

RESOLUÇÃO DE PROBLEMAS INVERSOS EM CINÉTICA QUÍMICA USANDO REDES NEURONAIIS ARTIFICIAIS

E. BORGES¹, D.C. MENEZES¹,

¹ Universidade Federal de Viçosa, Departamento de Química

E-mail: emilio.borges@ufv.br

RESUMO: *Todos os problemas físicos suscetíveis à modelagem matemática podem ser classificados como problemas diretos ou inversos. Os problemas diretos são os mais propalados; nesses, efeitos mensuráveis são determinados a partir de causas bem conhecidas. Os problemas inversos são aqueles em que se determinam causas a partir de efeitos. São tão importantes e possuem tantas aplicações científicas quanto os diretos embora seu tratamento matemático seja quase sempre mais complexo exigindo técnicas numéricas especiais. Uma dessas técnicas é um algoritmo baseado em redes neuronais artificiais que tem sido utilizado recentemente com grande sucesso em diferentes tipos de problemas inversos. Nesse artigo discute-se a filosofia do problema inverso em Cinética Química e aplica-se o método de redes neuronais artificiais para resolvê-lo. Dois modelos para mecanismos cinéticos, protótipos de grande relevância para uma série de reações químicas e biológicas são abordados. A eficiência da rede neuronal é comparada com aquela de outras técnicas numéricas de otimização disponíveis em pacotes comerciais, os algoritmos Levenberg-Maquardt e Simplex.*

PALAVRAS-CHAVE: Problemas Inversos; Cinética Química; Redes Neuronais Artificiais.

1. INTRODUÇÃO

1.1. Problemas Inversos em Ciência

Medidas experimentais são a base para a construção de modelos que possibilitam a obtenção de informações sobre a composição, estrutura e as diversas propriedades da matéria. Tais medidas, quando coletadas sob a rigorosa supervisão da metodologia científica constituem a essência das ciências naturais e são em grande parte responsáveis pela evolução tecnológica da civilização moderna levando o homem de suas primitivas experiências pré-históricas à descoberta de novos planetas potencialmente habitáveis fora de nosso sistema solar (Wittenmyer et al, 2014). A conexão entre diferentes campos de pesquisa, em especial teoria e experimento, torna imprescindível a análise correta e produtiva de informações experimentais o que envolve procedimentos numéricos desenvolvidos e executados computacionalmente. Dados experimentais podem ser pensados como o *efeito* produzido por determinada(s) *causa(s)* cuja identificação é uma tarefa indispensável, embora frequentemente intrincada, para a compreensão de um processo natural. A formulação quantitativa da causa envolve, em uma perspectiva científica, a formulação matemática do

problema, ou seja, alguma equação conectando causa e efeito deve ser encontrada. Por exemplo, algumas propriedades estruturais e elétricas de certos materiais, as quais são determinadas experimentalmente, podem ser pensadas como efeito cuja causa fundamental seria em uma primeira análise o tipo de interação potencial entre os átomos/moléculas constituintes do sistema (Borges et.al, 2008). A relação causa-efeito nos problemas científicos pode ser resumida de forma geral por equações como $K(f)=g$ em que g representa as medidas experimentais, f a causa física que origina os dados experimentais e K a equação de vínculo (linear ou não-linear) entre g e f . Geralmente são três possíveis tipos de equações K a conectar g e f ; equações diferenciais, integrais e matriciais. Nessa perspectiva é útil classificar os problemas causa-efeito em duas categorias; *diretos* ou *inversos*. A solução de um problema direto consiste em encontrar efeitos com base na descrição de suas causas. Por outro lado, um problema inverso seria a determinação de causas desconhecidas a partir da observação de seus efeitos (Borges, 2016; Groetsch, 1999). Pode-se resumir essas ideias da seguinte maneira:

Modelagem Numérica de um Problema Científico:

$$K(f) = g$$

- (a) determinar g a partir de K e f : problema direto;
- (b) determinar f a partir de K e g : problema inverso tipo 1;
- (c) obter K a partir de f e g : problema inverso tipo 2.

A resolução de problemas diretos é mais difundida, consistindo na maioria das abordagens realizadas por cientistas. Por exemplo, a partir da energia potencial entre núcleo-elétron (f - causa) pode-se resolver a equação de Schrödinger (K - equação diferencial) para o átomo de Hidrogênio (sistema) obtendo-se as energias possíveis do elétron (g - observado experimentalmente). Não há dificuldade numérica nesse método. Atualmente essa operação pode ser realizada em dispositivos computacionais portáteis. A resolução de problemas inversos dos tipos 1 e 2 também pode ser obtida sem muita dificuldade *desde que os problemas sejam bem colocados*. O matemático francês Jacques Hadamard (Hadamard, 1932) definiu um problema $K(f)=g$ como bem colocado quando;

- I.** Para cada dado g existe uma solução f ;
- II.** A solução f é única;
- III.** A dependência de f em g é contínua, isto é, quando o erro em g tende a zero, o erro na solução f também tende a zero.

Esses requisitos são conhecidos como **existência**, **unicidade** e **continuidade** das soluções f em relação aos parâmetros g e se forem obedecidos, técnicas simples de ajustes de dados tais como as baseadas em métodos de regressão linear ou não linear poderão ser utilizadas para a resolução do problema inverso. Todavia, se pelo menos uma das condições I, II e III for violada, o problema inverso será *mal colocado* e nesse caso mesmo o ruído experimental que está sempre presente em um processo de medição real dificultará a resolução podendo induzir importantes desvios na solução do problema inverso (Lemes et al,

2007). Diante disso, métodos numéricos especiais serão necessários para a determinação de soluções aceitáveis para os problemas inversos (Kirsch, 1996; Tarantola, 2004).

Um dos primeiros problemas inversos resolvidos se relaciona à notável descoberta de que nosso planeta é um mundo relativamente pequeno. O responsável foi o cientista e filósofo Eratóstenes. Consta em sua biografia que Eratóstenes lera em um livro de papiro sobre o fato aparentemente trivial de que em certa região da cidade egípcia de Siena (atualmente Aswan a 900km de Cairo) localizada ao sul de Alexandria, varetas verticais fixadas no solo não geravam sombra ao meio dia dos dias 21 de junho (Sagan, 1980). Eratóstenes descobriu intrigado que esse efeito não ocorria em Alexandria. Qual seria a causa para esse efeito? Nesse sentido Eratóstenes investigou um problema inverso. Começou, pois, a tentar identificar a causa; a não formação de sombras pelas duas varetas, a de Siena e a de Alexandria, só poderia ser explicada se a Terra fosse perfeitamente plana. Por outro lado, Eratóstenes concluiu que a única causa plausível para explicar sombra em Alexandria e ausência de sombra em Siena seria uma curvatura na superfície da Terra, ou seja, a mesma deveria ser aproximadamente esférica.

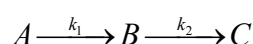
Quantificando sua análise (Francisco, 2010) Eratóstenes concluiu que a circunferência da Terra seria de 39772km . O valor aceito atualmente está entre 40030- 40075km muito próximo do que foi previsto por Eratóstenes (Sagan, 1980). A resolução desse problema inverso por Eratóstenes (*causa*-curvatura na superfície da Terra, *efeito*- diferentes comprimentos de sombras em objetos fixados em sua superfície em mesmas longitudes e diferentes latitudes) representou um grande incentivo para a exploração de regiões até então desconhecidas do planeta. Afinal, o valor encontrado para o diâmetro da Terra era relativamente modesto. Até mesmo a circum-navegação da Terra começou a ser imaginada e acabou sendo realizada por Fernão de Magalhães mais de mil anos depois. O planejamento de navegação das viagens de Cristóvão Colombo que culminaram com a descoberta das Américas se baseou no valor encontrado por Eratóstenes (Sagan, 1980).

1.2. Problemas Inversos em Cinética Química

O formalismo matemático em cinética química surgiu com o trabalho do alemão Ludwig Wilhelmy em 1850 sobre a hidrólise catalisada de sacarose produzindo glicose e frutose. Wilhelmy propôs uma lei de velocidades quantitativa para esse processo e também sugeriu de forma pioneira que uma reação química seria decorrente da interação entre as partículas e não das afinidades químicas entre espécies como pensando até aquele momento (Wilhelmy, 1850). Em seguida, os britânicos Augustus Harcourt e William Esson mostraram que uma reação química podia ocorrer através de uma série de etapas consecutivas e desenvolveram as primeiras equações empíricas em que as concentrações dos produtos e o tempo da reação eram conectados. A consequência desses trabalhos foi uma evolução na quantificação de processos cinéticos que consolidou a descrição de mecanismos de reações e modelos cinéticos em termos de equações diferenciais ordinárias ou parciais (Arnault, 2007).

Atualmente existe uma série de ferramentas computacionais apropriadas para a resolução de problemas diretos em processos cinéticos (Borges, 2016); o problema direto nesse âmbito seria a determinação de concentrações e/ou dados espectroscópicos a partir dos parâmetros cinéticos e do modelo cinético diferencial que controla a velocidade da reação. Em contrapartida o interesse pelo problema inverso da cinética química com a decorrente busca por métodos numéricos eficientes para resolvê-lo também se intensificou no século XX.

Essa abordagem trata da questão *quais parâmetros e/ou modelos cinéticos originam os dados experimentais reais?* Nesse quadro, de acordo com a notação $K(f)=g$ introduzida anteriormente, g representaria os dados experimentais referentes às concentrações (ou as próprias concentrações) dos participantes da reação, f seriam parâmetros cinéticos tais como constantes de velocidade, constantes de formação ou constantes de difusão e K prefigura o modelo diferencial consistente com o mecanismo da reação. O desafio consiste em encontrar os parâmetros cinéticos f a partir dos dados experimentais g utilizando-se o modelo diferencial K ou identificar ambos, tanto o modelo K e os parâmetros f , a partir dos dados experimentais (Lemes et al, 2009). Como exemplo, sejam reações cinéticas consecutivas de primeira ordem, compatíveis com o mecanismo



O modelo K é um sistema de equações diferenciais ordinárias do tipo:

$$\begin{aligned} \frac{dA}{dt} &= -k_1[A] \\ \frac{dB}{dt} &= k_1[A] - k_2[B] \\ \frac{dC}{dt} &= k_2[C] \end{aligned} \quad (1)$$

em que k_1 e k_2 são as constantes de velocidade para as etapas consecutivas do mecanismo; logo, f seria um vetor do tipo $f=(k_1, k_2)$. As concentrações das espécies, $[A]$, $[B]$ e $[C]$ são os dados experimentais, mas apenas $[C]$ é necessário para definir g que constitui assim um vetor com as concentrações do produto ao longo do tempo t . Esse problema inverso é mal colocado porque incertezas nas concentrações experimentais amplificam enormemente os erros na obtenção das constantes cinéticas devido à alta sensibilidade numérica das equações diferenciais a esses ruídos. Uma consequência disso é que passam a existir mais de um conjunto de constantes *aparentemente compatível* com os dados experimentais mas que não reproduzem bem tais concentrações. Em outras palavras não se encontra uma única e satisfatória solução. Esta dificuldade acarretou o desenvolvimento de técnicas matemáticas refinadas para a abordagem do problema inverso dentre as quais se destacam o método do gradiente conjugado, regularizações e redes neurais artificiais.

É interessante ressaltar que o mecanismo descrito pelo modelo representado pela Equação 1 pode caracterizar processos de diversos tipos e interesses, desde simples reações de hidrólise a reações biológicas ou nucleares, o que ilustra bem a importância e alcance dessa área. A Figura 1 ilustra claramente a ideia do problema inverso do tipo 1 em Cinética Química.

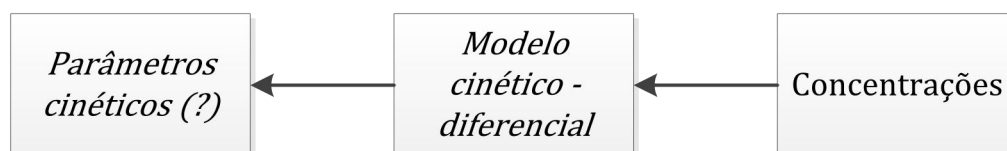


Figura 1: Problema inverso em cinética química.

2.METODOLOGIA

2.1 Redes Neurais Artificiais

Nos últimos anos uma metodologia computacional fundamentada em redes neuronais artificiais (RNA's) tem sido utilizada para a resolução de problemas inversos mal colocados em várias áreas. Esses algoritmos numéricos emulam o processo fisiológico de funcionamento de neurônios biológicos sendo adaptados para a resolução de problemas de otimização e controle. Os neurônios foram identificados como fibras nervosas organizadas em estruturas que passaram a ser denominadas redes neuronais (Cajal, 1909; Díaz, 1997).

Nos modelos que existem para a descrição do funcionamento de uma rede neuronal assume-se que determinado neurônio recebe sinais elétricos de seus neurônios vizinhos, os processa e gera novos sinais transmitindo-os através da rede. Os neurônios possuem um mecanismo de atribuição de pesos cujo efeito é a determinação da importância relativa dos sinais recebidos. Apenas os sinais julgados mais relevantes são propagados com amplitude significativa pela rede. Esse mecanismo constitui a aprendizagem neuronal e em um sentido matemático é considerado não linear, ou seja, as entradas dos sinais não variam linearmente com as saídas dos mesmos pois suas amplitudes são modificadas durante o processamento.

A causa biológica para essa não-linearidade é a taxa de variação da concentração de íons K^+ existentes dentro da célula neuronal em relação à concentração de íons Na^+ presentes fora da célula. Essa variação de concentração iônica origina uma diferença de potencial entre o interior e o exterior da membrana celular neuronal que acarreta um fluxo não-linear de corrente elétrica (Feng , 2004). Portanto, a maneira como a informação elétrica é transmitida é função da diferença de potencial entre os diferentes neurônios da rede. Um neurônio possui quatro estruturas principais, que são os dendritos, o corpo celular, os axônios e as sinapses. Uma ilustração dessas estruturas é mostrada na Figura 2.

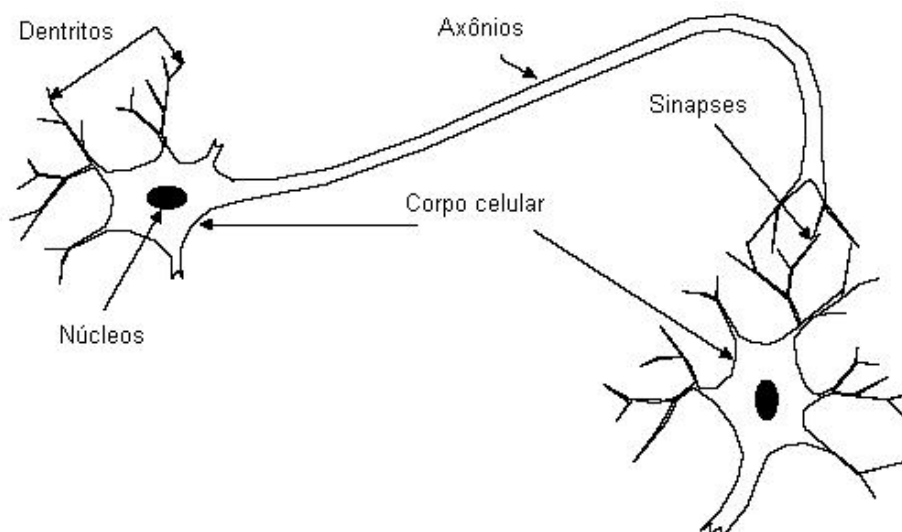


Figura 2. Componentes de um neurônio biológico.

Os dendritos são os receptores que captam sinais dos neurônios vizinhos; esses sinais são processados dentro de uma estrutura celular que a transmite através de uma extensa fibra, o axônio. No final da extremidade do axônio estão unidades inibitórias conhecidas como sinapses que controlam o curso de corrente elétrica na rede. As sinapses estabelecem pontos de contatos entre axônios e dendritos possibilitando as interconexões entre neurônios. Durante o desenvolvimento das pesquisas iniciais em neurociência a compreensão sobre o funcionamento de neurônios individuais ficou bem estabelecida.

A descrição de como ocorre o funcionamento conjunto dos neurônios que possibilita um nível de funcionalidades altamente complexas culminando com a consciência e a cognição era e ainda é um desafio. Warren S. McCulloch e Walter H. Pitts relataram em 1943 o primeiro modelo computacional para uma rede neuronal (McCulloch et al, 1943). Nesse modelo, caso a soma ponderada dos sinais de entrada de um neurônio ultrapassasse certo limite de corte conhecido como *threshold* então assumia-se valor unitário para a saída; se o valor de *threshold* não fosse ultrapassado então assumia-se valor zero para a saída.

Desenvolveu-se assim um procedimento de aprendizagem em que os pesos para os neurônios eram encontrados computacionalmente o que habilitava a RNA a realizar cálculos numéricos envolvendo funções matemáticas específicas. Demonstrou-se assim que um sistema constituído por neurônios artificiais organizados poderiam realizar cálculos úteis cientificamente desde que valores de pesos apropriados fossem atribuídos aos neurônios. Em 1957 um processo de aprendizagem iterativo para uma rede formada por uma camada única de neurônios conhecida como *single layer perceptron* foi desenvolvido por Franck Rosenblat em um laboratório da marinha americana em Cornell. O processo de aprendizagem desse algoritmo sempre convergia para um conjunto de pesos que possibilitava calcular a função desejada (Rosenblat, 1958).

Em uma conferência de 1958 Rosenblat fez declarações que levaram o *New York Times* a se referir a um *embrião de um computador que a Marinha espera ser capaz de caminhar, falar, ver, escrever, reproduzir a si mesmo e desenvolver autoconsciência* (Olazaran, 1996). Mas, em 1969, Marvin Minsky e Seymour Papert demonstraram a limitação da rede neuronal *single layer perceptron* a um pequeno número de funções e expressaram pessimismo quanto à eficiência mesmo de redes neuronais mais elaboradas, com múltiplas camadas, *multi layer perceptrons* (Minsky et al, 1969).

Em 1972, contudo, Stephen Grossberg publicou uma série de artigos introduzindo redes capazes de resolver vários tipos de funções matemáticas (Grossberg, 1972a, 1972b). Depois disso, apenas nos anos 80 o interesse pela inteligência artificial foi restabelecido. Um trabalho central nessa retomada foi o de John Hopfield em 1982 que sugeriu que uma rede neuronal com múltiplas camadas poderia ser descrita em termos de uma função energia construída a partir dos neurônios de tal maneira que possibilitaria qualquer tipo de cálculo desejado, envolvendo qualquer função, desde que um treinamento adequado para os neurônios fosse realizado (Hopfield, 1982). Esse treinamento foi eficientemente desenvolvido em 1986 em um método conhecido como *backpropagation* e a partir de então a inteligência artificial tornou-se um ativo campo de pesquisa com aplicações em diversas áreas da ciência e tecnologia.

De uma forma simplificada, as entradas dos neurônios artificiais correspondem aos sinais elétricos nos neurônios biológicos e os diferentes pesos para as entradas são determinados nos dendritos. A contribuição de cada entrada é estabelecida no corpo celular enquanto a rede onde ocorre a propagação da informação equivale aos axônios. Por fim, as

saídas de informações já moduladas por uma função não-linear na rede neuronal artificial representam as sinapses biológicas.

2.2 Como uma RNA pode resolver um problema inverso?

Uma RNA do tipo Hopfield é construída por uma camada de neurônios u interconectada. A saída de um neurônio u é função do sinal de entrada do mesmo o qual é convertido em um estado neuronal ativado por uma função $f(u)$. A função de ativação propaga a informação por um caminho descendente de energia e essa propagação é o processo de aprendizagem da rede neuronal artificial. A rede pode ser utilizada para a resolução de problemas inversos se uma função energia apropriada for definida (Borges, 2016). Por exemplo, pode-se construir a função energia como uma função erro específica para cada problema em análise. No âmbito de problemas inversos em que a obtenção de uma propriedade experimental é o objetivo, podem-se representar as propriedades experimentais e invertidas por P^{EXP} e P^{INV} definindo a função erro por

$$E = \frac{1}{2} \sum_{j=1}^m e_j^2 \quad (2)$$

em que $e_j = P_j^{INV} - P_j^{EXP}$ para m dados experimentais.

A Equação 2 será a energia da rede neuronal e a derivada dessa energia com relação ao tempo de aprendizagem da rede, τ , dará:

$$\frac{dE}{d\tau} = \sum_{i=1}^n \sum_{j=1}^m \left(e_j \frac{\partial P_j^{CAL}}{\partial f_i} \frac{\partial f_i}{\partial u_i} \frac{\partial u_i}{\partial \tau} \right) \quad (3)$$

em que n é a quantidade de neurônios envolvidos no processo. Esse número também é igual à quantidade de variáveis a serem invertidas. Enquanto o cérebro humano ativa bilhões de neurônios em um simples pensamento, as RNA's raramente ultrapassam mil neurônios na resolução mesmo de problemas físico-matemáticos complexos. A função erro irá sempre decrescer com o tempo de aprendizagem da rede neuronal pois duas condições são estabelecidas na rede de Hopfield; primeiramente, a condição:

$$\frac{du_i}{d\tau} = - \sum_{j=1}^m \frac{\partial P_j^{CAL}}{\partial f_i} e_j \quad (4)$$

é imposta, transformando (3) em

$$\frac{dE}{d\tau} = - \sum_{i=1}^n \frac{\partial f_i}{\partial u_i} \left(\frac{\partial u_i}{\partial \tau} \right)^2.$$

Em seguida, a relação:

$$\frac{\partial f_i}{\partial u_i} \geq 0$$

é aplicada, o que implica em,

$$\frac{dE}{d\tau} < 0 \quad (5)$$

Então a RNA de Hopfield tem a propriedade de sempre abaixar a função erro com o tempo de aprendizagem da rede. Enquanto os neurônios evoluem com o tempo de aprendizagem a função erro diminui até que um mínimo global seja obtido. As Equações 3 e 4 são integradas numericamente até que o tempo de aprendizagem, o qual define o critério de finalização da integração, seja alcançado. Nesse momento a função erro atinge seu valor mínimo. A inversão de parâmetros a partir de dados experimentais, isto é, a solução de um problema inverso, é obtida de acordo com as seguintes etapas:

i) A resolução do problema direto é requerida para se obter P_j^{EXP} . As equações que descrevem o problema direto podem ser do tipo integral, diferencial ou matricial.

ii) Os dados referentes à resolução do problema direto que podem ser pontos de discretização no caso integral, pontos de integração no caso diferencial e auto-valores no caso matricial, correspondem aos neurônios que são utilizados para a construção das equações (2)-(5).

iii) A taxa de variação dos neurônios em relação ao tempo de aprendizagem irá alcançar um valor constante em que as derivadas $\frac{du_i}{d\tau}$ se anulam pois a função erro sempre cai. A função erro atingirá seu menor valor quando essas derivadas finalmente forem nulas. Nesse momento os dados experimentais serão reproduzidos pelas propriedades invertidas.

iv) Quando iii) ocorrer o processo de aprendizagem da rede estará concluído, a função erro atingirá seu menor valor e $P_j^{INV} = P_j^{EXP}$.

Na rede neuronal como a função erro sempre diminui é possível obter dentre as inúmeras soluções para o problema inverso mal-colocado a que melhor reproduz os resultados experimentais. Essa solução é dita *regularizada*; portanto, a rede neuronal *regulariza* o problema inverso mal-colocado, o que métodos mais simples de otimização não conseguem fazer. Mais informações sobre a fundamentação teórica da rede de Hopfield podem ser encontrada nas referências (Hopfield,1982; Stuart et al, 1995).

3.RESULTADOS E DISCUSSÕES

Sistema 1

A rede neuronal artificial de Hopfield foi aplicada na resolução do problema inverso da cinética química em um sistema protótipo correspondente a uma reação consecutiva do tipo $A \xrightarrow{k_1} B \xrightarrow{k_2} C$ descrita pelo modelo diferencial dado pela Equação 1 como discutido anteriormente. Esse mecanismo cinético descreve vários tipos de reações químicas como a hidrólise de hidrocarbonetos, além de algumas reações de decaimento nuclear (Lemes et al, 2007). De acordo com a notação $K(f) = g$, K corresponde às equações diferenciais dadas na Equação 1 enquanto f e g são respectivamente o conjunto com as constantes de velocidade e o vetor com as concentrações de A, B ou C.

Na Equação 2, P_j^{EXP} será no presente exemplo, um vetor com concentrações experimentais do produto C e P_j^{INV} denotará o conjunto de concentrações teóricas obtidas por resolução da Equação 1 a partir de valores de k_1 e k_2 que serão determinados pela RNA. Portanto, as constantes de velocidade equivalem aos neurônios. Para inicializar a integração das Equações 3 e 4 são utilizadas condições iniciais para os neurônios, ou seja, supõem-se valores numéricos iniciais para as constantes de velocidade; esse procedimento é frequentemente guiado por alguma referência química, i.e, alguma reação já conhecida da qual se conheçam as constantes de velocidade. O número de camadas intermediárias da rede será função do tempo de aprendizagem. Assim para cada unidade desse tempo haverá uma camada intermediária.

A Equação 1 possui solução analítica (Arnault, 2007) dada por:

$$\begin{aligned} [A] &= [A]_0 e^{-k_1 t} \\ [B] &= [A]_0 \frac{k_1 (e^{-k_1 t} - e^{-k_2 t})}{k_2 - k_1} \\ [C] &= [A]_0 \frac{k_2 (1 - e^{-k_1 t}) - k_1 (1 - e^{-k_2 t})}{k_2 - k_1} \end{aligned} \quad (6)$$

Para o procedimento computacional a solução analítica para $[C]$ foi utilizada para simular os dados experimentais, i.e, o vetor P_j^{EXP} . Foram gerados dados para $t=0$ até $t=20$, ou seja um vetor com concentrações com dimensões 1×20 . Esses dados foram obtidos a partir de $k_1 = 0,5 \text{ min}^{-1}$ e $k_2 = 0,1 \text{ min}^{-1}$ que foram assumidas como valores experimentais; esses valores são da mesma ordem de grandeza de constantes de reações típicas de hidrólise de hidrocarbonetos. O vetor P_j^{EXP} assim criado foi introduzido na Equação 1 para inicializar o funcionamento da rede. Para verificar a capacidade da rede em reproduzir os valores experimentais foram testadas diferentes condições iniciais para os neurônios. Para todas essas diferentes condições iniciais fornecidas à rede, os estados neuronais sempre

convergir para os parâmetros experimentais, i.e, $k_1 = 0,5 \text{ min}^{-1}$ e $k_2 = 0,1 \text{ min}^{-1}$. Para ilustrar a convergência dos neurônios a Figura 3 mostra duas diferentes condições iniciais para as constantes.

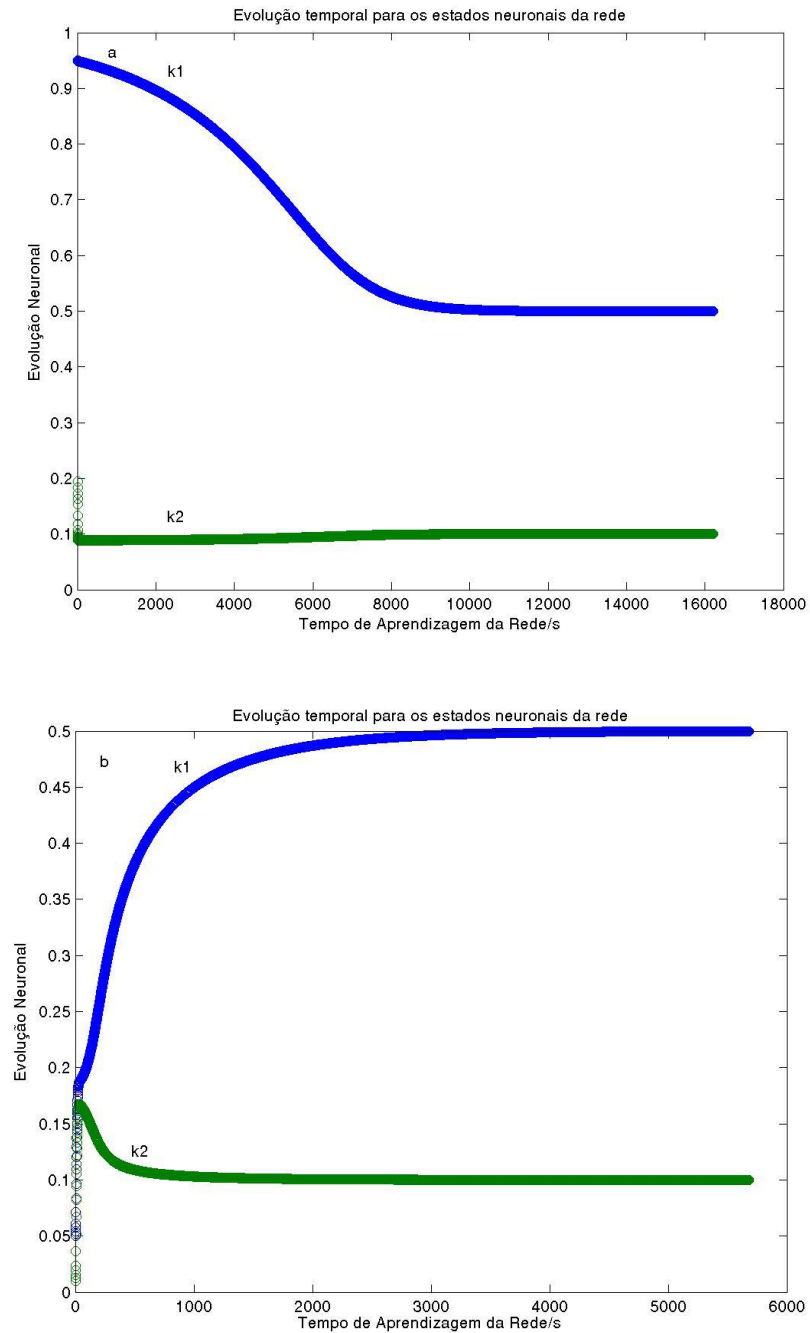


Figura 3: Evolução neuronal em diferentes condições iniciais para as constantes k_1 e k_2

Na figura 3a tem-se uma condição inicial com erro de +90% no valor de k_1 e k_2 enquanto em b tem-se uma condição inicial com erro de -90%. Pode-se visualizar a convergência para os valores experimentais ($k_1 = 0,5 \text{ min}^{-1}$ e $k_2 = 0,1 \text{ min}^{-1}$) nos dois casos.

É importante enfatizar que o presente método não requer soluções analíticas tais como aquelas da Equação 6 uma vez que soluções numéricas para a Equação 1 podem ser acopladas às Equações 2-4 para resolver o problema inverso. Dessa maneira a presente metodologia pode ser utilizada para qualquer problema cinético independentemente do modelo diferencial possuir solução analítica ou não. Para testar a estabilidade da RNA com respeito a erros experimentais, ruídos randômicos no intervalo 2-8% foram adicionados às concentrações experimentais (vetor P_j^{EXP}). Os ruídos foram controlados através de uma distribuição normal no intervalo $[-1,1]$, ou seja, foram adicionados randomicamente reproduzindo uma situação experimental real. Mesmo em um nível de ruído experimental de 8% o algoritmo mostrou-se estável apresentando um valor para a função erro, Equação 1, de 10^{-7} ; isso significa que as concentrações simuladas estão muito próximas das experimentais.

Uma comparação entre a eficiência do método da RNA e outros usualmente empregados em problemas de otimização envolvendo regressão não-linear pode ser realizada nesse ponto. Por exemplo, os algoritmos Simplex e Levenberg-Marquardt (Press et al, 1993) são comumente encontrados em softwares comerciais fechados e amplamente aplicados à minimização de funcionais como o da Equação 2. Se ruídos experimentais não são considerados, as técnicas Simplex e Levenberg-Marquardt (LM) têm eficiência similar fornecendo valores para a função erro muito próximos: Simplex, $E=10^{-28}$, Levenberg-Marquardt, $E=10^{-28}$. Mesmo nesse caso, a rede neuronal artificial é mais eficiente pois sua função erro é aproximadamente 1000 vezes menor, $E=10^{-31}$. Se ruídos experimentais forem considerados o desempenho da RNA continua melhor quando comparado com os métodos Simplex e LM como mostra a Tabela 1. No maior nível de ruído, 8% , a rede neuronal artificial consegue um valor para a função erro ainda 100 vezes menor do que os outros dois métodos.

Tabela 1: Robustez da RNA e dos algoritmos Simplex e LM. Sistema A-B-C

Ruídos em [C]/%	Função Erro Simplex	Função Erro LM	Função Erro RNA
0	10^{-28}	10^{-28}	10^{-31}
± 2	10^{-7}	10^{-7}	10^{-9}
± 4	10^{-6}	10^{-6}	10^{-8}
± 6	10^{-5}	10^{-5}	10^{-7}
± 8	10^{-5}	10^{-5}	10^{-7}

Sistema 2

O método RNA também foi usado para resolver um problema inverso em cinética química controlada pelo mecanismo de reações reversíveis consecutivas,



Novamente, A , B e C são, respectivamente, o reagente, o intermediário e o produto da reação e k_1, k_{-1}, k_2, k_{-2} são as constantes de velocidade da reação. O modelo diferencial agora será

$$\begin{aligned} \frac{dA}{dt} &= -k_1[A] + k_{-1}[B] \\ \frac{dB}{dt} &= k_1[A] - (k_{-1} + k_2)[B] + k_{-2}[C] \\ \frac{dC}{dt} &= k_1[B] - k_{-2}[C] \end{aligned} \quad (8)$$

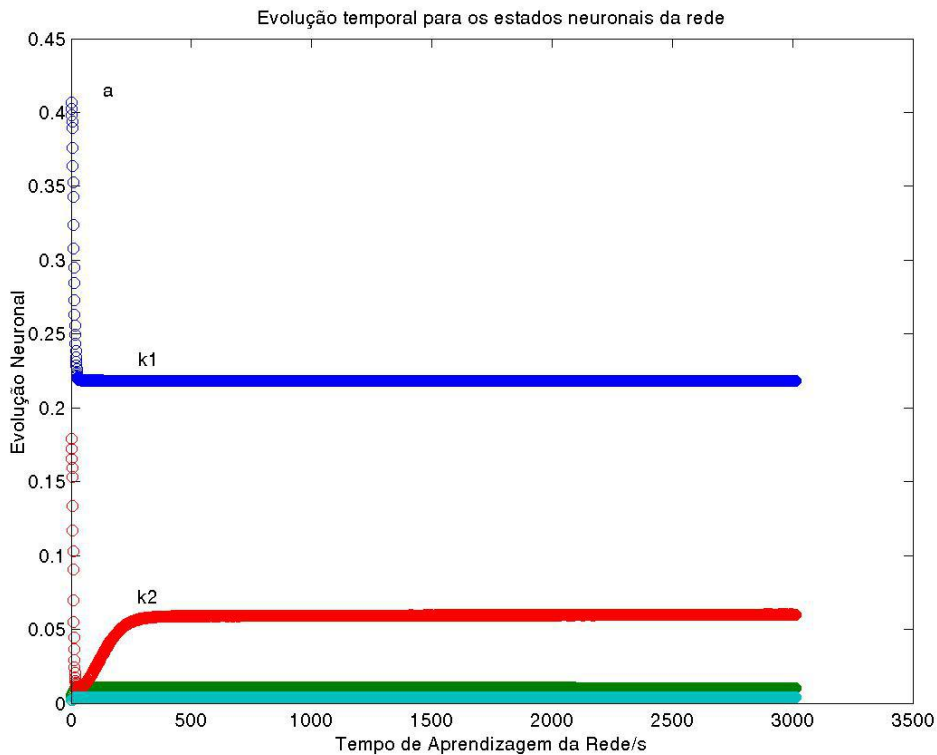
Um notável exemplo de aplicação para o mecanismo de reação em (7) é a cinética de inibição característica de algumas enzimas. Por exemplo, já foi relatada anteriormente (Borges et al, 2012; Menezes et al, 2012; Nazari et al, 2006) a maneira como o mecanismo de inibição de Urease por diferentes inibidores obedece às Equações 7 e 8. Nesse tipo de reação monitora-se a concentração do reagente A , ou seja, da enzima. Nessa perspectiva, a Equação 1, P_j^{EXP} corresponderá agora ao vetor com concentrações experimentais do reagente A e P_j^{INV} denotará o conjunto de concentrações teóricas obtidas por resolução de (7) a partir de valores de k_1, k_{-1}, k_2, k_{-2} que serão calculados pela RNA. As Equações 8 também possuem solução analítica:

$$\begin{aligned} [A] &= [A]_0 \left[\frac{k_{-1}k_{-2}}{ab} + \frac{k_{-1}k_{-2} - a(m-a)}{a(a-b)} e^{-at} - \frac{k_{-1}k_{-2} - b(m-a)}{b(a-b)} e^{-bt} \right] \\ [B] &= k_1[A]_0 \left[\frac{k_{-2}}{ab} + \frac{k_{-2} - a}{a(a-b)} e^{-at} - \frac{k_{-2} - b}{b(a-b)} e^{-bt} \right] \\ [C] &= k_1k_2[A]_0 \left[\frac{1}{ab} + \frac{1}{a(a-b)} e^{-at} - \frac{1}{b(a-b)} e^{-bt} \right] \\ m &= k_{-1} + k_2 + k_{-2} \\ a + b &= k_1 + k_{-1} + k_2 + k_{-2} \\ ab &= k_1k_2 + k_1k_{-2} + k_{-1}k_{-2} \end{aligned} \quad (9)$$

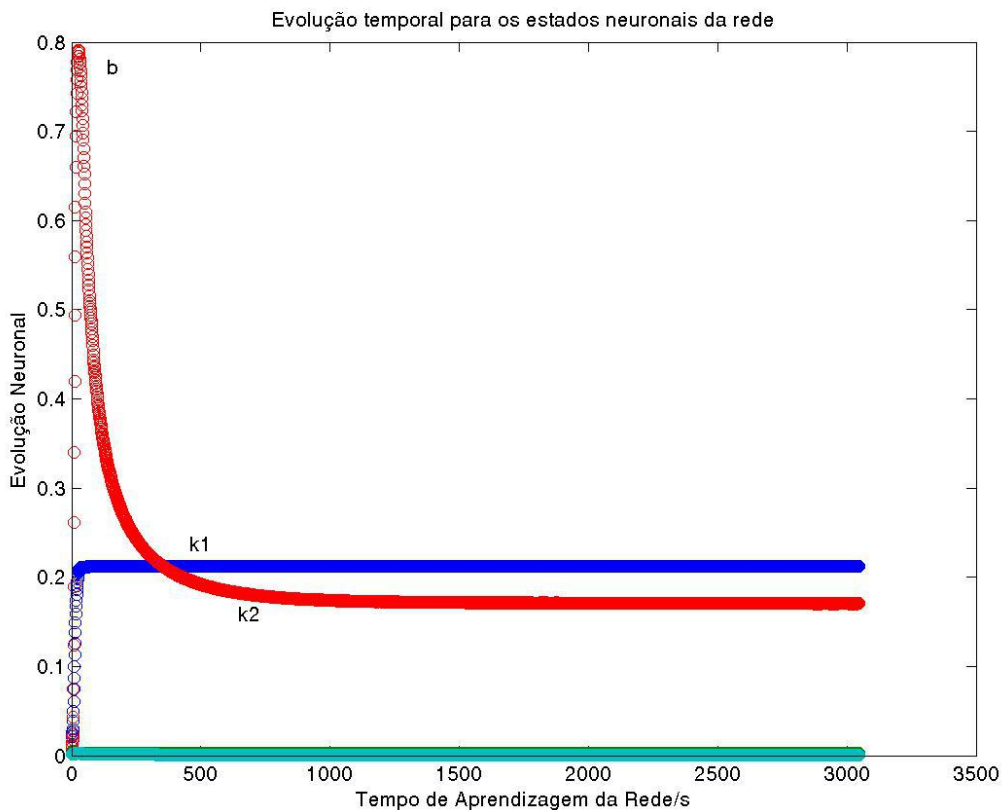
Novamente foram simulados dados experimentais para $t=0$ até $t=20$, a partir das constantes retiradas da referência (Nazari et al, 2006) correspondentes aos parâmetros para a inibição da enzima Jack-Bean Urease por materiais surfactantes.

$$\begin{aligned}
 k_1 &= 0,214\text{min}^{-1} \\
 k_{-1} &= 5,173 \times 10^{-3} \text{min}^{-1} \\
 k_2 &= 9,432 \times 10^{-2} \text{min}^{-1} \\
 k_{-2} &= 2,079 \times 10^{-3} \text{min}^{-1}
 \end{aligned}
 \tag{10}$$

O vetor P_j^{EXP} assim gerado foi introduzido na Equação 2 para inicializar o processamento da rede. A robustez da rede para recuperar os valores das constantes dadas na Equação 10 foi novamente posta em prova e sempre se verificou convergência. Um exemplo disso é mostrado na Figura 4 que ilustra a convergência de condições iniciais com erros de +90% (Figura 4 a) e -90% (Figura 4 b) para as constantes k_1 e k_2 . As constantes k_{-1} e k_{-2} foram mantidas fixas por todo o tempo de aprendizagem e também aparecem nos gráficos como parâmetros de resposta inalterada.



(a)



(b)

Figura 4: Evolução neuronal em diferentes condições iniciais para as constantes k_1 e k_2 .

As constantes k_1 e k_2 não evoluíram e aparecem como linhas retas.

Novamente a estabilidade da RNA com respeito a erros experimentais foi comparada com aquela das técnicas Simplex e LM em níveis de ruídos randômicos nas concentrações experimentais variando entre 2-8%. Os ruídos foram controlados através de uma distribuição normal como anteriormente. Os resultados estão apresentados na Tabela 2. Mesmo em um nível de ruído experimental de 8% a RNA mostrou-se estável, apresentando um valor para a função erro, Equação 2, de 10^{-10} . Se ruídos experimentais não são considerados, as técnicas Simplex e Levenberg-Marquardt têm eficiência similar fornecendo para a função erro: Simplex, $E=10^{-28}$, Levenberg-Marquardt, $E=10^{-28}$.

Como no exemplo anterior a rede neuronal artificial é mais eficiente apresentando função erro 100 vezes menor, $E=10^{-31}$. Se ruídos experimentais forem considerados o desempenho da RNA é ainda melhor quando comparado com os métodos Simplex e Levenberg-Marquardt. No maior nível de ruído, 8%, a rede neuronal artificial obtém um valor para a função erro cerca de 10000 vezes menor do que os outros métodos.

Tabela 2: Robustez da RNA e dos algoritmos Simplex e LM. Sistema A=B=C

Ruídos em [A]/%	Função erro Simplex	Função erro LM	Função erro RNA
0	10^{-28}	10^{-28}	10^{-34}
± 2	10^{-6}	10^{-6}	10^{-11}
± 4	10^{-6}	10^{-6}	10^{-10}
± 6	10^{-6}	10^{-6}	10^{-10}
± 8	10^{-5}	10^{-5}	10^{-10}

4. CONCLUSÕES

Nesse artigo são discutidos os princípios e características dos chamados problemas inversos mal colocados cujas aplicações abrangem inúmeras áreas da ciência e tecnologia. Especificamente, uma abordagem computacional baseada em redes neuronais artificiais foi apresentada como eficiente técnica para resolver problemas inversos em cinética química. Dois sistemas protótipos foram utilizados para ilustrar a aplicabilidade da rede neuronal para esse tipo de problema. Primeiramente, parâmetros cinéticos para um processo descrito por uma cinética consecutiva do tipo $A \xrightarrow{k_1} B \xrightarrow{k_2} C$, representativo de várias reações de interesse químico, foram invertidos a partir das concentrações do produto. Em uma segunda aplicação a metodologia numérica foi aplicada ao processo $A \xrightleftharpoons[k_{-1}]{k_1} B \xrightleftharpoons[k_{-2}]{k_2} C$ que descreve por exemplo inúmeros mecanismos de inibição enzimática. Novamente as constantes cinéticas foram eficientemente obtidas a partir de dados de concentração ao longo do tempo, para o reagente nesse caso.

A rede neuronal discutida aqui é geral e pode ser utilizada para a análise de qualquer tipo de sistema cinético passível de modelagem por um conjunto de equações diferenciais. Durante a propagação dos estados neuronais ocorre sempre a diminuição da função erro do problema a qual é definida pela diferença entre as normas euclidianas dos vetores com os dados experimentais e aqueles obtidos pela rede. Dessa maneira os parâmetros cinéticos obtidos reproduzem eficientemente dados experimentais reais. O método é robusto em relação às diferentes condições iniciais para os parâmetros procurados e em relação aos ruídos experimentais nas concentrações tendo desempenho superior ao de outros procedimentos usualmente encontrados em softwares comerciais, como o método de Levenberg-Maquardt e o algoritmo Simplex. Esse é um exemplo bem atual de como a área da inteligência artificial pode ser útil para as ciências naturais.

REFERÊNCIAS

- ARNAUT, L.; FORMOSINHO, S.; BURROWS, H.; **Chemical Kinetics: From Molecular Structure to Chemical Reactivity**, Elsevier, Amsterdam, 2007.
- BORGES, E. **Problemas Diretos e Inversos em Físico-Química: Aplicações em Dinâmica Molecular, Espectroscopia Vibracional, Cinética Química e Espalhamento Quântico**. 1.ED. Novas Edições Acadêmicas, OmniScriptum GmbH & Co. KG, Saarbrücken, Alemanha, 2016.
- BORGES, E.; MENEZES, D.C.; BRAGA, J.P.; An Ill-posed inverse problem in enzymatic kinetics: Jack-Bean urease denaturation by an anionic surfactant, **Int. J. Quantum Chem.**, v. 112, p. 3240–3245, 2012.
- BORGES, E.; LEMES, N.H.T.; SOUSA, R.V.; BRAGA, J.P.; Potential energy function from differential cross-section data: An inverse quantum scattering theory approach. **Int. J. Quantum. Chem.**, v. 108, p. 2623–2627, 2008.
- CAJAL, S.R.; **Histologie du système nerveux de l'homme et des vertébrés**, Maloine, Paris, 1909.
- DÍAZ, B.; ALFREDO, LUIS.; La obra neuro-embriológica de Santiago Ramón y Cajal. **Dynamis: Acta Hispanica ad Medicinae Scientiarumque. Historiam Illustrandam** Universidad de Granada, v. 17 p. 259-279, 1997.
- FENG, J (editor).; **Computational Neuroscience A Comprehensive Approach**, Chapman & Hall/Crc, Washington, D.C., 2004.
- FRANCISCO, S. P.; Medir el radio de la Tierra. Una experiencia educativa. **Revista de investigación en educación**, v. 1697, n. 8, p. 140 -146, 2010.
- GROSSBERG, S.; A neural theory of punishment and avoidance. **Mathematical Biosciences**, v. 15, p. 39–67, 1972a e Neural expectation: Cerebellar and retinal analogs of cells of fired by learnable or unlearned pattern classes. **Kybernetik**, v. 10, p. 49–57, 1972b.
- GROETSCH, C.W.; **Inverse problems: Activities for undergraduates**. Department Of Mathematical Sciences, University Of Cincinnati, The Mathematical Association Of America, 1999.
- HADAMARD, J.; **Le probleme de gauchy et les equations aux derivees partielles lineaires hyperboliques**, Hermann, Paris, 1932.
- HOPFIELD, J.J.; TANK, D.W.; Neural computation of decisions in optimization problems, **Biol. Cybern.**, v. 52, p. 141–152, 1982.

- KIRSCH, A.; **An Introduction to the Mathematical Theory of Inverse Problems**, Springer, Berlin, 1996.
- LEMES, N.H.T.; BORGES, E.; BRAGA, J.P.; A general algorithm to solve linear and nonlinear inverse problems. **J. Braz. Chem. Soc.**, v. 18, p. 1342–1347, 2007
- LEMES, N.H.T.; BORGES, E.; BRAGA, J.P. ; Rate constants and absorption coefficients from experimental data: An inversion procedure based on recursive neural networks, **Chemom. Intell. Lab. Syst.** v. 96, p. 84-87, 2009.
- LUDWIG, W.; Ueber das Gesetz, nach welchem die Einwirkung der Säuren auf den Rohrzucker stattfindet, **Annalen der Physik und Chemie**, v. 81, p. 413–433, 1850.
- McCULLOCH, W.S.; PITTS, W.H.; Neural computation of decisions in optimization problems, **Bull. Mathem. Biophys.** v. 5, p. 115–133, 1943.
- MENEZES, D.C.; BORGES, E.; TORRES, M.F.; BRAGA, J.P.; A kinetic study of jack-bean urease denaturation by a new dithiocarbamate bismuth compound, **Chem. Phys. Lett.**, v. 548, p. 85–89, 2012.
- MINSKY, M.; PAPERT, S.A.; **Perceptrons: an introduction to computational geometry**, M.I.T. Press, Cambridge, 1969.
- NAZARI, K.; MAHMOUDI, A.; ESMAEILI, N.; SADEGHIAN, L.; MOOSAVI-MOVAHEDI, A.A.; KHODAFARIN, R.; Denaturation of jack-bean urease by sodium *n*-dodecyl sulphate: A kinetic study below the critical micelle concentration, **Colloids Surf. B** v. 53, p. 139–148, 2006.
- OLAZARAN, M.; A Sociological Study of the Official History of the Perceptrons Controversy, **Social Studies of Science**, v. 26, p. 611–659, 1996.
- PRESS, W.H.; FLANNERY, B.P.; TEUKOLSKY, S.A.; VETTERLING, W.T.; **Numerical Recipes in Fortran 77: The Art of Scientific Computing**. Cambridge University Press, New York, 1993.
- ROSENBLAT, F.; The perceptron: A probabilistic model for information storage and organization in the brain, **Psychological Rev.**, v. 65, p. 386-408, 1958.
- SAGAN, C. **Cosmos**. Random House, New York, 1980.
- STUART, J.; NORVIG, R.; NORVIG, P.; **Artificial Intelligence A Modern Approach**, Prentice Hall, New Jersey, 1995
- TARANTOLA, A.; **Inverse problem theory and methods for model parameter estimation**, SIAM, Philadelphia, 2004.
- WITTENMYER, R.A.; TUOMI, M.; BUTLER, R.P.; JONES, H.R.A.; ANGLADA, E.G.;

HORNER, J.; TINNEY, C.G.; MARSHALL, J.P.; CARTER, B.D.; BAILEY, J.; SALTER, G.S.; O'TOOLE, S.J.; WRIGHT, D.; CRANE, J.D.; SCHECTMAN, S.A.; ARRIAGADA, P.; THOMPSON, I.; MINNITI, D.; JENKINS, J.S.; DIAZ, M. GJ 832C: A Super-Earth in the Habitable Zone. **The Astrophysical Journal**, v. 791, n. 2, p. 1-32, 2014.

RESOLUTION OF INVERSE PROBLEMS IN CHEMICAL KINETICS BY USING ARTIFICIAL NEURONAL NETWORKS

E. BORGES¹, D.C. MENEZES¹,

¹ Universidade Federal de Viçosa, Departamento de Química

E-mail: emilio.borges@ufv.br

ABSTRACT: *Physical problems that can be mathematically modeled are classified as direct or inverse problems. Direct problems are well-established. In these cases measurable effects are determined from identified causes. Inverse problems are those in which causes are found from their effects. Inverse problems are so important and have as many scientific applications as the direct ones although their mathematical approach is usually more complex requiring special numerical techniques. For instance, an algorithm based on artificial neural networks has been recently used with efficacy to handle inverse problems in several areas. In this paper the inverse problem in Chemical Kinetics is discussed and the artificial neural network method is applied to solve it. Two kinetic mechanisms which are relevant prototypes for several chemical and biological reactions are treated. The efficiency of the neural network technique is compared with that of Levenberg-Maquardt and Simplex algorithms, numerical optimization techniques available in commercial software.*

KEYWORDS: Inverse Problems; Chemical Kinetics; Artificial Neural Networks.