



**Simulação dos Níveis Freáticos em Poço Tubular**  
**Localizado em Aquífero Livre: uma Comparação de Técnicas Preditivas**  
Simulation of Water Table in Tubular Well  
Located in Unconfined Aquifer: a Comparison of Predictive Techniques

Thiago Boeno Patricio Luiz<sup>1</sup>; Thomas Schröder<sup>2</sup> & José Luiz Silvério da Silva<sup>3</sup>

<sup>1</sup>Universidade Federal de Santa Maria, Departamento de Engenharia Civil,  
Avenida Roraima, 1000, Camobi, 97105-900, Santa Maria, RS, Brasil

<sup>2</sup>Universidade Federal de Santa Maria, Departamento de Ciências Florestais,  
Avenida Roraima, 1000, Camobi, 97105-900, Santa Maria, RS, Brasil

<sup>3</sup>Universidade Federal de Santa Maria, Departamento de Geociências,  
Avenida Roraima, 1000, Camobi, 97105-900, Santa Maria, RS, Brasil

E-mails: thiagoboeno@hotmail.com; thomaschroder@gmail.com; silverioufsm@gmail.com

Recebido em: 16/07/2018      Aprovado em: 20/09/2018

DOI: [http://dx.doi.org/10.11137/2018\\_3\\_227\\_238](http://dx.doi.org/10.11137/2018_3_227_238)

## Resumo

A utilização de modelos autorregressivos para a predição e preenchimento de falhas a partir de uma série histórica de dados vem, notadamente, crescendo dentro da ciência hidrológica, especialmente para as atividades que envolvem o estudo de balanço hídrico e o gerenciamento da demanda hídrica em bacias hidrográficas. Este trabalho teve como objetivo examinar e comparar o uso potencial de diferentes técnicas preditivas em modelos autorregressivos para a simulação dos níveis freáticos em um poço tubular localizado em aquífero livre no Estado do Rio Grande do Sul, Brasil. Foram utilizados métodos Autorregressivos e de Médias Móveis (ARMA), Máquinas de Vetores de Suporte (MVS) e Redes Neurais Artificiais (RNA), buscando-se analisar o desempenho dessas metodologias para a simulação de níveis da água subterrânea em escala horária e diária. Para tanto, utilizaram-se séries históricas de nível estático e de precipitação pluviométrica, coletadas a cada hora por meio de uma estação de monitoramento automática com *data logger*. A metodologia baseada em Redes Neurais Artificiais apresentou a melhor performance, evidenciada pelo Coeficiente de Nash-Sutcliffe (CNS) na ordem de 0,99 na escala horária e, na ordem de 0,84 para a escala diária. A etapa de análise dos resíduos demonstrou a pequena margem de erro alcançada, permitindo validar o modelo para práticas e estudos em séries temporais de água subterrânea.

**Palavras-chave:** Água subterrânea; Modelagem hidrológica; Autorregressivo de Médias Móveis; Máquinas de Vetores de Suporte; Redes Neurais Artificiais

## Abstract

The use of autoregressive models for the prediction and gap filling from historical time-series has grown notably in the study of hydrological data, especially for activities that involves the study of water balance and hydric demand management in watersheds. This work aimed to examine and compare the potential use of different predictive techniques of autoregressive models for groundwater level simulation in tubular well located in unconfined aquifer in the State of Rio Grande do Sul, Brazil. Autoregressive Moving Averages (ARMA), Support Vector Machines (SVM) and Artificial Neural Networks (ANN) methods were used to analyze the performance of these methodologies for the prediction of groundwater levels on hourly and daily scale. Static water level and rainfall series were collected every hour by means of an automatic data logger monitoring station. The methodology based on Artificial Neural Networks presented the best performance, evidenced by the Nash-Sutcliffe Coefficient (NSC) in the order of 0,99 in hourly scale and, in order of 0,84 in daily scale. The residue analysis stage demonstrated the small margin of error achieved, allowing to validate the model for practices and studies in time series of groundwater levels.

**Keywords:** Groundwater; Hydrologic modeling; Autoregressive Moving Averages; Support Vector Machines; Artificial Neural Networks

## 1 Introdução

Estudos que englobam simulações de variáveis hidrológicas são de extrema importância para o planejamento de atividades de gestão dos recursos hídricos. O conhecimento da dinâmica de movimentação piezométrica de aquíferos em regiões onde se utiliza esse recurso para abastecimento é assunto de interesse, tanto do ponto de vista quantitativo, quanto para subsidiar estudos que tratam de avaliações ambientais e ecológicas.

A predição dos níveis de água subterrânea com bom grau de confiança pode trazer grandes vantagens para a elaboração de cenários e estratégias que auxiliem uma correta gestão da água, uma vez que possibilita realizar estimativas e preenchimento de falhas em dados de séries históricas dos níveis da água do meio subterrâneo. As variações dos níveis de água são, portanto, excepcionalmente importantes para os estudos que buscam estimar os volumes de recarga e/ou descarga hídrica nos aquíferos.

A utilização de enfoques estocásticos para a simulação de dados na hidrogeologia vem ganhando espaço dentro na academia devido a baixa necessidade de variáveis dependentes do fenômeno e ao desenvolvimento das linguagens de programação computacional. Dessa forma, são consideradas ferramentas que fornecem soluções empíricas, ou seja, são baseadas somente nos dados temporais de entrada e cujas variáveis seguem as leis estatísticas ali descritas (Manziona, 2015).

As vantagens da utilização de modelos baseados em séries temporais de longa duração vêm sendo verificadas em uma variedade de campos da hidrologia, principalmente para a previsão de processos hidrológicos, como o escoamento superficial, vazão, precipitação e a modelagem da qualidade da água. Assim, na maioria dos fenômenos hidrológicos é possível correlacionar seu estado em um determinado instante com seu estado no tempo passado, sendo este processo denominado autorregressivo (AR).

Sem embargo, diversos autores têm buscado analisar as flutuações dos níveis de água a fim de fazer previsões dos seus valores baseados em metodologias autorregressivas, ou seja, baseados em

séries históricas de dados, a saber Yoon *et al.* (2011), Shirmohammadi *et al.* (2012), Taormina *et al.* (2012), Shiri *et al.* (2013), Djurovic *et al.* (2015), Gong *et al.* (2016), Luiz *et al.* (2017). Dentre as abordagens utilizadas por estes autores, destacam-se utilizações com bons resultados de métodos Autorregressivos e de Média Móveis (ARMA), Máquinas de Vetores de Suporte (MVS) e Redes Neurais Artificiais (RNA) aplicados em diferentes meios porosos e com distintas escalas de tempo e estruturas autorregressivas.

A aplicação de métodos Autorregressivos e de Médias Móveis em séries temporais caracterizam-se por sua forte base estatística e analisam o sistema ou processo usando uma teoria probabilística. As metodologias que utilizam vetores de suporte baseiam-se em um conjunto de métodos de aprendizado máquinas supervisionadas usadas basicamente para classificação e regressão de dados. A partir de um conjunto utilizado para o treinamento, o algoritmo constrói um modelo que faz a predição de um valor dentro de uma determinada categoria.

As redes neurais são ferramentas que permitem a modelagem de dados estatísticos não lineares. Historicamente, foram concebidas de forma a mimetizar o funcionamento do sistema nervoso humano, sendo usadas para modelar relações complexas e para encontrar padrões estatísticos nos dados em uma distribuição de probabilidades entre variáveis observadas.

Visando difundir a utilização de técnicas autorregressivas em séries temporais de nível de água subterrânea e para contribuir no entendimento dos processos hidrodinâmicos da região das nascentes da Bacia Hidrográfica do Rio Vacacaí-Mirim, este trabalho objetiva avaliar o desempenho de simulações dos níveis de água em um aquífero livre, utilizando diferentes metodologias regressivas em dados de escala horária e diária. Dessa forma, avaliou-se a estrutura de entrada e o modelo que melhor representou as oscilações do lençol freático, a fim de fazer previsões dos seus níveis futuros com boa acurácia.

Com este trabalho, foi possível comparar as potencialidades das técnicas empregadas e aplicar ferramentas eficazes para a simulação do nível do lençol freático em aquífero livre, constituindo um

estudo de apoio para atividades que demandem o planejamento e o uso dos recursos hídricos subterrâneos bem como para subsidiar o preenchimento de possíveis falhas em medições de dados de séries históricas.

## 2 Material e Métodos

### 2.1 Localização e Descrição da Área do Estudo

A área de estudo localiza-se na Depressão Central do Estado do Rio Grande do Sul, ao sul do município de Itaara, na divisa com o município de Santa Maria. Está inserida na região hidrográfica do Guaíba e especificamente na Bacia Hidrográfica do Rio Vacacaí-Mirim. A estação de monitoramento utilizada foi construída próximo às nascentes do rio Vacacaí-Mirim (Figura 1) em uma altitude de cerca de 230 metros acima do nível do mar, sendo composta por um pluviôgrafo de caçamba e um poço tubular com 31 metros de profundidade equipado com um medidor do nível da água. Essa região apresenta grande importância local, uma vez que está localizada à montante da barragem do Departamento Nacional de Obras de Saneamento (DNOS) que é atualmente operada pela Companhia Rio Grandense de Saneamento (CORSAN). Essa barragem é responsável por parte do abastecimento público de água tratada do município de Santa Maria.

As águas subterrâneas da área de estudo foram inicialmente estudadas por Farias (2011), Silvério da Silva & Farias (2013) e Carmo (2014), os quais monitoraram parâmetros de nível de água, temperatura e condutividade elétrica em séries de longa duração. Os autores aplicaram metodologias baseadas nas variações diárias dos níveis de água para o estabelecimento de volumes anuais da recarga aquífera, chegando a valores que variaram entre 736 a 833 milímetros anuais.

Geomorfologicamente, a Bacia Hidrográfica do Rio Vacacaí-Mirim situa-se majoritariamente às bordas da Bacia Sedimentar do Paraná, caracterizada pela ocorrência de rochas vulcânicas da Formação Serra Geral na sua porção Noroeste que é mais elevada. Contudo, a maior parte da bacia abrange litologias sedimentares que formam aquíferos livres em suas cotas mais baixas.

A região onde se localiza o poço de monitoramento utilizado apresenta um perfil geológico de reservatório para aquíferos livres, caracterizado como uma localidade importante para recarga direta de águas subterrâneas (Farias, 2011), fazendo parte da Zona de Afloramentos do Sistema Aquífero Guarani (PSAG, 2009).

O perfil do poço apresenta um pacote sedimentar típico das formações do Sistema Aquífero

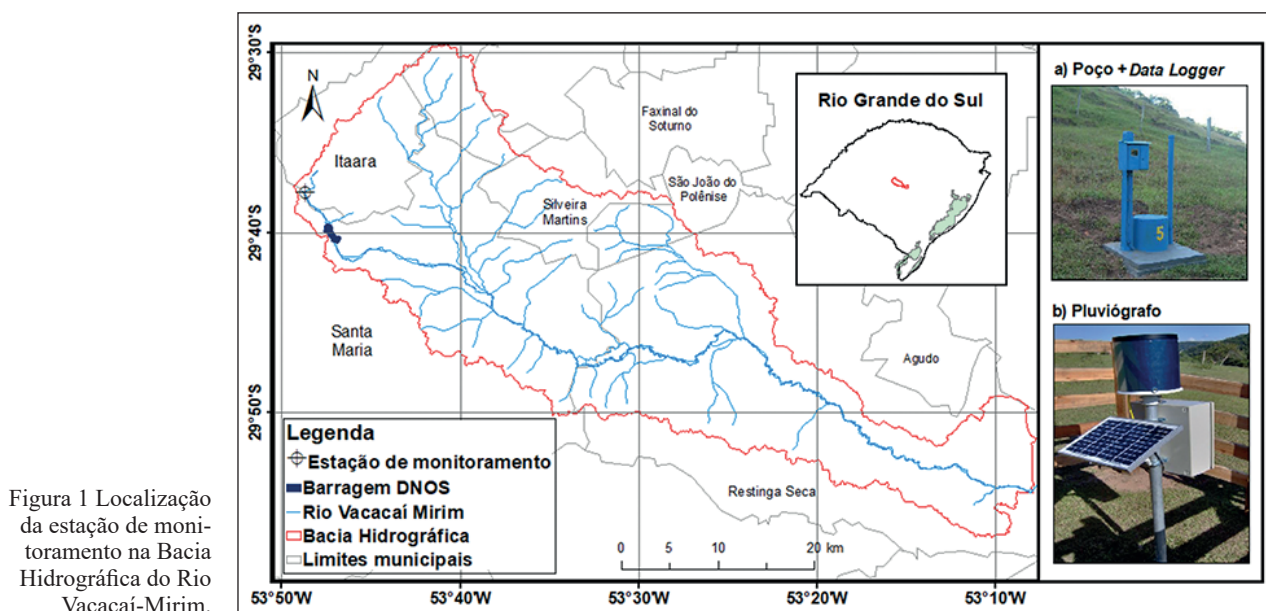


Figura 1 Localização da estação de monitoramento na Bacia Hidrográfica do Rio Vacacaí-Mirim.

Santa Maria, caracterizada por arenitos grossos a conglomeráticos na base, lamitos avermelhados, sil-titos e arenitos finos a médios no topo (Machado & Freitas, 2005).

O poço tubular foi construído obedecendo critérios técnicos da NBR 12244 (ABNT, 2006) com filtro de tubulação ranhurada para facilitar o fluxo de água proveniente do aquífero. O clima da área de estudo segundo a classificação de Köppen é do tipo Cfa, caracterizado como subtropical úmido, com chuvas bem distribuídas durante todos os meses do ano. Sua precipitação anual média está em torno de 1815 milímetros (Alvares *et al.*, 2013).

## 2.2 Coleta de Dados e Estruturação dos Modelos

Para a estruturação de um modelo para realizar as previsões dos níveis estáticos da água subterrânea, este trabalho utilizou, além do valor de nível, dados históricos das precipitações pluviométricas ocorridas no mesmo período, tendo em vista que essa variável influencia diretamente nas oscilações e nos picos de nível em aquíferos livres.

Estes dados foram monitorados por aparelhos automáticos de medição equipados com *data logger*, ajustados para registros horários de informações. O nível de água subterrânea foi monitorado por meio de sondas do tipo transdutoras de pressão, através de um *Levellogger Solinst* modelo 3001 submerso a 10 metros de profundidade, responsável pela medição da pressão da água. A pressão atmosférica foi monitorada por meio de um *Barologger Solinst* modelo LT F5/ML5 localizado ao lado do poço (Figura 1a). A partir da medição das pressões foi possível expressar a pressão absoluta e, posteriormente, determinar o nível estático da água subterrânea em metros. Esses equipamentos foram sincronizados e configurados para realizar leituras horárias, transferindo todos os dados armazenados em sua memória para um computador.

A precipitação pluviométrica foi registrada, de hora em hora, por pluviógrafo do tipo caçamba (Figura 1b), sendo de propriedade da Universidade Federal de Santa Maria (UFSM) e operado por grupos de pesquisa da universidade. Este equipamento

possui dispositivo automático para recarga de bateria por energia solar e está instalado próximo ao poço tubular em estudo. Ao total, durante um ano foram registrados 8784 dados de nível estático e de precipitação compreendidos no período de 07/02/2012 a 06/02/2013.

Além do pareamento horário dos dados de nível estático e da chuva, criou-se um conjunto de dados correspondente ao somatório dos valores de precipitação observados entre 0 e 23 horas enquanto que o nível estático foi observado às 23 horas, caracterizando um conjunto de observações diárias de nível de água subterrânea e precipitação acumuladas de um ano.

## 2.3 Métodos Utilizados

### 2.3.1 Autorregressivo e de Médias Móveis (ARMA)

Os métodos do tipo ARMA ( $p, q$ ) utilizam uma junção de modelos autorregressivos AR ( $p$ ) e de modelos de médias-móveis MA ( $q$ ), tendo notação ARMA ( $p, q$ ). O parâmetro  $p$  representa o número de termos autorregressivos e o parâmetro  $q$ , a ordem de médias móveis. O processo de MA ( $q$ ) refere-se aquele em que há forte presença de ruído branco, onde a correlação entre os dados sofre uma queda brusca devido a este efeito aleatório (Manziona, 2015).

Assim, neste modelo a relação entre as variáveis independentes é aditiva e linear em relação à variável dependente. Foram utilizados métodos baseados em algoritmos de otimização para a estimativa da relação. A metodologia padrão de análise e modelagem de séries temporais é descrita por Box *et al.* (2015) e está descrita na equação 1:

$$X_t = c + \varepsilon_t + \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (1)$$

Onde  $X_t$  é o valor linearmente estimado da variável dependente ( $X$ ) em um dado tempo ( $t$ ), com uma constante ( $c$ ) adicionada dos parâmetros autoregressivos ( $\varphi$ ) e de médias móveis ( $\theta$ ) e de um vetor de erros ( $\varepsilon$ ). Neste estudo, o valor dos parâmetros foi estimado através do método da máxima verossimilhança por meio do *software* R (R Core Team, 2016).

### 2.3.2 Máquina de Vetor de Suporte (MVS)

A metodologia de regressão que utiliza Máquina de Vetor de Suporte ou *Support Vector Machine* (SVM), introduzido por Vapnik (2000), é uma técnica de aprendizado de máquina supervisionada que possui a capacidade de “aprender” a partir de um conjunto de dados de treinamento e, com isso, através da construção de hiperplanos em um dado espaço altamente dimensional designar classes a um novo conjunto de dados. Uma das extensões deste método é chamado de Regressão de Vetor de Suporte. Nesta extensão, estimativas lineares, dada a variável  $Y$  e um conjunto de preditores  $X$ , podem ser obtidas pela minimização da seguinte função (Hastie *et al.*, 2009):

$$H(\beta, \beta_0) = \sum_{i=0}^N V(y_i - f(x_i)) + \frac{\lambda}{2} \|\beta\|^2 \quad (2)$$

Onde:

$$V_{\epsilon}(r) = \begin{cases} 0, & \text{se } |r| < \epsilon \\ |r| - \epsilon, & \text{se não} \end{cases}$$

Nesta fórmula é minimizado o somatório de uma dada função de perda ( $V$ , erros absolutos simples neste estudo) onde os erros ( $\epsilon$ ) menores que um dado valor (0,1 neste estudo) são igualados a zero e somente computados se maiores que este valor. Esta função é penalizada pela multiplicação da metade de uma constante ( $\lambda = 1$  neste estudo) pela norma do vetor dos coeficientes da função linear, ignorando a constante.

A estimativa dos valores ( $H$ ) é feita através da utilização do método de *kernel*, em que as variáveis independentes são implicitamente mapeadas em um espaço vetorial de alta dimensão. Para tanto, este método faz uso de funções de *kernel*, sendo que neste estudo foi utilizada a função específica denominada Função de Base Radial Gaussiana. Como forma de pré-processamento, todos os dados são escalados e padronizados para este método de estimativa. O ajuste aos dados das séries temporais foi feito através do pacote LIBSVM (Chang & Lin, 2011) disponível para a linguagem R de programação (R Core Team, 2016).

### 2.3.3 Redes Neurais Artificiais (RNA)

A RNA é uma estrutura matemática flexível com um grande número de parâmetros, capaz de aproximar qualquer tendência não linear. Estatisticamente, as redes neurais artificiais, em sua forma mais simples, podem ser representadas pela equação 3 (Venables & Ripley, 2013):

$$y_k = \phi_o(\alpha_k + \sum_h w_{hk} \phi_h(\alpha_h + \sum_i w_{ih} x_i)) \quad (3)$$

Esta fórmula representa as redes neurais com camada única, onde os preditores ( $x_i$ ) são ponderados por pesos ( $w_{ih}$ ) distribuídos entre neurônios (ou unidades escondidas) e somados a uma constante de viés (bias,  $\alpha_h$ ) comum a todos neurônios da camada. Estes valores são transformados por uma função ( $\phi_h$ ) não linear (neste caso a função logística), denominada função de ativação. Então uma combinação linear do resultado destas operações é novamente ponderada ( $w_{hk}$ ), adicionadas a uma constante ( $\alpha_k$ ) e, transformadas à escala original ( $\phi_o$ ).

Graficamente uma rede neural pode ser representada pelas relações entre as variáveis independentes e os neurônios da camada escondida e, pela relação destas unidades com a variável dependente (Figura 2). A influência das constantes em cada unidade (viés) é geralmente omitida nas representações gráficas de redes neurais. No caso de um modelo com seis variáveis independentes (uma estrutura autorregressiva de sexta ordem, por exemplo), uma rede neural artificial pode ser representada da seguinte forma:

Dado uma medida de erro ou função de perda (soma dos quadrados, neste estudo) e um controle de alteração dos pesos (*decay*, 0,01 neste estudo), este procedimento é repetido iterativamente até que um ponto de mínima seja encontrado. O grau de ajuste de uma rede neural aqui descrita (chamada de *feed-forward*) depende do tamanho, da camada escondida (número de neurônios) e do controle da alteração dos pesos. O número de neurônios escolhido neste estudo foi igual a cinco, resultando, no caso de seis variáveis independentes, na arquitetura neural descrita na Figura 2. O ajuste dos dados das séries temporais foi realizado utilizando-se o pacote do R *nnet* (Ripley & Venables, 2016).

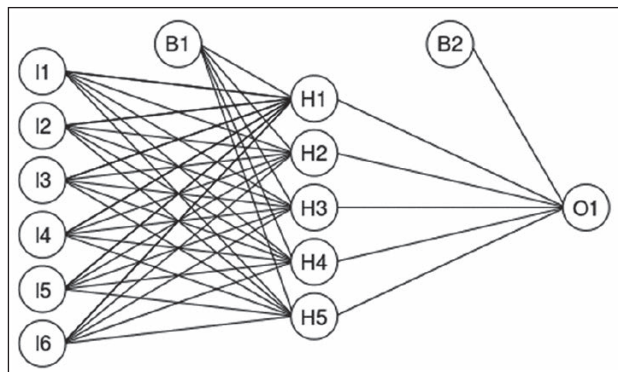


Figura 2 Estruturação do modelo RNA (*feed-forward*) para regressão com seis variáveis independentes.

## 2.4 Características das Séries Temporais

As séries temporais utilizadas para a modelagem foram sistematizadas no gráfico da Figura 3. A série de nível estático observada no poço teve um registro mínimo de 13,03 metros e um máximo de 8,87 metros, apresentando uma amplitude de cerca de 4,16 metros na coluna de água. Já a série de precipitação, apresentou um registro diário máximo de 55 milímetros e um acumulado anual de 2201 milímetros, demonstrando um valor acima da média anual registrada na região segundo Alvares *et al.* (2013). Visualmente é possível notar uma relação entre os volumes altos de precipitação com as oscilações mais acentuadas no nível estático do poço sob estudo.

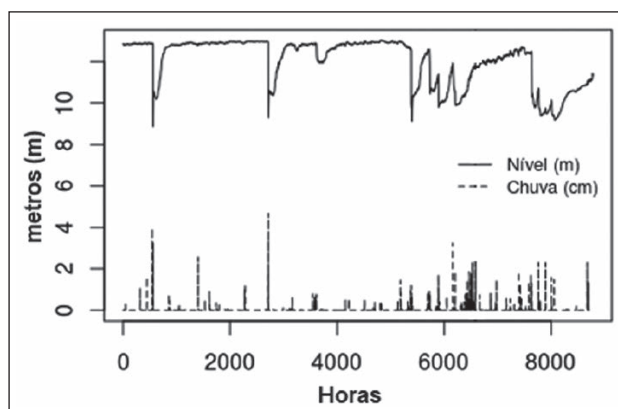


Figura 3 Séries temporais horárias de nível de água subterrânea (m) e de precipitação (cm).

A partir das séries históricas apresentadas, a fim de definir a melhor estruturação dos dados para fazer as simulações, realizou-se um diagnóstico para

definir a autocorrelação estatística da série de nível, aplicando-se para tanto, Funções de Autocorrelação (FAC) e Autocorrelação Parcial (FACP) em dados horários e diários do nível da água e de precipitação.

Segundo Manzione (2015), a FAC mede a correlação entre as observações que estão  $k$  períodos afastados, entretanto remove o efeito das correlações intermediárias. Dessa forma, a FAC avalia a correlação simples entre  $Y_t$  e  $Y_{t-k}$  em função da defasagem  $k$ , ou seja, mostra o quanto o processo é correlacionado com ele próprio em dois instantes de tempo diferentes, mostrando uma medida de dependência temporal entre os dados.

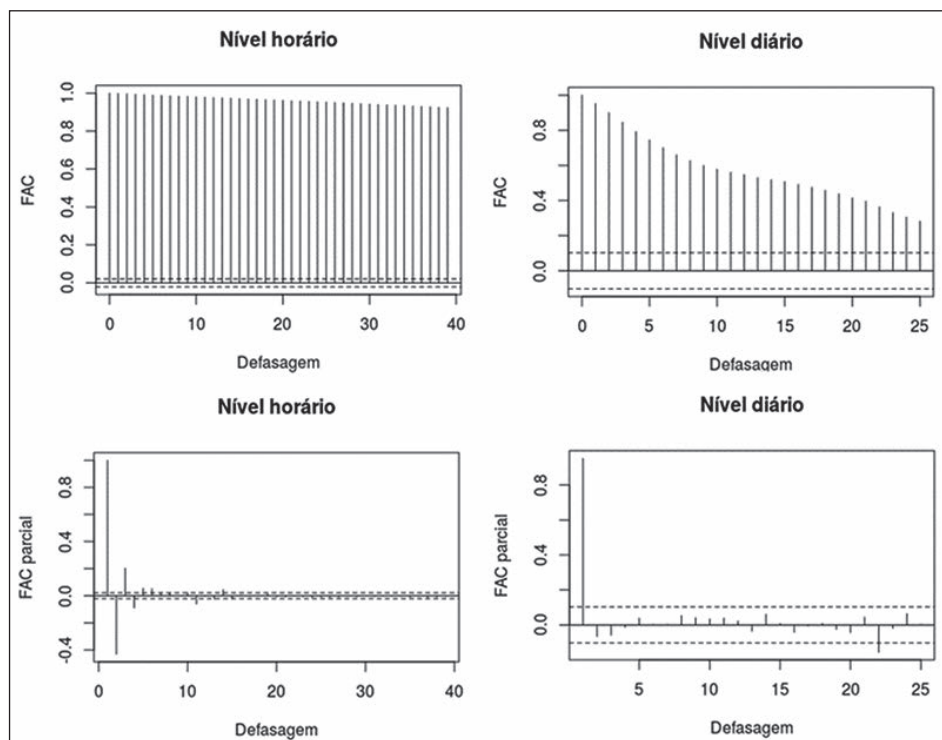
A aplicação da FACP mede o quanto há de correlação aleatória em função da autocorrelação entre os instantes  $t$  de uma série, confirmando ou refutando as premissas assumidas na análise da FAC (Manzione, 2015). Sendo assim, a intensidade da relação entre duas observações de uma mesma série é medida, ou seja, estima-se a autocorrelação entre diferentes conjuntos de observações (defasagens), controlando o efeito das demais e indicando o processo estocástico que pode ser assumido.

As defasagens utilizadas para a modelagem do processo de variação do nível da água subterrânea foram selecionadas de acordo com a função FAC e FACP (Figura 4) dos dados do nível. As funções de autocorrelação foram aplicadas para escalas horárias e diárias.

Dessa forma, a FAC fornece indicações sobre o quanto uma variável aleatória é capaz de influenciar seus vizinhos, ficando evidente, por meio dos gráficos da Figura 4, a existência de correlações significativas entre as observações, tanto horárias quanto diárias. Assim, a aplicação da FAC na série de dados sugere autocorrelações significativas para a série histórica estudada.

O gráfico de FACP aplicado a série histórica de níveis em escala horária indicou que uma relação de autorregressão com uma defasagem de seis períodos influenciam significativamente os valores subsequentes do lençol freático, ou seja, as 6 horas anteriores de nível da água possuem relação estatisticamente significativa com o período que se deseja

Figura 4 Funções de Auto-correlação (FAC) e Autocorrelação Parcial (FACP) para séries temporais horárias e diárias do nível da água em um poço tubular em aquífero livre. As linhas sólidas verticais indicam o grau de correlação entre as diferentes defasagens das séries temporais. As linhas tracejadas horizontais indicam um intervalo de confiança de 95% em relação à significância das correlações.



prever. Para os níveis diários, o gráfico de FACP indicou uma relação autorregressiva com uma defasagem de um período (autorregressivo de primeira ordem).

Apesar dos gráficos de FAC e FACP apresentarem boas aproximações das melhores estruturas de entrada a serem utilizadas na etapa de calibração dos modelos, seus resultados demonstram apenas aproximações de suas respectivas correlações entre seus vizinhos. Assim, testaram-se diferentes formas de estrutura até a sexta ordem de autocorrelação, as quais foram comparadas com base em estatísticas de ajuste.

## 2.5 Comparação dos Métodos Utilizados

Como forma de avaliar os diferentes métodos estatísticos utilizados, analisaram-se os erros a partir de uma escala aritmética e outra quadrática. O estimador de erros em escala aritmética utilizado foi o *bias*, que corresponde aos desvios absolutos médios entre as previsões e os valores observados. O estimador de erros em escala quadrática utilizado foi o Coeficiente de Nash-Sutcliffe (CNS) que correspon-

de a unidade menos a razão entre o somatório das diferenças ao quadrado das previsões e os valores observados, representando, assim, a eficiência do modelo em uma faixa que vai de  $-\infty$  (menor eficiência) a 1 (máxima eficiência). As modelagens bem como os cálculos para a avaliação das previsões dos métodos foram realizadas utilizando a linguagem R de programação (R Core Team, 2016).

Baseado nesta análise, testaram-se modelos ARMA até a sexta ordem de autocorrelação e primeira ordem de média móvel. Para as metodologias MVS e RNA, testaram-se a sua performance até a sexta ordem de autocorrelação. A inclusão da variável de precipitação dos seis períodos anteriores ao de previsão foi feita como adição aos modelos, objetivando uma representação com base física do fenômeno e a obtenção de melhorias nas previsões.

Os resultados obtidos na presente pesquisa são apresentados em 3 etapas. Na primeira etapa, realizou-se o ajuste (ou treinamento) dos modelos, que é a etapa responsável por ajustar as metodologias por meio da estatística intrínseca extraída da série histórica, avaliando seus ajustes através de escala aritmética e de escala quadrática. Para a segunda etapa,

escolheram-se os melhores modelos para a construção dos gráficos de previsões, comparando com os valores reais registrados em campo (etapa de teste). Na terceira e última etapa, realizou-se uma análise gráfica residual (erro previsto versus o registro observado) dividido pelo desvio padrão do modelo, buscando analisar os erros e sua magnitude bem como as tendências de imprecisão e acuracidade de cada metodologia a fim de validar seu uso.

### 3 Resultados e Discussão

#### 3.1 Avaliação de Ajuste dos Modelos

Inicialmente, para a etapa de ajuste, utilizou-se 75% da série histórica, ou seja, 6588 horas ou 274 dias ininterruptos para treinar os modelos. Assim, foram realizadas simulações empregando as metodologias aqui descritas até a sexta ordem de autocorrelação, usando os estimadores bias e CNS para avaliar suas performances.

Para a modelagem pelo método ARMA, testaram-se o poder preditivo de estruturas autorregressivas até a sexta ordem e de médias móveis até a primeira ordem. Dessa forma, são apresentados os valores dos estimadores calculados na Tabela 1, em escala horária e diária.

	Estrutura	Escala Horária		Escala Diária	
		bias	CNS	bias	CNS
ARMA (p, q)	ARMA (1, 0)	1,183	-1,063	1,314	-1,093
	ARMA (2, 0)	1,193	-1,068	1,391	-1,220
	ARMA (3, 0)	1,181	-1,073	1,356	-1,154
	ARMA (4, 0)	1,185	-1,076	1,348	-1,138
	ARMA (5, 0)	1,182	-1,075	1,349	-1,142
	ARMA (6, 0)	1,181	-1,069	1,350	-1,144
	ARMA (1, 1)	1,048	-0,285	2,460	-4,805
	ARMA (2, 1)	1,049	-0,292	2,473	-4,851
	ARMA (3, 1)	1,049	-0,292	2,624	-5,465
	ARMA (4, 1)	1,049	-0,260	2,669	-5,656
	ARMA (5, 1)	1,049	-0,204	2,502	-4,960
	ARMA (6, 1)	1,049	-0,154	1,036	-5,761

Tabela 1 Valores dos estimadores bias e CNS em escala horária e diária para a calibração do modelo ARMA de uma série temporal do nível estático e da precipitação em aquífero livre.

De modo semelhante ao observado nos gráficos de FACP, baseado nos valores dos estimadores bias e CNS apresentados na Tabela 1, escolheu-se a estrutura ARMA (6, 1) para explicar a relação dos

dados horários. A estrutura com melhor ajuste para os dados diários utilizados foi alcançada com o modelo ARMA (1, 0).

Para a calibração da metodologia MVS, da mesma forma que na metodologia ARMA, testaram-se estruturas autorregressivas até a sexta ordem. Assim, os valores bias e CNS para esse método são apresentados na Tabela 2.

	Estrutura	Escala Horária		Escala Diária	
		bias	CNS	bias	CNS
Máquina de Vektor de Suporte	AR1	0,120	0,950	0,393	0,816
	AR2	0,109	0,951	0,431	0,784
	AR3	0,112	0,947	0,438	0,776
	AR4	0,120	0,943	0,441	0,771
	AR5	0,126	0,939	0,443	0,766
	AR6	0,128	0,938	0,454	0,754

Tabela 2 Valores dos estimadores bias e CNS em escala horária e diária para calibração pelo modelo MVS de uma série temporal do nível estático e de precipitação em aquífero livre.

A partir da análise dos estimadores de veracidade e de precisão, evidencia-se que para a etapa de ajuste do modelo em MVS, as estruturas autorregressivas de segunda ordem em escala horária e a estrutura de primeira ordem em escala diária, foram as que apresentam menor erro e maior precisão. O alcance dos valores do CNS na escala horária e diária foram considerados como satisfatórios, uma vez que atingiram 0,95 e 0,81 respectivamente, mostrando o bom ajuste obtido na etapa de calibragem. Assim, para a realização das previsões, utilizaram-se essas estruturas nesse modelo.

Para a modelagem por meio de RNA, a exemplo das outras calibrações realizadas, testaram-se estruturas autorregressivas até a sexta ordem, para as variáveis nível estático e precipitação. Os resultados com os estimadores calculados estão apresentados na Tabela 3.

A partir da observação dos estimadores calculados na etapa de calibração da Tabela 3, considerou-se a estrutura autorregressiva de quarta ordem para a simulação dos níveis horários e estrutura autorregressivas de segunda ordem para a simulação da sé-



	Estrutura	Escala Horária		Escala Diária	
		bias	CNS	bias	CNS
Redes Neurais Artificiais	AR1	0,023	0,998	0,365	0,505
	AR2	0,018	0,998	0,303	0,846
	AR3	0,023	0,999	0,347	0,755
	AR4	0,017	0,999	0,399	0,848
	AR5	0,023	0,999	0,299	0,840
	AR6	0,018	0,999	0,587	0,821

Tabela 3 Valores dos estimadores bias e CNS de uma série temporal do nível estático e de precipitação em aquífero livre pelo método RNA.

rie dos níveis em escala diária. Em comparação com os valores obtidos com a metodologia MVS na Tabela 2, a metodologia baseada em RNA apresentou-se melhor em ambas as escalas testadas, atingindo um CNS de 0,99 e 0,84 respectivamente, demonstrando um ótimo ajuste dos dados de séries históricas de nível e de precipitação.

Ao comparar os valores do bias e do CNS nas três metodologias utilizadas, verificou-se que o ARMA não obteve bons ajustes nas estruturas testadas, demonstrando-se pouco eficiente para as escalas avaliadas. Isto não significa que não exista uma relação direta entre a precipitação e o nível da água, apenas que a forma de inclusão das médias móveis não gerou melhores ajustes entre as séries de dados. É possível que a utilização de escalas de tempo di-

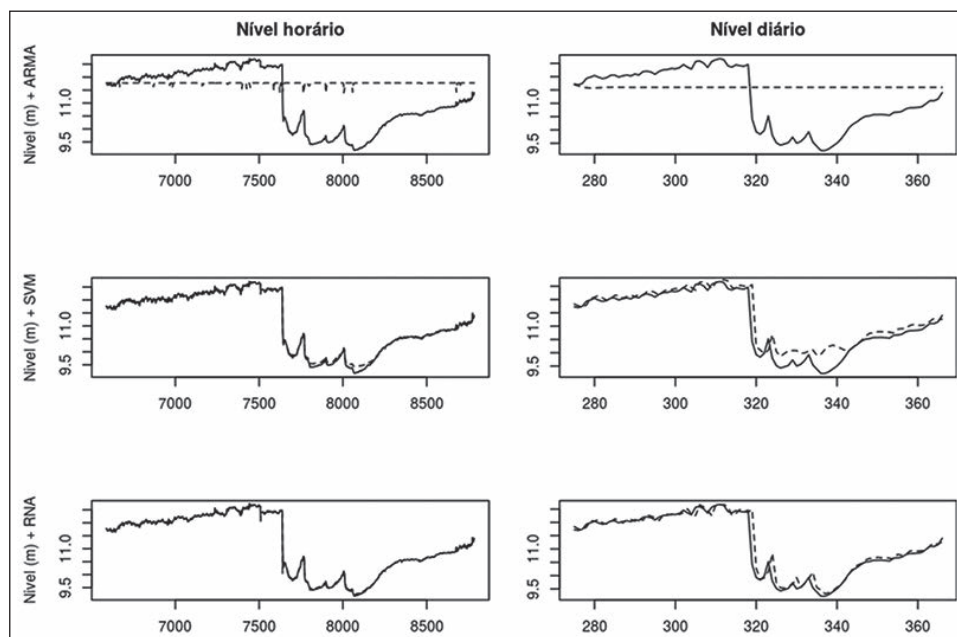
ferentes das testadas possam melhorar o ajuste dos modelos que utilizam essa metodologia.

### 3.2 Avaliação Gráfica das Previsões de Nível

Para a etapa de previsão dos níveis estáticos da água subterrânea, utilizaram-se as melhores estruturas autorregressivas segundo os valores dos estimadores, apresentados na etapa de ajuste. Assim, em escala horária, realizaram-se previsões dos níveis estáticos para 2196 horas e, para a escala diária, em 92 dias após a etapa de treinamento dos modelos. Ambas as escalas de tempo foram comparadas com o valor real medido do nível estático, obtidas por sonda nos trabalhos de campo. Os resultados das previsões para cada modelo estão dispostos em forma de gráfico, na Figura 5.

Em geral, na Figura 5 verificam-se que os modelos apresentaram bons ajustes, inferidos visualmente pela boa aderência entre as séries simuladas e as séries reais registradas em campo. A maior exceção foi a previsão realizada pelo modelo ARMA, que apresentou inconstâncias bastante significativas entre o simulado e o observado. Essa inexatidão pode estar associada ao fato de que essa metodologia é mais eficiente para descrever séries temporais estacionárias, uma vez que seu desenvolvimento se dá ao redor de uma média.

Figura 5 Observações (linhas sólidas) e previsões (linhas tracejadas) do nível estático horário e diário de um poço tubular em aquífero livre. As previsões são com base nos melhores modelos obtidos para cada método de estimativa.



Na escala horária, as previsões dos níveis estáticos usando os modelos com MVS e RNA apresentaram um excelente resultado, evidenciado pela aderência entre o observado e o simulado. Dessa forma, verifica-se que a utilização dessas metodologias nessa escala de avaliação demonstra-se extremamente eficaz e com alto poder preditivo agregado.

Na escala de modelagem diária, tanto na etapa de treinamento quanto na etapa de previsão, foi possível observar que a simulação dos níveis estáticos utilizando RNA obteve o melhor ajuste, dentre as metodologias testadas, quando comparadas com a série de dados obtidas em campo.

Dessa forma, observando os gráficos da Figura 5, é visível a superioridade dos modelos calibrados por MVS e RNA frente ao modelo calibrado por ARMA para a previsão do nível da água subterrânea no poço estudado e nas escalas testadas. A partir de um conjunto de dados composto pela série histórica de nível estático e pela série histórica da precipitação pluviométrica, foi possível estabelecer um modelo capaz de reproduzir as oscilações de níveis deste poço de monitoramento.

Para validar os modelos empregados, verificou-se os erros cometidos durante o processo de previsão através dos resíduos, que permitiram visualizar a magnitude de imprecisão.

### 3.3 Análise Gráfica Residual

Para a etapa de avaliação de erros, realizou-se a análise dos resíduos normalizados, objetivando validar os modelos utilizados para a previsão de níveis estáticos de água subterrânea. Assim, tendo em vista a necessidade de avaliar a variável resposta em relação a variável medida, construiu-se os gráficos de resíduos normalizados para os modelos, ilustrados nos gráficos da Figura 6.

Ao observar os gráficos de resíduos normalizados de cada previsão realizada, é possível notar os desvios gerados em cada previsão. Os erros cometidos pelo modelo ARMA foram considerados bem expressivos, uma vez que a plotagem entre a variável simulada e a observada se deram bastante fora da linha de resíduo nulo, constatando-se que essa metodologia foi a que apresentou os piores resultados neste estudo.

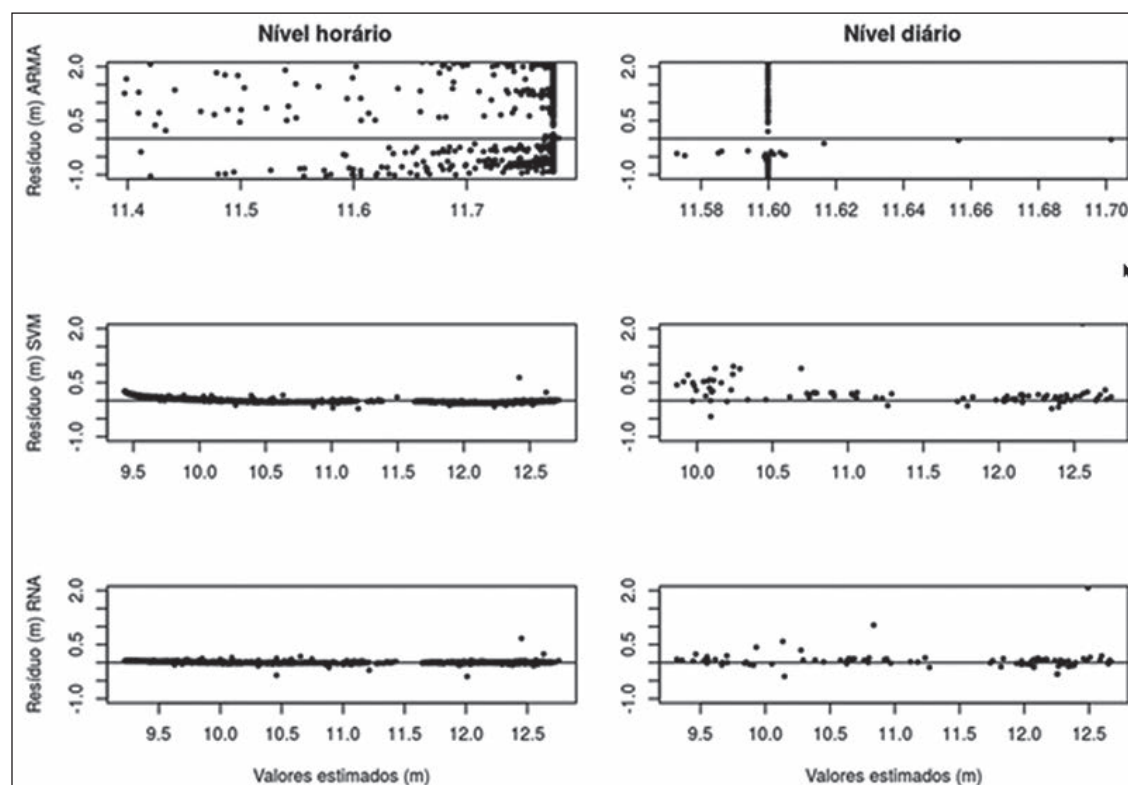


Figura 6 Erros das previsões do nível estático horário e diário de um poço tubular em aquífero livre. As previsões são com base nas melhores estruturas de entrada obtidas para cada método de estimativa.

Para os gráficos de resíduos em escala horária, os modelos de MVS e de RNA apresentaram bons ajustes a série de dados. No entanto, a metodologia MVS demonstrou inconsistências para valores de níveis estáticos entre 9,4 a 9,8 metros, com tendências a superestimar os valores preditos. Essa inconsistência está vinculada ao fato de que na etapa de treinamento, esses valores são pouco recorrentes no universo total de dados coletados. O modelo de RNA não apresentou essas inconsistências, evidenciado pelo ótimo ajuste à linha de resíduo zero dos valores preditos. Contudo, ressalta-se que o modelo de RNA apresentou um *outlier* conspícuo em comparação com o modelo MVS. No entanto, esse desvio representa uma minoria entre o total de valores previstos, identificando, assim, a RNA como a metodologia que apresentou a melhor performance para realizar-se a simulação da série de níveis estáticos em escala horária para os dados avaliados.

Para a modelagem em escala diária, percebe-se que o modelo de RNA também foi superior aos demais modelos testados. Assim como na escala horária, o modelo MVS em escala diária apresentou inconsistências de previsão para valores entre 9,8 e 10,3 metros de nível estático. Assim, ao analisar-se a gama total de resultados alcançados nas etapas deste estudo, o modelo composto por RNA demonstrou-se superior aos demais modelos testados para a simulação dos níveis freáticos de água subterrânea do poço de monitoramento.

#### 4 Conclusões

A partir dos resultados apresentados na etapa de ajuste e de previsão, apoiado nos valores dos estimadores utilizados, bem como na etapa de análise de erros, conclui-se que a metodologia que utiliza Redes Neurais Artificiais foi a que obteve melhor performance. Para a simulação da série de dados utilizados nessa pesquisa, a associação de estruturas de regressão na quarta ordem para dados horários e estrutura de segunda ordem para os dados diários demonstraram os melhores desempenhos.

Os resultados apresentados nas escalas propostas indicaram que o modelo que utiliza autorregressão e médias móveis não obteve bons ajustes e

previsões nas estruturas testadas. A metodologia de regressão MVS apresentou bons resultados, embora a metodologia de RNA tenha sido superior.

Dessa forma, a comparação de metodologias utilizadas para a simulação dos níveis a partir da calibração de modelos por séries históricas de nível estático e de precipitação pluviométrica, permitiu concluir que a utilização dos modelos MVS e RNA para os estudos de séries temporais para a hidrogeologia são de extrema eficiência e tendem a proporcionar bons resultados para estimativas dos níveis dos lençóis freáticos em aquíferos livres.

#### 5 Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. Os autores agradecem à UFSM pela estrutura fornecida e a equipe de pesquisa do Laboratório de Hidrogeologia (LABHIDROGEO) da UFSM.

#### 6 Referências

- ABNT – Associação Brasileira de Normas Técnicas. 2006. *NBR 12244 - Poço tubular Construção de poço tubular para captação de água subterrânea*. Rio de Janeiro, ABNT, 10 p.
- Alvares, C.A.; Stape, J.L.; Sentelhas, P.C.; De Moraes, G.; Leonardo, J. & Sparovek, G. 2013. Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, 22(6): 711-728.
- Box, G.E.; Jenkins, G.M.; Reinsel, G.C. & Ljung, G.M. 2015. *Time series analysis: forecasting and control*. New Jersey, John Wiley & Sons. 714p.
- Carmo, J.A.R. 2014. *Recarga subterrânea direta e sua influência sobre a temperatura e a condutividade elétrica em microbacias do rio Vacacaí-Mirim, RS*. Programa de Pós-Graduação em Engenharia Ambiental, Universidade Federal de Santa Maria, Dissertação de Mestrado, 132p.
- Chang, C.C. & Lin, C. 2011. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3): 1-27.
- Djurovic, N.; Domazet, M.; Stricevic, R.; Pocuca, V.; Spalevic, V.; Pivic, R.; Gregoric, E. & Domazet, U. 2015. Comparison of groundwater level models based on artificial neural networks and ANFIS. *The Scientific World Journal*, 15: 1-13.
- Farias, P.V.C. 2011. *Flutuação dos níveis de água subterrânea em microbacias hidrográficas do rio Vacacaí-Mirim*. Programa de Pós-Graduação em Engenharia Civil, Universidade Federal de Santa Maria, Dissertação

- de Mestrado, 102p.
- Gong, Y.; Zhang, Y.; Lan, S.; & Wang, H. 2016. A comparative study of artificial neural networks, support vector machines and adaptive neuro fuzzy inference system for forecasting groundwater levels near Lake Okeechobee, Florida. *Water resources management*, 30(1): 375-391.
- Hastie, T.; Tibshirani, R. & Friedman, J. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, Springer-Verlag. 745 p.
- Luiz, T.B.P.; Schröder, T. & Silvério-da-Silva. 2017. Previsão de níveis estáticos de água subterrânea em poço de monitoramento utilizando metodologia estocástica. *Anuário do Instituto de Geociências*, 40(2): 53-60.
- Machado, J.L.F. & Freitas, M.A. 2005. *Projeto Mapa Hidrogeológico do Rio Grande do Sul: relatório final*. Porto Alegre, CPRM. 71p.
- Manzione, R.L. 2015. Águas Subterrâneas: Conceitos e aplicações sob uma visão multidisciplinar. São Paulo, Paco Editorial. 388p.
- PSAG. 2009. *Projeto Sistema Aquífero Guarani. Projeto de proteção ambiental e desenvolvimento sustentável do Sistema Aquífero Guarani*. Disponível em: <http://www.mma.gov.br/publicacoes/agua/category/42-recursos-hidricos?limitstart=0>. Acesso em: 10 de maio de 2018.
- Silvério-da-Silva, J.L. & Farias, P.V.C. 2013. Estimativa de recarga de água subterrânea em microbacia hidrográfica do rio Vacacaí-Mirim. In: SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS, 20, Bento Gonçalves, 2013. *Anais*, Bento Gonçalves, ABRH, p. 1-8.
- Shiri, J.; Kisi, O.; Yoon, H.; Lee, K.K. & Nazemi, A.H. 2013. Predicting groundwater level fluctuations with meteorological effect implications. A comparative study among soft computing techniques. *Computers e Geosciences*, 56: 32-44.
- Shirmohammadi, B.; Vafakhah, M.; Moosavi, V. & Moghadamnia, A. 2012. Application of several data-driven techniques for predicting groundwater level. *Water Resources Management*, 27(2): 419-432.
- R Core Team. 2016. *R: A language and environment for statistical computing*. *R Foundation for Statistical Computing*, Vienna, Austria. Disponível em: <https://www.r-project.org>. Acesso em: 20 de outubro de 2016.
- Ripley, B. & Venables, W. 2016. *Package 'nnet'. R package version*. Disponível em: <ftp://ftp.usg.iu.edu/pub/CRAN/web/packages/nnet/nnet.pdf>. Acesso em: 10 de dezembro de 2016.
- Taormina, R.; Chau, K. & Sethi, R. 2012. Artificial neural network simulation of hourly groundwater levels in a coastal aquifer system of the Venice lagoon. *Engineering Applications of Artificial Intelligence*, 25(8): 1670-1676.
- Venables, W.N. & Ripley, B.D. 2013. *Modern applied statistics with S-PLUS*. New York, Springer-Verlag. 487p.
- Vapnik, V.N. 2000. *The nature of statistical learning theory*. Nova York, Springer-Verlag. 314 p.
- Yoon, H.; Jun, S. C.; Hyun, Y.; Bae, G.O. & Lee, K.K. A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of Hydrology*, 396(1): 128-138.