquelas que serão utilizadas para prever a precipitação de cada mês do ano.

A pesquisa se justifica na medida em que contribui para o reconhecimento dos fenômenos hidrológicos extremos que perturbam a RMBH permitindo, assim, um melhor planejamento das atividades econômicas e uma melhor proteção da população.

Este trabalho está organizado da seguinte forma: no capítulo 2 é feita uma breve revisão de modelos e métodos utilizados em hidrologia e de trabalhos relacionados; no capítulo 3 é realizada a análise exploratória dos dados e o algoritmo de previsão por conjuntos usando a técnica KNN é apresentado; no capítulo 4 os resultados obtidos são analisados; e, no capítulo 5, a conclusão e trabalhos futuros são apresentados.

2. Modelos e Métodos Utilizados em Hidrologia

Tanto a hidrologia quanto a meteorologia fazem uso intensivo de ferramentas matemático-computacionais para organizar, resumir e analisar a imensa quantidade de dados climatológicos, atuais e históricos, que são utilizados para prever as condições atuais e futuras do tempo, bem como para modelar e simular os fenômenos físico-químicos que determinam o seu comportamento.

A hidrologia, especificamente, vem, há muitos anos, procurando respostas satisfatórias para os diversos fenômenos hidrológicos, através de modelos conceituais, numéricos e de simulações [Todini, 2007]. Diversas técnicas têm sido utilizadas para prever o comportamento dos fenômenos hidrológicos, sendo que as mais relevantes são os métodos numéricos, os métodos estatísticos e os métodos de inteligência artificial.

Previsões hidrológicas geralmente são feitas com precisão de poucos dias, pois lidam com sistemas dinâmicos considerados caóticos. Ou seja, nestas previsões, o futuro possui um horizonte bem curto. Entretanto, a previsão de eventos extremos, como secas e inundações, exige um horizonte temporal maior e, portanto, de métodos específicos, para que se possa fazer um planejamento hídrico de médio e o longo prazo satisfatório.

Os modelos hidrológicos podem ser caracterizados da seguinte forma: os modelos contínuos simulam o ciclo da água completo e, os discretos, apenas partes do ciclo; os modelos agregados consideram a bacia hidrológica como um único elemento e, os distribuídos, a subdivide em elementos menores; os modelos empíricos utilizam equações de transformação de estado experimentais e, os conceituais, modelos abstratos; os modelos determinísticos utilizam uma sequência de dados históricos e, os estocásticos, dados sintetizados. Este trabalho utiliza um modelo discreto, distribuído, empírico e estocástico.

2.1. Métodos de Previsão em Hidrologia

Diferentes métodos podem ser utilizados para prever eventos hidrológicos, dependendo da experiência do hidrologista, da informação disponível, do nível de dificuldade que a previsão apresenta e do grau de precisão necessária na previsão. Duas abordagens muito utilizadas na previsão hidrológica são os métodos numéricos e os estatísticos. Os numéricos simulam modelos dinâmicos da atmosfera e, os estatísticos, tentam prever o comportamento dinâmico do sistema baseado nas informações passadas.

Os métodos numéricos apresentam alguns pontos fracos como, por exemplo, a imprecisão das equações utilizadas nos modelos, a ausência de dados iniciais de áreas de difícil observação, como montanhas e oceanos, e a necessidade de interpretação dos resultados por um meteorologista experiente.

2.2. Previsão por Conjuntos (Ensembles)

Evidências mostram que a frequência de eventos hidrológicos extremos irá aumentar nos próximos anos. Entretanto, a precisão dos métodos existentes de estimativa de inundação [Sraj et al., 2016] e de estiagem [Mishra & Desai, 2005] estão diminuído. Isto acontece porque as previsões hidrológicas são, geralmente, feitas com informações incompletas e incertas das variáveis de interesse, haja vista que qualquer medição hidroclimática, realizada por sensores, tem uma incerteza associada – além de não existir uma técnica de medição perfeita, sensores produzem erros sistemáticos e aleatórios [García-Marín et al., 2014].

Uma forma de considerar estas incertezas é o uso de previsões probabilísticas por conjuntos (*ensembles*), em que são gerados diversos cenários futuros possíveis, com o objetivo de quantificar as incertezas envolvidas no processo [Fan et al., 2015]. Na previsão por conjuntos, diferentes parametrizações, diferentes condições iniciais ou, ainda, diferentes modelos, são usados para gerar um conjunto de previsões tal que, cada um dos seus membros, representa uma possível trajetória do processo ao longo do horizonte de previsão (cenários equiprováveis) [Fan et al., 2015].

2.3. Trabalhos Relacionados

Caldeira et al. (2015) analisou o ajuste de diferentes modelos probabilísticos a séries de precipitação máxima diária anual no Rio Grande do Sul. As séries pluviométricas foram ajustadas às distribuições Log-Normal e Gumbel e a adequação verificada pelos testes de Kolmogorov-Smirnov e Qui-Quadrado. Franco et al. (2014) identificou, dentre as distribuições Gumbel, Gama e Generalizada de Valores Extremos, qual a mais indicada para aplicação às séries históricas de precipitação máxima diária anual na bacia hidrográfica do rio Verde, no Sul de Minas Gerais.

Para lidar com a incertezas das variáveis hidrológicas, Taillardat et al. (2016) propôs um método estatístico não paramétrico para pós-processamento de conjuntos (*ensembles*), baseado em florestas de regressão de quantil (QRF), para estimativa dos quantis desejados. Com este mesmo objetivo, Scheuerer (2014) apresentou um método de pós-processamento de conjunto (*ensembles*) em que a precipitação é modelada utilizando uma distribuição de valores extremos generalizada, que permite incorporar informações de vizinhança para melhorar o seu desempenho.

3. Metodologia

3.1. Considerações Iniciais

A hidrologia de uma região depende, em primeiro lugar, do seu clima, o que implica na necessidade de conhecer os processos meteorológicos que o determinam. Entre estes, a chuva se destaca como grandeza fundamental no equilíbrio da água. A complexidade dos processos naturais e o desconhecimento das causas que determinam eventos climáticos tem forçado a renunciar aos estudos analíticos e a usar as séries temporais de sucessivas observações de tais fenômenos [García-Marín et al., 2014].

Se a trajetória de uma variável é conhecida, pelos seus próprios fundamentos físicos, conceitos probabilísticos não precisam ser utilizados para a sua caracterização, mas, se não é este o caso, é preciso estimar não apenas o seu valor esperado mas, também, a sua função de densidade de probabilidade (pdf) [Todini, 2007].

O uso de distribuições estatísticas adequadas aos eventos hidrológicos extremos podem ser considerados ferramentas imprescindíveis para se obter uma gestão eficiente dos recursos hídricos, principalmente diante do aumento da demanda de água e o aumento das atividade em pequenas e medias bacias hidrográficas [Souza et al., 2012].

3.2. Análise Exploratória dos Dados

Este trabalho utiliza a série histórica de precipitações mensais da estação pluviométrica de Ponte Nova do Paraopeba, localizada no município de Juatuba/MG, que fica na RMBH. Trata-se de um registro completo de precipitações, com mais de 70 anos de observações. Para este trabalho, foram utilizados os dados correspondentes ao período de 01/01/1946 até 31/12/2015 (70 anos de observações), apresentados na Figura 1.

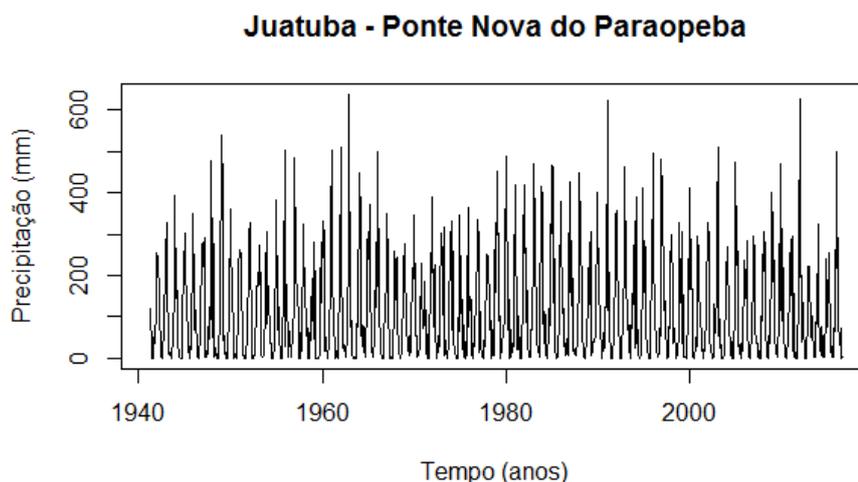


Figura 1 – Precipitações observadas em Ponte Nova do Paraopeba

A Figura 2 mostra o gráfico de barras e o *boxplot* da distribuição mensal de precipitação, por ano hidrológico (que começa em outubro de um ano e termina em setembro do próximo ano). A dispersão dos dados verificada nesta figura já sugere o tipo de distribuição de probabilidades que poderia ser ajustada para as precipitações.

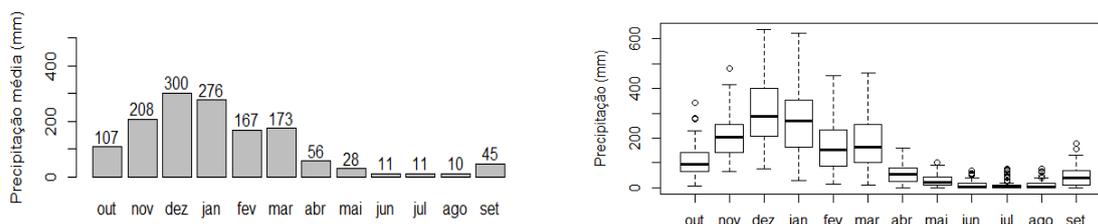


Figura 2 – Precipitação mensal media em Ponte Nova do Paraopeba

3.3. Método e Algoritmo Proposto

O método proposto neste trabalho é baseado no algoritmo estocástico desenvolvido por Nowak et al. (2010) para desagregação de vazões anuais em vazões diárias. Mas, diferente do algoritmo de Nowak, que utiliza a técnica de descoberta do vizinho mais próximo proposto por Lall & Sharma (1996) para gerar séries sintéticas em escalas menores (obtenção de dados diários a partir de dados anuais), o algoritmo aqui proposto modifica a mesma técnica para, entre várias distribuições de probabilidade candidatas, selecionar aquelas que melhor representam os dados e, a partir, delas fazer uma previsão por conjuntos. O algoritmo pode ser descrito da seguinte forma:

Passo 1: Como se trata de um processo estocástico, os dados históricos observados correspondem a uma única trajetória do sistema, sendo certo que, a partir destes dados, não é possível determinar com precisão a função de distribuição de probabilidades que melhor representa o sistema. Sendo assim, e para que se possa fazer uma previsão por conjuntos, para as precipitações observadas em cada mês são ajustadas N distribuições de probabilidades candidatas.

Passo 2: Uma vez ajustadas as N distribuições de probabilidades, as candidatas mais próximas (K vizinhos mais próximos) são identificados da seguinte forma: para cada distribuição de probabilidade candidata é gerada uma trajetória e calculado a correlação entre esta trajetória e os dados históricos. A ideia é que, quanto mais forte correlação entre a distribuição ajustada e a observação histórica, melhor é a representatividade da distribuição, tornando-a uma forte candidata. O número de vizinhos (candidatas) mais próximos é dado por $K = \sqrt{N}$, conforme heurística proposta por Lall & Sharma [1996] que, segundo estes autores, é muito eficaz em uma variedade de aplicações.

Passo 3: Para cada um dos K vizinhos mais próximos determinados no passo anterior é atribuído um peso baseado na função proposta por Lall e Sharma [3], dada por

$$W(i) = \frac{1}{i} / \sum_{i=1}^k \frac{1}{i}$$

onde K é o número de vizinhos mais próximos e i refere-se ao "índice vizinho", sendo que $i = 1$ é o vizinho mais próximo e assim por diante. Observe que, na equação acima, os pesos são normalizados de modo que somem uma unidade.

Passo 4: De posse dos K vizinhos mais próximos (melhores distribuições candidatas) e de seus respectivos pesos, a série de precipitações futuras é gerada, utilizando a previsão por conjuntos, da seguinte forma

$$S = T_1W_1 + T_2W_2 + \dots + T_NW_N,$$

em que T_i representa uma trajetória da candidate i , W_i é o seu peso e “+” é um operador de agregação, que não precisa ser, necessariamente, a soma.

4. Resultados e Discussão

Neste estudo preliminar, que tem por objetivo mostrar a viabilidade do algoritmo proposto, foram ajustadas cinco distribuições de probabilidade, cujas características podem ser encontradas em [Naghetini & Pinto, 2007], para as precipitações de Ponte Nova do Paraopeba: Exponencial, Gamma, Gumbel, Normal e Weibull.

Para ilustrar, a Figura 3 mostra 70 observações para cada mês das distribuições Gamma e Gumbel em relação a media mensal dos dados observados (correspondente a 70 anos). A Tabela 1 apresenta os resultados de ajuste das cinco distribuições candidatas para os meses de dezembro e julho (meses que apresentaram condições extremas) e, a Figura 4, mostra a dispersão das distribuições candidatas em relação aos dados observados para estes dois meses.

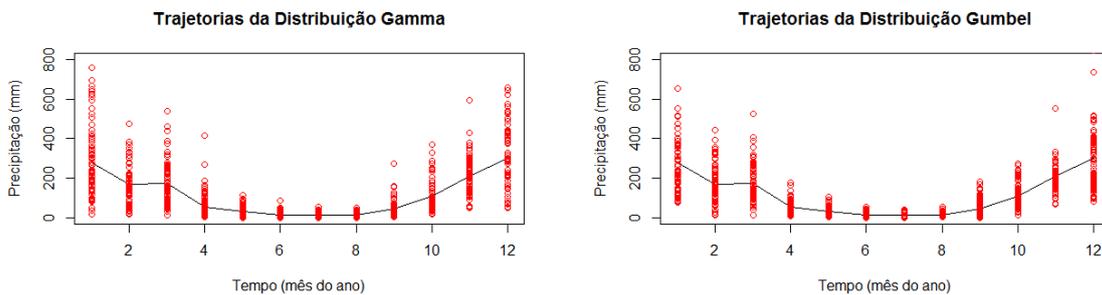


Figura 3 – Exemplos de resultados obtidos com as distribuições Gamma e Gumbel.

Tabela 1 – Resultados dos ajustes das distribuições para meses extremos.

Mês	Resultados dos ajustes					
Dezembro	Goodness-of-fit statistics					
		Gamma	Gumbel	Weibull	Normal Exponential	
	Kolmogorov-Smirnov statistic	0.06766349	0.07154201	0.06388267	0.08516653	0.3031188
	Cramer-von Mises statistic	0.04820282	0.05390731	0.03464616	0.06866842	2.0505039
	Anderson-Darling statistic	0.32993354	0.37254978	0.22800891	0.43106267	10.5725358
	Goodness-of-fit criteria					
		Gamma	Gumbel	Weibull	Normal Exponential	
Akaike's Information Criterion	876.6719	878.282	875.4924	879.2161	940.5455	
Bayesian Information Criterion	881.1689	882.779	879.9893	883.7131	942.7940	
Julho	Goodness-of-fit statistics					
		Gamma	Gumbel	Weibull	Normal Exponential	
	Kolmogorov-Smirnov statistic	0.4965571	0.2930485	0.4965571	0.2934709	0.4965571
	Cramer-von Mises statistic	5.0848553	1.8487063	5.0848553	2.2320688	5.0848552
	Anderson-Darling statistic	Inf	10.8470007	Inf	11.4511206	Inf
	Goodness-of-fit criteria					
		Gamma	Gumbel	Weibull	Normal Exponential	
Akaike's Information Criterion	473.8197	562.8927	473.8197	616.6062	471.8197	
Bayesian Information Criterion	478.3167	567.3897	478.3167	621.1032	474.0682	

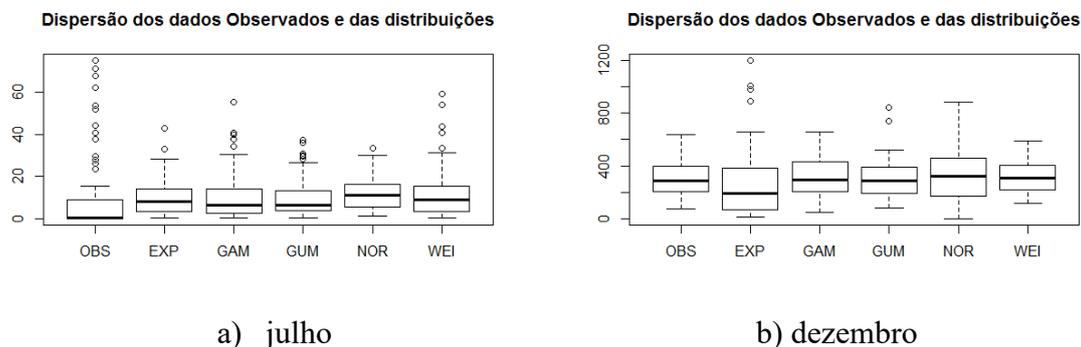


Figura 4 – Dispersão das distribuições para os meses de a) julho e b) dezembro.

O algoritmo proposto foi implementado em R (<https://cran.r-project.org/>), utilizando os diversos pacotes disponíveis para este *software*, entre os quais merecem destaque os seguintes: *actual*, *e1071*, *fitdistrplus*, *forecast*, *nnet*, *QRM* e *randomForest*.

A série histórica de 70 anos de precipitações (840 meses) foi dividida em uma série de treinamento de 60 anos (720 meses) e uma série testes de 10 anos (120 meses). Os resultados obtidos com o algoritmo proposto foram comparados com resultados obtidos pelas seguintes técnicas, que também estão implementadas no ambiente R: *Generalized Linear Models (GLM)*, *Support Vector Machines (SVM)*, *Random Forest (randomForest)* e *Neural Networks (NNET)*.

4.1. Análise Qualitativa dos Resultados

Foram feitos experimentos qualitativos com o intuito de validar o algoritmo proposto para, posteriormente, fazer os refinamentos necessários. Portanto, na sequência, serão apresentados apenas resultados qualitativos, haja vista que será necessário investir muito tempo e esforço computacional para apresentar resultados quantitativos relevantes.

A Tabela 2 mostra os resultados da escolha dos vizinhos mais próximos (distribuições candidatas) após uma rodada de execução do algoritmo. Os resultados estão dentro dos esperados, com as distribuições Gamma, Gumbel e Weibull, que reconhecidamente representam fenômenos hidrológicos [Naghetini & Pinto, 2007], aparecendo em primeiro lugar.

Tabela 2 – Distribuições selecionadas para cada mês

	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1	GAM	GAM	WEI	NOR	GAM	GUM	GUM	EXP	EXP	GUM	WEI	WEI
2	EXP	EXP	EXP	EXP	EXP	GAM	WEI	NOR	GAM	NOR	GUM	EXP
3	NOR	NOR	NOR	GAM	NOR	NOR	NOR	GAM	WEI	WEI	NOR	GUM

A Figura 05 mostra, à esquerda, os dados de precipitação observados e, à direita, os dados obtidos na previsão por conjuntos (por ano hidrológico). É possível observar, graficamente, que as precipitações previstas seguem o mesmo comportamento, no que se refere à distribuição de probabilidades, dos dados observados (conjunto de teste), inclusive em relação a máximos, mínimos e assimetria.

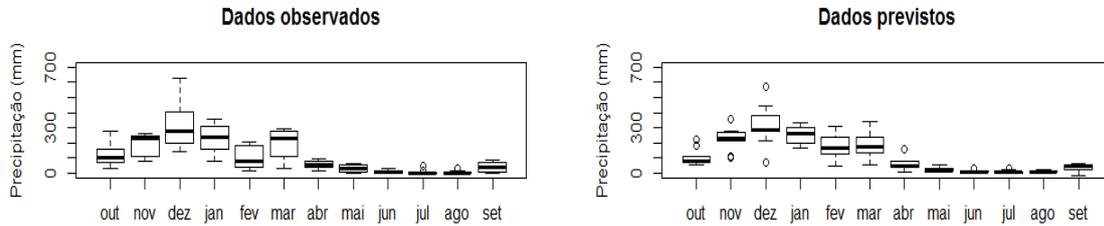
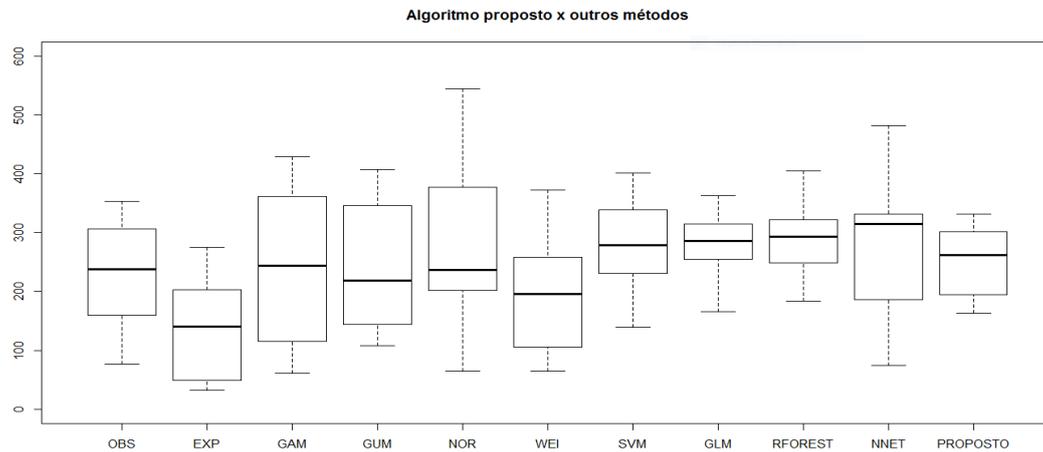
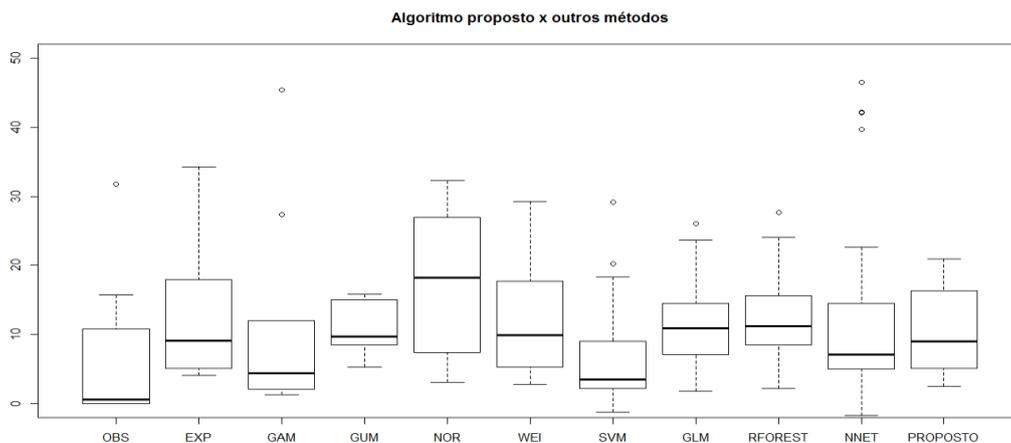


Figura 5 – Dados observados e dados obtidos na previsão por conjuntos

A Figura 6 mostra o resultado das previsões feitas, para os meses de janeiro e agosto, com o algoritmo proposto em comparação com os dados observados (OBS) e as previsões feitas pelas distribuições Exponencial (EXP), Gamma (GAM), Gumbel (GUM), Normal (NOR), Weibull (WEI) e pelos métodos de regressão linear *Support Vector Machine* (SVM), *Generalized Linear Model* (GLM), *RandomForest* (RFO) e Redes Neurais (NNET).



a) previsões para o mês de janeiro



b) previsões para o mês de agosto

Figura 6 – Comparação do algoritmo proposto com outros métodos

No geral, a análise qualitativa (gráfica) feita para os demais meses mostra que o algoritmo proposto gera uma distribuição de probabilidades que se aproxima bastante dos dados observados, apresentado resultados que se comparam ao método SVM.

5. Conclusão e Trabalhos Futuros

Este artigo apresentou um novo algoritmo para previsão de precipitações, baseado em previsões por conjuntos (*ensembles*) e que utiliza a técnica do vizinho mais próximo (KNN) para selecionar, entre várias distribuições de probabilidade, aquelas com melhor aderência aos dados históricos observados.

Foram utilizados os dados de precipitação da estação pluviométrica de Ponte Nova do Paraopeba, que fica em Juatuba, Região Metropolitana de Belo Horizonte. O algoritmo foi implementado utilizando o ambiente de programação R, com seus diversos pacotes. Os resultados preliminares mostraram que o algoritmo é capaz de gerar previsões com qualidade comparável a outros métodos, notadamente o SVM, GLM, randomForest e NNET.

Trata-se de uma pesquisa em estágio embrionário e, por isso, foram apresentados apenas resultados qualitativos que demonstram a viabilidade do algoritmo proposto. Na próxima fase, serão testados, quantitativamente, o desempenho e a eficácia do algoritmo proposto em relação aos demais métodos de previsão que foram utilizados.

Em trabalhos futuros, pretende-se, ainda, realizar as seguintes melhorias no algoritmo: no Passo 2, utilizar a Simulação de Monte Carlo para melhorar o cálculo da correlação entre as distribuições de probabilidades candidatas e os dados observados; no Passo 3, propor novas métricas para a determinação dos vizinhos mais próximos; e, no Passo 4, melhorar a previsão por conjuntos, o que poderá ser feito alterando o operador de agregação utilizado e, também, utilizando a Simulação de Monte Carlo, como propõe o trabalho Kottegoda (2014).

Agradecimentos

Ao Fundo Incentivo à Pesquisa e ao Programa Permanente de Capacitação Docente, da Pontifícia Universidade Católica de Minas Gerais (PUC Minas) pelo apoio dado através dos projetos FIP-2016/11052-2S e PPCD-Nº099/2015.

Referências

- Andrade, M. A., Mello, C. D., & Beskow, S. (2013). Simulação hidrológica em uma bacia hidrográfica representativa dos Latossolos na região Alto Rio Grande, MG. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 17(1), 69-76.
- Caldeira, T. L., Beskow, S., de Mello, C. R., Faria, L. C., de Souza, M. R., & Guedes, H. A. (2015). Modelagem probabilística de eventos de precipitação extrema no estado do Rio Grande do Sul. *R. Bras. Eng. Agríc. Ambiental*, 19(3), 197-203.
- Fan, F. M., Ramos, M.-H., & Collischonn, W. (2015). Sobre o uso de previsões hidrológicas probabilísticas para tomada de decisão. *Revista Brasileira de Recursos Hídricos*, 20(4), 914-926.

- Franco, C. S., Marques, R. F., Oliveira, A. S., & de Oliveira, L. F. (2014). Distribuição de probabilidades para precipitação máxima diária na Bacia Hidrográfica do Rio Verde, Minas Gerais. *Revista Brasileira de Engenharia Agrícola e Ambiental-Agriambi*, 18(7).
- García-Marín, A., Roldán-Cañas, J., Estévez, J., Moreno-Pérez, F., Serrat-Capdevila, A., González, J., ... & Giráldez, J. V. (2014). La hidrología y su papel en ingeniería del agua. *Ingeniería del agua*, 18(1), 1-14.
- Kottegoda, N. T., Natale, L., & Raiteri, E. (2014). Monte Carlo Simulation of rainfall hyetographs for analysis and design. *Journal of Hydrology*, 519, 1-11.
- Lall, Upmanu, and Ashish Sharma. "A nearest neighbor bootstrap for resampling hydrologic time series." *Water Resources Research* 32.3 (1996): 679-693.
- Mishra, A. K., & Desai, V. R. (2005). Drought forecasting using stochastic models. *Stochastic Environmental Research and Risk Assessment*, 19(5), 326–339. doi:10.1007/s00477-005-0238-4
- Naghetini, M. & Pinto, É. J. D. A. (2007). *Hidrologia Estatística*. Companhia de Pesquisa de Recursos Minerais (CPRM), 552 p.
- Nowak, K., Prairie, J., Rajagopalan, B., & Lall, U. (2010). A nonparametric stochastic approach for multisite disaggregation of annual to daily streamflow. *Water Resources Research*, 46(8).
- Souza, F. A. O. De, Silva, C. L. Da, Maggiotto, S. R., & Oliveira Júnior, M. P. De. (2012). Caracterização das vazões em uma pequena bacia hidrográfica do Distrito Federal, Brasil. *Revista Brasileira de Engenharia Agrícola E Ambiental*, 16(1), 10–17. doi:10.1590/S1415-43662012000100002
- Scheuerer, M. (2014). Probabilistic quantitative precipitation forecasting using ensemble model output statistics. *Quarterly Journal of the Royal Meteorological Society*, 140(680), 1086-1096.
- Sraj, M., Viglione, A., Parajka, J., & Blöschl, G. (2016). The influence of non-stationarity in extreme hydrological events on flood frequency estimation. *Journal of Hydrology and Hydromechanics*, 64(4), 426–437. doi:10.1515/johh-2016-0032
- Taillardat, M., Mestre, O., Zamo, M., & Naveau, P. (2016). Calibrated ensemble forecasts using quantile regression forests and ensemble model output statistics. *Monthly Weather Review*, 144(6), 2375-2393.
- Todini, E. (2007). Hydrological catchment modelling: past, present and future. *Hydrological Earth System Science*, 11(1), 468–482. doi:10.5194/hess-11-468-2007