

## Artigos

# Aplicação de Redes Neurais Artificiais Para Simulação de Fluxo e Transporte de Contaminantes Sob Incerteza de Condutividade Hidráulica

## Artificial Neural Network Application for Groundwater Flow and Contaminant Transport Under Uncertainty of Hydraulic Conductivity

Caetano Pontes Costanzo<sup>1</sup>; Alexandre Campane Vidal<sup>1</sup>; Michelle Chaves Kuroda<sup>1</sup>; Simony Yumi Sakamoto<sup>2</sup>✉

<sup>1</sup> Universidade Estadual de Campinas (Unicamp), Campinas, SP

<sup>2</sup> Walm Engenharia e Tecnologia Ambiental Ltda, Perdizes, São Paulo

✉ [caetano.costanzo@gmail.com](mailto:caetano.costanzo@gmail.com), [acvidal@gmail.com](mailto:acvidal@gmail.com), [mckuroda@gmail.com](mailto:mckuroda@gmail.com), [simony.yumi@gmail.com](mailto:simony.yumi@gmail.com)

### Palavras-chave:

Modelagem matemática.  
Águas Subterrâneas.  
Variabilidade espacial.  
Gerenciamento de áreas contaminadas.

### Keywords:

Mathematical modeling.  
Groundwater.  
Spatial variability.  
Environmental management of contaminated areas.

Revisado por pares.

Recebido em: 05/06/2019.

Aprovado em: 30/07/2019.

### Resumo

A incerteza relacionada à variabilidade espacial da condutividade hidráulica ( $K$ ) é um importante aspecto a ser considerado nas simulações da migração de plumas de contaminantes e, consequentemente no gerenciamento ambiental de áreas contaminadas. A utilização de modelos estocásticos de  $K$  em conjunto com a simulação de fluxo e o transporte de contaminantes é frequente em pesquisas para a previsão do comportamento da pluma, bem como para projetos de remediação sob a incerteza deste parâmetro hidrogeológico. No entanto, as técnicas tradicionais de modelagem de fluxo e de contaminantes, como o MODFLOW e o MT3D, baseadas no método de diferenças finitas, normalmente são lentas e exigem bastante esforço computacional. Dessa maneira, esta pesquisa aplicou redes neurais artificiais (RNA) em 100 cenários de condutividade hidráulica, realizados por métodos estocásticos, a fim de avaliar a incerteza de  $K$  na migração de uma pluma. Além disso, foi possível validar a utilização de RNA como ferramenta de decisão para a estimativa da concentração média de contaminantes nas águas subterrâneas ao longo de três anos de simulação. Com base nos resultados obtidos, foi possível avaliar possíveis técnicas de remediação para a área de pesquisa em função da migração da pluma sob a incerteza da condutividade hidráulica.

### Abstract

The uncertainty related to the spatial variability of hydraulic conductivity ( $K$ ) is an important aspect to be considered in simulations of contaminant plume migration and, consequently in the environmental management of contaminated areas. The use of stochastic  $K$  models with flow simulation and the transport of contaminants is frequent in researches to predict the behavior of the plume as well as to remediation projects under the uncertainty of this hydrogeological parameter. However, traditional flow modeling and contaminant techniques, such as MODFLOW and MT3D, based on the finite difference method, are usually slow and require considerable computational effort. Thus, this research applied artificial neural networks (ANNs) in 100 scenarios of hydraulic conductivity, performed by stochastic methods, in order to evaluate the uncertainty of  $K$  in the migration of the plume. In addition, it was possible to validate the use of ANNs as a decision tool for estimating the average concentration of contaminants in groundwater over three years of simulation. Based on the results obtained, it was possible to evaluate remediation techniques for the research area due to the migration of the plume under the uncertainty of the hydraulic conductivity.

DOI: <http://dx.doi.org/10.14295/ras.v33i3.29563>

## 1. INTRODUÇÃO

No cadastro de áreas contaminadas do Estado de São Paulo constam 6.110 áreas com problemas de contaminação, sendo que 72% dessas áreas são postos de combustíveis, 19% indústria, 5% comércio, 3% disposição de resíduos e 1% acidentes, agricultura e/ou fonte desconhecida (CETESB, 2018).

A contaminação das águas subterrâneas por meio de vazamentos, derramamentos e/ou despejos de substâncias químicas é considerada como um dos maiores desafios ambientais atualmente, sendo que a incerteza relacionada à variabilidade espacial do parâmetro de condutividade hidráulica ( $K$ ) é importante para o gerenciamento ambiental de áreas contaminadas (AHLFELD *et al.* 1988; JOHNSON e ROGERS, 2000; COSTANZO e VIDAL, 2014).

Ao influenciar o fluxo subterrâneo e o transporte de contaminantes em subsuperfície, a condutividade hidráulica é considerada uma variável relevante para as tomadas de decisões nos processos de investigação, remediação e monitoramento ambiental (BAYER e FINKEL, 2007).

A estimativa precisa deste parâmetro hidrogeológico é uma tarefa praticamente impossível, uma vez que ele possui alta variabilidade, sendo as maiores incertezas relacionadas à sua distribuição espacial e o impacto na migração das plumas de contaminação (RANJITHAN *et al.* 1993).

Dessa forma, o mapeamento dos valores de K, como o estabelecimento das direções de conectividade e continuidade em um aquífero contaminado, é importante para a simulação do fluxo e transporte de contaminantes (JOURNEL e ALABERT, 1988). Até mesmo quando a caracterização dessa variável espacial é adequada, obtida por meio de dados de campo, a dificuldade em representar a heterogeneidade espacial permanece, principalmente em função da incerteza associada às áreas não amostradas (RICCIARDI *et al.* 2007).

Com isso, a avaliação da incerteza de K em projetos ambientais pode ser realizada a partir de métodos de multicenários, como a abordagem pelo método Monte Carlo (estocástica). Nessa técnica, múltiplas realizações de campos de condutividade são geradas de forma randômica e o simulador de fluxo é aplicado em cada campo de K, sendo os resultados avaliados de forma conjunta no impacto de cada cenário da simulação de fluxo (LIU *et al.* 2004).

Diversos autores já adotaram a abordagem estocástica na tentativa de simular campos de condutividade hidráulica a fim de subsidiar o gerenciamento de aquíferos contaminados, como Costanzo e Vidal (2014); Alberto e Chang (2011); Flach *et al.* (2005); Demirci e Aksoy (2011); Lee *et al.* (2000); Liu *et al.* (2004). Esta abordagem representa realizações de campos de K para simular os possíveis cenários de condutividade e assim avaliar a incerteza deste parâmetro no deslocamento da pluma (RANJITHAN *et al.* 1993).

No entanto, nas últimas décadas, como forma de otimizar a aplicação das técnicas estocásticas, o método de redes neurais artificiais (RNA) vem sendo utilizado. A rede neural é um método que possui capacidade de aprendizagem relacionada à extração de regras e padrões do treinamento de grandes conjuntos de dados, com funcionamento inspirado em seres vivos (VINYALS e LE, 2015; HAYKIN, 1999; POULTON, 2001).

Ranjithan *et al.* (1993) utilizaram RNA para avaliar o impacto da variação espacial de K em um projeto de remediação da pluma por bombeamento e tratamento. Já Rogers e Dowla (1994) aplicaram redes neurais para prever os resultados do fluxo e transporte de contaminantes, bem como para encontrar um projeto de remediação eficiente em relação a custo e tempo.

Os autores Johnson e Rogers (2000) identificaram localizações ideais de poços de bombeamento e de injeção por meio da técnica de redes neurais artificiais. Outras pesquisas também utilizaram RNA em aquíferos contaminados, por exemplo, Maskey *et al.* (2000); Hassan *et al.* (2001); Rao *et al.* (2003); Coppola *et al.* (2003); Samani *et al.* (2007) e; Nikolos *et al.* (2007).

As técnicas de redes neurais artificiais estão em contínua evolução e são aplicadas em diversas áreas de pesquisas. A adoção desta técnica para a área de hidrogeologia permitirá ampliar a velocidade da análise de dados hidrogeológicos, os quais influenciam a tomada de decisão no gerenciamento de aquíferos contaminados.

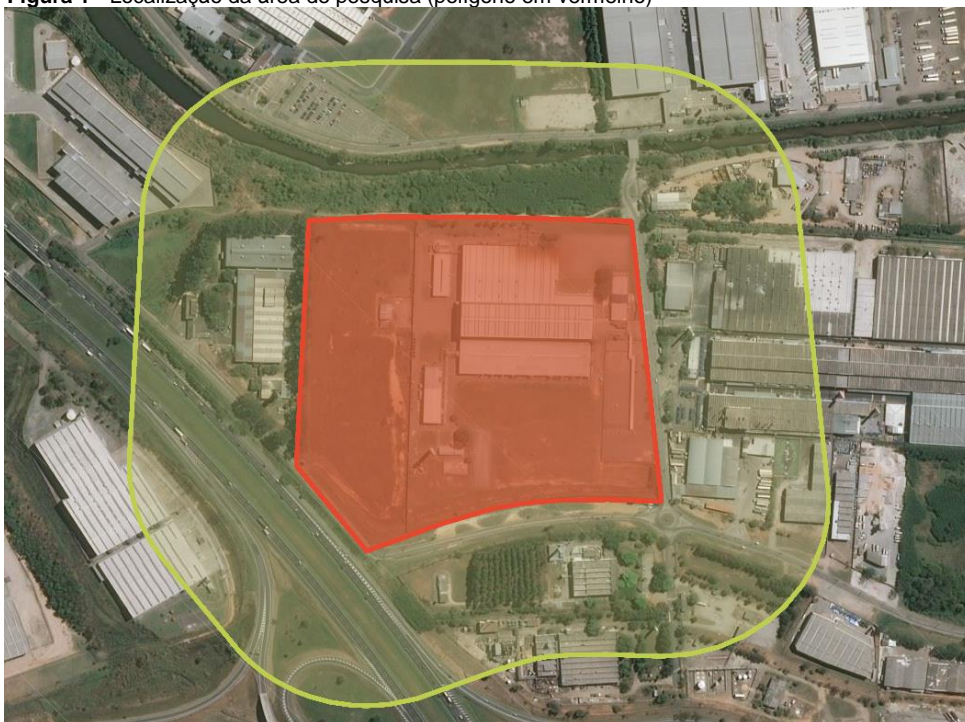
Portanto, esta pesquisa possui como objetivo principal a aplicação de redes neurais artificiais para a previsão do deslocamento de uma pluma de contaminação frente à incerteza relacionada à condutividade hidráulica (K).

Por meio da utilização de redes neurais espera-se uma maior velocidade na geração dos resultados a fim de subsidiar a tomada de decisão, bem como aprimorar o modelo conceitual hidrogeológico de uma área com problema de contaminação.

## 2. ÁREA DE PESQUISA

A área de estudo é a fábrica da Itaotec de Jundiaí/SP, a qual está exposta na página 420 do cadastro Áreas em Processo de Remediação no Estado de São Paulo da CETESB (2018). A área apresenta como principal foco de contaminação a presença de metais em águas subterrâneas, havendo indício também em subsolo, provenientes da antiga operação fabril, armazenamento e descarte de insumos e histórico de vazamento. O local vem sendo remediado por bombeamento e tratamento desde abril de 2014. A figura 1 contempla a localização da área de pesquisa.

**Figura 1** - Localização da área de pesquisa (polígono em vermelho)



### 2.1. Geologia Local

Foi possível caracterizar a geologia local, de acordo com WALM (2012), da seguinte maneira, da base para o topo da sequência estratigráfica alvo da pesquisa:

- 40 a 30 metros: rocha gnáissica pouco alterada cinza esverdeado e fraturado – Alteração do próprio Complexo Amparo;
- 30 a 0 metros: depósitos sedimentares quaternários relacionados aos processos de erosão e transporte do rio Jundiá situado a norte da área – Sedimentos Aluvionares do Cenozóico.

A base impermeável do modelo hidrogeológico conceitual da área de pesquisa são rochas granítico/gnáissicas do Complexo Amparo.

Durante as sondagens para a instalação dos poços de monitoramento, foram identificadas diversas litologias relacionadas aos sedimentos quaternários, sendo que estas litologias foram agrupadas em sedimentos argilosos e arenosos, os quais são condicionantes do fluxo subterrâneo local.

### 2.2. Hidrogeologia Local

A área de pesquisa possui profundidade média do nível d'água de 5,26m, mínima de 2,55m e máxima de 7,40m. O valor médio da profundidade do nível d'água está posicionado no intervalo dos sedimentos arenosos e nesta profundidade ocorrem as maiores concentrações da contaminação do aquífero não confinado.

O fluxo subterrâneo realizado através de observações de campo, e que serve como referência na validação dos modelos de fluxo, possui orientação preferencial de sul para norte, no sentido da zona de descarga local (rio Jundiá), com pequenas alterações de direção.

Dez ensaios *slug test* foram realizados para a avaliação dos dados de condutividade hidráulica do lençol freático local. Os resultados destes ensaios podem ser visualizados na Tabela 1. Estes ensaios foram realizados apenas na cobertura aluvionar da área.

O valor médio de condutividade foi de  $2,66 \times 10^{-4}$  m/s, mínimo de  $4,12 \times 10^{-5}$  m/s e máximo de  $8,32 \times 10^{-4}$  m/s.

**Tabela 1** - Ensaios *slug test* realizados em 10 poços de monitoramento

Poço	Condutividade (m/s)	Condutividade (m/ano)
PM-19	8,32E-04	262,38
PM-06/15	4,12E-05	12,99
PM-13C	7,07E-05	22,30
PM-13B	2,88E-04	90,82
PMm-19	1,62E-04	51,09
PMm-13C	6,99E-05	22,04
PMm-05	5,50E-04	173,45
PMm-17	3,58E-04	112,90
PMp-05	2,26E-04	71,27
PMp-13C	5,92E-05	18,67

### 2.3. Pluma de Contaminação

A pluma de contaminação foi delimitada segundo diretrizes da Companhia Ambiental do Estado de São Paulo – CETESB e com base nas concentrações do contaminante ferro dissolvido (Fe+2) (Figura 2).

Para gerar esta pluma inicial, objeto da simulação de fluxo e transporte de contaminantes, foi elaborada uma malha irregular de 52 x 53 blocos referentes a 430 x 540 m. Em profundidade, a pluma está numa camada de 5 m de espessura considerada como não confinada.

As direções de fluxo da água subterrânea são de sul para norte, por isso foi adotado condições de contorno com células de carga constante nas bordas sul e norte, seguindo as direções de fluxo no local. Na porção a norte do rio Jundiá as células foram consideradas nulas.

É possível perceber que as maiores concentrações da contaminação (*hot spots*) estão próximas ao PM-09. Para o melhor detalhamento, as células do grid foram refinadas nesta porção do terreno. Além disso, em função desta alta concentração diagnosticada nesta região, maior adensamento de poços foi instalado a jusante deste *hot spot* com o objetivo de delimitar a pluma em direção ao rio Jundiá.

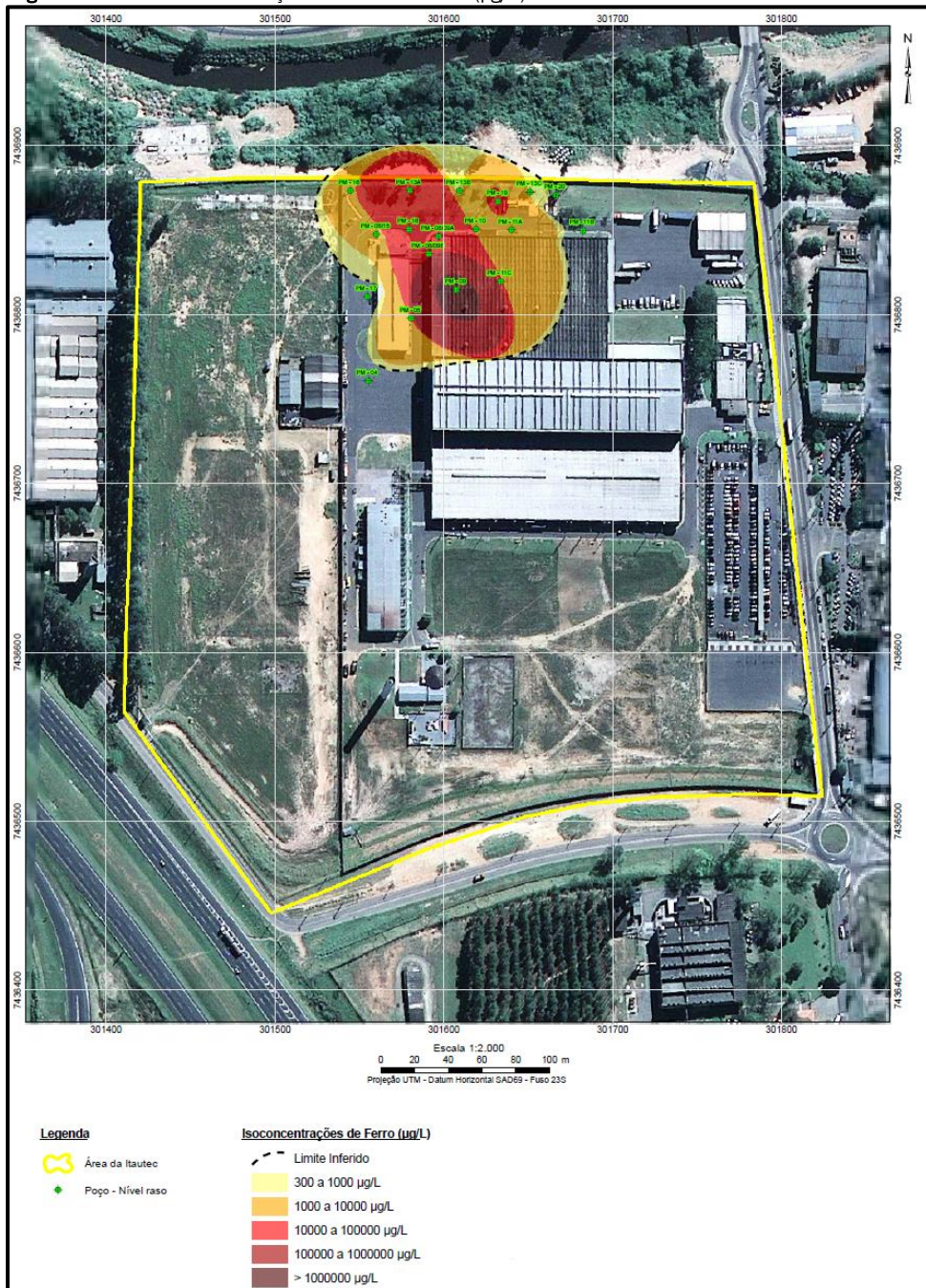
### 3. MÉTODOS

Foram realizadas simulações estocásticas para a geração de cem cenários com diferentes distribuições de condutividade hidráulica.

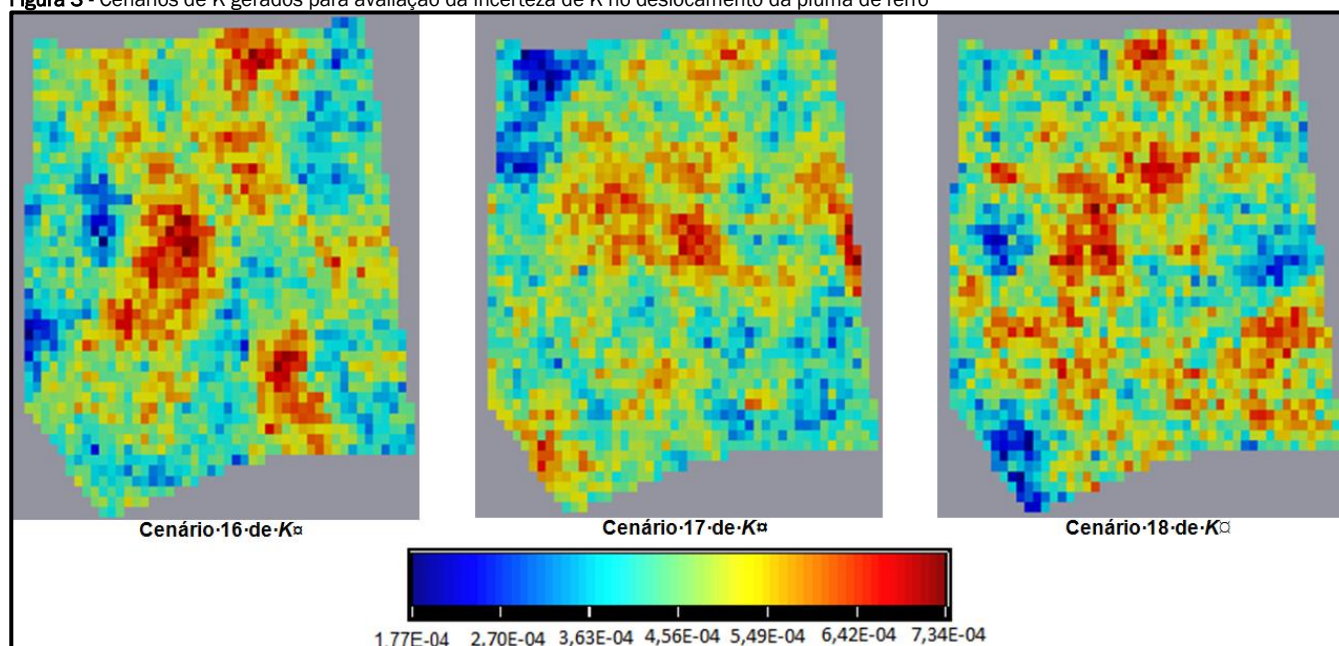
O método sequencial de simulação escolhido foi a simulação gaussiana sequencial (SGS). De acordo com Yamamoto e Landim (2013), a SGS é a aplicação de procedimento de simulação sequencial para funções aleatórias multigaussianas. Para tal, realiza-se a simulação de  $N$  variáveis aleatórias  $\{Z(x_i), i=1, N\}$  localizados sobre os nós de uma malha regular e condicionadas ao conjunto de  $n$  pontos de dados  $\{z(x_\alpha), \alpha=1, n\}$ .

Neste algoritmo, um valor aleatório é atribuído a cada célula que não possui dado experimental, definindo uma ordem aleatória para todas as células da malha. Para cada célula, uma função de densidade de probabilidade (fdp) é estipulada com base em um número de dados condicionantes vizinhos (dados iniciais e dados simulados precedentemente). O valor aleatório da fdp é atribuído estabelecendo-se uma continuidade espacial. A figura 3 exemplifica três cenários de  $K$  gerados com a metodologia proposta no presente trabalho.

Figura 2 - Pluma de contaminação dissolvida de ferro ( $\mu\text{g/L}$ )



**Figura 3** - Cenários de K gerados para avaliação da incerteza de K no deslocamento da pluma de ferro



A partir de 100 cenários de condutividade hidráulica, foram realizadas 30 simulações de fluxo e transporte de contaminantes utilizando-se da técnica tradicional por diferenças finitas (*MODFLOW* e *MT3D*). Os resultados destas trinta simulações foram utilizados para o treinamento das redes neurais artificiais. Desse modo, basicamente, a pesquisa se divide em cinco etapas metodológicas principais:

- i. Geração de 100 cenários de condutividade hidráulica por meio de métodos estocásticos (simulação gaussiana sequencial), sendo que quanto maior o número do cenário menor será a grandeza dos valores de K estipulados;
- ii. Simulação de fluxo e transporte de contaminantes por meio do *MODFLOW* para 30 cenários de K (cenário 1 ao cenário 30);
- iii. Com os resultados da simulação de fluxo, a rede neural foi treinada para prever os resultados do *MODFLOW*;
- iv. Simulação de fluxo e transporte de contaminantes por meio de redes neurais artificiais para 70 cenários de K (cenário 31 ao cenário 100) para a fase de avaliação;
- v. Avaliação da variabilidade da concentração de ferro em 3 anos do deslocamento da pluma para os 100 cenários modelados.

Para o treinamento da rede neural para a simulação da pluma de contaminação foram utilizados os parâmetros estatísticos

dos cenários de K e as concentrações de Fe nos poços de monitoramento.

A análise da incerteza relacionada à K no deslocamento da pluma de contaminação foi avaliada com base na correlação entre os valores médios dos cenários de condutividade hidráulica e as concentrações dissolvidas ( $\mu\text{g/L}$ ) de ferro que persistiram na área de pesquisa após três anos de simulação, ou seja, quanto maior o valor médio de K, menos ferro tende a ficar na área e vice e versa.

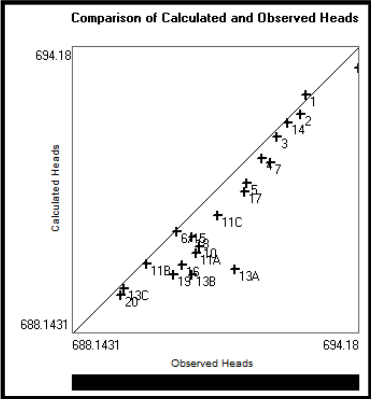
Os softwares utilizados nesta pesquisa foram: (i) *SGEMS* (REMY, 2004) para a geração dos 100 cenários de K; (ii) *MATLAB* e (iii) *MODFLOW* (U.S GEOLOGICAL SURVEY, 2011) para a simulação do fluxo e o transporte de contaminantes, todos os softwares foram utilizados em suas versões livres ou possuem licença de uso para a comunidade acadêmica.

### 3.1. Modelo de fluxo subterrâneo e de transporte de contaminante - *MODFLOW*

O modelo de fluxo subterrâneo e de transporte de contaminantes foi aplicado para o treinamento das redes neurais artificiais, bem como para prever a concentração da pluma de ferro em 3 anos de simulação.

A tabela 2 apresenta as principais características do modelo.

**Tabela 2** - Informações utilizadas para o modelo de fluxo e transporte de contaminantes.

Informações utilizadas		Justificativa da informação, caso necessária
Contaminante considerado	Ferro - Fe2+ (µg/L)	-
Coefficiente de distribuição	0,0	O coeficiente de distribuição para o ferro não é significativo, o que faz com que ele possua alta mobilidade.
Números de camadas	2: camada 1 - Topo, Não confinado camada 2 - Base, Confinado/ Não-confinado Transmissividades variadas.	Foram consideradas as litologias predominantes da área, ou seja, areia no topo e argila na base do modelo.
Número de células	Malha 16 x 20, com distribuição irregular e adensamento na região norte com maior número de poços.	Adensamento da malha na região crítica de contaminação - <i>hot spot</i> .
Condições de contorno	Carga hidráulica fixa durante a simulação na região sul da área de pesquisa e células nulas na região norte após o rio Jundiá	Com base no gradiente hidráulico.
Regime	Transiente	-
Tempo de simulação	3 anos	-
Calibração	Valores de nível d'água calculados e observados em campo apresentaram boa correlação	
Porosidade efetiva para a camada superior	4 e 6%	4% para a região de menor porcentagem e 6% para áreas de maior porcentagem de areias.
Porosidade efetiva para a camada inferior	2%	Constante para toda a camada em função do maior volume de sedimentos argilosos
Fator de retardamento	0,0	Fator de retardamento para o ferro não é significativo, o que faz com que ele possua alta mobilidade.
Coefficiente de decaimento de primeira ordem	Não foi utilizado	-
Concentrações iniciais utilizadas	Dados provenientes da pluma de contaminação delimitada a partir dos resultados das análises realizadas.	-
Cota do fundo e carga hidráulica do rio Jundiá	Levantamento topográfico	-
Valor de recarga do aquífero estudado	CEPAGRI - precipitação anual para a região da área de estudo (113 mm).	-
Evapotranspiração	Considerado nulo	Em função do tamanho da área (relativamente pequena) este parâmetro não foi considerado no modelo.
Condutividade hidráulica k	Cenários de K gerados por técnicas estocásticas - Camada superior - arenosa. Para a camada 2 (base do modelo - argila): 1,0x10 <sup>-6</sup> m/s - horizontal e 1,0x10 <sup>-7</sup> m/s - vertical.	Ensaio tipo <i>slug test</i> de campo.
Porosidade efetiva	4,5%	Média dos valores obtidos por análises de laboratório.
Densidade aparente	1,36 mg/m <sup>3</sup>	
Densidade seca	1,08 mg/m <sup>3</sup>	

### 3.2. Redes Neurais Artificiais - RNA

#### 3.2.1. Perceptron de Múltiplas Camadas (MLP)

Uma das técnicas mais utilizadas para predição de propriedades contínuas em inteligência artificial é o método Perceptron

de Múltiplas Camadas (MPL). A justificativa da ampla contribuição da técnica, especialmente em geociências, é justificada por Poulton (2001) pela sua extensa habilidade e flexibilidade de adaptação para solucionar problemas com dados escassos e de relações complexas, comuns nesta área.

MLP, como outras técnicas de redes neurais, é baseada no aprendizado por repetição ou treinamento. A rede é constituída de neurônios, ou nós, que estão conectados, imitando o funcionamento do cérebro humano (Figura 4). Desta forma, as amostras são recebidas pelos nós, como as sinapses atuam nos neurônios. Computacionalmente, esta equivalência é realizada por pesos associados a cada um dos nós.

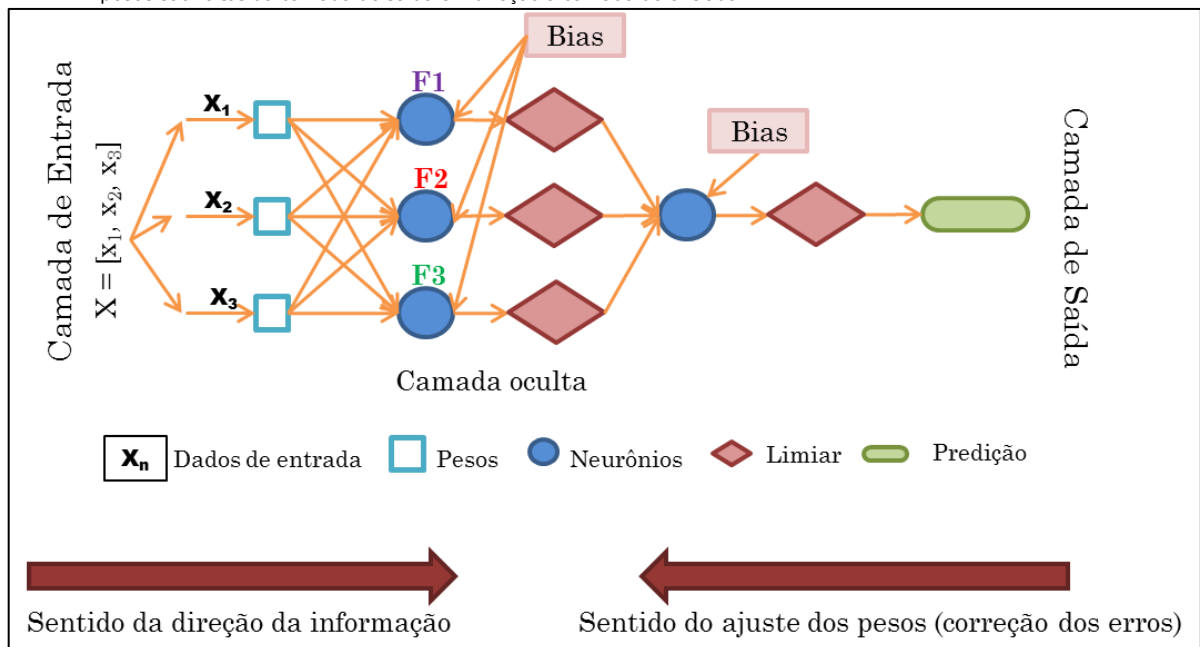
Tais pesos ponderam a influência das variáveis de cada amostra antes de transformá-las através da função de ativação, também associada a cada um dos nós. Em geral a saída desta transformação é vinculada a um valor constante. Este valor passa posteriormente por uma avaliação que condiciona se este pulso será inibido ou se será enviado para o próximo neu-

rônio ao qual está conectado, chamado de *threshold*, ou limiar.

A rede MLP é formada por três camadas distintas: camada de entrada, camada oculta e camada de saída. Na camada de entrada, os neurônios recebem as amostras de entrada. O treinamento da rede ocorre nas camadas ocultas e o resultado final é representado pela camada de saída.

Na figura abaixo (Figura 4) o exemplo simula uma amostra com 3 dimensões, ou 3 variáveis, cujos valores são ponderados por pesos e posteriormente transformados nos neurônios pelas funções de ativação.

**Figura 4** - Esquema do funcionamento da rede neural MLP. Formado por três camadas distintas (entrada, oculta e saída), a rede neural tem como principal objetivo adaptar os valores de pesos associados aos neurônios a fim de convergir o treinamento aos valores de saída desejados. Desta forma, depois de propagar o sinal de entrada até a camada de saída, os ajustes dos pesos são feitos da camada de saída em direção à camada de entrada



Note que, além do limiar que decidirá se o impulso será propagado, ainda há a adição de uma tendência (bias), representada por uma constante adicionada ao valor de saída da função de ativação. Na camada de saída, os valores limiares resultantes das funções de ativação são comparados e apenas o maior valor é propagado para a camada de saída, em um processo chamado competitivo.

Particularmente na rede neural MLP, o treinamento da rede acontece em duas etapas. Primeiro as amostras são recebidas pelos neurônios e depois acontece o ajuste da rede. Como já explicado, a cada amostra de entrada, a rede gera um resultado de saída. No segundo passo, este valor é comparado com aquele que se deseja prever, neste caso, valores de Fe, em um processo de treinamento supervisionado.

Durante o treinamento o objetivo principal da rede é adequar os valores de peso da rede a fim de convergir a série de cálculos não lineares às saídas desejadas. Desta forma, o sentido do cálculo de ajuste dos pesos é contrário ao caminho da in-

formação propagada pelas amostras de entrada na rede, pois o início dos cálculos de ajuste acontece a partir da camada de saída em direção à camada de entrada. Por isso este ajuste é conhecido como *back propagation*, ou retropropagação de erros.

Consequentemente, neste estudo, a função de ativação adotada foi a função sigmoide, caracterizada por ser continuamente diferenciável e positiva, garantindo a possibilidade de ter função inversa (ser ajustada) e apresentar resultados que não gerem saltos muito grandes dado um intervalo pequeno de dados de entrada, o que desestabilizaria a rede.

Para avaliação do desempenho da rede, os dados de entrada são divididos em três grupos: treinamento, validação e teste. Os dados de treinamento são utilizados para definir os parâmetros dos pesos, enquanto o conjunto de validação, em geral composto por 10 a 25% das amostras disponíveis, representa uma avaliação não enviesada dos resultados, pois utiliza um segundo conjunto de dados independente do treinamento



usado para predição (ZHAO *et al.* 2015). Através dos resultados dos dados de validação a arquitetura da rede (número de neurônios em cada camada) pode ser ajustada.

Por fim, utiliza-se o conjunto de dados de teste, que apresenta a mesma probabilidade de distribuição dos dados de treinamento, embora independente daquele. Este dado auxilia a identificação do *overfitting*, processo em que a rede passa a repetir resultados e perde a capacidade de aprender, ou generalizar os resultados para casos ainda não treinados.

Neste trabalho, o ajuste dos pesos foi aproximado pela função de Levenberg-Marquardt que apresenta bons resultados a um valor computacional menor do que o cálculo de inversão das funções, método matematicamente original. O erro dos ajustes foi medido a partir de valores do erro médio quadrático (MSE), critério usualmente adotado para as redes neurais supervisionadas.

Segundo Zhao *et al.* (2015) e Haykin (1999) os maiores desafios da rede neural são associados a adequada seleção de dados para o treinamento da rede, que deve compor uma gama extensa de possibilidades de valores e ocorrências (também para não serem considerados pontos *outliers*) e o correto ajuste da rede.

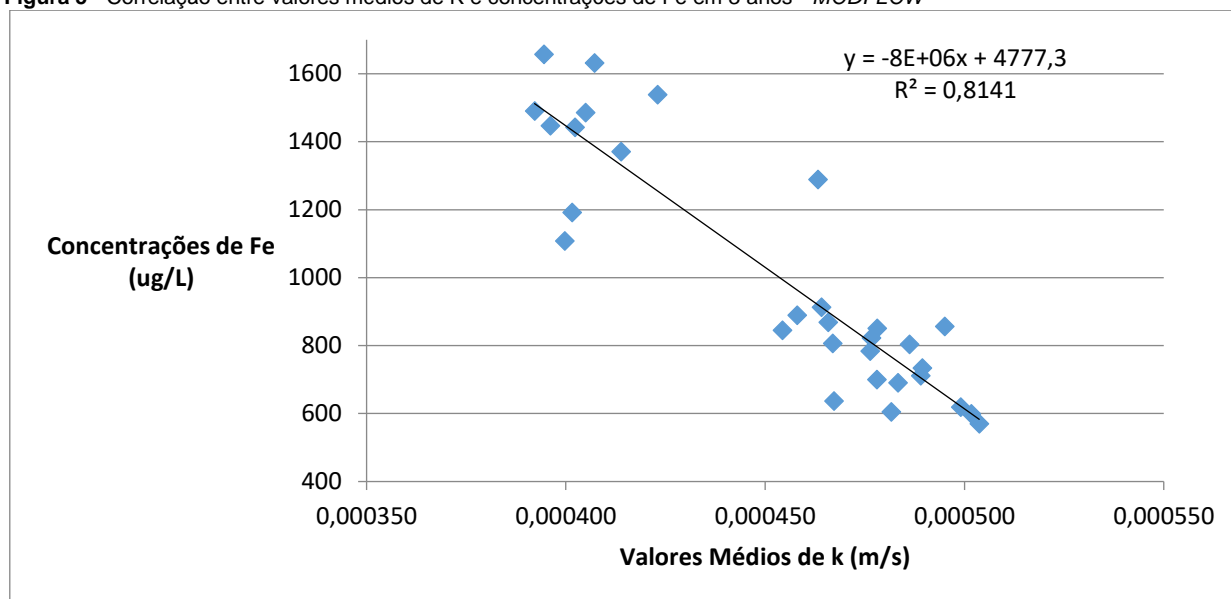
## 4. RESULTADOS E DISCUSSÕES

### 4.1. Aplicação do MODFLOW – 30 Cenários de K

As simulações de fluxo e do transporte de contaminantes, utilizando-se do MODFLOW, foram realizadas para 30 cenários de condutividade hidráulica, atribuindo-se para cada célula um valor de K e sua média estipulada como o valor K médio para cada cenário. A modelagem do fluxo de transporte considerou apenas as componentes de advecção e dispersão, não sendo considerados os processos de adsorção e decaimento. A predição dos valores de Fe com o método tradicional durou cerca de 1 hora por modelo, sendo necessárias 30 horas para simular todos os modelos. Vale ressaltar, contudo, que o MODFLOW gera informações de concentração para cada célula da malha modelada em cada passo de tempo *t* estipulado (1, 2 e 3 anos).

Com o MODFLOW foi possível verificar a correlação esperada, ou seja, quanto maior o valor médio de K de cada cenário, menores as concentrações finais de Fe dissolvido na área de pesquisa, e logo menor a massa dos contaminantes remanescente, e, conseqüentemente, maior o aporte da pluma para o rio Jundiá, conforme pode ser visto na figura 5.

Figura 5 - Correlação entre valores médios de K e concentrações de Fe em 3 anos - MODFLOW



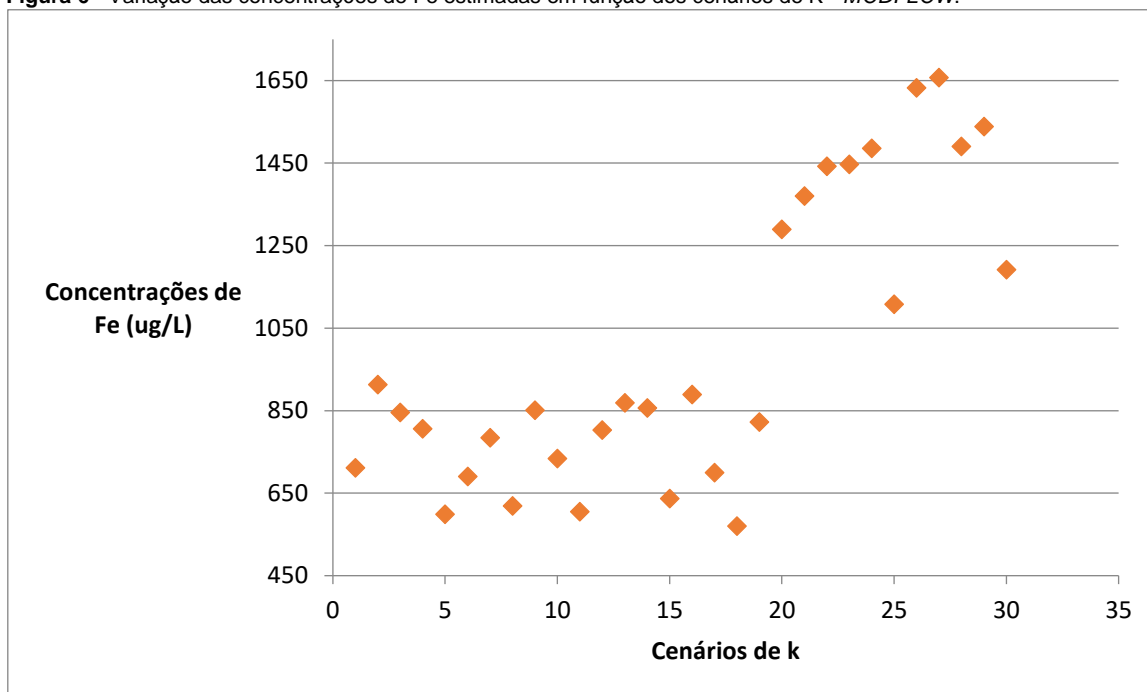
A concentração de Fe obtida variou de 570,36 µg/L a 1657,38 µg/L. Para as modelagens realizadas com o MODFLOW, a variabilidade encontrada nas concentrações de Fe na água subterrânea frente aos cenários de condutividade hidráulica influencia de forma significativa a tomada de decisão em relação ao gerenciamento da área contaminada.

O valor mínimo é cerca de 3 vezes menor que o máximo e essa variação pode levar às considerações distintas sobre a utilização das águas subterrâneas, bem como às interpretações diferentes sobre o aporte da contaminação no rio Jundiá. A figura 6 apresenta a variabilidade das concentrações estimadas de Fe para cada cenário de K.

O grupo de resultados com as concentrações médias (média aritmética das concentrações da malha) mais elevadas de Fe (entre 1050 µg/L e 1660 µg/L) significa menor aporte de contaminante para o rio, enquanto que os resultados com concentrações médias mais baixas de Fe (entre 570 µg/L e 900 µg/L) indicam maior aporte da pluma de contaminação à jusante da área de pesquisa.

A maioria dos poços de monitoramento da área possui valores negativos de potencial redox (18 dos 33 poços), característico de ambiente redutor, o que pode indicar pouca perda de massa ao longo do fluxo.

Figura 6 - Variação das concentrações de Fe estimadas em função dos cenários de K - MODFLOW.



Percebe-se que a variabilidade encontrada nas concentrações de Fe, em 3 anos de simulação, foi significativa e que a mesma é capaz de influenciar tomadas de decisões opostas no gerenciamento da área, referente à concentração máxima remanescente na área de estudo e à relação entre as águas subterrâneas contaminadas e seu possível aporte em corpos hídricos.

#### 4.2. Treinamento de RNA para os 30 cenários

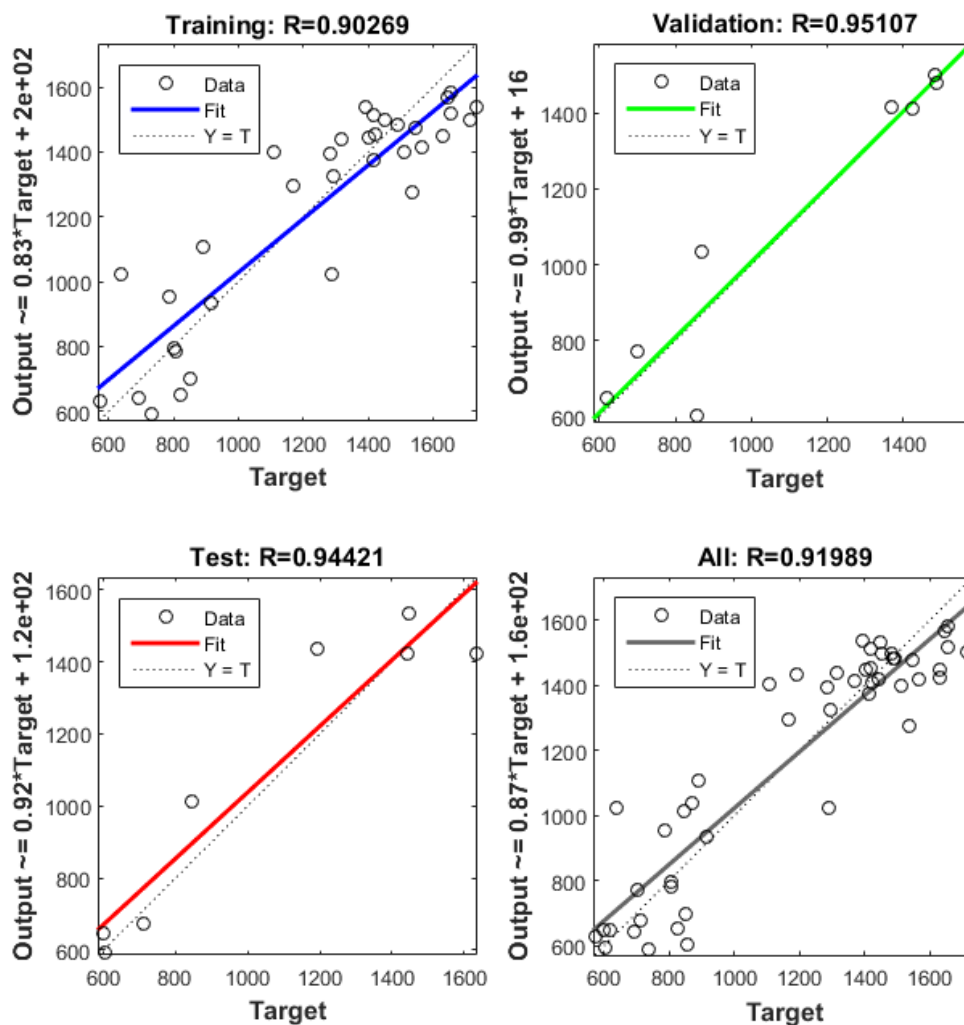
O conjunto de dados utilizados para o treinamento da rede MLP foi calculado a partir de 30 cenários de K (cenário 1 ao cenário 30), dos quais foram utilizados como variáveis os valores médios, o desvio padrão, e os percentis de 25, 50 e 75%, obtidos dos mapas de K com o propósito de prever a concen-

tração de Fe.

Para esta aplicação, a configuração da arquitetura da rede foi definida com 5 neurônios na camada oculta, que convergiu em apenas 6 iterações, em menos de 1 segundo, tendo como resultado final a concentração média final de Fe restante na área de estudo. A correlação entre os valores reais e esperados foi de 92%, valor médio da correlação dos dados utilizados para treinar e validar a técnica (Figura 7).

Note que com esta estrutura de rede, os valores de correlação dos dados de treinamento, teste e validação ficaram todos acima de 90%, demonstrando que sua arquitetura está adequada para esta predição.

**Figura 7** - Os gráficos mostram os resultados obtidos do treinamento da técnica MLP. O valor médio do desempenho da rede neural foi de 92% para a predição de Fe. Todas as correlações foram acima de 90%, demonstrando que a arquitetura da rede está bem definida



#### 4.3. 100 Cenários de K e Possíveis Técnicas de Remediação

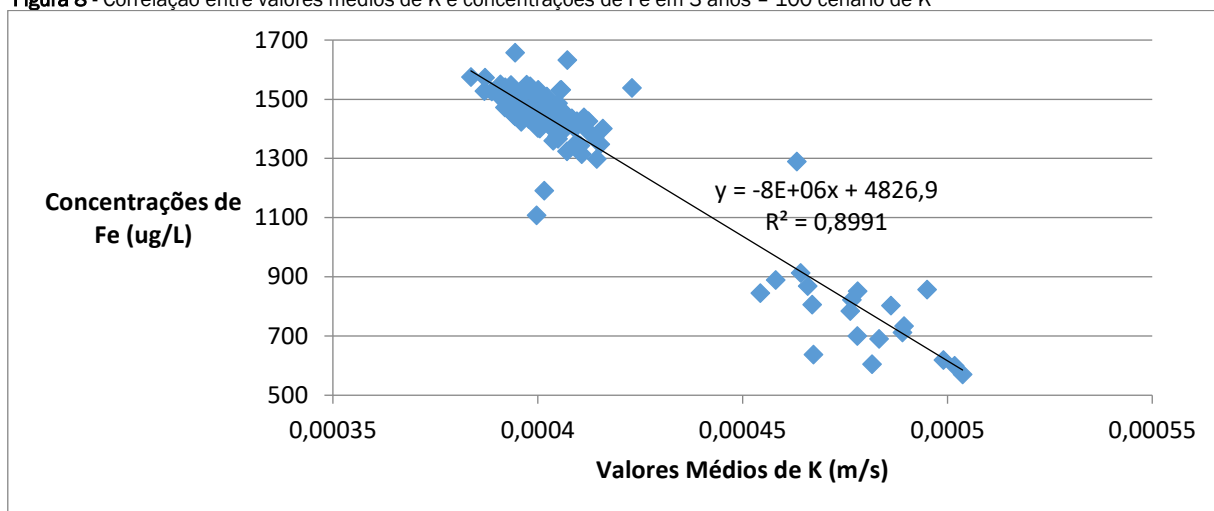
A partir do treinamento com MLP, 70 pontos de controle de concentração média de Fe foram estimados, baseados no mesmo conjunto estatístico de valores de K, para os quais os valores de Fe eram desconhecidos.

A figura 8 apresenta a correlação entre os valores médios dos cenários de condutividade hidráulica e as concentrações médias de Fe resultantes. Como esperado, a correlação se mantém e quanto menor os valores médios de condutividade, maiores as concentrações médias de Fe que permanecem na

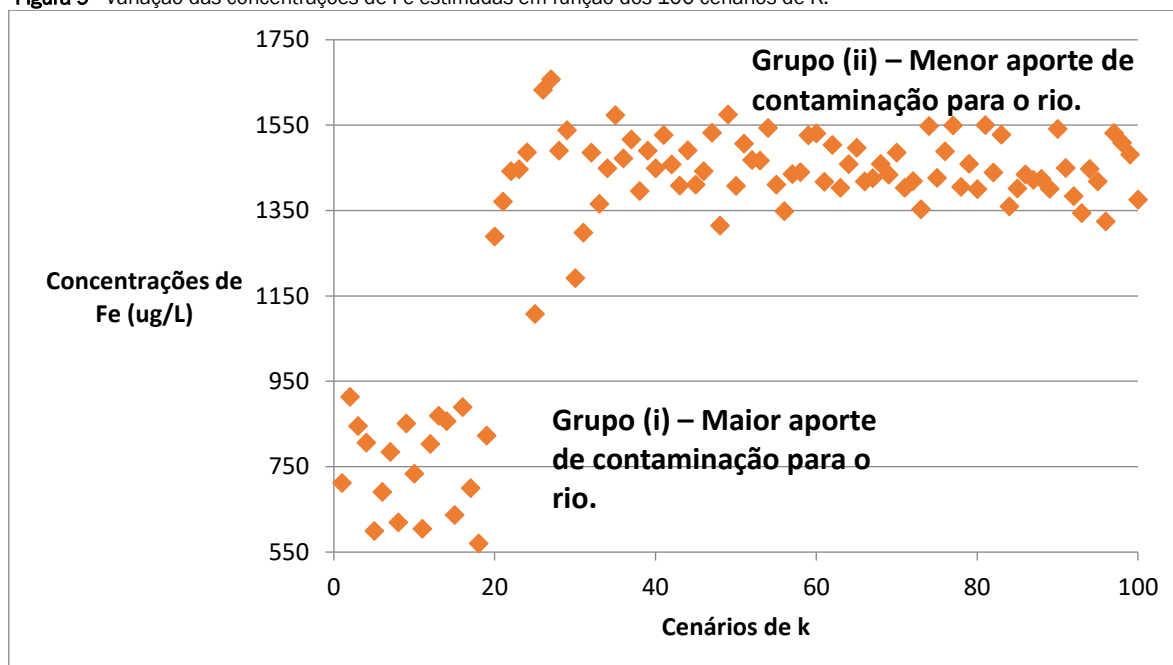
área após 3 anos de simulação.

Na figura 9 é possível verificar a variabilidade das concentrações médias de Fe resultantes para os 100 cenários de condutividade gerados na pesquisa. Notam-se dois grupos de resultados: (i) um com valores menores de concentrações de Fe (entre 550 µg/L e 950 µg/L) e (ii) outro com valores maiores (entre 1350 µg/L e 1700 µg/L). O primeiro indica maior aporte de contaminação para o rio Jundiáí, enquanto que o segundo menor migração da pluma para o mesmo em 3 anos de simulação.

**Figura 8** - Correlação entre valores médios de K e concentrações de Fe em 3 anos – 100 cenários de K



**Figura 9** - Variação das concentrações de Fe estimadas em função dos 100 cenários de K.



A variabilidade dos resultados de Fe dissolvido em função dos cenários de K (Figura 9) pode ser considerada alta frente à tomada de decisão no decorrer do gerenciamento ambiental da área contaminada, quer dizer, a depender do uso do solo pretendido para a área, a concentração mínima encontrada de 570,36  $\mu\text{g/L}$  pode ser avaliada como legal para uso industrial, por exemplo, enquanto a máxima, de 1657,38  $\mu\text{g/L}$ , não. Dessa forma, nota-se a influência da variabilidade de K na tomada de decisão para a gestão dos recursos hídricos subterrâneos contaminados.

O grupo (i) indica a necessidade de implantação de uma contenção hidráulica à jusante da pluma para evitar o aporte de contaminação no rio Jundiá, enquanto que o grupo (ii) indica a necessidade de uma medida de remediação com o objetivo apenas de abater a concentração do contaminante na área *in situ*, sem a obrigação de contenção da pluma, podendo ser um bombeamento e tratamento no centro de maior concen-

tração da contaminação.

A variabilidade na migração da pluma em função dos cenários de condutividade hidráulica é significativa podendo comprometer a tomada de decisão em relação à recuperação ambiental da área.

A pluma de Fe utilizada nessa pesquisa é apenas um exemplo, podendo ser considerada hipotética, sendo que a variabilidade de K é a principal variável que controla a migração de plumas de contaminação e influencia diretamente o comportamento da contaminação no decorrer do tempo e, conseqüentemente, na tomada de decisão para o bom gerenciamento ambiental da área.

Cada cenário de K modelado por redes neurais para a migração da pluma de Fe em 3 anos de simulação durou um tempo de apenas 2 segundos por cenário de condutividade hidráulica.

lica, enquanto que por meio do MODFLOW cada simulação levou cerca de 1 hora. Dessa forma, a técnica de RNA para prever o cenário ambiental de uma pluma de contaminantes se mostrou significativamente mais rápida em relação aos métodos tradicionais de simulação de fluxo e transporte de contaminantes.

## 5. CONCLUSÕES

A utilização de RNA para a simulação de fluxo e o transporte de contaminantes se mostrou adequada e confiável para a área de pesquisa. As correlações encontradas foram validadas com base nos resultados do treinamento da rede neural a partir das concentrações de Fe estimadas pelo MODFLOW, bem como no conceito hidrogeológico baseado em hidrofácies (NEUMAN, 1990), o qual relaciona diretamente litologia, condutividade hidráulica e transporte de contaminantes.

Todos os 100 cenários simulados por RNA foram realizados em um total de apenas 2 minutos e 20 segundos para a obtenção das concentrações de Fe em 3 anos de simulação. O MODFLOW demorou 1 hora por cenário de K, ou seja, 30 horas no total para a mesma simulação. Numa situação de maior complexidade, a técnica RNA passa a ser fundamental para subsidiar a tomada de decisão mais rápida.

Os resultados das concentrações do contaminante indicaram tomadas de decisões distintas para o gerenciamento ambiental do local, ratificando a importância da variabilidade de K no comportamento da migração de uma pluma de contaminação, bem como de que modo remediá-la.

A presente pesquisa abordou a utilização de RNA para a previsão da migração de pluma de contaminação, técnica esta, ainda, pouca utilizada no Brasil, tanto na academia como no mercado ambiental.

No entanto, trabalhos futuros podem utilizar RNA para a previsão de técnicas de remediação de maneira quantitativa com a previsão e estimativa de taxas de bombeamento e/ou injeção, bem como de localização ideais de poços para a boa remediação da área otimizando tempo e consequentemente custos.

## REFERÊNCIAS

AHLFELD, D.P., MULVEY, J.M., and PINDER G.F., Contaminated Groundwater Remediation Design Using Simulation, Optimization, and Sensitivity Theory 2, Analysis of a Field Site. *Water Resources Research*, v. 24, n. 3, p. 443-452p. 1988.

ALBERTO, C.M; CHANG, H.K. Os Desafios da Remediação por Bombeamento e Tratamento em Domínios Geológicos Complexos: O Uso de Modelos Estocásticos Para Auxiliar na Determinação da Zona de Captura. *Águas Subterrâneas*, v.25, n.1, p.121-142, 2011.

BAYER, P., and FINKEL, M., Optimization of concentration control by evolution strategies: Formulation, application, and assessment of remedial solutions. *Water Resources Research*, v.. 43, W02410, 2007.

CETESB – Companhia Ambiental do Estado de São Paulo. *Texto explicativo. Relação de áreas contaminadas e reabilitadas no Estado de São Paulo*. 2018.

CHEN-WUING, L., JANG CHENG-SHIN, J., CHUNG-MIN, L., Evaluation of arsenic contamination potential using indicator kriging in the Yun-Lin aquifer (Taiwan). *Science of the Environment*, v. 321, p. 173-188, 2004.

COPPOLA, E., SZIDAROVSKY, F., POULTON, M., AND CHARLES, E., Artificial neural network approach for predicting transient water levels in a multilayered groundwater system under variable state, pumping, and climate conditions. *Journal of Hydrologic Engineering*, v. 8, n. 6, p. 348– 359, 2003.

COSTANZO; VIDAL. *Análise de Incerteza de Cenários de Bombeamento e Tratamento em Áreas Contaminadas*. *Águas Subterrâneas* v. 29, n. 1, p. 13-29, 2014.

FLACH, G. P, HARRIS, M. K., SMITS, A. D, SYMS, F.H. *Modeling aquifer heterogeneity using cone penetration testing data and stochastic upscaling methods*. The American Association of Petroleum Geologists/Division of Environmental Geosciences. *Environmental Geosciences*, v.12, n. 1, P. 1-15, 2005.

GORELICK, S. M., *Sensitivity analysis of optimal groundwater contaminant capture curves: spatial variability and robust solutions*, paper presented at the Conference and Exposition: Solving Groundwater Problems with Models, Natl. Water Well Assoc., Denver, Colo., Feb. 10-12, 1987.

GUNGOR-DEMIRCI, G.; AKSOY, A. Change in optimal pump-and-treat remediation design and cost for different correlation lengths os spatially variable hydraulic conductivity field. *Quarterly Journal of Engineering and Hydrogeology*, v. 44, p. 469-480, 2011.

HASSAN, A.E., POHLMANN, K.F., AND CHAMPAN, J.B., Uncertainty analysis of radionuclide transport in a fractured coastal aquifer with geothermal effects. *Transport in Porous Media* v. 43, p. 107–136, 2001.

HAYKIN, S. *Neural Networks: a comprehensive foundation*. 2. ed., Hamilton, Ontario, Canada: Bookman, McMaster University, 2008. 900p

HAYKIN, S. *Neural Networks*. Virginia, USA: Prentice Hall. 1999. 842p.

JOHNSON, V.M.; ROGERS, L.L. Accuracy of Neural Network Approximators in Simulation Optimization, *J of Water Res. Planning & Mgt.*, ASCE, v. 126, n. 2, p. 48-56, 2000.

JOURNEL, A.G.; ALABERT, F.G., *Focusing on spatial connectivity of extreme-valued attributes: Stochastic indicator models of reservoir heterogeneities*. *SPE Paper N° 18324 presented at the 63rd Annual Technical Conference and Exhibition of the Society of Petroleum Engineers*. Houston, TX, 1988.

LEE, M.K.; SAUNDERS, J.A; WOLF, L.W. Effects of geologic heterogeneities on pump-and-treat and in situ bioremediation: A stochastic analysis. *Environmental Engineering Science*, v. 17, p. 183–189, 2000.

LIU CHEN-WUING; JANG CHENG-SHIN; LIAO CHUNG-MIN. *Evaluation of arsenic contamination potential using indicator kriging in the Yun-Lin aquifer (Taiwan)*. *Science of the Environment*, v. 321, p. 173-188, 2004.

- MASKEY, S., DIBIKE, Y.B., JONOSKI, A. ; SOLOMATINE, D.P., *Groundwater model approximation with artificial neural network for selecting optimal pumping strategy for plume removal, in AI Methods in Civil Engineering Applications*, edited by O. Schleider and A. Zijderveld, 67 – 80p, Cottbus. 2000.
- MASSMANN, J. ; FREEZE, R.A. Groundwater Contamination from Waste Management Sites: The Interaction between Risk-Based Engineering Design and Regulatory Policy, 1, *Methodology. Water Resources Research*, v.32, n. 2, p. 351-367, 1987a.
- MASSMANN, J. ; FREEZE, R.A. *Groundwater Contamination from Waste Management Sites: The Interaction between Risk-Based Engineering Design and Regulatory Policy, 2, Results*”, *Water Resources Research*, v. 32, n. 2, p. 368-380, 1987b.
- MEYER, P.D. ; BRILL, E.D. *A Method for Locating Wells in a Groundwater Monitoring Network Under Conditions of Uncertainty*, *Water Resources Research*, v. 24, n. 8, p. 1277-1282, 1988.
- MORGAN, D.R. *Decision making under uncertainty using a new chance constrain programming technique: A groundwater reclamation application*. Ph.D. (Dissertation) - Dep. of Civ. Eng., Univ. of Ill. at Urbana-Champaign. 1990.
- NEUMAN, S.P. Universal scaling of hydraulic conductivities and dispersivities in geological media. *Water Resour. Res.*, v. 26, n. 8, p. 1749-1758, 1990.
- NIKOLOS, I.K., STERGIADI, M., PAPAPOPOULOU, M.P., KARATZAS, G.P. Artificial neural networks an alternative approach to groundwater numerical modeling and environmental design. *Hydrol Process*, v. 22, n. 17, p. 3337–48, 2007.
- POULTON, M.M., Multi-Layer Perceptrons and Back-Propagation Learning. *Handbook of Geophysical Exploration, Seismic Exploration. Computational Neural Networks for Geophysical Data Processing*, v. 30, p. 27-53, 2001.
- RANJITHAN, S.; EHEART, J.W. Neural Network-Based Screening for Groundwater Reclamation Under Uncertainty. *Water Resources Research*, v. 29, n. 3, p. 563-574, 1993.
- RAHMAN, M.R., EHEART, J.W. ; VALOCCHI, A.J. *Aquifer remediation design under parameter uncertainty, paper presented at the Spring Meeting, AGU, Baltimore, 1990.*
- RAO, S., HANDAVESWARA, B.S., BHALLAMUDI, S.M., SRIVIVASULU, V. Optimal groundwater management in deltaic regions using simulated annealing and neural networks *Water Resour. Manage*, v. 17, p. 409-428, 2003.
- REMY, N., BOUCHER, A., WU, J. *Applied Geostatistics with SGeMS*, Cambridge University Press, 2004. 264 p.
- RICCIARDI, K.L.; PINDER, G.F.; KARATZAS, G.P. Efficient groundwater remediation system design subject to uncertainty using robust optimization. *Journal of Water Resources Planning Management*, ASCE, v. 133, p. 253–263, 2007.
- ROGERS, L.L., AND DOWLA, F.U. Optimization of Groundwater Remediation Using Artificial Neural Networks With Parallel Solute Transport Modeling. *Water Res. Research*, v.30, n. 2, p. 457-481, 1994.
- SAMANI, M., GOHRI-MOGHADAM, M., SAFAVI, A.A. A simple neural network model for the determination of aquifer parameters. *Journal Hydrology*, v. 340, n. 1/2, p. 1–11, 2007.
- USGS - United States Geological Survey, *Office of Groundwater: Status of MODFLOW Versions and MODFLOW-Related Programs Available on USGS Web Pages*. Disponível em: <http://water.usgs.gov/nrp/gwsoftware/modflow-status-2011Jan.pdf>. 2011.
- VINYALS, O.; LE, Q.V., *A neural conversational model. In: 31st International Conference on Machine Learning*, Lille, France. JMLR: W&CP, vol. 37. 2015.
- WAGNER, B.J.; GORELICK S.M., Optimal groundwater quality management under parameter uncertainty. *Water Resources Research*, v. 23, n. n. 7, p. 1162-1174, 1987.
- WAGNER, B.J. ; GORELICK, S. M. Reliable aquifer remediation in the presence of spatially variable hydraulic conductivity: from data to design. *Water Resour. Res.*, v. 25, n. 10, p. 2211-2226, 1989.
- WALM Engenharia e Tecnologia Ambiental. *Investigação Ambiental Detalhada da Fábrica da Itaútec*, 2012.
- YAMAMOTO, J.K., LANDIM, P.M. B. *Geoestatística: conceitos + aplicações*. São Paulo: Oficina de Textos. 2013.
- Zhao T., Jayaram V., Roy A., Marfurt K.J., *A comparison of classification techniques for seismic facies recognition. Interpretation*, v.3, n. 4, 2015.