

Modelo Baseado na Técnica de Redes Neurais para Previsão de Vazões na Bacia do Rio São Francisco

Luana Ferreira Costa Gomes e Suzana Maria Gico Lima Montenegro

Departamento de Engenharia Civil – UFPE

luanafcg@yahoo.com.br, suzanam@ufpe.br

Mêuser Jorge da Silva Valença

Departamento de Engenharia de Computação - UPE (Universidade de Pernambuco)

meuserv@yahoo.com.br

Recebido: 29/11/07 - revisado: 05/06/08 - aceito: 23/02/10

RESUMO

Um dos principais insumos na definição da geração das usinas hidrelétricas é a previsão de vazões. Na elaboração dessas previsões, diversos modelos podem ser utilizados. Pode-se citar como exemplo os modelos físicos, estatísticos e os baseados na técnica de redes neurais. O uso da técnica de redes neurais tem se intensificado cada vez mais, uma vez que, modelos baseados nessa técnica são de fácil aplicação e têm proporcionado resultados satisfatórios. A análise prévia das informações que serão usadas na calibração e utilização da rede neural pode trazer ganhos significativos no desempenho da mesma. Assim, esse trabalho apresenta a análise dos dados pluviométricos e fluviométricos da área a montante do reservatório de Três Marias, no rio São Francisco, bem como a calibração de um modelo baseado na técnica de redes neurais para a previsão de vazões naturais afluentes. São apresentados o processo de preenchimento de falhas históricas, análise de consistência e análise geostatística, como ferramenta para seleção de postos pluviométricos e análise de precipitação média da área. Os resultados obtidos mostraram que o modelo calibrado com a técnica de redes neurais teve um desempenho superior ao modelo estocástico PREVIVAZ e, entre as redes neurais analisadas, a NSRBN teve um desempenho um pouco superior a MLP.

Palavras-chave: redes neurais, previsão de vazões, Reservatório de Três Marias.

INTRODUÇÃO

Os modelos hidrológicos são importantes no planejamento e gestão de recursos hídricos. Dentre eles, destacam-se modelos determinísticos, estocásticos e os baseados na técnica de redes neurais. A implementação e a calibração de vários desses modelos apresentam normalmente várias dificuldades, requerendo sofisticadas ferramentas matemáticas, quantidades significativas de parâmetros de calibração, além de vasta experiência com o modelo.

Por esse motivo, o uso da técnica de redes neurais tem se intensificado, uma vez que, modelos baseados nessa técnica são de fácil aplicação e têm apresentado resultados satisfatórios.

As redes neurais artificiais (RNAs) são uma técnica integrante dos Sistemas Inteligentes, inspirados no comportamento humano ou da natureza. As RNAs vêm sendo sugeridas para análise de séries

temporais devido à sua capacidade de tratar com relações de entrada-saída de dados não-lineares, destacando sua habilidade de aprendizado e capacidade de generalização, associação e busca paralela. Estas qualidades as tornam capazes de identificar e assimilar as características mais marcantes das séries, tais como sazonalidade, periodicidade, tendência, entre outras, muitas vezes camufladas por ruídos.

Em todos os processos de modelagem, a familiaridade com os dados disponíveis é importante. Os modelos de redes neurais não são exceção, e o pré-processamento das informações pode ter um efeito significativo no desempenho do modelo.

No âmbito do Setor Elétrico, a obtenção de previsões de vazões permite que os diferentes modelos de operação, otimização e simulação energética possam ser utilizados para avaliar as condições operacionais das usinas hidrelétricas e termelétricas no futuro, propiciando subsídios para a tomada de decisões em relação à melhor forma de operar o Sistema Interligado Nacional - SIN, explorando toda

sua capacidade hidrológica e proporcionando ganhos sinérgicos. Estas previsões de vazões podem permitir também tomadas de decisões que minimizem os efeitos de cheias excepcionais, desde que realizadas com relativa precisão e antecedência.

Nos processos de planejamento e programação da operação do SIN, um dos principais produtos é o PMO – Programa Mensal de Operação. Este documento se constitui numa referência técnica mensal, com revisão semanal, para as metas de geração de todas as usinas hidrelétricas e termelétricas integrantes do SIN. Além disso, são estabelecidas também as metas de intercâmbio entre os subsistemas elétricos e outras diretrizes operativas. Para a elaboração do PMO e suas revisões, um dos principais insumos é a previsão de vazões para todos os locais de aproveitamentos hidrelétricos.

Este trabalho apresenta os resultados da calibração de dois modelos de previsão de vazões naturais médias diárias, 12 dias à frente, utilizando a técnica de redes neurais, e a análise e tratamento de informações advindas de diferentes postos pluviométricos e fluviométricos, a montante do aproveitamento hidrelétrico de Três Marias, bem como a informação das vazões naturais afluentes a esse reservatório. Nessa calibração são utilizadas duas metodologias de redes neurais distintas: MLP (Multi-Layer Perceptrons) e NSRBN (Non-linear Sigmoidal Regression Blocks Networks).

DESCRIÇÃO DA ÁREA EM ESTUDO

A Agência Nacional de Energia Elétrica - ANEEL divide a bacia do rio São Francisco em nove sub-bacias, sendo a região estudada a que corresponde à sub-bacia 40. Esta sub-bacia representa a área de drenagem do reservatório de Três Marias e parte da área de drenagem do Alto São Francisco.

O clima predominante na região é o Tropical Úmido e Temperado de Altitude, com insolação média anual de 2400h e evaporação média anual de 1000mm (ANA (2004)). Com relação às precipitações, os totais pluviométricos anuais variam, de forma geral, entre 1700mm nas cabeceiras da bacia a 1000mm na confluência do rio São Francisco com rio das Velhas (ANA, 2004).

A usina hidrelétrica de Três Marias é de propriedade da CEMIG (Companhia Energética de Minas Gerais), tem uma potência instalada de 396 MW, área de drenagem de 50.732 km² e vazão média de longo termo (MLT) de 689 m³/s.

REDES NEURAIS

As variantes de uma rede neural são muitas (ex: n° de neurônios na camada de entrada, escondida e de saída, função de transferência), e combinando-as, pode-se mudar a arquitetura conforme a necessidade da aplicação. Com relação ao número de neurônios nas camadas escondidas, este é geralmente definido empiricamente, sendo que os melhores resultados são obtidos através de um processo de tentativa e erro (Lima et al., 2005).

A topologia define a maneira com que os vários neurônios constituintes de uma rede neural são interconectados.

Redes MLP com retropropagação

Em uma rede neural MLP os neurônios são arranjados em camadas. A estrutura deste modelo é composta por uma ou mais camadas de neurônios entre a camada de entrada e a camada de saída.

O algoritmo amplamente conhecido na literatura como retropropagativo ("*backpropagation*") define uma maneira sistemática de atualização dos pesos das camadas, a partir da idéia que os erros dos neurônios da camada escondida são determinados pela propagação reversa dos erros dos neurônios da camada de saída (Faraco et al., 1998).

Algoritmo de Levenberg-Marquardt (LM)

Um dos métodos mais rápidos para o treinamento de redes neurais feedforward de tamanho moderado (até algumas centenas de pesos sinápticos) é o método de Levenberg-Marquardt, que é uma variação do retropropagativo de modo a obter maior velocidade de convergência (Pacheco et al, 2005).

Enquanto o retropropagativo padrão utiliza a descida de gradiente como método de aproximação do mínimo da função erro, o algoritmo de Levenberg-Marquardt (LM) utiliza uma aproximação pelo método de Newton (Hagan e Menhaj, 1994).

Redes NSRBN (Adaptado de Valença, 2005)

As redes neurais construtivas, ao contrário das redes de topologia fixa, constroem sua própria estrutura, incluindo unidades e/ou camadas intermediárias, conforme a demanda do processo de treinamento. Dentre estas, uma que tem sido aplicada na área de recursos hídricos são as redes neurais compostas com blocos de regressões sigmóides

não-lineares (NSRBN), proposta por Valença (1999).

Dentre as características principais das redes NSRBN duas merecem destaque: uma estrutura compacta, por possuir apenas uma camada escondida, e um algoritmo construtivo baseado em um método de aprendizado construtivo que é responsável pela definição de sua complexidade ótima.

A proposta do modelo NSRBN é de construir uma rede por blocos de polinômios homogêneos, utilizando uma função de ativação adequada (ex: tangente hiperbólica), para os neurônios escondidos. Quanto aos neurônios de saída, estes têm uma ativação mais complexa, uma vez que realizam uma regressão logística não-linear com relação à saída dos neurônios escondidos.

ANÁLISE DOS DADOS

Neste trabalho foram usadas informações de estações pluviométricas, fluviométricas e de vazão natural do reservatório de Três Marias, do período de janeiro/1987 a dezembro/2004. Todos os dados pluviométricos e fluviométricos utilizados foram da base de dados da CEMIG, com complementação de dados da ANEEL e da ANA. A vazão natural do reservatório de Três Marias corresponde à existente na base de dados do ONS (Operador Nacional do Sistema Elétrico).

Informações pluviométricas

Na seleção dos postos pluviométricos buscou-se aqueles que estão atualmente em operação e com histórico sem muitas falhas, tendo sido selecionados inicialmente 18 de um total de 40 disponíveis.

As informações foram submetidas a um preenchimento de falhas, quando necessário, pelo método da ponderação regional, e a uma análise de consistência, pelo método do vetor regional.

Observou-se que, com exceção do posto de Arcos, todos os postos têm correlação com os demais postos acima de 0,50, o que foi considerado suficiente neste trabalho.

A análise de consistência dos dados mensais de precipitação, através do vetor regional e de gráficos dupla-massa, mostrou uma boa consistência dessas informações. Contudo, os postos de Ponte dos Vilelas e Arcos apresentaram uma mudança de declividade no gráfico dupla-massa, causada pelo preenchimento de falhas das informações. Vale realçar, como comentado no parágrafo anterior,

que o posto de Arcos não teve uma boa correlação com os demais postos pluviométricos da bacia.

Análise geoestatística

Objetivando verificar a variabilidade espacial da precipitação nos postos selecionados foi feita uma análise geoestatística. Essa análise contemplou informações médias mensais agrupadas das seguintes formas: período úmido (novembro a abril), período seco (maio a outubro), total anual e total mensal.

Foram ajustados os parâmetros do variograma teórico a partir de um variograma experimental, considerado neste estudo isotrópico, para os modelos Esférico, Exponencial e Gaussiano, e na interpolação usou-se a Krigagem Ordinária.

Maiores informações sobre a metodologia podem ser obtidos em Holawea & Dutterb (1999) e Vieira (2000).

Foram analisados mapas com as isolinhas da precipitação total anual e do erro de interpolação, para o semivariograma esférico, que foi o melhor ajustado para este caso.

Observou-se que o valor do alcance é aproximadamente o mesmo para todos os variogramas ajustados e que a configuração dos mapas de isolinhas também se assemelha para todas as situações. Por outro lado, a ordem de grandeza dos valores é alterada, ou seja, a estrutura de correlação é semelhante em todos os casos.

Segundo classificação proposta por Cambardella et al. (1994), os semivariogramas da precipitação total anual, mensal e do período úmido, podem ser classificados como de moderada dependência espacial. Apenas o da precipitação do período seco pode ser classificado como de forte dependência espacial.

Como era de se esperar, para todas as situações estudadas, a análise dos erros obtidos com a interpolação através da Krigagem mostra um incremento do erro nas áreas desprovidas de postos pluviométricos. Então, partiu-se para analisar as estações pluviométricas de responsabilidade da ANA/CPRM (Serviço Geológico do Brasil).

Com a análise da localização dos postos pluviométricos da ANA/CPRM e dos erros obtidos na Krigagem, foram selecionadas as áreas onde poderiam ser colocados postos pluviométricos adicionais, visando reduzir os erros de interpolação obtidos, com uma conseqüente melhora na determinação da chuva média da área. Analisando a disponibilidade de informações, foram selecionados mais 10 postos pluviométricos. A figura 1 ilustra a disposição desses

postos juntamente com os da CEMIG, ou seja, todos os postos selecionados.

Mais uma vez, observou-se que o valor do alcance é aproximadamente o mesmo para todos os variogramas ajustados e que a configuração dos mapas de isolinhas também se assemelha para todas as situações. Por outro lado, a ordem de grandeza dos valores é alterada, ou seja, a estrutura de correlação é semelhante em todos os casos. Também como no caso anterior, os semivariogramas da precipitação total anual, mensal e do período úmido, podem ser classificados como de moderada dependência espacial e o da precipitação do período seco pode ser classificado como de forte dependência espacial.

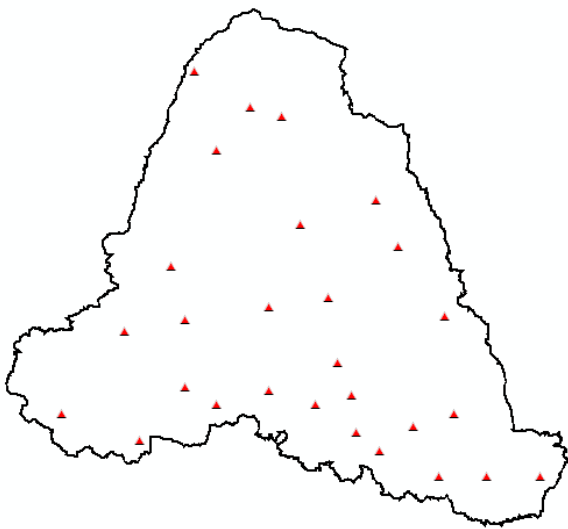


Figura 1 – Postos pluviométricos.

Observou-se que a validação cruzada produziu um bom resultado.

Como, com exceção da média da precipitação total anual, todos os semivariogramas das informações mensais tiveram um melhor ajuste através do modelo esférico, também foi aplicada a Krigagem usando um semivariograma esférico para os dados diários, para estimativa da precipitação média espacial.

Informações fluviométricas

Primeiramente foram usados os postos fluviométricos da CEMIG atualmente em operação na área de drenagem do reservatório de Três Marias. A figura 2 ilustra as sub-bacias da área de drenagem do reservatório de Três Marias.

Analisando os coeficientes de correlação dos dados dos postos fluviométricos selecionados, observa-se que os mesmos têm uma correlação razoável, com coeficientes de correlação sempre superiores a 0,69, o que leva a concluir que essas informações são bem correlacionadas.

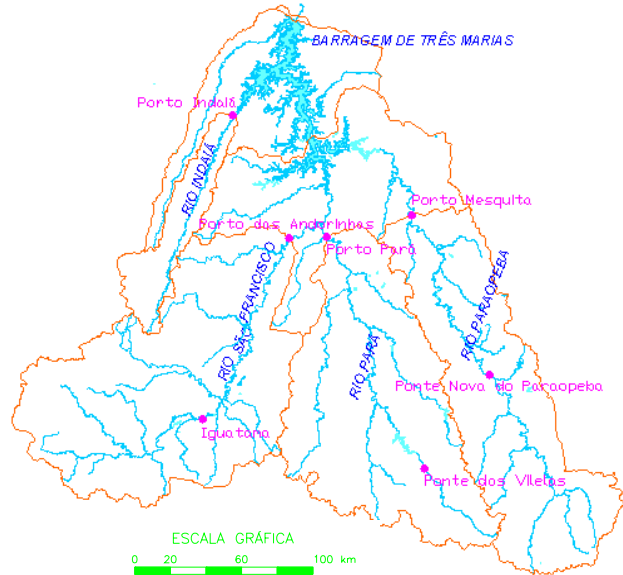


Figura 2 – Sub-bacias da área de drenagem de Três Marias com localização dos postos fluviométricos.

Os postos de Porto das Andorinhas e Porto Mesquita não possuem nenhuma falha nesse período, os postos de Porto Indaia, Porto Pará Igatama, Ponte Nova do Paraopeba e Ponte dos Vilelas Juntamente tiveram 0,1%, 0,1%, 40,5% e 0,6% de falhas, respectivamente. O preenchimento de falhas dessas séries foi feito a partir de uma regressão linear simples com o posto mais bem correlacionado e a partir de uma correlação múltipla, com os demais postos da bacia.

Comparando as correlações dos dados após o preenchimento de falhas pelos dois métodos, observa-se que há um ganho em se usar a regressão múltipla, e assim foi feito neste trabalho.

Análise Estatística

Na análise estatística foi estudada a matriz de correlação, a ACP (Análise de Componentes Principais), a função de auto-correlação (FAC) e a função de correlação cruzada.

Nos resultados da FAC para as informações de vazão e de precipitação média espacial da área observou-se que a partir de 0,6 a mesma começa a ficar constante com a variação do lag (representa a defasagem temporal da variável, em dias). Assim, considerando esse valor de 0,6 como limite da FAC, a tabela 1 mostra as defasagens obtidas.

Tabela 1 – defasagens obtidas para FAC máxima de 0,6.

Variável		defasagem
Postos Fluviométricos	Porto das Andorinhas	27
	Porto Indaiá	8
	Porto Mesquita	10
	Porto Pará	11
	Iguatama	26
	Ponte Nova do Paraopeba	10
	Ponte dos Vilelas Jusante	11
Postos Pluviométricos	Chuva Média por Thiessen	7

Observou-se que as séries de vazões naturais apresentam altos valores de correlação para defasagens temporais de até 17 dias, o que mostra que a série não pode ser considerada como uma seqüência aleatória. Nas simulações com redes neurais, a série deve ser analisada em termos de tendência e sazonalidade. Os gráficos de auto-correlação servem também para verificar a existência de sazonalidade na série. Neste caso, o gráfico de auto-correlação aparece na forma de uma senóide. No caso das séries de vazões, observou-se que não há sazonalidade marcante, para o lag analisado. Observou-se que a série de precipitação é menos auto-correlacionada que a série de vazão, mas também não pode ser considerada como uma seqüência aleatória.

Os gráficos da FAC podem ser analisados para verificar se a série pode ser adequadamente modelada por um modelo auto-regressivo linear. Ou seja, a previsão de um registro pode ser realizada como uma combinação linear dos registros anteriores e o gráfico de auto-correlação aparece como uma reta. Este não é o caso para as séries de vazão, indicando que a modelagem adequada para essas séries podem ser não linear.

A plotagem da correlação cruzada de duas variáveis com a defasagem representa como varia a correlação entre duas séries de acordo com a defasagem de uma delas. A defasagem está relacionado

com a representatividade e a área de drenagem do posto fluviométrico. Assim, considerando um limite da correlação cruzada máxima de 0,6, observa-se que o posto de Porto Indaiá, que é o posto fluviométrico mais próximo do reservatório e tem menor área de drenagem, teve a menor defasagem com a vazão natural de Três Marias, que depende de uma vazão de até 8 dias deste posto.

Os postos de Porto Pará e Ponte dos Vilelas Jusante, Porto Mesquita e Ponte Nova do Paraopeba, estão, aos pares, na mesma sub-bacia (como mostrado na figura 2) e tiveram o mesmo defasagem. Ou seja, mesmo estando mais próximo ao reservatório, os postos de Porto Pará e Porto Mesquita não proporcionaram uma maior defasagem com a vazão natural que o posto de Ponte dos Vilelas Jusante e Ponte Nova do Paraopeba.

Observa-se que as defasagens obtidos para os postos de Porto das Andorinhas e de Iguatama, que também estão na mesma sub-bacia, são bem próximos, ficando o primeiro com uma defasagem maior, por estar mais próximo ao reservatório, conseqüentemente, com uma área de drenagem maior. Esses postos estão no curso principal do rio São Francisco e são os que têm maior representatividade e defasagem temporal com a vazão natural de Três Marias.

Observou-se também que a vazão natural de Três Marias tem uma relação temporal com até 7 dias de precipitação média da área.

A matriz de correlação das informações foi calculada usando-se o preenchimento de falhas dos postos fluviométricos através da regressão linear simples e a chuva média, onde se observou uma correlação positiva entre a vazão natural de Três Marias e a vazão do posto fluviométrico de Porto das Andorinhas. Também se observou uma grande correlação negativa entre a vazão natural e a vazão dos postos fluviométricos de Porto Indaiá, Ponte Nova do Paraopeba, Ponte dos Vilelas Jusante e entre a precipitação média da área.

Com os postos da CEMIG e da ANA, foi obtida a matriz de correlação usando-se o preenchimento de falhas através da regressão linear múltipla e a chuva média calculada por Thiessen. Destaca-se a melhor correlação obtida entre a vazão natural ao reservatório de Três Marias e o posto de Iguatama, que tinham uma correlação, com a regressão simples, no intervalo de -0,25 a -0,50 e passaram a ter uma correlação no intervalo de -0,75 a -1,0.

Substituindo-se a chuva diária média espacial calculada por Thiessen pela calculada por Kriging, não se observou diferença na matriz de correlação.

Uma Análise de Componentes Principais - ACP foi feita usando-se o preenchimento de falhas através da regressão linear simples e a chuva média calculada com os postos da CEMIG. A componente CP 1 (96,84%), a CP2 (1,94%) e a CP3 (0,51%) juntas explicam 99,29% da variância total da série (e 0,71% representa o resíduo, ou seja, é explicada por nenhuma das três componentes). A CP1 é composta, principalmente, pela informação da vazão natural de Três Marias, da vazão no posto fluviométrico de Ponte dos Vilelas e Porto Indaiá, e da precipitação média da área. A CP2 é formada pela transformação linear das vazões dos postos fluviométricos de Porto das Andorinhas, Porto Indaiá, Porto Mesquita, Porto Pará, Iguatama e Ponte Paraopeba. Observou-se que a vazão natural é a variável mais importante para explicar os eventos de grande variação, sendo seguida, respectivamente, pela precipitação média da área e pelas informações dos postos fluviométricos de Ponte dos Vilelas, Porto Indaiá e Ponte Nova do Paraopeba.

Também se efetuou uma ACP usando-se o preenchimento de falhas dos postos fluviométricos através da regressão linear múltipla e a chuva média calculada por Thiessen, com os postos da CEMIG e da ANA. A componente CP 1 (97,68%), a CP2 (1,51%) e a CP3 (0,43%) juntas explicam 99,62% da variância total da série (e 0,38% representa o resíduo, ou seja, é explicada por nenhuma das três componentes). Observa-se que a utilização da regressão linear múltipla no preenchimento de falhas e a o cálculo da chuva média também com os postos da ANA explicaram melhor a variância total da série (de 99,29 para 99,62%).

Na ACP efetuada usando-se o preenchimento de falhas dos postos fluviométricos através da regressão linear múltipla e a chuva média calculada por Kriging, com os postos da CEMIG e da ANA, obteve-se a componente CP 1 (97,66%), a CP2 (1,51%) e a CP3 (0,43%) que juntas explicam 99,60% da variância total da série. Observa-se que a utilização da chuva média por Kriging praticamente não alterou a explicabilidade da variância total da série, de 99,62 para 99,60%.

Foram feitas análises da correlação semanal da vazão natural de Três Marias com a precipitação média da área por Thiessen (usando só os postos da CEMIG) e com a calculada por Kriging. Observa-se que o coeficiente de determinação do primeiro caso foi de 0,279 e do segundo 0,323 e, com a chuva média por Thiessen (com os postos da ANA e CEMIG) o coeficiente de determinação foi de 0,327, ou seja, esta última opção apresentou uma melhor correlação com a vazão natural de Três Marias.

ATUAIS ERROS DE PREVISÕES DE VAZÕES AFLUENTES AO RESERVATÓRIO DE TRÊS MARIAS

O PMO é elaborado nas terças-feiras e suas revisões semanais, às quartas-feiras, e compreende sempre informações médias semanais operativas de sábado a sexta. Atualmente essas previsões semanais são obtidas pelo modelo PREVIVAZ (Cepel, 1997).

Para completar a semana operativa (terça ou quarta à sexta) necessita-se de vazões diárias que são solicitadas ao Agente de Geração responsável pelo aproveitamento, no caso do reservatório de Três Marias, a CEMIG, ou obtida através do modelo PREVIVAZH (Carvalho, 2001). Nessas previsões diárias a CEMIG utiliza o modelo IPH III (Mota Júnior, 1982). Nas tabelas 2 e 3 são apresentados os atuais erros de previsão de vazões naturais diárias e semanais do reservatório de Três Marias, onde MAPE representa o erro percentual médio relativo (eq.1).

$$MAPE = \frac{1}{N} \left[\sum_{p=1}^N \frac{|Z_p - Z_o|}{Z_o} \right] * 100 \tag{1}$$

Tabela 2 – Erros das previsões de vazões diárias ao reservatório de Três Marias no PMO.

Ano	MAPE (%)	
	Cemig	PrevivazH
2001	18,1	46,9
2002	21,5	50,1
2003	16,7	27,1
Média:	18,8	41,4

Tabela 3 – Erros das previsões de vazões semanais ao reservatório de Três Marias no PMO – 1ª semana operativa (sábado a sexta).

Ano	MAPE (%)		
	úmido	seco	Média
1999	-	-	34,2
2000	-	-	36,8
2001	-	-	44,7
2002	44,0	40,2	42,1
2003	40,6	22,4	31,5
2004	35,2	10,8	23,0
Média:	40,0	24,5	35,4

SIMULAÇÕES COM AS REDES NEURAIS

Neste trabalho, a previsão de vazões médias diárias ao reservatório de Três Marias, 12 dias à frente, foi feita através de duas metodologias. Na primeira usou-se uma rede MLP. Na segunda, testou-se uma rede NSRBN.

Como dados de entrada para as redes neurais, foram consideradas as informações dos postos pluviométricos (precipitação média da área), fluviométricos (7 postos) e a vazão natural verificada no reservatório de Três Marias, todas variáveis com diversas defasagens temporais (de 1 a 5 dias).

No processo de treinamento das redes neurais foi usado o método da validação cruzada, que utiliza um conjunto de dados independente para determinar o ponto de parada ótimo durante o treinamento, de forma a minimizar especialmente os riscos de superajustamento (Osório e Bittencourt, 2000). Desta forma, o conjunto de dados foi dividido em três sub-conjuntos:

- treinamento: amostra de dados usada para modificar os pesos (de 1987 a 1995). Cumpre salientar que os anos envolvidos no treinamento abrangeram secas e cheias da bacia do rio São Francisco, garantindo a representatividade da amostra;
- validação: padrões usados para evitar principalmente o problema de superajustamento (de 1996 a 1999);
- teste: padrões para testar o desempenho do modelo (de 2000 a 2004).

A seguir são descritas as particularidades de cada metodologia de rede neural utilizada.

Rede MLP

Foi usada uma rede MLP com uma camada de entrada, uma camada intermediária e uma camada de saída, treinadas com algoritmo de aprendizado supervisionado Levenberg-Marquardt (LM). Os pesos iniciais foram gerados aleatoriamente.

Como funções de transferência foram usadas as funções tansigmóide, figura 3, na camada intermediária, linear, figura 4, na camada de saída.

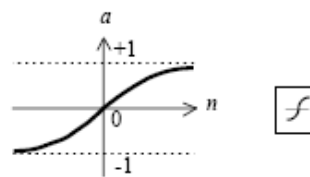


Figura 3 - Função de transferência tan-sigmoide

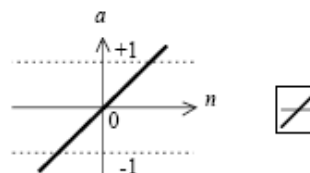


Figura 4 - Função de transferência linear

Os dados de entrada e os de saída foram normalizados, de forma a terem média zero e desvio padrão unitário e, através de uma ACP, foram desconsideradas as informações que contribuíram com menos de 0,1% da variação total dos dados.

O modelo foi configurado para interromper o treinamento quando uma das condições seguintes fosse satisfeita:

- O número máximo de épocas fosse alcançado (100). Uma época corresponde a um ciclo em que todos os dados de entrada são apresentados à rede neural;
- O desempenho fosse minimizado até a meta (zero);
- O gradiente de desempenho ficasse abaixo do gradiente mínimo ($e-10$);
- O parâmetro de ajustamento de Marquardt ($\mu=0,005$) excedesse $\mu_{\max}(e-10)$;

Visando obter melhores resultados, foi ajustado uma rede neural para o período úmido (novembro a abril) e outra para o período seco (maio a outubro).

Rede NSRBN

Foram estudadas diversas configurações de redes neurais, alterando-se as variáveis de entrada, os parâmetros do algoritmo (taxa de aprendizagem, semente para geração e momentum) e os critérios de parada (época mínima e época máxima).

Para reduzir a variabilidade das informações e dar a mesma importância a todas, os dados de entrada da vazão natural foram logaritmados.

Como no caso da rede MLP, foi ajustada uma rede para o período úmido da área em estudo (novembro a abril) e outra para o período seco (maio a outubro).

Como a rede NSRBN usada visa minimizar o erro de previsão para a média dos primeiros 7 dias, foram ajustadas duas redes neurais: uma para esses primeiros 7 primeiros dias e outra para os últimos 5 dias, totalizando os 12 dias de previsão, tanto no período úmido como no período seco. A rede ótima foi tomada como a composição das duas redes.

RESULTADOS E DISCUSSÕES

São mostrados resumidamente alguns resultados encontrados. Maiores detalhes podem ser obtidos em Gomes (2006).

A seguir são mostrados os resultados da melhor rede neural MLP do período úmido, em termos de MAPE dos dados de teste. Nas diversas simulações, foram alteradas as variáveis de entrada e suas defasagens temporais.

- O melhor modelo, em termos de MAPE médio dos 12 dias de previsão, é a simulação que considera apenas informações dos postos fluviométricos, pluviométricos e da vazão natural ocorrida, do dia em que se está fazendo a previsão.
- A precipitação é importante nas simulações.
- A informação da vazão natural pode contribuir ou não para a melhora dos resultados da simulação;
- Na simulação onde foram consideradas apenas as informações de vazão natural do reservatório de Três Marias com um, dois e três dias de defasagem, os resultados da previsão foram de péssima qualidade.

A seguir são mostrados os resultados da melhor rede neural MLP do período seco, em termos de MAPE dos dados de teste.

- O melhor modelo, em termos de MAPE médio dos 12 dias de previsão, é a simulação que considera informações dos postos fluviométricos e da vazão natural ocorrida, para diversas defasagens.

- A informação da vazão natural piorou os resultados da simulação.

Selecionando a melhor rede neural MLP para os períodos úmido e seco e colocando na mesma base de dados do PMO, as tabelas 4 e 5 mostram o desempenho das previsões diárias e semanais, respectivamente.

Tabela 4 – Erros da previsão de vazões diárias com os melhores modelos da Rede MLP.

Ano	MAPE (%)
2000	11,8
2001	17,9
2002	18,6
2003	14,1
2004	22,3
Média:	16,9

Tabela 5 – Erros da previsão de vazões semanais com os melhores modelos da Rede MLP.

Ano	MAPE (%)		
	úmido	seco	Média
2000	22,5	18,7	20,6
2001	45,3	27,0	36,1
2002	39,9	29,4	34,6
2003	34,2	19,5	26,9
2004	41,4	12,5	27,0
Média:	36,7	21,4	29,0

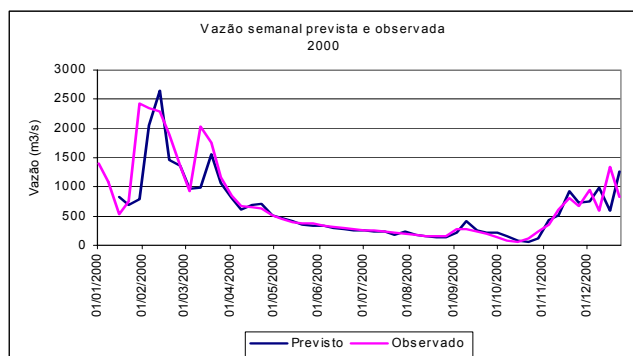


Figura 5 – Vazão prevista e verificada – rede MLP.

A Figura 5 mostra um exemplo dos resultados obtidos com a rede MLP, para o ano de 2000.

Já nas simulações com a rede NSRBN, observou-se que a alteração das variáveis básicas de entrada (alfa, semente, beta, épocas mínima e época máxima) não produziu alteração significativa, o que facilitou a simulação com esse modelo, visto que não são necessárias diversas simulações para ajuste desses parâmetros.

Também se observou que a calibração de uma rede neural para os 7 primeiros dias de previsão, leva a resultados superiores aos obtidos com a calibração de 12 dias, visto que para os primeiros 7 dias a rede NSRBN utilizada não otimiza o EMQ e sim a distância entre os centros de gravidade do hidrograma previsto e observado.

Os resultados das simulações do período úmido com a rede NSRBN para os 7 primeiros dias e para os 12 dias de previsão mostram:

- A informação da vazão natural piorou os resultados da simulação.
- A precipitação é importante nas simulações.
- A ACP ajuda na seleção das variáveis.
- A análise de correlação cruzada ajuda na seleção das variáveis.
- Considerando como critério de seleção o erro mínimo do conjunto de validação, a melhor simulação para os 7 primeiros dias foi a simulação que considera apenas informações fluviométricas, pluviométricas e de vazão natural do dia em que está fazendo a previsão, e, para os últimos 5 dias de previsão, a que, em relação a melhor rede para a previsão dos primeiros 7 dias, adiciona-se apenas mais 2 dias de precipitação.

Considerando os dados de validação e juntando as duas melhores redes NSRBN do período úmido, com um MAPE de 20,2% para os 7 primeiros dias e 36,5% para os 5 últimos dias de previsão, obteve-se um MAPE de **27,0%** para os 12 dias de previsão.

Os resultados das simulações do período seco com a rede NSRBN para os 7 primeiros dias e para os 12 dias de previsão mostram:

- A informação da vazão natural melhorou os resultados da simulação..
- A análise de Componentes Principais – ACP ajuda na seleção das variáveis.

- A análise de correlação cruzada ajuda na seleção das variáveis.
- Considerando como critério de seleção o erro mínimo do conjunto de validação, a melhor simulação para os 7 primeiros dias e para os 5 últimos dias foi a que considera informações fluviométricas com 3 dias de antecedência e do dia que se está fazendo a previsão dos postos fluviométricos de Porto das Andorinhas, Porto Indaiá, Porto Mesquita e Porto Pará, da chuva média e de vazão natural com 3, 2 e 1 dia de antecedência.

Considerando os dados de validação e juntando as duas melhores redes NSRBN do período seco, com um MAPE de 13,0% para os 7 primeiros dias e 25,4% para os 5 últimos dias de previsão, obteve-se um MAPE de **18,2%** para os 12 dias de previsão.

Selecionando o melhor modelo para os períodos úmido e seco e colocando na mesma base de dados do PMO, as tabelas 6 e 7 mostram o desempenho das previsões diárias e semanais, respectivamente.

Tabela 6 – Previsão de vazões diárias com os melhores modelos da Rede NSRBN.

Ano	MAPE (%)
2000	12,9
2001	17,1
2002	20,8
2003	12,7
2004	21,8
Média:	17,1

Tabela 7 – Previsão de vazões semanais com os melhores modelos da Rede NSRBN.

Ano	MAPE (%)		
	úmido	seco	Média
2000	24,4	18,9	21,6
2001	46,4	18,6	32,5
2002	43,6	28,9	36,3
2003	32,7	15,8	24,2
2004	34,6	12,3	23,5
Média:	36,3	18,9	27,6

A Figura 6 mostra um exemplo dos resultados obtidos com a rede NSRBN, para o ano de 2000.

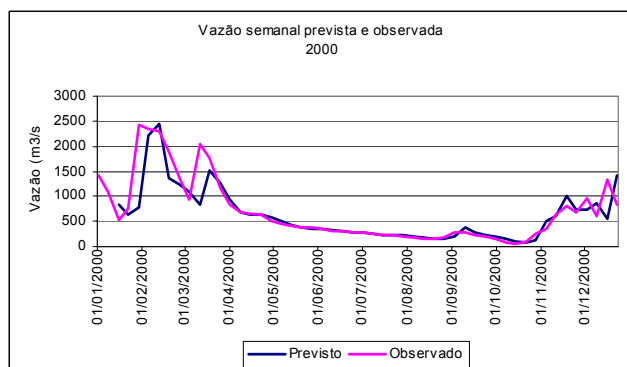


Figura 6 – Vazão prevista e verificada – rede NSRBN

Considerando os dados de validação das redes MLP, obteve-se um MAPE de **28,2%** no período úmido e **15,9%** no seco, para os 12 dias de previsão. Também considerando os dados de validação das redes NSRBN, obteve-se um MAPE de **27,0%** no período úmido e de **18,2%** no seco.

Colocando na mesma base dos dias do PMO, obteve-se um MAPE da previsão diária média dos anos de 2001 a 2003 de 16,9% para a rede MLP, de 16,9% para a rede NSRBN, enquanto a atual previsão do PREVIVAZH é de 41,4% e da CEMIG 18,8%. Já para as vazões semanais, considerando a média dos anos de 2000 a 2004, obteve-se um erro de previsão de 29,0% para a rede MLP, de 27,6% para a rede NSRBN, enquanto a atual previsão do PREVIVAZ é de 35,2%.

CONCLUSÕES

- A Função de Auto-Correlação mostrou que, para as defasagens temporais (lags) analisadas, não há efeito de sazonalidade nas séries de vazões. A informação da precipitação média espacial da bacia também não apresentou efeito de sazonalidade, mas se mostrou bem menos auto-correlacionada que as séries de vazões.
- A inclusão de postos pluviométricos nas áreas com grande erro de interpolação, apontadas no estudo geoestatístico, mostrou um deslocamento da área de precipitação máxima da bacia.

- A chuva média diária calculada com a metodologia de Kriging mostrou-se bem semelhante à calculada com o polígono de Thiessen, o que mostra que para uma rede densa de pluviômetros, o método de Thiessen gera bons valores.
- Os resultados das redes MLP e NSRBN foram bem próximos, tendo a segunda um desempenho melhor para as previsões semanais e um pouco pior para as previsões diárias.
- A melhor simulação para a rede NSRBN do período seco foi formada pelos postos selecionados através da análise de correlação cruzada, reforçando a importância da análise dos dados na definição dos variáveis de entrada da rede neural.
- Tanto nas simulações com a rede MLP quanto com a rede NSRBN, ficou clara que a adição de vazões naturais do reservatório de Três Marias de dias passados piora o desempenho das previsões para os 12 dias considerados. Também se observou a importância das informações de precipitação, principalmente nas simulações do período úmido.
- A metodologia de redes neurais possibilita a obtenção de uma previsão de vazões superiores às obtidas com os modelos estatísticos, tanto a nível diário quanto semanal.

AGRADECIMENTOS

Os autores agradecem o apoio prestado pelo Operador Nacional do Sistema Elétrico (ONS) e pela Companhia Energética de Minas Gerais (Cemig).

REFERÊNCIAS

- ANA, 2004. Plano de Recursos Hídricos da Bacia do São Francisco. Resumo Executivo.
- CEPEL, 1997. Modelo de Previsão de Vazões Semanais Aplicado ao Sistema Hidroelétrico Brasileiro – PREVIVAZ, Relatório Técnico DPP/PEL 125/97, Rio de Janeiro, Brasil.
- CAMBARDELLA, C. A., Moorman T. B., Novak J. M., Parkin T. B., Karlen, D. L., Turco, R. F., Kornopka, A. E. 1994. Fieldscale variability of soils properties in Central lo-

- wa Soils. In: Soil Science Society of America Journal, v. 58, p. 1501-1511.
- CARVALHO, A. R. L. de. 2001. Previsão de Afluências Diárias para Horizontes de 1 a 7 dias Através de uma Abordagem Estocástica – Modelo PREVIVAZH. Universidade Federal do Rio de Janeiro – UFRJ, Departamento de Engenharia Civil. Dissertação de Mestrado.
- FARACO, R. A. L.; Costa JR., P. P. da; Cruz, F. R. B. da, 1998. Minimização do erro no algoritmo back-propagation aplicado ao problema de manutenção de motores. Pesquisa Operacional, vol. 18, nº1.
- GOMES, L. F. C., 2006. Previsão de Vazões Naturais Diárias ao Reservatório de Três Marias usando a técnica de redes neurais. Universidade Federal de Pernambuco – UFPE, Departamento de Engenharia Civil. Dissertação de Mestrado.
- HAGAN, M. T.; Menhaj, M. B. 1994. Training feedforward networks with the Marquardt algorithm. IEEE Transactions on Neural Networks, 5 (6): 989-993.
- HOLAWEA, F; DUTTERB, R. 1999. Geostatistical study of precipitation series in Austria: time and space. In: Journal of Hydrology 219.
- LIMA, H. V. C.; Lima, L. C. T. M.; Lima, C. H. R. 2005. Modelagens hidrológica conceitual e por redes neurais artificiais aplicadas ao processo chuva-vazão: uma análise comparativa. In: Anais do XVI Simpósio Brasileiro de Recursos Hídricos. João Pessoa – PB.
- MOTA JÚNIOR, J. C. 1982. Simulação Hidrológica de uma bacia urbana: Arroio Dilúvio. Porto Alegre: UFRGS – Curso de Pós-Graduação em Recursos Hídricos e Saneamento. 177f. Dissertação de Mestrado.
- OSÓRIO, F.; Bittencourt, J.R., 2000. Sistemas Inteligentes baseados em RNAs aplicados ao Processamento de Imagens, I Workshop de Int. Artificial, UNISC.
- PACHECO, M. A. C.; VELLASCO, M. M; SOUZA, R. C.; 2005. Modelo de Previsão de Vazões com Incorporação de Informações de Precipitação. Relatório 4 do Projeto PreVIP entre ONS e PUC-Rio.
- VALENÇA, M. J. S. 1999. Análise e projeto de redes neurais com abordagem construtiva para modelagem de sistemas complexos. Tese de Doutorado em ciência da computação. Departamento de Informática – UFPE.
- VALENÇA, M. J. S. 2005. Aplicando redes neurais: um guia completo. Editora Livro rápido. Olinda, PE: Ed. Do Autor, 284p.
- VIEIRA, F.R. 2000. Tópicos em ciência do solo. Campinas, São Paulo – SP. Vol.1. Págs 1 a 43.

Runoff Forecasting Model Based on The Neural Networks Technique For The São Francisco River Basin

ABSTRACT

One of the main issues in hydropower plant operation planning is the runoff forecast. Different types of models can be used to perform this forecast. Examples can be cited, such as physical models, statistical models and models based on the neural network technique. The use of the latter technique has lately been intensified, because models based on this technique are easily applied and yield satisfactory results. The previous analysis of the information that will be used in the calibration and use of the neural network may have significant positive effects on its performance. This work was to show analyses of rainfall and runoff data of Três Marias reservoir, in the São Francisco River, and the calibration of a model based on the neural network technique to forecast of natural inflows. The paper presents general results of the rainfall data analyses, including estimation of missing values, consistency analyses and Geostatistics as a tool for selecting rainfall stations and estimating the average precipitation of the area. The results showed that the model calibrated using the neural network technique performed better than model PREVIVZ,, and among the neural network types analyzed, NSRBN performed slightly better than MLP.

Keywords: *neural network, runoff forecast, Três Marias reservoir.*