

SISTEMA DE APOIO À DECISÃO PARA CONTROLE DE OUTORGAS: UMA SOLUÇÃO BASEADA EM REDES NEURAI ARTIFICIAIS

Márian da Costa Rohn¹; Marco Costa²; Luciana Vaz do Nascimento²;

Rosana Colaço Gibertoni¹ & Eloy Kaviski¹

Resumo – Este artigo apresenta os primeiros resultados obtidos na realização de um sistema de apoio à decisão para controle de outorgas, utilizando uma solução baseada em redes neurais artificiais. O objetivo principal do projeto consiste na implementação de um sistema computacional capaz de avaliar novos pedidos de outorga de atividades industriais e de saneamento básico, tanto de captação de água quanto de lançamento de efluentes, verificando a compatibilidade com o tipo de atividade realizada pela empresa. Este artigo apresenta somente as estimativas da vazão captada para as atividades industriais. As redes neurais artificiais conseguiram identificaram relações entre dados característicos das indústrias e o consumo de água.

Abstract – This article presents the first results obtained after the implementation of a decision support system for water licensing control employing a method based on artificial neural networks. The project's main objective is the implementation of a computing system that evaluates new requests for water licensing made by industrial activities and basic sanitation works, as well as their compatibility with the type of activity of the company. This article refers to the inception part of the project where the estimates of water volume for industrial use are computed. The artificial neural networks identified the relations between the company's characteristic data and water consumption.

Palavras-Chave – outorga, vazão captada, redes neurais artificiais

¹ LACTEC CEHPAR - Instituto de Tecnologia para o Desenvolvimento – LACTEC - Centro de Hidráulica e Hidrologia Prof. Parigot de Souza – CEHPAR - Caixa Postal 1309, CEP 80011-970, Curitiba – PR - Fone: (0xx41) 361-6314, fax: (0xx41) 267-1754, e-mail: marian@lactec.org.br

² Montgomery Watson Brasil

INTRODUÇÃO

Vários estados brasileiros, com base em experiências similares realizadas em outros países, têm implementado programas de automonitoramento industrial e de autodeclaração para controle ambiental dos efluentes lançados em corpos hídricos. Estes programas prevêm que a empresa realize periodicamente análises e medições de vazões, informando ao órgão de controle o resultado obtido. Inspiradas no mesmo critério, algumas instituições ligadas ao gerenciamento dos recursos hídricos pretendem implantar modelos similares para o controle das outorgas de captação e lançamento de efluentes.

Este tipo de automonitoramento tem a vantagem de suprir parte da carência de recursos humanos e financeiros pela qual passam os órgãos de controle, impossibilitados de realizar uma fiscalização constante. A coleta, a análise físico-química das amostras e as medições de vazão de água captada e de efluentes, neste caso, são realizadas diretamente pelas indústrias geradoras, sem custo para o órgão de controle. Por outro lado, quem está fornecendo dados para o órgão de controle é quem deveria ser controlado, o que pode sugerir eventuais declarações incorretas.

Para que seja bem sucedido, um programa de autocontrole deveria estar apoiado a ações de auditoria, envolvendo coletas periódicas e aleatórias de efluentes, realizadas pelo órgão competente. Mas este processo, muitas vezes, esbarra nas mesmas carências financeiras citadas anteriormente. Sendo assim, é necessário aproveitar as técnicas mais arrojadas envolvendo *Sistemas de Informações* que permitam a utilização, por parte do órgão de controle, de um sistema computacional que auxilie no controle dos dados recebidos junto ao automonitoramento e às autodeclarações.

O conceito é simples: com base num banco de dados de indústrias brasileiras são calculadas, levando em consideração a tipologia de atividade industrial, estimativas do consumo de água e cargas lançadas. Estes valores estimados são comparados aos declarados pela empresa e é verificada a compatibilidade entre o declarado e o esperado. Quando os valores declarados pela indústria ultrapassam o que foi estimado, a empresa poderá ser questionada sobre os resultados. Nesta fase poderá ser realizada uma hierarquização das intervenções partindo das empresas que mais se distanciaram do esperado e realizar fiscalizações em campo direcionadas exclusivamente a essas empresas.

A SUDERHSA (Superintendência de Desenvolvimento de Recursos Hídricos e Saneamento Ambiental) contratou o consórcio de empresas, Montgomery Watson Brasil e COBRAPE, para desenvolver e implantar o *Sistema de Declaração de Uso dos Recursos Hídricos e de Automonitoramento* e elaborar o *Manual de Outorgas* para o Estado do Paraná. Um dos objetivos

desta contratação é a implementação das ferramentas de automonitoramento e autodeclaração como base para o controle de outorga de direito de uso do Estado do Paraná. Dentre as ferramentas que estão sendo desenvolvidas, destaca-se o modelo matemático de estimativa da vazão captada, vazão lançada e carga poluidora dos usuários outorgados de recursos hídricos.

Neste contexto, o presente artigo descreve exclusivamente as atividades relacionadas ao desenvolvimento de um modelo matemático baseado em redes neurais artificiais (RNA) para estimar vazão captada das bacias hidrográficas, assim como vazão lançada e alguns parâmetros referentes à qualidade da água. O modelo está sendo desenvolvido pela Montgomery Watson Brasil junto ao Centro de Hidráulica e Hidrologia Professor Parigot de Souza – CEHPAR, do Instituto de Tecnologia para o Desenvolvimento – LACTEC.

As redes neurais artificiais são uma técnica integrante dos Sistemas Inteligentes, inspirados no comportamento humano ou da natureza. São alternativas às técnicas usuais e já consagradas da estatística, pesquisa operacional e modelagem numérica, e vêm ganhando destaque na aplicação de problemas de classificação de dados.

Na fase inicial do projeto, somente foram realizadas estimativas da vazão captada. Estas estimativas foram efetuadas em função de parâmetros característicos do usuário dos recursos hídricos. Está em fase experimental o modelo para estimar outras variáveis, tais como vazão efluente e carga poluidora, através de parâmetros como DBO (demanda bioquímica de oxigênio), DQO (demanda química de oxigênio) e sólidos suspensos.

O sistema ora proposto se configura como uma importante ferramenta de planejamento e assistência na tomada de decisões sobre o controle ambiental dos recursos hídricos disponíveis para o setor industrial. É importante ressaltar que os dados coletados para a Bacia do Alto Iguaçu terão condição de ser extrapolados para todo o Estado do Paraná e poderão, desde que devidamente complementados com dados locais, ser utilizados para estimativas em outros estados brasileiros.

DADOS UTILIZADOS NO MODELO

O modelo matemático desenvolvido estima a vazão captada do corpo hídrico. Estas estimativas foram efetuadas com base em parâmetros característicos do usuário dos recursos hídricos.

Os estudos foram conduzidos com um conjunto de dados de indústrias da bacia do Alto Iguaçu na região Metropolitana de Curitiba, Paraná, que realizam captação de água nos mananciais da região. Foram também considerados dados de indústrias que utilizam, para abastecimento, exclusivamente água da rede pública.

O modelo possui restrições a algumas atividades que não fazem parte da amostra ou que ao menos não são semelhantes com as disponíveis. As atividades mais relevantes do ponto de vista da

quantidade de água captada e da quantidade e qualidade dos efluentes lançados estão integralmente representadas no banco de dados atual. Futuramente, realizada a coleta de novos dados sobre atividades ora não presentes no sistema, o modelo poderá ser estendido ao universo industrial do país.

O modelo desenvolvido é válido para usuários situados em qualquer bacia hidrográfica, pois os parâmetros pesquisados dependem principalmente do nível tecnológico e da produção atingida. Na escolha das variáveis utilizadas na elaboração do modelo foram selecionadas aquelas que fossem de fácil acesso e que a empresa não tivesse dificuldade em declarar. Foram evitadas variáveis tais como quantidade e tipologia de matéria prima ou de produto acabado por serem informações que nem sempre as indústrias têm interesse em divulgar.

As variáveis utilizadas na elaboração do modelo são:

- a) CNAE (Classificação Nacional de Atividades Econômicas), que é um instrumento de identificação econômica das unidades produtivas, uniformizado nacionalmente, seguindo padrões internacionais definidos no âmbito da ONU. O CNAE é composto por 5 dígitos (<http://www.receita.fazenda.gov.br/srf.www>);
- b) número de empregados da empresa;
- c) energia elétrica consumida por mês (kWh/mês);
- d) vazão captada (m³/mês);

O código CNAE é a variável que auxiliou na classificação de dados necessária na calibração do modelo. Este código identifica os segmentos produtivos através de uma estrutura composta por 3 níveis hierárquicos. O primeiro nível é formado pelos dois primeiros dígitos, formando a Divisão. O Grupo corresponde ao segundo nível e é composto pelo terceiro dígito. O quarto dígito compõe o terceiro nível constituindo as Classes.

Os dados foram divididos em vários conjuntos, sempre levando-se em consideração o tipo de indústria a que pertenciam. O número de funcionários e o consumo de energia são variáveis fortemente associadas ao nível de produção. Até certo ponto, estas variáveis utilizadas em conjunto relacionam-se também com o nível tecnológico das indústrias. Portanto, as variáveis explicativas na estimativa da vazão captada são: número de funcionários da empresa e consumo mensal de energia.

REDES NEURAIS

As redes neurais artificiais são modelos matemáticos que procuram imitar o funcionamento do cérebro humano, formado por milhões de neurônios que se interconectam entre si. O cérebro humano é um sistema de processamento de informação altamente complexo, não-linear e paralelo que tem a capacidade de organizar os neurônios de tal forma que seu desempenho na execução de

certas funções (reconhecimento de exemplos) dá-se de forma mais rápida que os computadores atuais.

Recentemente as redes neurais têm mostrado um bom desempenho como uma ferramenta de regressão, especialmente se utilizada para o reconhecimento de padrões e aproximação de funções. Uma importante característica das redes neurais reside no fato que não é necessário um conhecimento detalhado sobre as relações entre as variáveis envolvidas no problema.

As redes necessitam de uma quantidade considerável de dados históricos para que consigam extrair satisfatoriamente as características relevantes existentes no conjunto de dados. Se treinada corretamente, a rede é capaz não somente de aproximar qualquer função, mas também de generalizar, proporcionando saídas corretas para entradas não apresentadas anteriormente.

Fundamentos

O neurônio possui um ou mais sinais de entrada e um sinal de saída. Sua semelhança com o neurônio biológico é muito grande, pois ambos possuem disparos de saída podendo receber muitas entradas.

As entradas de um neurônio artificial podem ser comparadas exatamente como estímulos para o neurônio natural. Todos estes estímulos são trazidos até o neurônio simultaneamente, ou seja, se um neurônio possuir cinco entradas, os sinais das cinco entradas deverão chegar até o núcleo de processamento ao mesmo tempo, isto é, paralelamente.

Os pesos são um atributo importantíssimo dos neurônios e podem ser comparados com os dendritos realizando as suas sinapses em outros neurônios. Por esta razão, eles são chamados de pesos sinápticos. Os pesos (w) são valores que representam o grau de importância que determinada entrada possui em relação a um determinado neurônio. Um peso, quando bastante estimulado, passa a ter cada vez mais influência no resultado do sinal de saída. O objetivo principal é, portanto, determinar o melhor conjunto de pesos capaz de representar um dado problema.

Os pesos podem ser vistos, matematicamente, como um vetor de valores (w_1, w_2, \dots, w_n). As conexões com pesos positivos representam conexões excitatórias (reforço na ativação do neurônio) e as com pesos negativos, conexões inibitórias (inibição na ativação do neurônio). Havendo mais de um neurônio na rede, tem-se uma coleção de vetores, ou seja, uma matriz de pesos na qual cada vetor corresponde a um neurônio. Quando as entradas (x_1, x_2, \dots, x_n) são apresentadas para o neurônio, elas são multiplicadas pelos pesos e a soma destes resultados é o sinal de excitação no neurônio. As entradas multiplicadas pelos pesos recebem, depois dessa operação, o nome de entradas ponderadas.

A rede pode aprender qual das combinações de pesos é a mais relevante para uma dada situação. A combinação mais relevante irá enviar o sinal mais forte, isto é, o valor mais alto, para as

camadas seguintes e terá, conseqüentemente, um maior efeito na solução final, do que as demais combinações menos relevantes para um caso particular.

A função de ativação antecede a função de transferência e tem por atribuição repassar o sinal para a saída do neurônio. A função de ativação é uma função de ordem interna, cuja atribuição é fazer acontecer um nível de ativação dentro do próprio neurônio, ou seja, é uma decisão tomada pelo neurônio sobre o que fazer com o valor resultante do somatório das entradas ponderadas. Essa decisão terá um efeito restrito ao próprio neurônio. Usualmente, a função de ativação é a própria função de soma das entradas ponderadas do neurônio.

A função de transferência pode ter muitas formas e métodos. Dentre as mais utilizadas podem ser citadas as funções linear, rampa e sigmóide. A função de transferência define e envia para fora do neurônio o valor passado pela função de ativação.

As relações entre as entradas, saídas e os pesos são similares a uma regressão linear, mas se uma função de transferência não-linear for aplicada, ela permitirá que a rede aproxime, de maneira mais próxima do que os métodos puramente lineares, relações não-lineares entre as variáveis de entrada e saída.

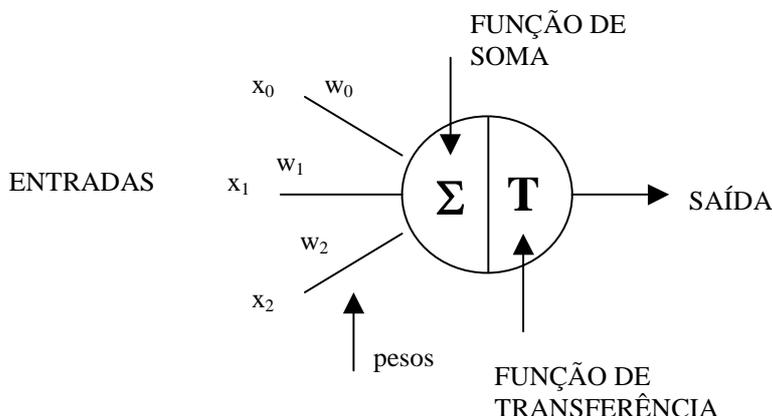


FIGURA 1 – NEURÔNIO ARTIFICIAL

Fonte: TAFNER et al. (1996)

A combinação de diversos neurônios é chamada de rede neural. As entradas podem ser conectadas em muitos neurônios com vários pesos, resultando assim em uma série de saídas, uma para cada neurônio. Essas conexões, em comparação com o sistema biológico, representam o contato dos dendritos com outro neurônio formando a sinapse.

A camada de entrada tem esse nome apenas como figurativo, pois sua única função é armazenar a informação de entrada para ser passada para a camada seguinte de neurônios. Os

elementos dessa camada normalmente possuem um nome diferente de neurônio para não serem confundidos com neurônios de processamento.

As redes neurais podem possuir camadas escondidas (*hidden*), também chamadas de intermediárias ou ocultas. Essas camadas intermediárias situam-se entre a camada de entrada e a camada de saída da rede neural. Não existe uma regra que defina o número de camadas escondidas, podendo variar de 0 a n camadas. Essas camadas são compostas por neurônios que possuem exatamente a mesma estrutura que os neurônios da camada de saída. Sua única diferença é não ter contato com o mundo externo e assim, os sinais são passados para os outros neurônios obedecendo as regras de transferência de cada neurônio. O número de neurônios da camada escondida também não obedece a nenhuma regra específica. São os neurônios das camadas escondidas que permitem que a rede detecte e capture as características relevantes nos dados, e realize um complexo mapeamento não-linear entre as variáveis de entrada e saída.

A arquitetura da rede neural é livre, podendo ser modificada conforme a necessidade do projetista. A escolha da rede reside em sua habilidade ou não de realizar o mapeamento da função em questão. A performance de um modelo de RNA depende da adequação da sua arquitetura em simular determinado problema. A arquitetura ótima é encontrada através de um processo de tentativa e erro, que por vezes pode levar um tempo considerável até ser encontrado.

O aprendizado de uma rede neural envolve a modificação do seu padrão de interconexão (pesos). Neste trabalho foi utilizado o aprendizado supervisionado, realizado da seguinte maneira: a rede deverá possuir pares de entrada e saída, ou seja, um conjunto de entradas e um conjunto com as saídas desejadas para cada entrada. Toda vez que uma entrada for apresentada à rede, deverá se verificar se a saída obtida (gerada a partir dos cálculos efetuados com os pesos que a rede possui) confere com a saída desejada para aquela entrada. Se for diferente, a rede deverá ajustar os pesos de forma que armazene o conhecimento desejado. Essa iteratividade do treinamento deverá ser repetida com todo o conjunto de treinamento (entradas e saídas) a fim de que as respostas da rede estejam mais próximas das saídas desejadas.

O ajuste sináptico é o aprendizado em cada neurônio do fato apresentado, ou seja, cada neurônio, em conjunto com todos os outros, representa a informação que passou pela rede. Nenhum neurônio guarda em si todo o conhecimento, mas faz parte de uma malha que retém a informação graças a todos os seus neurônios. O conhecimento dos neurônios reside nos pesos sinápticos. O ajuste é realizado em função de um cálculo que aponta a quantidade de erro do resultado (saída). Esse ajuste procura corrigir os pesos de modo que se produza a saída desejada diante da respectiva entrada. O ajuste sináptico é resultado de um cálculo matemático. Esse cálculo visa somar ao peso atual um valor que corresponda a quantidade de erro gerada pela rede, e desta forma, corrigir o valor do peso.

A rede utilizada neste estudo é chamada de rede de retropropagação ou *feedforward*, pois o processamento da informação dá-se no sentido progressivo, através das interconexões entre os neurônios das camadas adjacentes. Cada unidade de uma camada é conectada para a frente a cada unidade da camada seguinte. As ativações fluem da camada de entrada para a camada oculta, e daí para a camada de saída.

Esta rede geralmente é treinada com a regra delta generalizada, que consiste na aplicação do gradiente decrescente ao problema de múltiplas camadas e também é conhecido como algoritmo de *backpropagation* (GALVÃO, 1999). Essa regra consiste na aplicação do gradiente descendente ao problema de múltiplas camadas, utilizando-se o algoritmo de *backpropagation*. O gradiente descendente tem como objetivo a busca de um mínimo global, através de incrementos nos valores dos pesos por uma quantidade que é proporcional à primeira derivada (o gradiente) do erro entre o valor da saída desejada e o valor atual de saída fornecido pelo elemento de processamento em relação ao valor do peso. O mínimo global é caracterizado como uma solução teórica ideal que apresenta o menor erro possível.

O termo *backpropagation* está relacionado, na realidade, com a propagação recursiva dos erros, ou seja, a importância do algoritmo está em fornecer um método eficiente de cálculo das derivadas da função erro. Portanto, o treinamento pode ser dividido em duas etapas: a primeira em que os erros são propagados recursivamente através da rede, de tal forma que o termo *backpropagation* é aqui utilizado para representar a determinação das derivadas da função erro, e uma segunda etapa na qual estas derivadas são utilizadas para realizar o ajuste nos pesos.

Existem vários outros algoritmos para a minimização do erro. No entanto, diversos autores concordam que o método Levenberg-Marquardt apresenta melhores resultados no tratamento de problemas não-lineares (SPERB et al., 1999), razão pela qual este algoritmo de treinamento foi utilizado neste estudo.

O método Levenberg-Marquardt (LM) é um algoritmo do tipo “quase Newton”, que utiliza a informação das segundas derivadas, ao contrário do gradiente descendente, que utiliza apenas a primeira. Por este motivo, o algoritmo LM tende a ser mais eficiente que o gradiente descendente. Contudo, há a necessidade de se armazenar a matriz Hessiana, que contém as segundas derivadas, o que resulta num uso significativamente maior da memória do computador, podendo comprometer sua eficiência no treinamento de redes muito grandes.

Modelo Proposto

Determinados problemas são bem apropriados para a capacidade de reconhecimento de padrões de uma rede neural, ao passo que outros são melhor resolvidos a partir de métodos

tradicionais. As redes neurais são largamente indicadas, por exemplo, em casos nos quais há pouco conhecimento das relações entre as variáveis de entrada e saída.

A maioria dos modelos de redes neurais possui alguma regra de treinamento, na qual os pesos de suas conexões são ajustados de acordo com os padrões apresentados. Em outras palavras, elas aprendem através de exemplos. Deste modo, fica evidente a potencialidade das redes na resolução de problemas a partir de experiências anteriores, desde que estas experiências estejam representadas por um número considerável de dados (ROHN,2002).

Na construção deste modelo, foi realizada uma classificação prévia com o propósito de agrupar os dados pertencentes a uma mesma divisão, ou seja, unir indústrias de tipos similares.

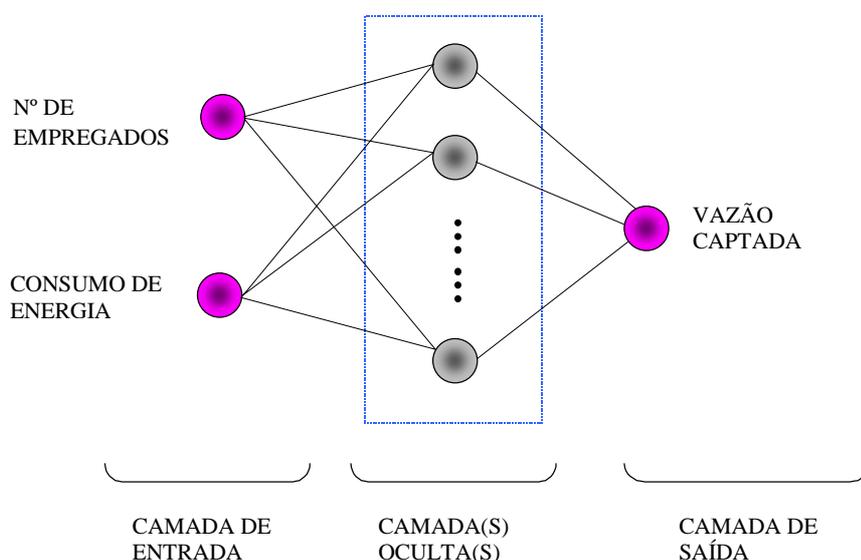


Figura 2 – Modelo de rede proposto para estimativa da vazão captada

Como o desempenho das redes é extremamente dependente do seu conjunto de treinamento, foram realizadas simulações para tipos de indústrias com pelo menos 8 pontos observados.

Para cada tipo de indústria para a qual havia dados disponíveis, foi desenvolvida uma rede nos moldes da figura 2. Conforme mostra esta figura, a camada de entrada é composta inicialmente por duas variáveis: o número de empregados da empresa e o consumo de energia elétrica. O número de camadas ocultas, que define a complexidade da rede, não tem um valor definido, sendo, portanto, um importante parâmetro a ser determinado no decorrer do estudo.

Freqüentemente é útil adaptar as entradas e saídas a uma escala mais apropriada, de modo que estes valores estejam sempre em um mesmo intervalo. Para atender ao modelo proposto, os dados passaram por um processo de normalização, já citado anteriormente, e que consiste na divisão dos dados pelo valor máximo de cada variável, de modo que a magnitude dos dados fique restrita ao

intervalo $[0,1]$. Este procedimento de adaptar as entradas e saídas a uma mesma escala visa simplificar o processo de treinamento.

Concluída esta etapa, os dados foram separados em dois conjuntos: conjunto de treinamento, que foi utilizado para o treinamento da rede e conjunto de teste, que foi utilizado para verificar o desempenho da rede sob condições reais de utilização. Esta tarefa requer uma análise cuidadosa sobre o problema para minimizar ambigüidades e erros nos dados. Além disso, os dados coletados devem ser significativos e cobrir amplamente o domínio do problema.

A etapa sucessiva foi a definição da configuração da rede, que pode ser dividida em três sub-etapas: seleção do tipo neural apropriado à aplicação, determinação da topologia da rede a ser utilizada (o número de camadas, o número de unidades em cada camada, etc.), e a determinação de parâmetros do algoritmo de treinamento e funções de transferência. Estas etapas têm um grande impacto no desempenho do sistema resultante. Esta fase de definição é bastante delicada, pois envolve a obtenção do conjunto de variáveis significativas para a resolução do problema. Esta obtenção implica na identificação das variáveis relacionadas ao problema, bem como na remoção de variáveis não confiáveis para o processo.

A arquitetura de uma rede consiste na descrição de quantas camadas esta rede possui, do número de neurônios em cada camada, das funções de transferências e da maneira pela qual estas camadas estão conectadas. A melhor arquitetura depende do tipo de problema a ser representado pela rede. Neste trabalho, todas as redes criadas possuem somente uma camada escondida.

Fletcher e Goss (apud KIM e BARROS, 2001) propuseram um intervalo de valores para determinar um número mais adequado de neurônios na camada escondida. Este intervalo vai de $(2n+1)$ até $(2n^{0.5}+m)$, sendo n o número de neurônios da camada de entrada e m o número de neurônios da camada de saída. Esta fórmula, no entanto, só serve como um guia. Os melhores resultados são obtidos através de um processo de tentativa e erro.

A escolha de função de transferência também exerce um papel importante no desempenho do sistema. Neste estudo uma função sigmóide foi adotada para as camadas ocultas, ao passo que uma função linear foi empregada na camada de saída. A utilização de uma função de transferência não-linear permite que a rede armazene relações não-lineares entre os dados de entrada e saída. Esta função, contudo, não foi adotada na camada de saída porque impõe um resultado entre 0 e 1. Isto obriga que a saída da rede esteja limitada a um valor máximo conhecido. A fim de contornar esta situação, foi utilizada uma função de transferência linear.

De maneira geral, pode-se dizer que qualquer função pode ser representada por uma rede com duas camadas (oculta e de saída), desde que suas funções de transferência sejam, respectivamente, sigmóide e linear.

O quarto passo foi o treinamento da rede. O treinamento inicia quando um exemplo é mostrado à rede e termina quando todas as conexões estão devidamente ajustadas para reconhecer os exemplos mostrados à rede. Quanto mais completo o conjunto de dados submetidos à rede, melhor será o seu treinamento.

Para se obter um bom conjunto de dados, deve-se levar em conta certos aspectos. Primeiramente, deve-se buscar dados que tenham um grau de pertinência elevado em relação ao escopo do problema. Deve-se também obter um número elevado de conjuntos de dados para que, submetidos à rede, efetuem um bom treinamento. É válida a regra que diz que, quanto maior a experiência, melhor o desempenho. Se a rede for treinada por uma quantidade considerável de dados, ela será não somente capaz de aproximar qualquer função, como também estará apta a generalizar, proporcionando resultados coerentes para entradas jamais vistas.

Nesta fase, seguindo o algoritmo de treinamento escolhido, foram ajustados os pesos das conexões. Neste trabalho, foi utilizado o algoritmo de treinamento de Levenberg-Marquardt, que é tido como o algoritmo que apresenta melhores resultados no tratamento de problemas não-lineares.

É importante também que sejam considerados alguns aspectos tais como a inicialização da rede, o modo de treinamento e o tempo de treinamento.

Uma boa escolha dos valores iniciais dos pesos da rede pode diminuir o tempo necessário para o treinamento. Neste caso, os pesos da rede foram inicializados randomicamente dentro do intervalo $[-1,1]$.

Quanto ao tempo de treinamento, vários fatores podem influenciar a sua duração, porém sempre será necessário utilizar algum critério de parada. Um dos problemas que ocorrem durante o treinamento da rede é chamado *overfitting*. Quando isto ocorre, o erro no conjunto de treinamento é muito pequeno, mas quando a rede é testada com um novo conjunto de dados, o erro é grande. A rede memoriza os exemplos do conjunto de treinamento, mas não aprende a generalizar para novas situações.

A tabela 1 resume as quantidades de pontos utilizadas para treinamento e teste, para cada rede. O número de identificação do teste corresponde aos dois primeiros dígitos do código CNAE, ou seja, estão relacionados ao tipo de atividade da indústria.

Tabela 1 – Divisão dos pontos

Variável estimada	Número do teste	Conjunto treinamento	Conjunto teste
Vazão captada	Atividade15	30	5
	Atividade17	9	1
	Atividade20	23	3
	Atividade21	11	2
	Atividade22	7	1
	Atividade24	24	3
	Atividade25	15	2
	Atividade26	9	1
	Atividade28	18	3
	Atividade29	22	3
	Atividade31	9	1
	Atividade34	10	1
	Atividade36	18	2

De maneira geral, o treinamento deve ser interrompido quando a rede apresentar uma boa capacidade de generalização e quando a taxa de erro for suficientemente pequena, ou seja, menor que um erro admissível. Assim, o objetivo é encontrar um ponto ótimo de parada com erro mínimo e capacidade de generalização ótima.

Para encontrar este ponto ótimo, foram simulados, para cada modelo, vários testes variando-se o número de épocas. O número de épocas corresponde ao número de vezes em que os dados são apresentados à rede.

RESULTADOS

De maneira geral, pode-se perceber que com o aumento do número de épocas, o erro no conjunto de treinamento diminui consideravelmente. Isso pode ser explicado pelo fato que a rede torna-se especialista somente para aquele conjunto de dados específico e perde sua capacidade de generalização, tendo em vista os resultados inferiores para o conjunto de teste. Por outro lado, se o número de épocas for muito pequeno, a rede não conseguirá extrair as características relevantes intrínsecas ao conjunto de dados. Por isso, vários testes foram realizados para cada rede criada, variando-se o número de épocas, com o intuito de encontrar um ponto ótimo de treinamento, ou seja, que corresponda a uma máxima capacidade de generalização com um erro mínimo.

A tabela 2 traz todas as redes criadas com os respectivos número de épocas e número de neurônios na camada escondida.

Tabela 2 – Configuração das redes

Variável estimada	Número do teste	Nº de neurônios	Nº de épocas
Vazão captada	Atividade15	17	20
	Atividade17	2	40
	Atividade20	15	50
	Atividade21	14	30
	Atividade22	7	5
	Atividade24	7	11
	Atividade25	10	18
	Atividade26	5	20
	Atividade28	10	100
	Atividade29	10	48
	Atividade31	10	50
	Atividade34	10	100
	Atividade36	5	31

A fim de medir o desempenho da rede, e também como forma de selecionar a rede com melhores resultados, procurou-se verificar o seu comportamento durante o treinamento e também através de simulações com o conjunto de teste. Para o conjunto de treinamento, levou-se em consideração o coeficiente R, da regressão realizada entre os dados observados e os valores estimados. Se R for igual a unidade, isto indica uma perfeita correlação entre os valores. Para o conjunto de teste, foi obtido o erro médio quadrático, calculado da seguinte maneira:

$$e = \frac{1}{n} \sum_{k=1}^n e(k)^2 = \frac{1}{n} (y(k) - \hat{y}(k))^2 \quad (1)$$

sendo y a grandeza observada, e \hat{y} a grandeza estimada.

A tabela 3 mostra estes coeficientes para todas as redes.

Tabela 3 – Medidas do desempenho dos modelos

Variável estimada	Número do teste	R (treinamento)	Erro médio quadrático (teste)
Vazão captada	Atividade15	0,9301	$5,13 \times 10^7$
	Atividade17	0,9901	488,4809
	Atividade20	0,9822	$5,015 \times 10^3$
	Atividade21	1	$1,48 \times 10^8$
	Atividade22	0,9909	$7,31 \times 10^2$
	Atividade24	0,8323	$1,32 \times 10^6$
	Atividade25	0,9336	$9,16 \times 10^3$
	Atividade26	0,9978	$5,571 \times 10^5$
	Atividade28	0,9998	178,361
	Atividade29	0,9324	$2,27 \times 10^4$
	Atividade31	0,9982	228,416
	Atividade34	1	$6,1021 \times 10^3$
	Atividade36	0,8220	$1,18 \times 10^4$

As tabelas 4 e 5 trazem os resultados das estimativas de vazão captada para dois grupos.

Tabela 4 – Estimativa de vazão captada para o grupo 17 (Fabricação de produtos têxteis)

Tipo de dado	Vazão captada estimada (m ³ /mês)	Vazão captada observada (m ³ /mês)
Treinamento	14858	14860
	472	2256
	432	53
	37	13
	322	58
	426	40
	466	16
	180	23
	356	251
	Teste	184,89

Tabela 5 – Estimativa de vazão captada para o grupo 34 (Fabricação e montagem de veículos automotores, reboques e carrocerias)

Tipo de dado	Vazão captada estimada (m ³ /mês)	Vazão captada observada (m ³ /mês)
Treinamento	5998	6000
	54593	54600
	11990	11997
	2766	2735
	16301	16300
	53	19
	47	58
	80	16
	399	523
	31026	31024
Teste	349,12	271

Pode-se perceber que na grande maioria dos testes para estimativa da vazão captada, o modelo obedeceu a tendência de superestimar ou subestimar demasiadamente as vazões inferiores. As vazões superiores, por outro lado, foram melhor estimadas. Para o fim a que se destina o modelo, pode-se considerar que o fato do modelo estimar bem as vazões superiores, mesmo em condições adversas, representa um ponto positivo que estimula a continuidade deste trabalho.

CONCLUSÕES

O trabalho apresentou os primeiros resultados obtidos na realização de um sistema de apoio à decisão para controle de outorgas, utilizando uma solução baseada em redes neurais artificiais. Durante o desenvolvimento do modelo, foi testada também a análise estatística multivariada, implementando a análise de agrupamentos pela lógica difusa. Contudo, as redes neurais artificiais apresentaram resultados superiores para esta aplicação.

O objetivo da primeira fase do projeto foi a implementação de um modelo matemático com capacidade de avaliar novos pedidos de outorga de captação de água para atividades industriais e se estes pedidos são compatíveis com o tipo de atividade realizada pela empresa.

O sistema se configura como uma importante ferramenta de planejamento e assistência na tomada de decisões sobre o controle ambiental dos recursos hídricos disponíveis para o setor industrial.

Está em fase experimental um modelo que permita a estimativa de outras variáveis, tais como vazão efluente e carga poluidora, através de parâmetros como DBO (demanda bioquímica de oxigênio), DQO (demanda química de oxigênio) e sólidos suspensos.

É importante ressaltar que os dados coletados para a Bacia do Alto Iguaçu podem ser extrapolados para todo o Estado do Paraná e sucessivamente poderão, se devidamente complementados com dados locais, ser utilizados para estimativas em outros estados brasileiros.

REFERÊNCIAS BIBLIOGRÁFICAS

- GALVÃO, C. de O. Controle de processos e sistemas de apoio à decisão. In: GALVÃO, C. de O. e VALENÇA, M. J. S. Sistemas inteligentes: aplicações a recursos hídricos e sistemas ambientais. Porto Alegre: Ed. Universidade/UFRGS/ABRH, pp. 225-246 (Coleção ABRH de Recursos Hídricos; 7.), 1999.
- KIM, G.; BARROS, A. P. Quantitative flood forecasting using multisensor data and neural networks. **Journal of Hydrology**, Amsterdam, v.246, p.45-62, 2001.
- ROHN, M. da C. Uma aplicação das redes neurais artificiais à previsão de chuvas de curtíssimo prazo. Dissertação de Mestrado, UFPR, Programa de Pós-Graduação em Engenharia de Recursos Hídricos e Ambiental, Curitiba: novembro, 2002.
- SPERB, R. M. et al. Prevent: protótipo de um sistema de previsão de enchentes baseado em redes neurais. In: SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS, 13., 1999, Belo Horizonte-MG. **Anais...** Porto Alegre : ABRH, 1999. 1 CDRom.
- TAFNER, M. A.; XEREZ, M. de; RODRIGUES FILHO, I. W. **Redes neurais artificiais: introdução e princípios de neurocomputação**. Blumenau: EKO, 1996. 202 p., 1 disquete.