

**UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA DE PRODUÇÃO**

Marina Fernandes Branco Pitanga Lopes

**Análise comparativa entre o desempenho dos métodos de regressão
não linear e redes neurais artificiais através do planejamento de
experimentos**

Itajubá, março de 2020

**UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS GRADUAÇÃO EM
ENGENHARIA DE PRODUÇÃO**

Marina Fernandes Branco Pitanga Lopes

Análise comparativa entre o desempenho dos métodos de regressão não linear e redes neurais artificiais através do planejamento de experimentos

Dissertação submetida ao programa de Pós-Graduação em Engenharia de Produção como parte dos requisitos para obtenção do Título de Mestre em Engenharia de Produção.

Área: Engenharia de Produção

Orientador: Prof. Dr. Pedro Paulo Balestrassi

Itajubá, março de 2020

Folha de aprovação

DEDICATÓRIA

Aos meus pais Maricy e Sérgio, aos meus queridos avós, Aarão e Marina, ao André e a todos os meus amigos e familiares.

AGRADECIMENTOS

Agradeço primeiramente a Deus. Aos meus pais, Maricy e Sérgio, por sempre estarem ao meu lado, acreditando nas minhas escolhas, sonhando junto comigo e me ensinando a ser uma pessoa melhor a cada dia.

Aos melhores avós do mundo, Aarão e Marina, por serem meu porto seguro em Minas e por todo o precioso tempo que passamos juntos durante esses anos de mestrado. Ao meu namorado André, por todo companheirismo, apoio e incentivo durante todas as etapas da minha pesquisa.

Aos meus amigos do mestrado Natália, Estevão e José Gíglío por todo trabalho em equipe e ajuda com críticas e sugestões, tanto no trabalho como em artigos.

Aos professores, em especial ao meu orientador e amigo Pedro Paulo Balestrassi, que desde o começo incentivou minha pesquisa e confiou no meu trabalho, dando todo o suporte e atenção para a criação da minha pesquisa, sempre disposto a tirar todas as minhas dúvidas e a sugerir melhorias para que esse trabalho fosse o melhor possível.

À UNIFEI e à CAPES pelo apoio financeiro e estrutural à pesquisa brasileira em engenharia de produção.

A todos vocês meu muito obrigada, vocês foram responsáveis pela conclusão de mais uma etapa incrível da minha vida.

RESUMO

A modelagem de dados é uma técnica que auxilia a tomada de decisão e a resolução de diversos tipos de problemas em sistemas reais. Entretanto, a escolha da melhor técnica para modelar um sistema real nem sempre é uma tarefa fácil, pois cada sistema possui suas características específicas e não é possível inferir que o melhor método para determinada situação será o melhor em outros contextos. Devido a essa necessidade de generalização, o presente trabalho tem como objetivo apresentar uma análise comparativa entre o desempenho dos métodos de regressão não linear (RNL) e redes neurais artificiais (RNAs) na modelagem de conjuntos de dados gerados através de um planejamento de experimentos (DOE), de modo a simular diferentes cenários nos quais os métodos serão aplicados. Dessa forma, os conjuntos de dados foram modelados por cada um dos métodos e suas respostas foram avaliadas através do coeficiente de determinação (R^2). Os resultados mostraram que os métodos estudados apresentam diferenças estatisticamente significativas, com um nível de significância de 5%. Além disso, concluiu-se nesta pesquisa que se os experimentadores conhecem a priori o modelo não linear que define a relação entre as variáveis de entrada e a variável resposta, a regressão não linear supera a rede neural artificial em termos de R^2 .

Palavras-chave: Modelagem de dados, regressão não linear, redes neurais artificiais, planejamento de experimentos.

ABSTRACT

Data modeling is a technique that assists decision making and the resolution of several types of problems in real systems. However, choosing the best technique for modeling a real system is not always an easy task, as each system has its specific characteristics and it is not possible to infer that the best method for a given situation will be the best in other contexts. Due to this need for generalization, the present work aims to present a comparative analysis between the performance of nonlinear regression (NLR) and artificial neural networks (ANNs) methods in modeling data sets generated through a design of experiments (DOE), in order to simulate different scenarios in which the methods will be applied. Thus, the data sets were modeled by each of the methods and their responses were assessed using the coefficient of determination (R^2). The results showed that the studied methods present statistically significant differences, with a significance level of 5%. In addition, in this research it was concluded that if the experimenters know a priori the nonlinear model that defines the relationship between the input variables and the response variable, nonlinear regression outperforms artificial neural network in terms of R^2 .

Key words: Data modeling, nonlinear regression, artificial neural networks, design of experiments.

LISTA DE FIGURAS

Figura 1 - Publicações (a) e citações (b) na <i>Web of Science</i>	14
Figura 2 - Modelo não linear de um neurônio	22
Figura 3 - Arquitetura de rede <i>feedforward</i>	24
Figura 4 - Arquitetura de rede <i>feedback</i>	24
Figura 5 - Arquitetura de uma rede MLP com duas camadas ocultas	25
Figura 6 - Modelo experimental de um sistema genérico	31
Figura 7 - Modelo experimental de um sistema de simulação	32
Figura 8 - Classificação da pesquisa	33
Figura 9 - Curva do <i>Power and Sample Size for paired t</i>	44
Figura 10 - ANOVA RNL	46
Figura 11 - ANOVA RNA.....	46
Figura 12 - Gráfico de Pareto para R ² RNL	48
Figura 13 - Gráfico e Pareto para R ² RNA	49
Figura 14 - Gráfico de Efeitos principais para R ² RNL.....	50
Figura 15 - Gráfico de Efeitos principais para R ² RNA	50
Figura 16 - Gráfico de Iterações (AB) para R ² RNA.....	51
Figura 17 - Gráfico de Iterações (CD) para R ² RNA.....	51

LISTA DE TABELAS

Tabela 1 - Resumo das pesquisas comparativas anteriores	29
Tabela 2 - Níveis dos fatores analisados no DOE	37
Tabela 3 - Arranjo Experimental	41
Tabela 4 - Respostas do DOE.....	43
Tabela 5 - Coeficientes estimados	45
Tabela 6 - Análise de resíduos.....	47
Tabela 7 - Respostas do ajuste do modelo de regressão.....	47

SUMÁRIO

1. INTRODUÇÃO.....	12
1.1 Contexto.....	12
1.2 Relevância.....	13
1.3 Justificativa	14
1.4 Objetivo geral	16
1.4.1 Objetivos específicos.....	16
1.5 Estrutura do trabalho.....	16
2. FUNDAMENTAÇÃO TEÓRICA	17
2.2 Regressão não linear	17
2.2.1 Métodos de estimação	18
2.2.2 Método dos mínimos quadrados	18
2.2.3 Métodos Iterativos.....	19
2.3 Redes Neurais artificiais	20
2.3.1 Componentes de uma rede	21
2.3.2 Arquiteturas de uma rede	23
2.3.3 Redes neurais <i>perceptrons</i> de múltiplas camadas	25
2.3.4 Processos de aprendizagem.....	26
2.3.5 O algoritmo de retropropagação do erro (<i>error back-propagation</i>)	26
2.4 Pesquisas comparativas anteriores (RNL x RNA).....	28
2.5 Planejamento de experimentos (DOE)	30
2.5.1 DOE para simulação.....	32
3. METODOLOGIA DE PESQUISA	33
3.1 Classificação da pesquisa.....	33
3.1.1 Experimentação	34
3.1.2 Modelagem e simulação.....	35
3.2 Procedimento	36
4. APLICAÇÃO	37
4.2.1 Conceitualização e definição do problema	37
4.2.2 Escolha dos fatores e níveis de trabalho	37
4.2.3 Escolha do projeto experimental.....	39
4.2.4 Modelagem	41

4.2.5 Solução dos modelos	41
4.2.6 Análise estatística dos resultados.....	42
4.2.7 Interpretação dos resultados.....	48
5. CONCLUSÕES E RECOMENDAÇÕES	52
REFERÊNCIAS	54
APENDICE A – Banco de dados 2	58
APÊNDICE B – Resolução pelo <i>software</i> Minitab.....	63
APÊNDICE C – Resolução pelo <i>software</i> Statistica - <i>Intelligent Problem Solver</i> (IPS)	66

1. INTRODUÇÃO

1.1 Contexto

A modelagem de dados é uma técnica importante que auxilia a tomada de decisão e a resolução de problemas em sistemas reais. De acordo com Xu, Wong e Chin (2013), as técnicas de previsão e modelagem de dados podem ser classificadas como lineares ou não lineares. Os métodos lineares podem ser classificados como métodos menos exigentes em termos de técnicas, enquanto os métodos não lineares são capazes de modelar o relacionamento entre as variáveis de entrada e saída de qualquer sistema complexo e dinâmico.

Existem inúmeros métodos não lineares disponíveis na literatura que diferem entre si devido aos seus conceitos, suposições subjacentes e procedimentos de modelagem (KHATIBI *et al.*, 2014). Entretanto, a escolha da melhor técnica de modelagem para um sistema real nem sempre é uma tarefa fácil, pois geralmente esses sistemas são complexos e possuem a influência de um conjunto de fatores externos e internos.

A regressão não linear (RNL) é um dos métodos estatísticos clássicos mais utilizados na modelagem de dados não lineares. A popularidade desse método é resultado de diversos fatores, como por exemplo: sua metodologia estabelecida, histórico de aplicação, disponibilidade de *softwares* e aceitação no meio acadêmico (RAZI; ATHAPPILY, 2005).

De acordo com Yagiz, Sezer e Gokceoglu (2012), o método de regressão não linear busca encontrar um relacionamento não linear entre a variável dependente (ou variável resposta) e o conjunto de variáveis independentes (ou variáveis de entrada) de um determinado sistema. Portanto, a aplicação desse método exige que algumas suposições sobre as relações funcionais entre essas variáveis sejam feitas. Contudo, alguns problemas não lineares são altamente complexos e as técnicas clássicas de regressão não são capazes de resolvê-los. Sendo assim, com intuito de solucionar esses casos, novas técnicas de modelagem vêm sendo desenvolvidas.

O método denominado de redes neurais artificiais (RNAs) é uma dessas técnicas e tem se tornando cada vez mais popular nos últimos anos, aumentando sua aplicabilidade em diferentes disciplinas. A rede neural artificial (RNA) pode ser definida como um modelo de raciocínio baseado no cérebro humano, sendo este composto por um conjunto densamente interconectado de células nervosas ou unidades de processamento de

informações básicas, chamadas de neurônios. Sendo assim, o cérebro humano pode ser descrito como um sistema de processamento de informações altamente complexo, não linear e paralelo, e as RNAs podem ser consideradas como uma reprodução simplificada desse complexo sistema (YAGIZ; SEZER; GOKCEOGLU, 2011).

De acordo com Basheer e Hajmeer (2000) já existem muitos pesquisadores que concordam sobre a superioridade da rede neural artificial quando comparada com os métodos estatísticos clássicos. A superioridade das RNAs aumenta à medida que a dimensionalidade e não linearidade dos problemas crescem pois as técnicas clássicas de regressão geralmente não conseguem produzir aproximações tão precisas nesses casos (YAGIZ; SEZER; GOKCEOGLU, 2012).

Os estudos que comparam o desempenho de modelagem e previsão de dados a partir de métodos estatísticos clássicos e redes neurais artificiais, geralmente fazem essa comparação em apenas um cenário específico. Sendo assim, não é possível inferir que o melhor método para essa situação será o mesmo para outros contextos.

Uma abordagem alternativa a esses estudos é fazer uma análise comparativa onde os métodos possam ser aplicados em diversos cenários. Esse tipo de análise pode ser feita através da técnica de planejamento de experimentos.

Montgomery (2005) define planejamento de experimentos, que vem do inglês *Design of Experiments* (DOE), como um processo onde se planejam experimentos para que dados apropriados sejam coletados e depois analisados por métodos estatísticos, resultando em conclusões válidas e objetivas. Neste trabalho, o DOE será usado para simular conjuntos de dados sintéticos por meio da combinação de diferentes fatores, criando diferentes cenários para os métodos de modelagem serem aplicados.

Portanto, esta pesquisa tem como objetivo apresentar uma análise comparativa entre os métodos de regressão não linear e redes neurais artificiais, utilizando a técnica de planejamento de experimentos. Dessa forma, será possível simular conjuntos de dados sintéticos com intuito de comparar o desempenho dos métodos, identificar quais fatores e interações mais afetam os resultados e definir qual o melhor método em termos de precisão de modelagem.

1.2 Relevância

Já existem diversos trabalhos disponíveis na literatura que apresentam um estudo comparativo entre métodos estatísticos clássicos e as redes neurais artificiais, com o objetivo de demonstrar qual dos métodos possui o melhor desempenho. Entretanto, ao

realizar uma análise bibliométrica é possível identificar uma deficiência na literatura em relação a estudos que fazem esse tipo de comparação envolvendo o método de regressão não linear.

Sendo assim, foi feita uma pesquisa por título de trabalhos nas bases de dados “Scopus” e “Web of Science”, alternando os seguintes conjuntos de palavras: *comparison nonlinear regression artificial neural networks*; *forecasting nonlinear regression artificial neural network*; e *nonlinear regression artificial neural networks*. A partir dessa busca obteve-se um total de 13 artigos que fazem a comparação entre esses dois métodos nas duas bases de dados, permitindo identificar que a quantidade de publicações sobre esses dois temas é relativamente pequena. Entretanto, o número de citações sobre esses temas tem crescido muito nos últimos anos, o que confirma a ampla utilização desses métodos. A Figura 1 ilustra o cenário descrito anteriormente.

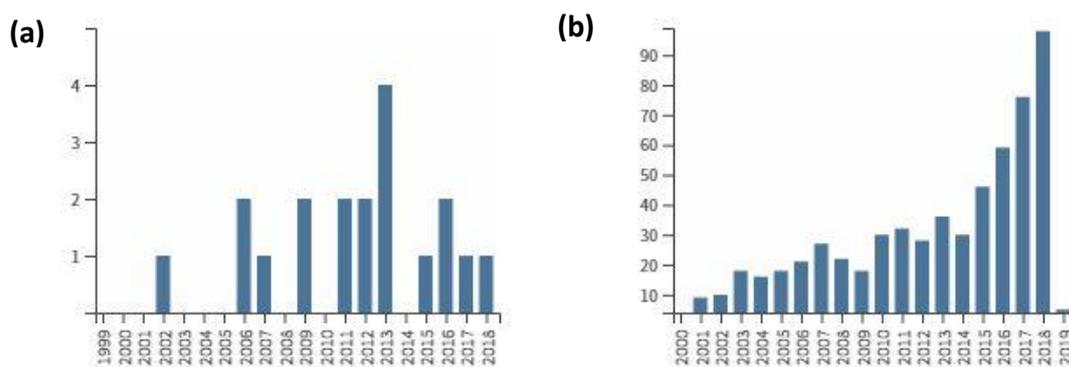


Figura 1 - Publicações (a) e citações (b) na *Web of Science*

1.3 Justificativa

A modelagem de dados é uma técnica bastante relevante dentro de diversas áreas da pesquisa e também nas práticas organizacionais. A escolha adequada desse tipo de técnica proporciona um aumento na qualidade das informações disponíveis, auxiliando de maneira rápida e eficaz, a tomada de decisão e a resolução de problemas dentro de sistemas reais.

Dentro da área da engenharia de produção é possível fazer a modelagem de diferentes tipos de dados, tais como: demandas na cadeia de suprimentos, vendas de novos produtos no mercado, tempo de ciclo de máquinas, produtividade, custo de manutenção de máquinas, risco financeiro, planejamento da produção, dentre outros. Tendo em vista

a importância desses dados para as organizações, a precisão da modelagem é um fator muito importante dentro das empresas.

As técnicas de modelagem e previsão são áreas ativas de investigação entre os pesquisadores, pois disponibilizam modelos e métodos úteis para fornecer suporte à tomada de decisão (JONES *et al.*, 2009). Os métodos utilizados para desenvolver a modelagem de dados são inúmeros. A regressão não linear e as redes neurais artificiais são exemplos desses métodos e já estão sendo representados nas pesquisas de alguns autores conforme os exemplos do próximo parágrafo.

Adamowski *et al.* (2012) utilizaram esses métodos para fazer a previsão da demanda de água urbana em Montreal, Benbouras *et al.* (2018) preveem parâmetros geotécnicos a partir de um teste de penetração em todos os tipos de solos, Hatua, Majumdar e Das (2013) desenvolveram modelos de previsão para o fator de proteção ultravioleta do tecido, Yu, Leung e Bienfang (2006) tem como finalidade em seu estudo prever o crescimento de camarões em um ambiente comercial, Esfe *et al.* (2015) desenvolveram modelos de previsão para a condutividade térmica de nanofluidos, Mahmoodi e Naderi (2016) fizeram a previsão das propriedades mecânicas de um processo de deformação plástica severa. Sendo assim, foi possível observar que esses dois métodos têm sido utilizados em diversas áreas da pesquisa com o objetivo de se obter o melhor e mais eficaz método de previsão e modelagem de dados.

De acordo com Benbouras *et al.* (2018), o método de análise de regressão é um processo estatístico que visa estimar o relacionamento entre duas ou mais variáveis e inclui diversas técnicas de modelagem para o ajuste de uma equação linear ou não linear. Já a rede neural artificial é uma técnica de estimativa não linear que processa a informação simulando o comportamento do cérebro humano. A RNA “aprende” os padrões de um conjunto de dados históricos e generaliza as relações matemáticas entre as variáveis de entrada e a variável de saída (variável resposta) correspondente (YU; LEUNG; BIENFANG, 2006).

Devido ao constante desenvolvimento em *hardware* e *software*, a coleta de bancos de dados tem sido cada vez mais facilitada, tornando o método de RNAs mais comumente utilizado que os métodos estatísticos clássicos, o que contribuiu para aumentar a utilização desse método em uma ampla quantidade de disciplinas.

Portanto, com aprimoramento das redes neurais artificiais, essa ferramenta tem se tornado uma abordagem alternativa ao método estatístico clássico de regressão não linear.

Dessa forma, surgiu a necessidade de entender qual desses métodos possui o melhor desempenho de modelagem e obtém o melhor resultado em diferentes situações.

1.4 Objetivo geral

O trabalho em questão tem como objetivo comparar os métodos de RNL e RNAs através da modelagem dos conjuntos de dados de diversos tipos de cenários que serão determinados por um planejamento de experimentos.

1.4.1 Objetivos específicos

De modo a cumprir com o objetivo geral desse trabalho, surgem os seguintes objetivos específicos:

1. Gerar conjuntos de dados sintéticos para diferentes cenários por meio do uso do planejamento de experimentos (DOE).
2. Realizar a modelagem desses conjuntos de dados utilizando os métodos de regressão não linear e redes neurais artificiais.
3. Analisar a precisão dos resultados através do cálculo do coeficiente de determinação para as respostas de modelagem obtidas por ambos os métodos.
4. Utilizar um teste de hipótese para verificar se há uma diferença estatística significativa no desempenho desses métodos.

1.5 Estrutura do trabalho

A estrutura do trabalho tem início com a contextualização do assunto que será abordado, seguida pela relevância, justificativa e objetivo geral, itens já apresentados anteriormente.

No capítulo 2, o trabalho apresenta a fundamentação teórica, abordando os assuntos necessários para o desenvolvimento da pesquisa. Inicialmente, explicam-se alguns conceitos sobre os métodos de modelagem e assuntos estatísticos necessários na pesquisa. Em seguida, apresentam-se alguns trabalhos relacionados ao tema. A fundamentação teórica também contém uma breve explicação sobre simulação por DOE, visto que este será o processo utilizado para obtenção dos bancos de dados.

O capítulo 3 apresenta o método de pesquisa utilizado, abrangendo a classificação da pesquisa e o procedimento realizado. Já o capítulo 4 faz o detalhamento e aplicação das etapas do procedimento de pesquisa apresentadas no capítulo anterior.

E, por fim, o capítulo 5 apresenta as considerações finais e sugestões para trabalhos futuros.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 Considerações iniciais

As duas técnicas de modelagem de dados utilizadas nesse trabalho são normalmente consideradas como advindas de diferentes campos do conhecimento. A regressão não linear normalmente é descrita como originária do campo de técnicas estatísticas clássicas, enquanto as redes neurais artificiais são abordadas em materiais referentes à mineração de dados, principalmente naqueles que abordam o aprendizado por máquinas. Primeiramente, esse capítulo explora a definição do método de regressão não linear e suas aplicações. Em seguida, apresenta os conceitos sobre o método de redes neurais artificiais que são devidamente detalhados e exemplificados. A seguir, é realizada a revisão de artigos que fazem uma análise comparativa entre o desempenho desses métodos. Por fim, são apresentados os conceitos sobre a técnica de planejamento de experimentos (DOE) que será utilizada para gerar os conjuntos de dados sintéticos, representando diferentes cenários onde os métodos de modelagem poderão ser aplicados.

2.2 Regressão não linear

Existem problemas presentes nas mais diversas áreas da pesquisa que envolvem a exploração das relações entre duas ou mais variáveis. Segundo Montgomery e Runger (2011) a análise de regressão é uma das técnicas estatísticas mais utilizadas para modelar e investigar essas relações.

A análise de regressão é usada para modelar a relação entre as variáveis de entrada e a variável resposta de um determinado sistema. As variáveis de entrada são denotadas por x_i e estão sob controle ou são observadas pelo experimentador. Já a variável resposta é denotada por y . O objetivo da regressão é prever ou classificar a resposta y em relação as variáveis de entrada x_i (WARNER; MISRA, 1996).

De acordo com Mattos (2013), o modelo clássico de regressão, também denominado modelo normal linear, é uma das técnicas estatísticas clássicas mais utilizadas para a estimação de dados. Entretanto, algumas de suas suposições como, por exemplo, a normalidade e a linearidade nos parâmetros, não são satisfeitas em determinadas situações práticas. Então, nesses casos, é necessário aplicar um modelo de regressão não linear.

Esse modelo é muito utilizado por pesquisadores em diversas áreas do conhecimento como: engenharia, ecologia, química, economia, agricultura, dentre outras.

Bates e Watts (1988) descrevem o modelo de regressão não linear de acordo com a Equação 1.

$$Y_n = f(x_n, \theta) + Z_n \quad (1)$$

Onde f é a função expectativa, x_n é o vetor das variáveis independentes para n casos, θ representa os parâmetros desconhecidos e z_n é o vetor dos erros aleatórios. Sendo assim, para que um modelo seja considerado não linear é necessário que ao menos uma das derivadas da função expectativa em relação aos parâmetros, dependa de pelo menos um dos parâmetros.

2.2.1 Métodos de estimação

A obtenção das estimativas dos parâmetros do modelo descrito pela Equação 1 pode ser feita através de vários métodos de estimação, dentre eles está o método dos mínimos quadrados, que será utilizado no presente trabalho.

2.2.2 Método dos mínimos quadrados

Assumindo o modelo descrito pela Equação 1, a estimativa de mínimos quadrados de θ , denotado por $\hat{\theta}$, que minimiza a soma dos quadrados dos erros é dada pela Eq. 2.

$$SQRes(\theta) = \sum_{i=1}^n [y_i - f(x_i, \theta)]^2 \quad (2)$$

Segundo Mattos (2013), para determinar o estimador dos mínimos quadrados deve-se derivar a Equação 2 em relação a cada um dos parâmetros e igualar as equações à zero, conforme mostram as Equações 3 e 4.

$$\left. \frac{\partial SQRes(\theta)}{\partial \theta_r} \right|_{\theta = \hat{\theta}} = 0 \quad (r = 1, 2, \dots, p) \quad (3)$$

$$\sum_{i=1}^n [y_i - f(x_i, \theta)] \left. \frac{\partial f(x_i, \theta)}{\partial \theta_r} \right|_{\theta = \hat{\theta}} = 0 \quad (r = 1, 2, \dots, p) \quad (4)$$

Desta forma, é possível determinar um sistema de p equações normais (4) para o modelo não linear. Na maioria dos modelos não lineares não é possível resolver esse

sistema de equações analiticamente sendo necessário o uso de métodos iterativos nesses casos.

2.2.3 Métodos Iterativos

De acordo com Mazucheli e Achcar (2002), são propostos na literatura vários métodos iterativos que podem ser utilizados para a obtenção das estimativas dos mínimos quadrados dos parâmetros de um modelo de regressão não linear, como por exemplo os métodos de *Newton-Raphson*, *Gauss-Newton*, *Steepest-Descent* e *Marquard*. De acordo com Bates e Watts (1988) o método de *Gauss-Newton* é o algoritmo de estimativa mais popular entre eles e, portanto, será o método utilizado no presente trabalho.

2.2.3.1 Método Gauss-Newton

Bates e Watts (1988) definem o método de *Gauss-Newton* como uma aproximação linear da função expectativa que tem como objetivo melhorar iterativamente o palpite inicial $\theta^{(a)}$ para θ e continuar melhorando as estimativas até que não haja mais mudança. Ou seja, a função expectativa será expandida em uma série de Taylor de primeira ordem como mostra a Equação 5.

$$f(x, \theta) \approx f(x, \theta^{(a)}) + F(\theta^{(a)})(\theta - \theta^{(a)}) \quad (5)$$

Onde $F(\theta) = \frac{\partial f(x, \theta)}{\partial \theta} = \left[\left(\frac{\partial f(x_i, \theta)}{\partial \theta_j} \right) \right]$. Definindo $r(\theta)$ como sendo um vetor de resíduos, pode-se escrever a Equação 6.

$$\begin{aligned} r(\theta) &= y - f(x, \theta) \\ &\approx r(\theta^{(a)}) - F^{(a)}(\theta - \theta^{(a)}) \end{aligned} \quad (6)$$

Onde $F^{(a)} = F(\theta^{(a)})$. Substituindo $r'(\theta)r(\theta)$ em $SQRes(\theta)$ obtém-se a Eq. 7.

$$SQRes(\theta) \approx r'^{(\theta^{(a)})}r(\theta^{(a)}) - 2r'^{(\theta^{(a)})}F^{(a)}(\theta - \theta^{(a)}) + (\theta - \theta^{(a)})'F'^{(a)}F^{(a)}(\theta - \theta^{(a)}) \quad (7)$$

Logo, $SQRes(\theta)$ será minimizada de acordo com a Equação 8.

$$\theta - \theta^{(a)} = [F'^{(a)}F^{(a)}]^{-1}F'^{(a)}r(\theta^{(a)}) \quad (8)$$

Então, devido à aproximação $\theta^{(a)}$, a próxima aproximação é dada pela Eq. 9.

$$\theta^{(a+1)} = \theta^{(a)} + [F'^{(a)}F^{(a)}]^{-1}F'^{(a)}r(\theta^{(a)}) \quad (9)$$

Por fim, a Equação 9 resulta no método iterativo conhecido como Método de *Gauss-Newton*. O processo descrito na Equação 9 é repetido até que a convergência seja obtida.

2.3 Redes Neurais artificiais

Conforme foi visto no item anterior, a aplicação do método de regressão não linear exige que uma relação funcional conhecida seja imposta ao conjunto de dados. Portanto, se este modelo assumido não estiver correto, irão ocorrer erros de modelagem. Uma abordagem alternativa a este caso é não assumir qualquer relação funcional entre as variáveis e deixar que os dados definam essa forma por si mesmos. Essa abordagem é a base do poder das redes neurais artificiais.

De acordo com Adamowski *et al.* (2012), as RNAs podem ser definidas como um processo orientado por dados que possui um algoritmo matemático flexível capaz de modelar complexas relações não lineares entre as variáveis de entrada e a variável resposta de um determinado sistema sem qualquer conhecimento prévio desse conjunto de dados.

As RNAs foram desenvolvidas a partir do estudo do funcionamento da rede de neurônios do cérebro humano, devido a facilidade e eficácia com que o cérebro realiza tarefas difíceis e complexas e principalmente pela sua capacidade de aprender. Haykin (2008) define a rede neural como um processador que possui uma estrutura maciçamente e paralelamente distribuída, constituída de unidades de processamento simples, que possuem propensão natural para armazenar conhecimento experimental e disponibilizá-lo para o uso. A RNA se assemelha ao cérebro humano em dois aspectos, sendo eles:

- O conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem;
- Forças de conexão entre os neurônios, também conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.

Segundo Basheer e Hajmeer (2000) embora as RNAs tenham sido desenvolvidas com base no funcionamento do cérebro humano, a ideia principal dessa técnica não é replicar o funcionamento desse sistema, mas sim fazer uso do que é conhecido sobre a funcionalidade das redes biológicas para resolver problemas complexos. Portanto, a atratividade na utilização desse método vem das suas características notáveis de processamento que se assemelham com as do sistema biológico, como por exemplo: a não linearidade, alto paralelismo, robustez, tolerância ao erro e a falhas, aprendizagem, possibilidade de lidar com informações imprecisas e capacidade de generalização.

É de suma importância que os modelos artificiais possuam essas características pois, a não linearidade proporciona um melhor ajuste aos dados, a insensibilidade a ruídos fornece uma previsão precisa mesmo na presença de dados imprecisos ou erros de medição, o alto paralelismo implica em um rápido processamento e tolerância a falhas de *hardware*, a capacidade de aprendizagem e adaptação permitem que o sistema possa atualizar sua estrutura em resposta a mudanças no ambiente, e pôr fim a generalização permite que o modelo seja aplicado em dados nunca antes vistos (BASHEER; HAJMEER, 2000).

2.3.1 Componentes de uma rede

A rede neural artificial pode ser descrita como uma rede de nós de processamento simples ou neurônios, interconectados entre si em uma ordem específica, executando manipulações numéricas simples. A Figura 2 representa o modelo de um neurônio que serve de base para o projeto de redes neurais artificiais.

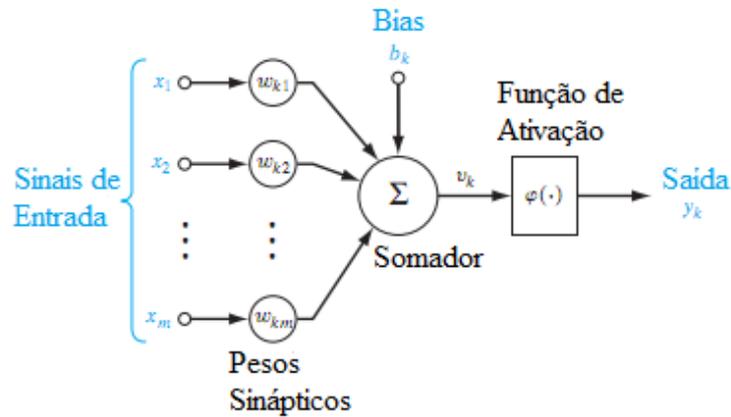


Figura 2 - Modelo não linear de um neurônio

Fonte - Adaptado de Haykin (2008)

Haykin (2008) identificou a partir desse modelo a presença de três elementos básicos, sendo eles:

- Um conjunto de sinapses, cada uma delas caracterizada por um peso característico, onde o sinal x_j na entrada j conectada ao neurônio k é multiplicado pelo peso sináptico w_{kj} ;
- Um combinador linear para somar os sinais de entrada, ponderados pelas respectivas sinapses do neurônio;
- Uma função de ativação para limitar a amplitude de saída do neurônio. A função de ativação limita a faixa de amplitude permitida do sinal de saída a um valor finito.

Sendo assim, um neurônio k pode ser descrito matematicamente através das Equações 10 e 11 abaixo.

$$u_k = \sum_{j=0}^m \omega_{kj} x_j \quad (10)$$

$$y_k = \varphi(u_k + b_k) \quad (11)$$

Onde x_1, x_2, \dots, x_m são os sinais de entrada, w_1, w_2, \dots, w_m são os respectivos pesos sinápticos do neurônio k , u_k é a saída do combinador linear devido aos sinais de entrada, b_k é o bias, $\varphi(\cdot)$ é a função de ativação e y_k é o sinal de saída do neurônio (HAYKIN, 2008).

Sendo assim, o neurônio k recebe sinais de entrada (x_j), que podem ser informações de outros neurônios ou de estímulos externos, através de conexões simuladas pelos pesos sinápticos (w_{kj}). O efeito de um sinal é determinado pela multiplicação do valor do sinal pelo peso da conexão correspondente ($x_j \times w_{kj}$). Em seguida, é efetuada a soma dos valores ($x_j \times w_{kj}$) de todas as conexões, resultando em um valor u_k . Esse valor é então somado a um elemento polarizador b_k , (bias), que possui o efeito de aumentar ou diminuir o argumento da função de ativação. O valor resultante é processado por uma função de ativação, que produz um sinal de saída (y).

2.3.2 Arquiteturas de uma rede

A arquitetura da rede refere-se à disposição dos neurônios, um em relação ao outro. Além disso, a arquitetura de uma rede neural artificial é responsável por restringir o tipo de problema no qual a rede poderá ser utilizada, e é definida pelo número de camadas e pelo tipo de conexão entre os nós (HAYKIN, 2008).

Existem diferentes tipos de arquiteturas nas quais as redes neurais artificiais podem ser estruturas. Em geral, podemos identificar duas categoriais fundamentais de arquitetura de rede de acordo com o tipo de conexão entre os nós: rede *feedforward* (rede alimentada adiante) e rede *feedback* (rede recorrente).

Segundo Haykin (2008), a rede alimentada adiante ou acíclica, na forma mais simples de uma rede em camadas, possui uma camada de entrada de nós de fonte que se projeta sobre uma camada de saída de neurônios (nós computacionais), mas não vice e versa, conforme mostra a Figura 3. Esse tipo de rede também pode apresentar em sua estrutura a presença de uma ou mais camadas ocultas, cujos nós computacionais correspondem aos chamados neurônios ocultos, como mostra a Figura 5 no item 2.3.3.

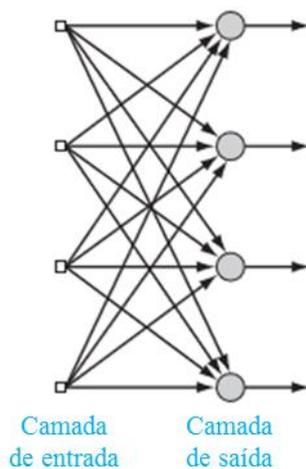


Figura 3 - Arquitetura de rede *feedforward*

Fonte – Adaptado de Haykin (2008)

Já a rede recorrente ou cíclica possui pelo menos um laço de realimentação. Nesse caso, a rede pode consistir, por exemplo, de uma única camada de neurônios com cada neurônio alimentando seu sinal de saída de volta para a entrada de todos os outros neurônios, conforme mostra a Figura 4 (HAYKIN, 2008).

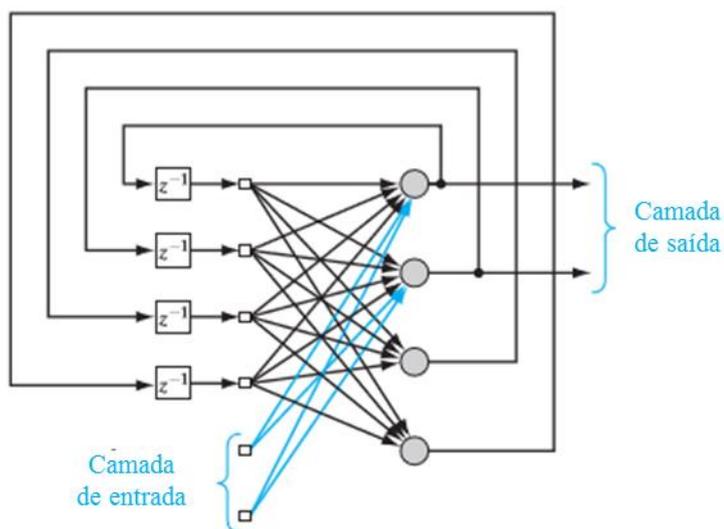


Figura 4 - Arquitetura de rede *feedback*

Fonte - Adaptado de Haykin (2008)

2.3.3 Redes neurais *perceptrons* de múltiplas camadas

As redes neurais *perceptrons* de múltiplas camadas (*Multi-Layer Perceptron* – MLP), objeto do presente estudo, é do tipo de rede *feedforward* e é a arquitetura mais popular de redes neurais artificiais devido a sua poderosa capacidade de modelagem (ADAMOWSKI *et al.*, 2012). A rede MLP é composta por uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída, conforme mostra a Figura 5. Além disso, seus neurônios de entrada convertem seus pesos e bias (viés) em pesos e bias de saídas através de uma função de ativação. Dentre as diversas funções de ativação utilizadas, a função sigmoide, representada na Equação 16, é a que possui um maior número de aplicações na prática. (AMEDI; BAGHBAN; AHMADI, 2016).

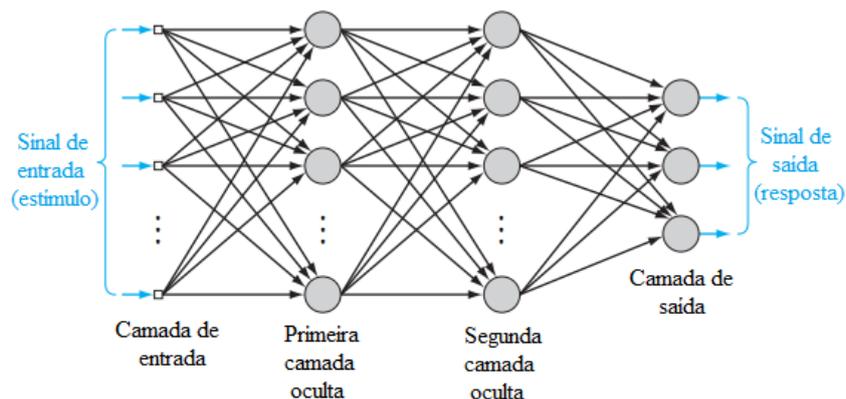


Figura 5 - Arquitetura de uma rede MLP com duas camadas ocultas

Fonte - Adaptado de Haykin (2008)

A arquitetura MLP é capaz de modelar relacionamentos complexos entre variáveis de entrada e saída, onde os seus elementos não lineares (neurônios) estão dispostos em camadas sucessivas e o fluxo de informação ocorre de maneira unidirecional da camada de entrada para a camada de saída através das camadas ocultas. A função dos neurônios presentes nas camadas ocultas é intervir entre as camadas externas de entrada e saída de maneira útil. Ou seja, quando uma ou mais camadas ocultas são adicionadas, a rede é capaz de extrair estatísticas de ordem elevada de suas entradas (HAYKIN, 2008).

De acordo com Haykin (2008), a forma em que os neurônios são estruturados em uma arquitetura de rede está diretamente relacionada ao seu algoritmo de aprendizagem. O algoritmo de aprendizagem mais popular para as redes do tipo MLP é o algoritmo de retropropagação do erro (*error back-propagation*) (VINOOTH; RAJARATHIAN;

BARGAVI, 2016). Sendo assim, este será o algoritmo utilizado no presente trabalho e será descrito no item 2.3.5.

2.3.4 Processos de aprendizagem

De acordo com Basheer e Hajmeer (2000), o processo de aprendizagem das redes neurais artificiais tem como função atualizar a representação interna do sistema em resposta a estímulos externos, para que uma tarefa específica possa ser executada. O algoritmo de aprendizagem é o procedimento necessário para que essa função seja realizada, pois tem a função de ajustar os pesos sinápticos de acordo com um objetivo de projeto. Existe uma grande variedade de algoritmos de aprendizagem disponíveis na literatura que se diferem entre si pela forma com que o ajuste do peso sináptico de um neurônio é formulado. Dessa forma, as cinco regras básicas de aprendizagem são: aprendizagem por correção e erro, aprendizagem baseada em memória, aprendizagem hebbiana, aprendizagem competitiva e aprendizagem de Boltzmann (HAYKIN, 2008).

A partir dessas regras de aprendizagem são apresentados dois tipos de paradigmas principais: aprendizado supervisionado e aprendizado não supervisionado. No primeiro, a rede recebe conjuntos de exemplos de padrões de entrada e seus correspondentes padrões de saída, tais exemplos podem ser modelos existentes ou dados históricos; a resposta fornecida pela rede é comparada com a saída esperada e o erro verificado é informado à rede para que o algoritmo seja ajustado visando uma melhor aproximação nas respostas futuras. Já o aprendizado não supervisionado, não utiliza conjuntos previamente conhecidos, ou seja, apenas os padrões de entradas são inseridos na rede, a qual tenta progressivamente codificar características dos dados para classificá-los automaticamente. O presente trabalho utiliza o aprendizado supervisionado.

2.3.5 O algoritmo de retropropagação do erro (*error back-propagation*)

Dos algoritmos de aprendizagem que utilizam o aprendizado supervisionado, o mais popular entre eles é o algoritmo de retropropagação do erro. De acordo com Monjezi, Ghafurikalajahi e Bahrami (2011), nesse algoritmo o processo começa na fase de treinamento onde cada padrão de entrada do conjunto de treinamento é apresentado as camadas de entrada da rede e depois propagado para as camadas seguintes (adiante). Assim que esse processo de propagação chega à camada de saída, é calculado o sinal do erro a partir da comparação entre os valores de saída obtidos pela rede e os valores de

saída desejados. O sinal de erro para cada padrão de saída desejado é então retropropagado da camada de saída para a camada de entrada a fim de ajustar ou alterar os pesos de cada camada de forma adequada. Portanto, esse algoritmo tem como função minimizar o erro obtido ajustando os pesos e limiares de ativação em busca dos pontos mínimos da superfície de erro.

A lógica usada no desenvolvimento do algoritmo de retropropagação do erro está apresentada nas Equações de 12 a 16.

$$x_i = X_i, 1 \leq i \leq m \quad (12)$$

$$net_i = \sum_{j=1}^{i-1} W_{ij} x_j, m < i \leq N + n \quad (13)$$

$$x_i = s(net_i), m < i \leq N + n \quad (14)$$

$$Y_i = x_{i+N}, 1 \leq i \leq n \quad (15)$$

Onde a função s representada em (14) é normalmente a função sigmoide (16).

$$S(Z) = \frac{1}{1 + e^{-Z}} \quad (16)$$

O valor de N é uma constante inteira maior que m , $N + n$ determina a quantidade de neurônios na rede, net_i é o nível de voltagem excitando um neurônio e x_i é a intensidade do resultado da saída do neurônio (WERBOS, 1990). O peso W_{ij} é escolhido para minimizar o erro no conjunto de treinamento (Equação 17).

$$E = \sum_{t=1}^T E(t) = \sum_{t=1}^T \sum_{i=1}^n \left(\frac{1}{2}\right) (\hat{Z}_i(t) - Z_i(t))^2 \quad (17)$$

Para calcular $\hat{Z}_i(t)$ e os erros $E(t)$, iniciam-se os pesos W_{ij} com valores arbitrários. Em seguida, é feito o cálculo das derivadas de E em relação à todos os pesos. No caso de o erro aumentar ao aumentar os pesos, é necessário que o peso seja ajustado para baixo. Entretanto, se o erro diminuir, o ajuste deve ser para cima. Após esses ajustes de pesos, inicia-se esse processo novamente até o momento em que o erro estabilize (WERBOS, 1990).

2.4 Pesquisas comparativas anteriores (RNL x RNA)

Conforme foi dito anteriormente, os métodos de RNL e RNAs são duas técnicas de modelagem que têm sido amplamente utilizadas em diversas áreas da pesquisa, além de serem exploradas nos estudos de muitos autores que tem como objetivo verificar qual desses possui o melhor desempenho em problemas específicos.

Yu, Leung e Bienfang (2006) avaliaram o potencial das RNAs como uma alternativa à regressão não linear com o objetivo de fazer a previsão do crescimento de camarões em uma fazenda comercial no Havaí. Dados empíricos dessa fazenda foram coletados e analisados. Em seguida, o desempenho dos métodos foi avaliado através de quatro medidas: raiz do erro quadrático médio (RMSE), coeficiente de determinação (R^2), porcentagem de pontos de viragem errados e porcentagem de valores previstos com tolerância de 5% aos valores reais correspondentes. Depois de comparar os resultados, os autores concluíram que ambos os métodos obtiveram bons ajustes ao conjunto de dados, porém o método de RNAs teve um desempenho ligeiramente melhor que os melhores modelos não lineares, fazendo a previsão mais precisa do conjunto de dados.

Razi e Athappily (2005) utilizaram os métodos de redes neurais artificiais, regressão não linear, classificação e árvore de regressão (CART) em seu estudo. Dessa forma, um grande conjunto de dados sobre os hábitos de pessoas fumantes foi utilizado para prever a resposta “Dias na cama devido à doença” através desses métodos. A precisão da previsão foi medida a partir de três métricas de erro: erro percentual absoluto médio (MAPE), erro quadrático médio (MSE) e *large prediction error* (LPE). Os autores concluíram que os modelos desenvolvidos pelos métodos de RNAs e de classificação e árvore de regressão forneceram uma previsão mais precisa ao conjunto de dados quando comparados com o modelo da regressão não linear.

Adamowski *et al.* (2012) compararam o desempenho dos métodos de regressão linear múltipla (RLM), regressão não linear múltipla (RNLM), autoregressivo integrado de médias móveis (ARIMA), redes neurais artificiais (RNAs) e transformada discreta de wavelet combinada com RNAs (WA-RNA) na previsão da demanda de água urbana em Montreal no Canadá. As métricas utilizadas para comparar o desempenho desses métodos foram: o coeficiente de determinação, o erro quadrático médio, raiz do erro quadrático médio relativo e o índice de eficiência. Os modelos que utilizaram o método de WA-RNAs forneceram previsões mais precisas para a demanda de água urbana em relação aos modelos RLM, RNLM, ARIMA e RNAs. Os resultados deste estudo indicam que os

modelos WA-RNAs são um método potencialmente promissor para a previsão da demanda de água urbana e por isso devem ser desenvolvidas novas pesquisas envolvendo essa nova técnica.

Ozçelik *et al.* (2013) fizeram uma análise comparativa utilizando os métodos de redes neurais artificiais e regressão não linear para estimar a altura da árvore de zimbro a partir de seu diâmetro. Os critérios de avaliação do desempenho desses métodos foram: raiz do erro quadrático médio, coeficiente de correlação, viés médio e critério de informação de Akaike. O estudo concluiu que ambas as técnicas de modelagem são ferramentas úteis para a previsão da altura a partir do diâmetro do tronco das árvores. No entanto, os autores ressaltaram que existem limitações de cada um dos métodos que devem ser levadas em consideração, a fim de selecionar a técnica mais adequada para cada contexto. Por fim, os autores concluíram que o método de RNAs pode ser considerado uma abordagem alternativa e eficaz para a regressão não linear, podendo fornecer ao sistema de modelagem de crescimento florestal melhores informações, aumentando a capacidade de gerenciar adequadamente os ecossistemas florestais.

Os trabalhos descritos acima estão resumidos na Tabela 1 abaixo, de acordo com os métodos de modelagem utilizados.

<i>Autores</i>	<i>Aplicação</i>	<i>Métodos de modelagem</i>	<i>Resultados</i>
Yu, Leung e Bienfang (2006)	Previsão do crescimento de camarões em uma fazenda comercial no Haváí	RNAs e RNL	RNAs superou a RNL.
Razi e Athappily (2005)	Previsão de dados sobre fumantes	RNAs, RNL e CART	RNAs e CART superaram a RNL.
Adamowski et al. (2012)	Previsão da demanda de água urbana em Montreal	RLM, RNLM, ARIMA, RNAs e WA-RNAs	WA-RNAs superou os demais métodos.
Ozçelik et al. (2013)	Modelagem da altura da árvore de Zimbro a partir do seu diâmetro	RNAs e RNL	Ambos os métodos apresentaram bons resultados.

Tabela 1 - Resumo das pesquisas comparativas anteriores

Em geral, esses estudos concluíram que o desempenho do método de RNAs é geralmente melhor que o da regressão não linear em relação à precisão da modelagem e previsão dos dados. No entanto, esses trabalhos avaliaram os métodos apenas no cenário específico em que o estudo estava sendo realizado, portanto, suas conclusões não podem ser generalizadas para outros contextos. Então, uma abordagem alternativa a esses estudos é avaliar os métodos em mais de um cenário. Essa análise pode ser feita através da técnica de planejamento de experimentos.

Sendo assim, o presente trabalho tem como objetivo utilizar o DOE para simular conjuntos de dados sintéticos por meio da combinação de diferentes fatores descritos no item 4.2.2. Esses conjuntos de dados representam diferentes cenários onde os métodos serão aplicados. Dessa forma, será possível desenvolver uma análise mais profunda sobre quais fatores e interações são significativos e podem influenciar a precisão de ambos os métodos, além de apoiar estatisticamente todas as conclusões tiradas pelo pesquisador.

2.5 Planejamento de experimentos (DOE)

O Planejamento de experimentos, termo que vem do inglês *Design of Experiments* (DOE), é uma metodologia amplamente utilizada por pesquisadores que lidam com experimentos em situações práticas, devido a uma enorme quantidade de histórias de sucesso a partir de sua aplicação. Montgomery (2005) define o DOE como um processo de planejamento de experimentos onde dados apropriados são coletados e depois analisados através de métodos estatísticos, resultando em conclusões válidas e objetivas.

Um experimento pode ser definido como um teste ou uma série de testes em que o experimentador faz alterações no conjunto de variáveis ou fatores de entrada (x) de maneira controlada (c) a fim de identificar e observar como as respostas (y) desse sistema se comportam diante dessas mudanças, conforme mostra a Figura 6 (GOMES, 2010).

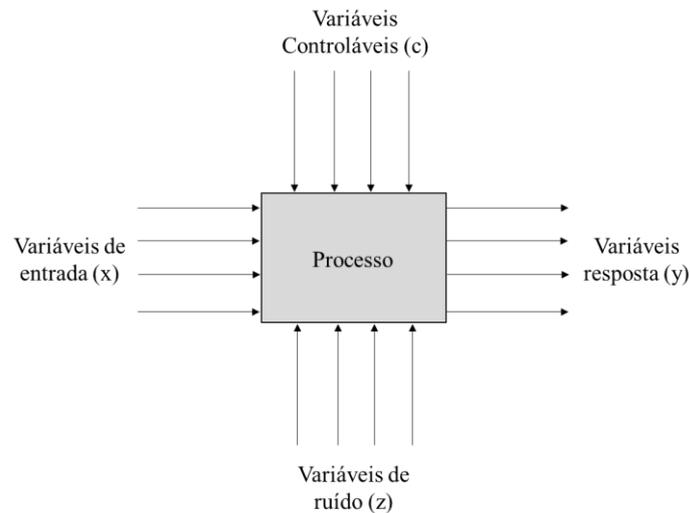


Figura 6 - Modelo experimental de um sistema genérico

Fonte - Adaptado de STAICULESCU *et al.* (2005)

A elaboração do DOE pode ser feita através de três tipos de arranjos principais: fatorial, misturas e RSM (*Response Surface Methodology*). De acordo com Montgomery (2005), quando o experimento envolve o estudo dos efeitos de dois ou mais fatores em uma resposta o arranjo fatorial é o mais eficiente nesse caso.

No arranjo fatorial, em cada ensaio completo ou réplica do experimento todas as combinações possíveis dos níveis dos fatores são investigadas, para verificar o efeito em conjunto desses fatores em uma variável resposta. Existem diversos casos especiais de arranjos fatoriais e o mais importante entre eles é o arranjo fatorial do tipo 2^k , com k fatores onde cada fator possui dois níveis. Esses níveis, -1 e +1, representam os limites inferior e superior para o intervalo em que a variável é analisada e podem ser caracterizados como quantitativos (como valores de temperatura ou tempo) ou como qualitativos (como uma máquina com desempenho nos níveis “alto” e “baixo”). Sendo assim, uma réplica completa desse tipo de arranjo requer $2 \times 2 \times \dots \times 2 = 2^k$ observações (MONTGOMERY, 2005).

A medida que o número de fatores em um arranjo do tipo 2^k aumenta, o número de execuções necessárias para uma réplica completa do projeto experimental supera os recursos da maioria dos experimentadores. Nesse caso, se o experimentador puder assumir que certas iterações de ordem alta entre os fatores não são significativas, informações sobre os efeitos principais e iterações de ordem baixas podem ser obtidas executando apenas uma fração do arranjo fatorial completo. Experimentos desse tipo são denominados arranjos fatoriais fracionados e estão entre os tipos de projeto mais

utilizados para o *design* de produto e processo, melhoria de processo e experimentação industrial/empresarial (MONTGOMERY, 2005).

2.5.1 DOE para simulação

O DOE é uma metodologia conhecida pelo seu histórico de sucesso em aplicações industriais. No entanto, essa ferramenta também pode ser utilizada para outros fins, por exemplo, pode ser aplicada em problemas de simulação. De acordo com Kleijnen *et al.* (2005), o uso do DOE no desenvolvimento e simulação de modelos permite que informações detalhadas sobre o comportamento do modelo sejam descobertas, faz com que a equipe de modelagem discuta em detalhes as implicações de várias premissas do modelo, auxilia a formulação de perguntas quando os analistas não sabem antecipadamente quais perguntas devem ser feitas, desafia ou confirma expectativas sobre a direção e a importância relativa dos efeitos dos fatores e até mesmo a descoberta de problemas na lógica do programa. Dessa forma, o DOE deve ser considerado uma parte importante e relevante dentro da análise de simulação.

Ao usar o DOE para simulação o principal objetivo não é a otimização, nesse caso os esforços são dedicados a encontrar políticas ou decisões robustas que auxiliem a tomada de decisões (BALESTRASSI *et al.*, 2009). Neste trabalho, o DOE será utilizado para simular conjuntos de dados sintéticos através da combinação de diferentes fatores que estão descritos no item 4.2.2, a fim de avaliar como esses fatores e suas interações afetam o desempenho de modelagem de cada um dos métodos analisados neste estudo. Nesse contexto, a Figura 6 pode ser representada analogamente pela Figura 7.

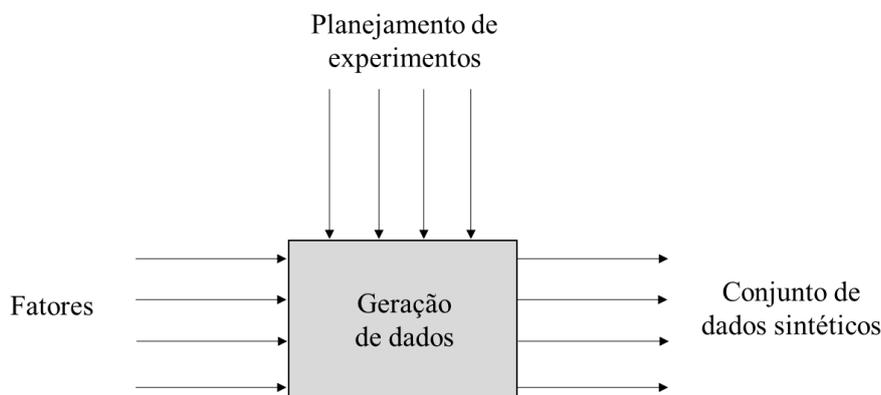


Figura 7 - Modelo experimental de um sistema de simulação

3. METODOLOGIA DE PESQUISA

Com base nos conceitos apresentados no capítulo anterior, este capítulo tem como objetivo apresentar a classificação da pesquisa, o método adotado e a sequência das etapas executadas.

3.1 Classificação da pesquisa

A classificação da presente pesquisa foi feita de acordo com Miguel *et al.* (2014). Estes autores sugerem a classificação da pesquisa científica segundo os critérios de natureza, objetivos, abordagem e métodos utilizados, Figura 8.

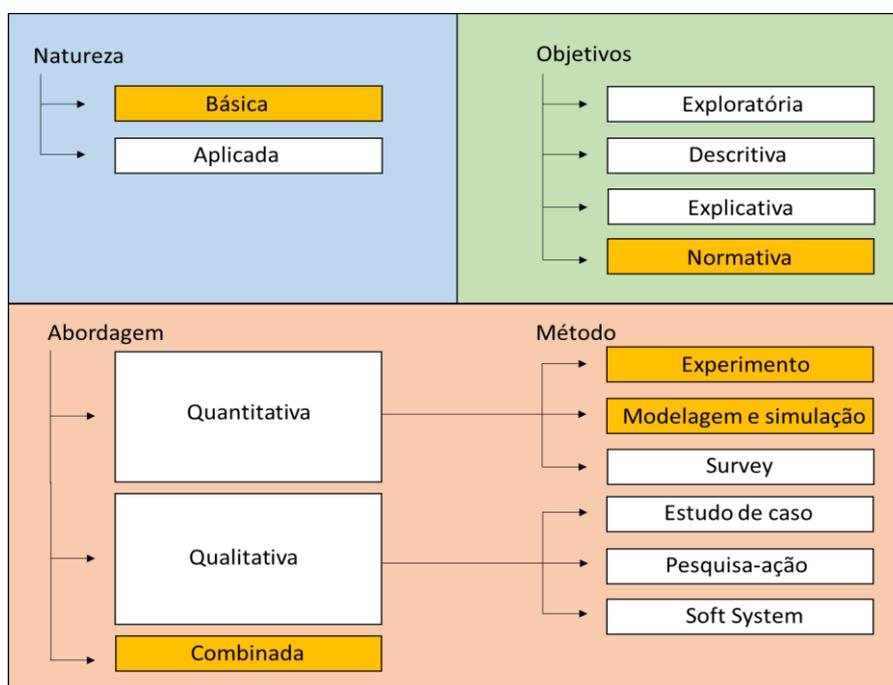


Figura 8 - Classificação da pesquisa

Fonte - Adaptado de Miguel et al. (2014)

Portanto, essa pesquisa será classificada da seguinte forma:

- Quanto à natureza: Uma pesquisa é classificada como básica quando ela está ligada ao conhecimento científico sem o objetivo comercial. Já a pesquisa aplicada se caracteriza por ser uma pesquisa de caráter prático, onde os resultados devem ser aplicados (APPOLINÁRIO, 2006). Portanto, este trabalho se classifica como pesquisa básica, pois busca desenvolver conhecimento através de conjuntos de dados sintéticos e disponibilizá-lo eventualmente para o uso de outros pesquisadores e empresas na resolução de problemas.

- Quanto ao objetivo: De acordo com Bertrand e Fransoo (2002), a pesquisa normativa está interessada principalmente no desenvolvimento de políticas, estratégias e ações para aprimorar os resultados disponíveis na literatura, encontrar uma solução ótima para um problema recém-definido ou comparar diferentes estratégias na resolução de um problema específico. Dessa forma, esta pesquisa se classifica como normativa em relação aos seus objetivos, pois busca desenvolver novas estratégias e ações para resolver um determinado problema.
- Quanto à abordagem: Na abordagem qualitativa, o pesquisador busca entender os fenômenos, observando-os, interpretando-os e descrevendo-os (MELLO, et al., 2012). Já a abordagem quantitativa, representa as informações de forma numérica e as analisa por meio de técnicas estatísticas (MORESI, 2003). Então, quanto à abordagem, esta pesquisa trará uma abordagem combinada dos métodos de experimentação e modelagem e simulação. Primeiramente será empregado o método de experimentação durante a geração dos conjuntos de dados sintéticos e, posteriormente, a modelagem e simulação, pela aplicação dos métodos de modelagem aos conjuntos de dados e avaliação de seus desempenhos.
- Quanto aos métodos: De acordo com Köche (2013), no método de experimentação, o pesquisador faz a análise do problema, formula as hipóteses e manipula as variáveis de entrada selecionadas, de maneira que uma relação entre tais variáveis e os resultados sob investigação seja estabelecida. Já o método de modelagem e simulação deve ser usado quando o objetivo é prever o efeito de alterações no sistema ou avaliar seu comportamento ou desempenho (BERTRAND, J. W. M., FRANSOO, J. C., 2012).

3.1.1 Experimentação

De modo geral, a pesquisa experimental consiste em determinar um objeto de estudo, selecionar as variáveis de entrada que seriam capazes de influenciá-lo e definir as formas de controle e de observação dos efeitos que essas variáveis podem produzir no objeto. Portanto, na pesquisa experimental, o pesquisador utiliza técnicas estatísticas para demonstrar as relações causais entre as variáveis de entrada e a variável resposta.

De acordo com Köche (2013), na pesquisa experimental o pesquisador é capaz de analisar o problema, formular hipóteses e manipular as variáveis de entrada selecionadas para que sejam estabelecidas relações entre tais variáveis e os resultados que estão sob investigação.

No presente trabalho a técnica de experimentação utilizada foi o DOE do tipo fatorial fracionado 2_5^{5-1} com duas réplicas, resultado em 32 experimentos, seus fatores e níveis serão devidamente detalhados no item 4.2.2.

Conforme foi explicado no item 2.5, essa técnica é um processo de planejamento de experimentos onde dados apropriados são coletados e depois analisados através de métodos estatísticos, resultando em conclusões validas e objetivas (MONTGOMERY, 2005). Assim, Montgomery (2005) propõe que o emprego do planejamento de experimentos deve considerar as seguintes etapas:

1. Definição do problema;
2. Escolha dos fatores e definição dos níveis de trabalho;
3. Seleção das variáveis de resposta;
4. Escolha do projeto experimental;
5. Execução dos experimentos;
6. Análise estatística dos dados;
7. Conclusões e recomendações.

3.1.2 Modelagem e simulação

Chung (2004) define o processo de modelagem e simulação como a criação e experimentação de um modelo matemático computadorizado de um sistema físico, que auxilia na análise de sistemas e na tomada de decisão.

Portanto, esse processo tem os seguintes propósitos: Obter *insights* sobre a operação de um sistema, desenvolver políticas operacionais ou de recursos para melhorar o desempenho do sistema, testar novos conceitos e/ou sistemas antes da sua implementação e obter informações sem perturbar o sistema real (CHUNG, 2004).

Então, devido aos propósitos descritos anteriormente a técnica de Modelagem e Simulação, se aplicada de acordo com uma metodologia apropriada, disponibilizará resultados estatisticamente confiáveis para direcionar os pesquisadores a identificarem os melhores caminhos possíveis durante o processo de tomada de decisões.

De acordo com Mitroff *et al.* (1974), a pesquisa de modelagem e simulação é dividida em 4 etapas, sendo elas:

1. Conceitualização;
2. Modelagem;
3. Solução pelo modelo;
4. Implementação.

3.2 Procedimento

O procedimento de pesquisa desse trabalho baseou-se na combinação entre a metodologia proposta por Montgomery (2005), descrita no item 3.1.1, e na metodologia introduzida por Mitroff *et al.* (1974), descrita no item 3.1.2. Essa combinação resultou em um procedimento de pesquisa combinado que irá seguir as seguintes etapas:

1. Conceitualização e definição do problema;
2. Escolha dos fatores e definição dos níveis de trabalho;
3. Escolha do projeto experimental;
4. Modelagem;
5. Solução pelo modelo;
6. Análise estatística dos dados;
7. Interpretação dos resultados.

4. APLICAÇÃO

Neste capítulo as etapas do procedimento de pesquisa descritas no item 3.2 serão devidamente detalhadas e aplicadas a este estudo.

4.2.1 Conceitualização e definição do problema

Nesta etapa inicial é fundamental que o escopo do problema e os objetivos da pesquisa sejam definidos. Neste trabalho o problema de pesquisa é determinar qual dos métodos de modelagem possui o melhor desempenho. Os métodos de modelagem estudados serão: Regressão Não Linear (RNL) e Redes Neurais artificiais (RNAs).

Os objetivos específicos dessa pesquisa, assim como mencionado no item 1.4.1, são: gerar conjuntos de dados sintéticos através do planejamento de experimentos, aplicar os métodos de modelagem nesses conjuntos de dados e comparar os resultados obtidos para extrair conclusões válidas e objetivas.

4.2.2 Escolha dos fatores e níveis de trabalho

Nessa etapa, os fatores a serem estudados e seus níveis foram selecionados. Com intuito de gerar diferentes cenários para serem modelados através dos dois métodos, primeiramente foram identificados possíveis fatores de influência no desempenho desses métodos, sendo eles: o tamanho do conjunto de dados (N), o número de variáveis de entrada (x_i), o tipo de função não linear usada para gerar a variável resposta (f), o valor da correlação entre as variáveis de entrada (ρ) e o desvio padrão do erro aleatório adicionado a variável resposta (σ). Esses fatores foram avaliados em dois níveis cada um conforme mostra a Tabela 2.

	<i>FATORES</i>	<i>NÍVEIS</i>	
1	N	80	320
2	x_i	2	4
3	f	<i>Rosenbrock</i>	<i>Ackley N.4</i>
4	ρ	0,35	0,90
5	σ	$0,2 * \sigma_y$	$0,8 * \sigma_y$

Tabela 2- Níveis dos fatores analisados no DOE

Com intuito de verificar a influência do tamanho do conjunto de dados no desempenho dos métodos, optou-se por trabalhar com os valores 80 e 320. Já os valores definidos para a correlação entre as variáveis de entrada, utilizados na Equação 18, foram

de 0,35 para uma fraca correlação e 0,90 para uma forte correlação. É importante destacar que em ambos os casos foram considerados valores de *p-value* significativos de forma a garantir que a correlação fosse significativa, entretanto podendo ser forte ou fraca.

A fim de obter esses valores de correlação para as variáveis de entrada do conjunto de dados, foi utilizada a função de gerar dados aleatórios do *software* Minitab, considerando uma distribuição normal multivariada com média zero e desvio padrão igual a 1. Dessa forma, esses dados foram gerados a partir de um vetor de médias zero e uma matriz 2×2 de variância-covariância composta pelos valores mostrados na Equação 18 no caso dos cenários com 2 variáveis de entrada. O procedimento para os cenários com 4 variáveis de entrada é o mesmo só que resulta em uma matriz 4×4 .

$$\begin{bmatrix} \sigma_{x_1}^2 & \rho_{12} \\ \rho_{12} & \sigma_{x_2}^2 \end{bmatrix} \quad (18)$$

Onde $\sigma_{x_1}^2$ and $\sigma_{x_2}^2$ são as variâncias das variáveis x_1 e x_2 , respectivamente, e ρ_{12} é a correlação entre as duas variáveis de entrada.

As duas funções utilizadas para modelar a variável dependente (y) em relação as variáveis independentes (x_i) foram escolhidas na plataforma virtual *BenchmarkFcns*. Essas funções foram escolhidas com intuito de verificar se o número de parâmetros não lineares presentes na equação influencia o desempenho dos métodos, visto que a função *Rosenbrock* possui um parâmetro não linear e a função *Ackley N. 4* possui três parâmetros não lineares, conforme mostram as Equações 19 e 20, respectivamente (os valores destacados em negrito representam os parâmetros não lineares).

$$f(x) = \sum_{i=1}^n [100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2] \quad (19)$$

$$f(x) = \sum_{i=1}^{n-1} (e^{-0,2} \sqrt{x_i^2 + x_{i+1}^2} + 3(\cos(2x_i) + \text{sen}(2x_{i+1}))) \quad (20)$$

Para aproximar os valores das respostas (y) de um cenário mais próximo da realidade, foi adicionado a essas funções um erro aleatório gerado a partir de uma distribuição normal com média zero e um desvio padrão que variou em cada um dos

experimentos no *software* Minitab, resultando no valor final da variável resposta a ser modelada (y_{real}). Esse desvio foi baseado nas considerações de Cohen (1988) a respeito do chamado *Effect Size* (ES). Assim, um erro aleatório pequeno é aquele cujo desvio padrão é $0,2 \times \sigma_y$, sendo que σ_y representa o desvio padrão da variável resposta (y) antes de ser acrescido o erro aleatório. Enquanto que um erro aleatório grande apresenta um desvio igual a $0,8 \times \sigma_y$.

Existem diversas medidas de erro disponíveis na literatura que tem como objetivo avaliar o desempenho de modelagem de métodos estatísticos. São exemplos dessas medidas: Coeficiente de determinação (R^2), Raiz do Erro Quadrático Médio (RMSE), Erro Percentual Absoluto Médio (MAPE), Erro Quadrático Médio (MSE), Raiz do Erro Quadrático Médio Relativo (RRMSE), dentre outras. O coeficiente de determinação (R^2) foi a métrica escolhida para avaliar o desempenho dos métodos estudados na presente pesquisa. A fórmula matemática utilizada para calcular o R^2 está representada na Equação 21. O valor excelente para essa métrica é 1.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (21)$$

Onde:

n : Tamanho do conjunto de dados

\hat{y}_i : Valores ajustados (modelados)

y_i : Valores observados

\bar{y} : Média de y_i

4.2.3 Escolha do projeto experimental

Considerando os cinco fatores apresentados na Tabela 2, foi possível gerar o DOE utilizando o *software* Minitab. O arranjo utilizado foi do tipo 2_V^{5-1} com duas réplicas, que é um fatorial fracionado com resolução V, resultando em 32 experimentos representados na Tabela 3.

Como esses experimentos serão utilizados apenas para a simulação de dados, não foi necessário randomizá-los. A matriz de *design* é um guia que indicará a combinação de fatores utilizados para gerar os bancos de dados que serão modelados

através dos métodos de regressão não linear e redes neurais artificiais.

Sendo assim, de acordo com cada linha do arranjo experimental, foi possível gerar um conjunto de dados sintéticos contendo diferentes níveis para cada um dos fatores. Por exemplo, o ‘Banco de dados 1’ é composto por um conjunto de dados de 80 valores, 2 variáveis de entrada, utilizou a função *Rosenbrock* para gerar os valores da resposta (y), a correlação entre as variáveis de entrada foi de 0,35 e um grande efeito representado por $0,8 \times \sigma_y$.

<i>Run</i>	<i>N</i>	x_i	<i>f</i>	ρ	σ
1	80	2	Rosenbrock	0,35	$0,8 * \sigma_y$
2	320	2	Rosenbrock	0,35	$0,2 * \sigma_y$
3	80	4	Rosenbrock	0,35	$0,2 * \sigma_y$
4	320	4	Rosenbrock	0,35	$0,8 * \sigma_y$
5	80	2	Ackley N.4	0,35	$0,2 * \sigma_y$
6	320	2	Ackley N.4	0,35	$0,8 * \sigma_y$
7	80	4	Ackley N.4	0,35	$0,8 * \sigma_y$
8	320	4	Ackley N.4	0,35	$0,2 * \sigma_y$
9	80	2	Rosenbrock	0,90	$0,2 * \sigma_y$
10	320	2	Rosenbrock	0,90	$0,8 * \sigma_y$
11	80	4	Rosenbrock	0,90	$0,8 * \sigma_y$
12	320	4	Rosenbrock	0,90	$0,2 * \sigma_y$
13	80	2	Ackley N.4	0,90	$0,8 * \sigma_y$
14	320	2	Ackley N.4	0,90	$0,2 * \sigma_y$
15	80	4	Ackley N.4	0,90	$0,2 * \sigma_y$
16	320	4	Ackley N.4	0,90	$0,8 * \sigma_y$
17	80	2	Rosenbrock	0,35	$0,8 * \sigma_y$
18	320	2	Rosenbrock	0,35	$0,2 * \sigma_y$
19	80	4	Rosenbrock	0,35	$0,2 * \sigma_y$
20	320	4	Rosenbrock	0,35	$0,8 * \sigma_y$
21	80	2	Ackley N.4	0,35	$0,2 * \sigma_y$
22	320	2	Ackley N.4	0,35	$0,8 * \sigma_y$
23	80	4	Ackley N.4	0,35	$0,8 * \sigma_y$
24	320	4	Ackley N.4	0,35	$0,2 * \sigma_y$

25	80	2	Rosenbrock	0,90	$0,2 * \sigma_y$
26	320	2	Rosenbrock	0,90	$0,8 * \sigma_y$
27	80	4	Rosenbrock	0,90	$0,8 * \sigma_y$
28	320	4	Rosenbrock	0,90	$0,2 * \sigma_y$
29	80	2	Ackley N.4	0,90	$0,8 * \sigma_y$
30	320	2	Ackley N.4	0,90	$0,2 * \sigma_y$
31	80	4	Ackley N.4	0,90	$0,2 * \sigma_y$
32	320	4	Ackley N.4	0,90	$0,8 * \sigma_y$

Tabela 3 - Arranjo Experimental

4.2.4 Modelagem

Na etapa de modelagem, o modelo conceitual é convertido no modelo computadorizado. Os modelos computadorizados neste estudo são os bancos de dados gerados a partir da matriz de experimentos representada na Tabela 3. Para cada teste (*Run*) foi gerado um banco de dados utilizando o *software* Minitab, resultando em 32 bancos de dados no total. No Apêndice A, está representado um exemplo de como esses bancos de dados foram gerados. Este exemplo, denominado “Banco de dados 2” é resultado da combinação dos níveis dos fatores definidos pelo experimento 2, sendo eles: conjunto de dados de 320 valores, 2 variáveis de entrada, utilizou a função *Rosenbrock* para gerar os valores da resposta (y), a correlação entre as variáveis de entrada foi de 0,35 e um pequeno efeito representado por $0,2 \times \sigma_y$.

4.2.5 Solução dos modelos

Após gerar o conjunto de dados sintéticos, com os valores das variáveis de entrada (x_i) e suas respectivas respostas (y_{real}) para cada um dos cenários determinados pelo DOE, foi possível fazer a comparação entre o desempenho dos dois métodos. Para isso, foi feita a modelagem das respostas (y_{real}) obtidas através de cada um dos 32 bancos de dados utilizando os métodos de RNL e RNAs. A modelagem feita pelo método de regressão não linear utilizou o *software* Minitab e seu passo a passo está devidamente detalhado no Apêndice B. Já a modelagem feita pela técnica de redes neurais artificiais utilizou o *software* Statistica e seguiu as etapas descritas no Apêndice C. Os resultados de modelagem desses métodos foram representados nos bancos de dados através das colunas “ y_{RNL} ” e “ y_{RNA} ”, conforme mostra o exemplo no Apêndice A.

4.2.6 Análise estatística dos resultados

A análise estatística dos resultados é a etapa onde é realizada a modelagem matemática das respostas, análise de resíduos, adequação e ajuste dos modelos.

Após a modelagem das respostas dos 32 bancos de dados através dos métodos de regressão não linear e redes neurais artificiais, foi possível calcular os valores de R^2 , a fim de verificar a precisão de modelagem desses métodos. O cálculo dessa métrica foi feito através da Equação 21, conforme foi descrito no item 4.2.2, onde os valores observados (y_i) correspondem a variável resposta (y_{real}) e os valores ajustados (\hat{y}_i) correspondem aos resultados de modelagem (y_{RNL} e y_{RNA}). As respostas estão descritas na Tabela 4.

<i>Run</i>	<i>Respostas</i>	
	<i>R²RNL</i>	<i>R²RNA</i>
1	68,36%	66,49%
2	96,16%	93,31%
3	97,12%	93,51%
4	61,57%	57,29%
5	96,58%	94,61%
6	60,91%	57,30%
7	62,69%	43,43%
8	96,07%	88,16%
9	95,94%	96,09%
10	62,53%	61,86%
11	67,92%	52,78%
12	96,54%	95,01%
13	57,67%	62,71%
14	96,69%	94,86%
15	96,19%	88,13%
16	59,32%	57,64%
17	60,98%	58,95%
18	96,33%	92,50%
19	96,80%	85,92%
20	63,13%	59,26%
21	96,20%	92,96%
22	59,10%	56,62%
23	64,55%	49,01%
24	96,45%	81,17%
25	97,17%	96,78%
26	54,29%	53,87%
27	64,47%	59,95%
28	96,50%	92,83%

29	66,95%	66,27%
30	95,86%	91,48%
31	96,14%	90,87%
32	66,41%	63,06%

Tabela 4 - Respostas do DOE

A partir dos resultados observados na Tabela 4, é possível verificar qual dos métodos apresenta o melhor desempenho e estabelecer relações entre os parâmetros de entrada e as respostas analisadas.

Sendo assim, foi possível realizar o teste *paired-t*, no *software* Minitab, com intuito de verificar se a diferença entre os métodos de modelagem é de fato significativa com um nível de significância de $\alpha = 5\%$. A seguir, são apresentadas as hipóteses nula e alternativa, respectivamente, para esse teste.

H_0 : O desempenho da RNA **não** é significativamente diferente da RNL.

H_1 : O desempenho da RNA **é** significativamente diferente da RNL.

No teste *paired-t* a estatística normalmente utilizada é o valor de “*p-value*”. Sendo assim, quando seu valor é menor que o nível de significância α , a hipótese nula é rejeitada (MONTGMERY, 2005). Comparando as respostas R^2 RNL e R^2 RNA, o valor de *p-value* obtido foi de 0,00 ($< 0,05$), o que nos leva a rejeitar a hipótese nula e aceitar a hipótese alternativa de que os métodos possuem desempenhos diferentes, com nível de significância de 5%. Portanto, é possível concluir que para esse conjunto de dados a RNL superou as RNAs em termos de R^2 .

Esse resultado se justifica, pois, ao gerar o conjunto de dados foi estabelecida uma relação previamente conhecida entre as variáveis de entrada e a variável resposta. Nesse caso, como as respostas foram obtidas a partir das funções *Rosenbrock* e *Ackley N.4*, o pesquisador conhecia os modelos não lineares que forneciam o melhor ajuste ao conjunto de dados, o que favoreceu o desempenho de modelagem da regressão não linear.

Para garantir que esse teste tem um tamanho de amostra adequado para alcançar um poder de teste aceitável, realizamos o teste *Power and Sample Size for paired t* no *software* Minitab. Nesta análise, o objetivo é verificar se o teste *paired-t* pode detectar uma diferença de 0,05 entre as respostas (R^2 RNL e R^2 RNA) para uma amostra de 32 valores. Um valor de poder de 0,90 é geralmente considerado adequado. A partir da figura 9 podemos verificar que, o valor de poder obtido foi 0,99, indicando uma chance de 99%

de detectar uma diferença entre as duas respostas analisadas quando essa diferença realmente existe.

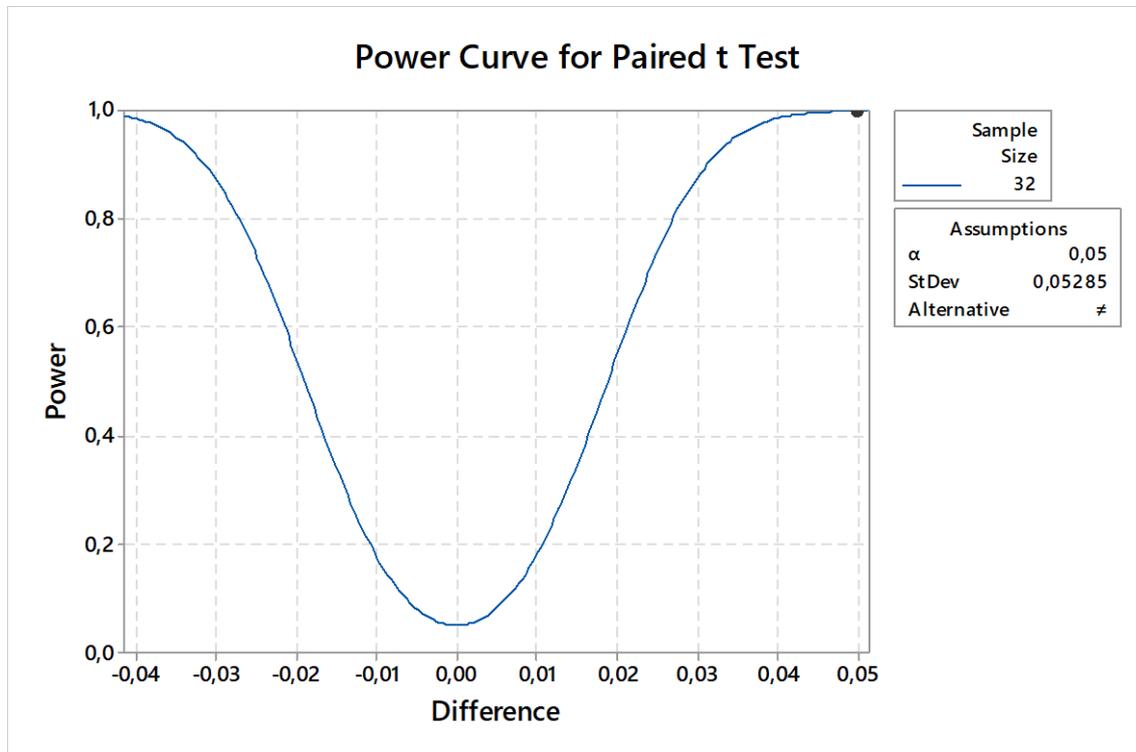


Figura 9 - Curva do Power and Sample Size for paired t

A fim de compreender quais os fatores e interações influenciam o desempenho de cada método, utilizou-se o *software* Minitab para fazer a análise das respostas do DOE, representadas na Tabela 4. Dessa forma, foi possível estabelecer relações matemáticas entre as respostas analisadas e os parâmetros de entrada usando o Método dos Mínimos Quadrados Ordinários (OLS) e a Análise de Variância (ANOVA). Os coeficientes estimados no OLS estão indicados na Tabela 5, onde os coeficientes em negrito indicam termos significativos (valor de $p < 0,05$). Os coeficientes são codificados pelas letras A, B, C, D e E. Todas as análises a seguir consideraram um nível de significância de 5%.

<i>Coeficientes</i>		<i>Respostas</i>	
<i>Fator</i>	<i>Nome</i>	<i>RNL</i>	<i>RNA</i>
Constante	Constante	0,795	0,748
A	N	-0,009	-0,001
B	x_i	0,006	-0,025
C	f	-0,003	-0,012
D	ρ	-0,001	0,017
E	σ	-0,169	-0,169
AB	$N \times x_i$	0,003	0,020
AC	$N \times f$	0,005	0,002
AD	$N \times \rho$	0,000	-0,001
AE	$N \times \sigma$	-0,008	0,005
BC	$x_i \times f$	-0,001	-0,010
BD	$x_i \times \rho$	0,004	0,010
BE	$x_i \times \sigma$	0,006	-0,001
CD	$f \times \rho$	0,002	0,016
CE	$f \times \sigma$	-0,001	0,003
DE	$\rho \times \sigma$	0,000	0,002

Tabela 5 - Coeficientes estimados

Em seguida, foi realizada a Análise de Variância (ANOVA). Essa análise também foi realizada para cada resposta, R^2 RNL e R^2 RNA conforme mostram as Figuras 10 e 11, respectivamente. Os resultados da ANOVA indicam que todos os modelos tiveram bons ajustes, indicando grande confiabilidade e preditividade.

Analysis of Variance						
Source	DF	Adj SS	Adj MS	F-Value	P-Value	
Model	15	0,926280	0,061752	69,96	0,000	
Linear	5	0,921556	0,184311	208,82	0,000	
N	1	0,002427	0,002427	2,75	0,117	
Xi	1	0,001269	0,001269	1,44	0,248	
f	1	0,000202	0,000202	0,23	0,639	
ρ	1	0,000018	0,000018	0,02	0,888	
σ	1	0,917640	0,917640	1039,65	0,000	
2-Way Interactions	10	0,004725	0,000472	0,54	0,841	
N*Xi	1	0,000205	0,000205	0,23	0,637	
N*f	1	0,000756	0,000756	0,86	0,369	
N*ρ	1	0,000002	0,000002	0,00	0,965	
N*σ	1	0,001920	0,001920	2,18	0,160	
Xi*f	1	0,000061	0,000061	0,07	0,795	
Xi*ρ	1	0,000498	0,000498	0,56	0,463	
Xi*σ	1	0,001057	0,001057	1,20	0,290	
f*ρ	1	0,000189	0,000189	0,21	0,650	
f*σ	1	0,000033	0,000033	0,04	0,848	
ρ*σ	1	0,000003	0,000003	0,00	0,951	
Error	16	0,014122	0,000883			
Total	31	0,940403				

Model Summary				
S	R-sq	R-sq(adj)	R-sq(pred)	
0,0297093	98,50%	97,09%	93,99%	

Figura 10 - ANOVA RNL

Analysis of Variance						
Source	DF	Adj SS	Adj MS	F-Value	P-Value	
Model	15	0,978200	0,065213	54,40	0,000	
Linear	5	0,949891	0,189978	158,46	0,000	
N	1	0,000016	0,000016	0,01	0,910	
Xi	1	0,019326	0,019326	16,12	0,001	
f	1	0,004541	0,004541	3,79	0,069	
ρ	1	0,009012	0,009012	7,52	0,014	
σ	1	0,916997	0,916997	764,89	0,000	
2-Way Interactions	10	0,028309	0,002831	2,36	0,061	
N*Xi	1	0,012752	0,012752	10,64	0,005	
N*f	1	0,000146	0,000146	0,12	0,731	
N*ρ	1	0,000043	0,000043	0,04	0,853	
N*σ	1	0,000888	0,000888	0,74	0,402	
Xi*f	1	0,003208	0,003208	2,68	0,121	
Xi*ρ	1	0,003069	0,003069	2,56	0,129	
Xi*σ	1	0,000068	0,000068	0,06	0,815	
f*ρ	1	0,007756	0,007756	6,47	0,022	
f*σ	1	0,000270	0,000270	0,23	0,641	
ρ*σ	1	0,000108	0,000108	0,09	0,768	
Error	16	0,019182	0,001199			
Total	31	0,997382				

Model Summary				
S	R-sq	R-sq(adj)	R-sq(pred)	
0,0346247	98,08%	96,27%	92,31%	

Figura 11 - ANOVA RNA

Posteriormente, para garantir que o modelo de regressão obtido pelo OLS tem um bom ajuste ao conjunto de dados original, foi realizada a análise de resíduos. Segundo Montgomery (2005), os resíduos devem ser não correlacionados (C), normais (N) e distribuídos aleatoriamente (R).

Na primeira análise, os resultados mostraram que todos os métodos apresentaram resíduos não correlacionados (baixo valor de *Pearson* e valor de $p > 0,05$). Já os resíduos da RNL não apresentaram distribuição normal (valor de $p < 0,05$ e *Anderson-Darling (AD)* > 1), o que implica em intervalos de confiança imprecisos e valores de p imprecisos. Os resíduos da RNA também não apresentaram distribuição aleatória, apresentando causas especiais de “Mistura” e “Oscilação” (valor de $p < 0,05$).

Portanto, a fim de atingir as condições residuais ideais e garantir que o modelo ajuste bem os dados, foi necessário realizar uma verificação da adequação dos modelos através da redução do modelo, ou seja, eliminando alguns de seus termos e/ou interações não significativos. A Tabela 6 apresenta a análise residual, a coluna “antes” representa a primeira análise e a coluna “depois” representa o modelo reduzido.

	<i>Método</i>	<i>RNL</i>		<i>RNA</i>	
		<i>antes</i>	<i>depois</i>	<i>antes</i>	<i>depois</i>
(N)	<i>AD</i>	1,470	0,703	0,351	0,270
	<i>P-value</i>	< 0,005	0,059	0,447	0,654
(C)	<i>Pearson</i>	-0,012	0,045	-0,129	-0,109
	<i>P-value</i>	0,949	0,816	0,481	0,553
(R)	<i>Agrupamento</i>	0,764	0,717	0,984	0,764
	<i>Mistura</i>	0,236	0,283	0,016	0,236
	<i>Tendência</i>	0,667	0,325	0,985	0,902
	<i>Oscilação</i>	0,333	0,675	0,015	0,098

Tabela 6 - Análise de resíduos

Em seguida, foi possível concluir que todos os modelos apresentaram bons ajustes, uma vez que os valores de R^2 (*adj.*) e R^2 (*pred.*) foram maiores que 80%, indicando grande confiabilidade e previsibilidade, como mostra a Tabela 7.

<i>Método</i>	<i>RNL</i>	<i>RNA</i>
R^2 (<i>adj.</i>) (%)	98,45	96,51
R^2 (<i>pred.</i>) (%)	98,04	95,19

Tabela 7 - Respostas do ajuste do modelo de regressão

Assim, os modelos apresentam os formatos descritos pelas Eqs. 22 e 23.

$$R^2 NLR = 0,7927 - 0,0001 A + 0,0060 B - 0,1550 E - 0,0001 AE \quad (22)$$

$$R^2 ANN = 0,8849 - 0,0005 A - 0,0579 B - 0,0473 C + 0,0610 D - 0,1693 E + 0,0002 AB + 0,0566 CD \quad (23)$$

4.2.7 Interpretação dos resultados

Tendo em vista a modelagem matemática desenvolvida na etapa anterior será possível descrever o comportamento das variáveis, a relação entre elas e estimar os efeitos produzidos nas respostas observadas. Sendo assim, nesta etapa será feita a interpretação dos resultados obtidos através da análise do comportamento das respostas devido às alterações nos parâmetros de entrada.

Para melhor compreender o comportamento dos termos indicados nas Equações 22 e 23, geramos os Gráficos de Pareto, Figuras 12 e 13, com intuito de visualizar os fatores significativos de cada método.

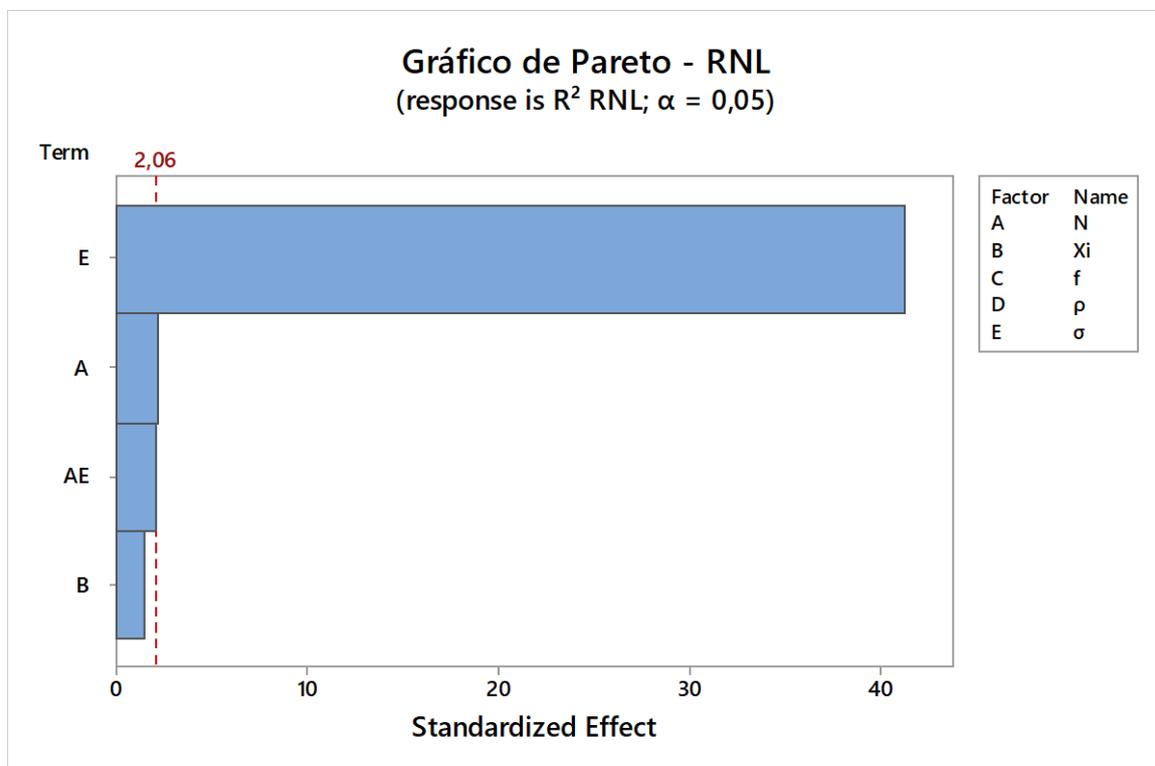


Figura 12 – Gráfico de Pareto para R^2 RNL

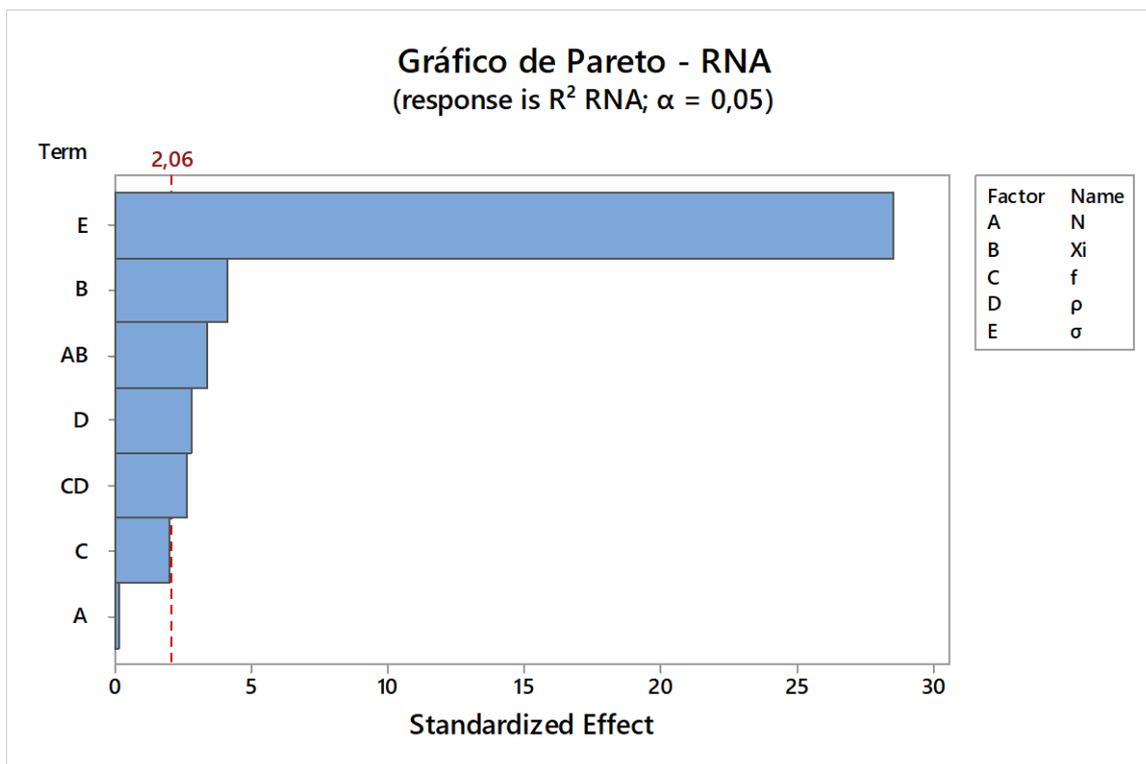
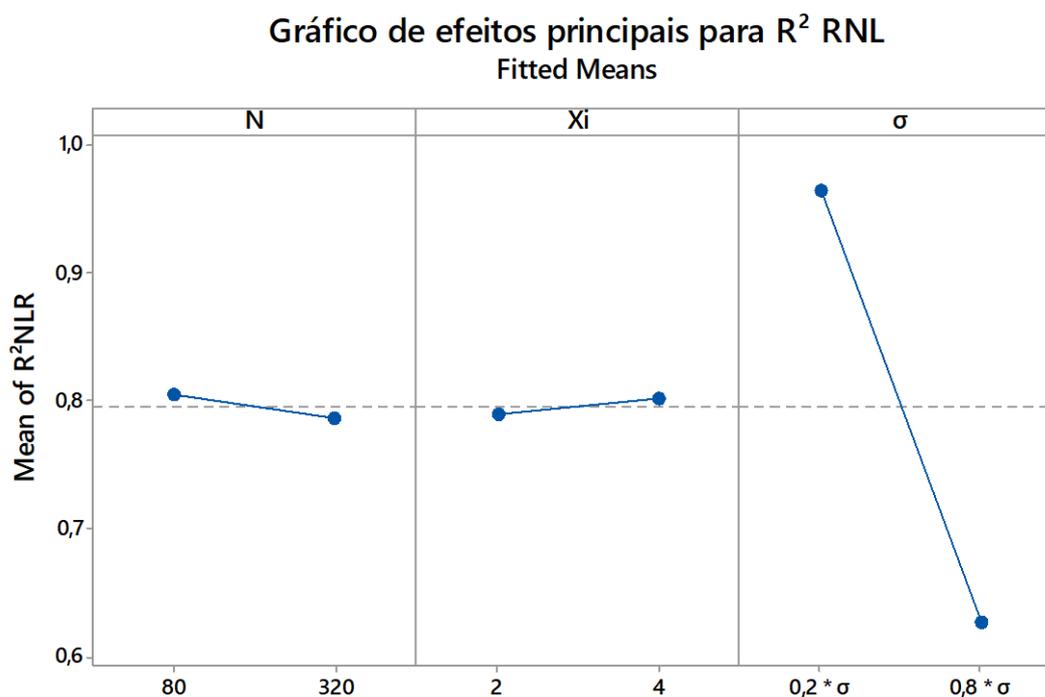
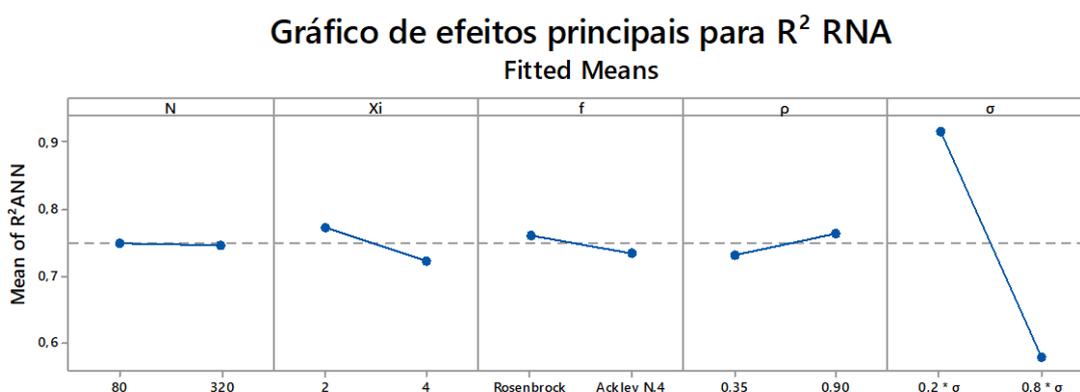


Figura 13 - Gráfico e Pareto para R^2 RNA

A fim de entender como esses fatores significativos afetam o desempenho dos métodos de modelagem, foram elaborados os gráficos de efeitos principais para todas as respostas analisadas (Figuras 14 e 15).

Então, a partir das Figuras 14 e 15, foi possível inferir que o desvio padrão do erro aleatório adicionado a variável resposta (E) com um pequeno efeito implica em um aumento no desempenho de ambos os métodos. A RNL teve um aumento de 62,65% para 96,45% e a RNA de 57,90% para 91,76%. Um menor número de variáveis de entrada (B) implica em um aumento no desempenho da RNA e não é significativo para a RNL. O desempenho da RNA aumentou de 72,38% para 77,29%. A correlação entre as variáveis de entrada (D) é significativa apenas para RNA, onde uma alta correlação produz melhores resultados para esse método, com uma taxa de recuperação de 76,51%.

De acordo com Yagiz, Sezer e Gokceoglu (2012), a superioridade das RNAs aumenta à medida que a dimensionalidade do conjunto de dados aumenta. Porém, neste estudo, devido aos parâmetros selecionados, o tamanho da amostra (A) não foi um fator significativo para as RNAs. Já para a RNL, esse fator foi significativo, visto que a menor dimensionalidade do conjunto de dados implicou em um aumento no desempenho desse método de 78,65% para 80,46%, conforme mostra a Figura 14.

Figura 14 – Gráfico de Efeitos principais para R² RNLFigura 15 - Gráfico de Efeitos principais para R² RNA

Além da influência dos fatores, a influência das interações AB e CD também é significativa para o desempenho da RNA. Conforme foi mostrado na Figura 16, o melhor desempenho da RNA ocorre quando o conjunto de dados possui 80 valores e o número de variáveis de entrada é igual a 2. Então, o valor médio mais baixo (70,45%) e o mais alto (79,36%) ocorrem quando B é igual a 4 e 2, respectivamente, neste contexto. Além disso, o desempenho da RNA aumenta quando a função usada para gerar y é a *Ackley N.4* e a correlação entre as variáveis de entrada é alta, como mostra a Figura 17. Assim,

o valor médio mais baixo (70,41%) e o mais alto (76,88 %) ocorrem quando D é igual a 0,35 e 0,90, respectivamente.

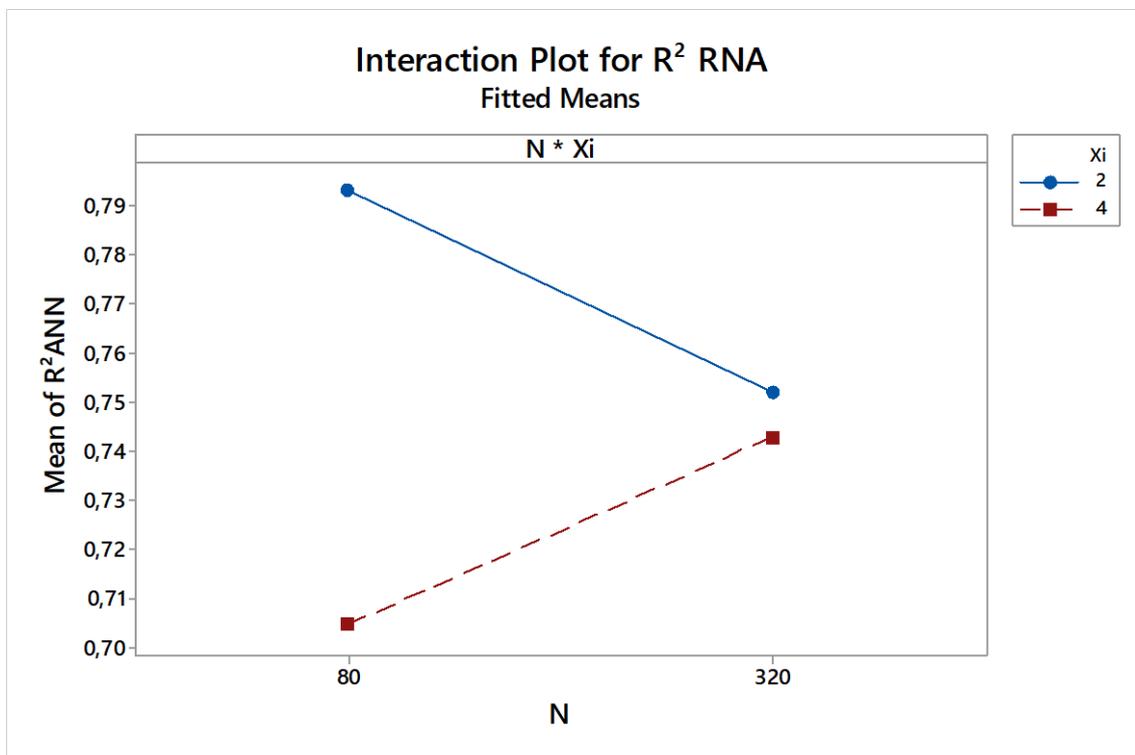


Figura 16 - Gráfico de Iterações (AB) para R^2 RNA

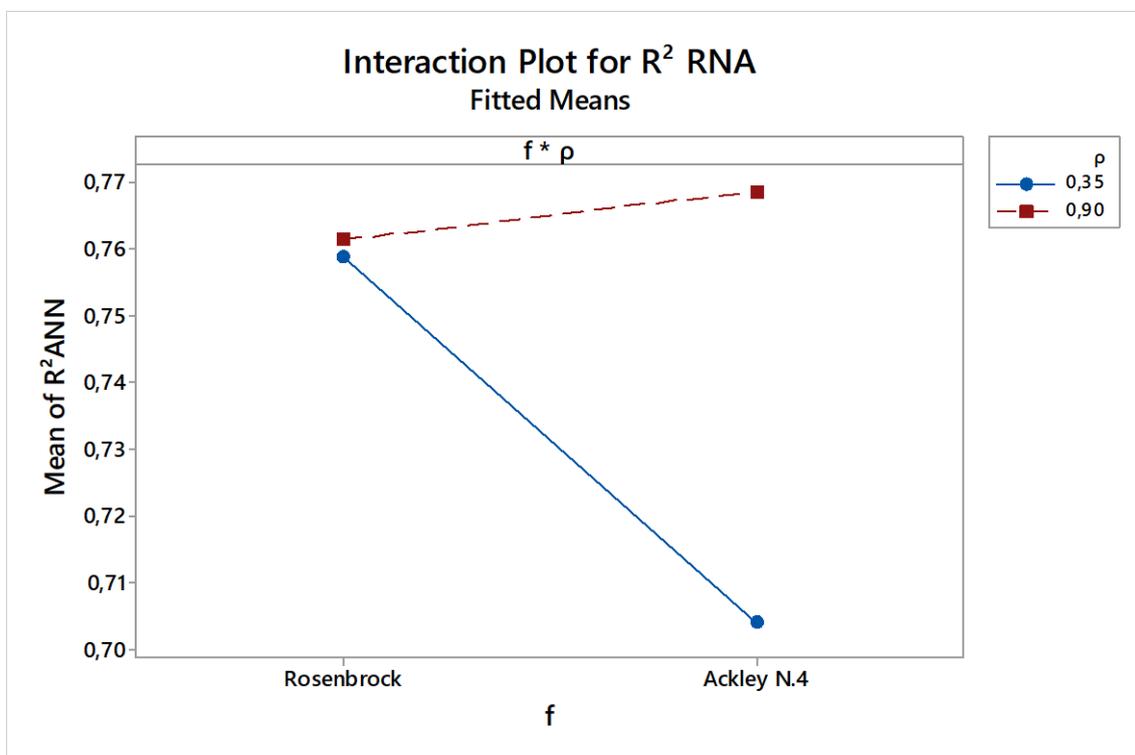


Figura 17 - Gráfico de Iterações (CD) para R^2 RNA

5. CONCLUSÕES E RECOMENDAÇÕES

A modelagem de dados é uma técnica muito importante e amplamente utilizada em diversos tipos de sistemas do mundo real, podendo ser feita através de vários tipos de métodos. No entanto, a escolha do melhor método a ser utilizado para cada tipo de problema nem sempre é uma tarefa fácil. Já existem alguns estudos disponíveis na literatura que fazem a comparação entre a precisão da modelagem a partir de técnicas estatísticas clássicas e redes neurais artificiais. Entretanto, estes estudos fazem essa comparação em apenas um tipo de cenário específico, o que impossibilita concluir que o melhor método neste cenário será o melhor em outros contextos.

Sendo assim, devido a essa necessidade de generalização, o presente estudo desenvolveu uma análise comparativa entre os métodos de regressão não linear e redes neurais artificiais com o objetivo de avaliar o desempenho de modelagem desses métodos em uma grande variedade de cenários.

Essa análise foi baseada em conjuntos de dados sintéticos, gerados através da técnica de planejamento de experimentos (DOE), cujos fatores de influência definidos foram: tamanho do conjunto de dados, número de variáveis de entrada, tipo de função não linear usada para gerar a variável resposta, valor da correlação entre as variáveis de entrada e desvio padrão do erro aleatório adicionado a variável resposta. A modelagem pelo método de RNL foi feita através do *software* Minitab e, para o método de RNAs, utilizou-se o *Intelligent Problem Solver* (IPS) do *software* Statistica. A precisão de modelagem desses métodos foi comparada através do coeficiente de determinação (R^2) e de técnicas estatísticas.

Embora ambos os métodos de modelagem tenham apresentado um bom ajuste para os dados analisados, a partir da análise das respostas dos 32 experimentos, foi possível concluir que a regressão não linear forneceu resultados relativamente mais precisos em comparação com as redes neurais artificiais, em termos de R^2 . Esse resultado é justificado porque ao gerar os conjuntos de dados foi estabelecida uma relação previamente conhecida (a partir das funções *Rosenbrock* e *Ackley N.4*) entre as variáveis de entrada e a variável resposta o que favoreceu o desempenho da RNL. Portanto, os resultados mostraram que o desempenho dos métodos de RNL e RNAs tem uma diferença estatística significativa, sendo possível afirmar que a técnica de RNL superou as RNAs nesse contexto.

Além disso, a partir da análise das respostas do DOE foi possível concluir que há uma melhora no desempenho dos dois métodos quando o erro aleatório adicionado à variável resposta possui um pequeno efeito. Já o número de variáveis de entrada e o valor da correlação entre elas são fatores significativos apenas para a RNA e, o tamanho do conjunto de dados é significativo apenas para a RNL. Além da influência dos fatores, a influência das interações AB e CD também é significativa para o desempenho da RNA, ou seja, o melhor desempenho desse método ocorre quando o conjunto de dados possui 80 valores e o número de variáveis de entrada é igual a 2 e, quando a função usada para gerar a variável resposta é a *Ackley N.4* e a correlação entre as variáveis de entrada é alta.

Esse estudo difere dos outros principalmente no que diz respeito ao método utilizado para generalizar e simular os conjuntos de dados que permitiu criar contextos diferentes nos quais os métodos puderam ser aplicados. Além disso, o uso da técnica de DOE realiza uma análise mais profunda sobre como os fatores e interações significativos podem influenciar a precisão de modelagem dos dois métodos e apoia estatisticamente o pesquisador na escolha do método de modelagem mais adequado às características do cenário em que o estudo será desenvolvido. Outra diferença é a aplicação do método de RNL que foi implementado usando o *software* Minitab e a RNA que utilizou o *software* Statistica com os parâmetros de treinamento cuidadosamente definidos.

Portanto, os dois métodos analisados são ferramentas úteis para a modelagem de dados não lineares, desde que o pesquisador conheça as principais características do conjunto de dados e aplique o método corretamente. No entanto, o método de RNAs é uma abordagem alternativa eficaz para a regressão não linear, fornecendo uma modelagem precisa para o conjunto de dados sem a necessidade de que sejam feitas suposições a priori sobre as relações funcionais entre as variáveis de entrada e saída.

Como trabalhos futuros, propõem-se: um estudo comparativo entre redes neurais artificiais e outros métodos estatísticos clássicos usando a técnica de planejamento de experimentos; um estudo comparativo entre RNL e RNAs usando outras métricas que não o coeficiente de determinação ajustado como por exemplo a Raiz do Erro Quadrático Médio (RMSE), o Erro Percentual Absoluto Médio (MAPE), o Erro Quadrático Médio (MSE) e a Raiz do Erro Quadrático Médio Relativo (RRMSE); e um estudo de caso usando o resultado obtido no presente trabalho.

REFERÊNCIAS

ADAMOWSKI, J.; CHAN, H. F.; PRASHER, S. O.; OZGA-ZIELINSKI, B.; SLIUSARIEVA, A. Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. **Water Resources Research**, v. 48, n. 1, p. 1–14, 2012.

AMEDI, H. R.; BAGHBAN, A.; AHMADI, M. A. Evolving machine learning models to predict hydrogen sulfide solubility in the presence of various ionic liquids. **Journal of Molecular Liquids**, n. 216, p. 411–422, 2016.

APPOLINÁRIO, F. **Metodologia da ciência: filosofia e prática da pesquisa**. São Paulo: Pioneira Thomson Learning, 2006.

BALESTRASSI, P. P.; POPOVA, E.; PAIVA, A. P.; LIMA, J. W. M. Design of experiments on neural network's training for nonlinear time series forecasting. **Neurocomputing**, v. 72, p. 1160-1178, 2009.

BASHEER, I. A.; HAJMEER, M. Artificial neural networks : fundamentals , computing, design , and application. **Journal of Microbiological Methods** v. 43, p. 3–31, 2000.

BATES, D. M.; WATTS, D. G. **Nonlinear Regression Analysis and Its Applications**, John Wiley & Sons, 1988.

BENBOURAS, M. A. *et al.* Comparing nonlinear regression analysis and artificial neural networks to predict geotechnical parameters from standard penetration test. **Urbanism. Architecture. Constructions**, v. 9, n.3, p. 275-288, 2018.

BERTRAND, J. W. M., FRANSOO, J. C. Modelling and Simulation: operations management research methodologies using quantitative modeling. **International Journal of Operations & Production Management**, v. 22, p. 241-264, 2002.

CHUNG, C. A. **Simulation Modeling Handbook: a practical approach**. Washington, D. C: CRC Press, 2004.

COHEN, J. **Statistical Power Analysis for the Behavioral Sciences**, 2a.ed., Lawrence Erlbaum Associates, Publishers, New York, 1988.

ESFE, M. H.; AFRAND, M.; YAN, W.; AKBARI, M. Applicability of artificial neural network and nonlinear regression to predict thermal conductivity modeling of Al₂O₃-water nanofluids using experimental data. **International Communications in Heat and Mass Transfer**, v. 66, p. 246-249, 2015.

GOMES, J. H. F. **Análise e otimização da soldagem de revestimento de chapas de aço ABNT 1020 com utilização de arame tubular inoxidável austenítico**. Dissertação submetida ao Programa de Pós Graduação em Engenharia de Produção, Universidade Federal de Itajubá, Itajubá, 2010.

HATUA, P.; MAJUMDAR, A.; DAS, A. Predicting the ultraviolet radiation protection by polyester–cotton blended woven fabrics using nonlinear regression and artificial neural network models. **Photodermatology, Photoimmunology & Photomedicine**, v. 29, p. 182-189, 2013.

HAYKIN, S. **Neural Networks and Learning Machines**. 3a. ed. Prentice Hall, 2008.

JAIN, A. K.; MAO, J. Artificial Neural Networks: A Tutorial. **IEEE**, v. 29, n. 3, p. 31-44, 1996.

JONES, S. S.; EVANS, R. S.; ALLEN, T. L.; THOMAS, A.; HAUNG, P. J.; WELCH, S. J.; SNOW, G. L. A multivariate time series approach to modeling and forecasting demand in the emergency department. **Journal of Biomedical Informatics**, v. 42, n. 1, p. 123-139, 2009.

KLEIJNEN, J. P. C.; SANCHEZ, S. M.; LUCAS, T. W. CIOPPA, T. M. A User's Guide to the Brave New World of Designing Simulation Experiments. **Journal on Computing**, v. 17, n. 3, p. 263-289, 2005.

KHATIBI, R. ; GHORBANI, M. A. ; NAGHIPOUR, L.; JOTHIPRAKASH, V. ; FATHIMA, T. A. ; FAZELIFARD, M. H. Inter-comparison of time series models of lake levels predicted by several modeling strategies. **Journal of Hydrology**, v. 511, p. 530-545, 2014.

KÖCHE, J. C. **Fundamentos de metodologia científica: teoria da ciência e iniciação à pesquisa**. 26. ed. Petrópolis, RJ: Vozes, 2013.

MAHMOODI, M.; NADERI, A. Applicability of Artificial Neural Network and Nonlinear Regression to Predict Mechanical Properties of Equal Channel Angular Rolled Al5083 Sheets. **Latin American Journal of Solids and Structures**, v. 13, n. 8, p. 1515-1525, 2016.

MATTOS, T. B. **Modelos Não Lineares e suas Aplicações**. Monografia apresentada ao Curso de Estatística, Universidade Federal de Juiz de Fora, 2013.

MAZUCHELI, J.; ACHCAR, J. A. Algumas considerações em regressão não linear. **Acta Scientiarum**, v. 24, n. 6, p. 1761-1770, 2002.

MELLO, C. H. P. et al. Pesquisa-ação na engenharia de produção: proposta de uma estruturação para sua condução. **Produção**, v. 22, n. 1, p. 1-13, 2012.

MIGUEL, P. A. C.; FLEURY, A.; MELLO, C. H. P.; NAKANO, D. N.; TURRIONI, J. B.; LEE HO, L.; MORABITO, R.; MARTINS, R. A.; PUREZA, V. **Metodologia de pesquisa em engenharia de produção e gestão de operações**. 2ª. Ed. Rio de Janeiro: Elsevier, 2014.

MITROFF I. I., BETZ F., PONDY L. R., SAGASTI F. On managing science in the system age: two schemas for the study of science as a whole system phenomenon. **Interfaces**, v.4, n.3, p.46-58, 1974.

MONJEZI, M.; GHAFURIKALAJAHI, M.; BAHRAMI, A. Prediction of blast-induced ground vibration using artificial neural networks. **Tunnelling and Underground Space Technology**, v. 26, p. 46-50, 2011.

MONTGOMERY, D. C. **Design and Analysis of Experiments**. 6a. ed. John Wiley & Sons, 2005.

MONTGOMERY, D. C.; RUNGER, G. **Applied Statistics and Probability for Engineers**. 5a. ed. John Wiley & Sons, 2011.

MORESI, E.A.D. **Apostila de metodologia da pesquisa**. Brasília: Universidade Católica de Brasília, 2003.

ÖZÇELİK, R.; DIAMANTOPOULOU, M. J.; CRESCENTE-CAMPO, F.; ELER, U. Estimating Crimean juniper tree height using nonlinear regression and artificial neural network models. **Forest Ecology and Management**, n. 306, p. 52-60, 2013.

RAZI, M. A.; ATHAPPILY, K. A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models. **Expert Systems with Applications**, v. 29, n. 1, p. 65-74, 2005.

STAICULESCU, D.; BUSHYAGER, N.; OBATOYINBO, A.; MARTIN, L. J.; TENTZERIS, M. M., Design and Optimization of 3-D Compact Stripline and Microstrip Bluetooth/WLAN Balun Architectures Using the Design of Experiments Technique, **IEEE Transactions on Antennas and Propagation**, v.53, n.5, p. 1805–1812, May 2005.

VINOTH, B.; RAJARATHIAN, A.; BARGAVI, S. K. M. Nonlinear Regression and Artificial Neural Network Based Model for Forecasting Paddy (*Oryza Sativa*) Production in Tamil Nadu. **Journal of Mobile Computing & Application**, v. 3, n. 3, p. 01-06, 2016

WARNER, B.; MISRA, M. Understanding Neural Networks as Statistical Tools. **The American Statistician**, v. 50, n. 4, p. 284-293, 1996.

WERBOS, P. J.; Backpropagation Through Time: What It Does and How to Do It. **IEEE**, v. 78, n. 10, p. 1550-1560, 1990.

XU, M.; WONG, T. C.; CHIN, K. S. Modeling daily patient arrivals at Emergency Department and quantifying the relative importance of contributing variables using artificial neural network. **Decision Support Systems**, v. 54 , n. 3 , p. 1488-1498, 2013.

YAGIZ, S.; SEZER, E. A.; GOKCEOGLU, C. Artificial neural Networks and nonlinear regression techniques to assess the influence of slake durability cycles on the prediction of uniaxial compressive strength and modulus of elasticity for carbonate rocks. **International Journal for Numerical and Analytical Methods in Geomechanics**, v. 36, n. 14, p. 1636-1650, 2012.

YU, R.; LEUNG, P.; BIENFANG, P. Predicting shrimp growth: Artificial neural network versus nonlinear regression models. **Aquacultural Engineering**, v. 34, n. 1, p. 26-32, 2006.

APENDICE A – Banco de dados 2

<i>StdOrder</i>	<i>N</i>	<i>Xi</i>	<i>f</i>	<i>p</i>	<i>σ</i>
2	320	2	Rosenbrock	0,35	0,2 * σ _y

<i>Dados experimentais</i>		<i>Dados Modelados</i>			<i>Medidas de precisão</i>	
<i>x1</i>	<i>x2</i>	<i>yreal</i>	<i>yRNL</i>	<i>yRNA</i>	<i>R² RNL</i>	<i>R² RNA</i>
-0,80	-0,29	-44,18	89,84	104,71	96,16%	93,31%
-0,78	1,49	141,22	78,37	-202,12		
-0,96	0,61	342,25	12,25	-15,89		
-0,53	0,61	-66,00	12,11	-92,81		
-0,66	0,28	483,44	4,00	-33,66		
0,45	2,13	737,59	371,52	-5,66		
0,11	2,16	370,67	465,61	-130,20		
-0,29	-0,05	-60,98	2,47	-12,20		
-0,93	0,60	191,78	9,19	-24,82		
-0,01	0,21	441,65	4,60	-7,30		
-0,99	-1,39	563,51	570,58	504,09		
0,78	0,13	172,20	23,73	225,75		
-0,64	2,11	226,91	293,64	-284,92		
0,81	1,82	-107,00	136,75	158,88		
0,23	0,05	319,06	0,11	54,27		
-0,21	0,69	603,55	42,37	-83,26		
0,32	0,70	15,97	35,70	37,64		
-0,87	-0,71	-69,39	221,45	229,18		
0,63	1,11	229,92	50,58	123,28		
-0,70	-0,97	312,33	215,80	198,02		
-0,54	0,65	-784,55	14,29	-97,55		
1,00	0,44	239,91	31,49	302,53		
-2,53	0,25	4129,64	3810,53	3782,55		
-0,50	0,22	167,98	1,25	-42,22		
1,57	1,41	-20,45	115,93	539,81		
-0,19	-0,02	495,53	0,86	-9,97		
0,81	-0,27	422,88	86,77	253,40		
0,38	-1,44	477,80	251,98	182,36		
0,33	0,28	191,22	2,81	67,48		
1,67	-1,36	2006,09	1741,66	632,82		
0,12	-0,02	-394,32	0,29	34,99		
-2,37	-0,33	3504,37	3544,11	3651,52		
0,55	1,38	-318,76	117,42	74,48		
0,10	0,04	-122,43	0,32	25,49		
-0,49	2,51	-46,45	517,07	-312,03		
-1,00	0,48	338,02	28,91	24,59		
-0,27	0,34	585,92	8,19	-54,15		
1,53	-0,36	350,94	727,15	559,45		
1,30	0,78	18,44	83,90	425,94		
0,80	-0,29	-201,90	88,10	249,49		
0,47	0,01	177,50	4,33	123,75		
-2,05	-0,47	2315,83	2213,96	2303,44		
-0,53	-0,34	136,09	39,58	39,54		
1,43	-0,47	80,00	642,04	517,01		
-0,75	1,94	214,14	191,77	-266,97		
0,61	-1,03	226,43	198,41	217,02		
-0,93	-0,79	-12,18	278,17	283,72		
-0,02	-0,86	215,68	74,23	84,69		
0,39	0,80	-154,51	41,78	53,09		
-0,74	0,00	121,86	31,27	28,67		
-0,72	-0,03	-229,88	31,88	29,55		
-1,96	0,74	1059,16	981,93	1070,59		
1,45	1,29	552,75	66,62	481,91		
0,36	0,68	-331,30	30,17	51,21		
-0,53	2,31	281,46	416,14	-295,98		
-0,53	1,22	-81,59	90,61	-170,46		
-0,86	-1,80	744,42	655,23	507,22		
-2,20	0,26	1818,12	2097,51	2151,82		
-0,72	0,60	-77,44	2,33	-74,51		
0,23	-0,76	44,14	67,34	111,45		
-1,91	-0,07	1194,83	1397,69	1481,24		

<i>Dados Experimentais</i>		<i>Dados Modelados</i>		
<i>x1</i>	<i>x2</i>	<i>yreal</i>	<i>yRNL</i>	<i>yRNA</i>
0,13	0,61	-476,82	36,22	-10,98
0,30	-0,40	-20,51	24,49	101,06
-1,28	-1,54	1149,31	1020,25	943,17
0,12	0,13	-389,93	1,58	22,92
1,51	1,49	-195,75	63,13	504,38
0,86	0,22	199,26	27,14	251,52
-0,81	0,10	-599,23	32,89	31,79
2,47	1,17	2482,59	2447,25	1015,35
0,11	-0,57	20,88	34,19	76,34
0,41	1,59	97,15	202,62	10,87
-0,07	-0,23	-340,80	5,90	21,75
1,74	0,95	188,97	436,77	639,10
-0,17	-0,45	396,35	23,33	35,30
-0,80	-0,95	266,71	257,64	243,48
0,00	0,25	269,15	6,75	-10,34
0,03	0,89	463,89	80,36	-54,85
0,19	0,81	455,91	59,43	-7,22
2,55	0,70	3705,30	3364,92	1068,64
0,34	1,32	-49,75	147,42	2,77
1,45	-0,81	727,16	850,64	528,03
-1,26	-0,67	451,34	517,00	549,64
0,38	0,27	-266,49	1,63	81,00
1,88	0,57	638,74	891,67	721,35
-1,01	-0,53	-2,80	243,44	269,98
0,27	1,12	475,33	109,37	-5,21
0,68	0,95	190,26	23,87	150,41
-0,82	-1,01	-203,40	286,66	267,65
0,42	0,65	-118,36	22,13	73,06
1,34	0,77	182,21	107,59	446,99
-0,82	0,23	-61,13	21,46	10,65
0,31	0,05	37,66	0,28	75,72
1,54	-1,86	1745,53	1810,53	573,20
0,79	1,71	556,54	118,22	157,08
0,58	0,71	89,62	13,42	124,91
0,42	2,10	912,24	374,02	-17,94
-1,85	-0,29	1859,40	1403,25	1481,84
0,32	0,22	422,67	1,48	67,43
-0,99	0,46	-3,14	30,60	28,03
-0,36	-1,12	101,60	158,74	126,33
0,97	0,30	362,11	41,30	294,38
0,06	0,15	-160,08	2,50	9,84
-1,08	0,34	116,51	70,50	94,99
-2,93	-0,79	8538,49	8831,14	8779,97
1,33	0,09	394,75	283,61	459,38
1,23	0,75	-240,49	58,25	393,13
-0,83	0,79	48,37	2,77	-84,02
-0,79	1,97	-70,61	184,23	-270,56
0,34	0,80	-247,48	46,66	38,01
-2,09	-1,53	3903,84	3492,81	3729,41
1,29	0,86	-251,38	65,36	418,28
-0,25	-0,63	80,07	49,49	54,67
0,59	0,62	-461,29	7,34	132,03
-0,63	0,58	-59,64	4,92	-82,67
0,18	0,20	226,62	3,03	31,99
-0,23	-0,40	-94,40	20,81	27,99
0,37	-0,06	-410,02	4,07	98,35
-0,66	1,40	241,81	96,21	-195,44
0,94	-0,58	-265,13	214,68	310,68
-0,44	0,31	-328,17	2,24	-55,77
0,92	-0,64	-519,23	222,60	306,28
-0,09	0,09	22,85	1,09	-10,10
0,83	-0,14	-197,49	68,90	254,97
0,34	0,97	872,72	74,08	26,02
-1,75	-1,34	1481,12	1946,03	1974,27
-0,70	0,11	419,12	16,24	1,30
-0,83	-0,42	536,15	125,06	141,44
0,05	0,38	115,75	14,60	-11,05

<i>Dados Experimentais</i>		<i>Dados Modelados</i>		
<i>x1</i>	<i>x2</i>	<i>yreal</i>	<i>yRNL</i>	<i>yRNA</i>
-0,78	0,13	283,80	24,11	16,20
-0,35	-1,69	627,69	333,34	207,81
0,83	0,21	-62,57	23,81	242,02
-0,30	0,40	65,27	10,22	-62,31
-0,20	0,94	-463,28	81,45	-106,45
1,35	0,16	733,82	281,59	469,24
-0,67	0,07	151,70	16,48	2,38
-0,46	0,25	-317,52	1,17	-47,71
0,86	1,64	-152,29	82,33	189,58
-0,47	-1,06	-30,34	165,35	137,89
0,00	0,36	-118,64	13,15	-17,78
0,69	-0,43	326,50	81,96	215,52
-0,19	0,17	-441,94	2,43	-29,86
-1,10	-1,95	685,76	1010,27	840,44
-1,25	-0,70	836,46	522,68	551,90
-2,73	0,11	5238,47	5469,45	5359,95
-0,77	-0,14	223,23	54,98	62,74
-1,26	-0,23	128,34	333,68	391,98
0,25	-0,29	-481,91	12,90	83,63
-0,08	-0,54	-97,37	30,27	49,97
0,78	1,36	-37,37	55,19	171,95
-2,02	1,52	323,40	658,71	729,88
0,59	0,00	37,57	11,92	161,47
1,04	-0,43	-323,40	232,94	348,05
-0,66	-1,16	585,00	258,34	218,82
0,10	-0,16	237,58	3,26	41,54
1,04	-1,39	319,18	610,86	367,47
-1,34	0,16	558,06	273,97	346,32
-0,35	1,51	144,55	193,31	-185,10
1,04	1,93	-154,08	73,24	257,79
0,24	0,28	82,33	5,31	40,30
0,28	0,97	-36,48	80,05	7,42
-1,26	-1,83	675,22	1178,42	1047,24
0,87	1,04	292,36	8,80	220,92
-0,83	-1,63	234,37	540,06	427,29
-0,37	-0,20	749,63	12,50	5,90
-0,61	-0,83	-14,50	147,03	138,04
-0,43	-0,32	-121,72	27,50	25,92
1,09	-0,49	217,72	281,97	368,04
-0,07	-0,45	21,60	20,90	42,26
0,72	-0,32	-0,67	71,64	223,19
-1,82	-0,96	2020,25	1838,18	1894,57
-0,12	-0,06	483,96	1,01	-0,15
0,95	1,69	7,93	61,67	229,22
1,56	-0,11	486,01	646,83	571,04
-0,21	1,20	166,55	134,68	-133,00
-0,64	-0,53	-9,42	90,15	94,66
0,55	0,86	-741,71	32,00	102,99
-0,12	1,06	283,74	110,97	-102,41
-1,84	-1,28	3024,20	2195,70	2255,49
1,24	1,96	65,32	19,39	349,20
-0,15	-1,70	446,47	297,86	173,61
-1,49	-0,71	862,09	860,98	901,12
0,73	-1,40	-193,38	375,66	265,72
-0,97	-0,76	-56,89	291,95	301,19
-0,31	-1,49	515,19	252,74	168,19
-3,15	-1,73	13726,47	13683,07	13093,39
-2,06	-1,41	3586,78	3229,15	3419,60
-0,34	-0,49	79,29	37,37	40,25
-0,21	-0,69	294,60	54,89	60,75
-1,50	0,49	557,09	312,30	398,95
0,72	1,16	-204,91	42,09	153,68
0,96	0,10	77,65	69,25	298,16
-2,98	-1,65	10983,64	11224,48	11501,35
-0,42	-0,88	-330,92	113,46	101,54
0,29	1,47	308,76	192,76	-24,05
-1,77	-0,79	1525,09	1559,09	1613,99

<i>Dados Experimentais</i>		<i>Dados Modelados</i>		
<i>x1</i>	<i>x2</i>	<i>yreal</i>	<i>yRNL</i>	<i>yRNA</i>
-2,06	0,29	1378,98	1592,34	1666,09
-1,52	-1,23	1254,66	1272,42	1262,70
-0,07	-0,86	-26,61	76,43	82,70
-0,48	0,79	287,23	33,00	-116,34
-0,81	1,68	-493,32	109,14	-229,45
-0,02	-0,71	-124,25	51,71	71,53
-0,25	-1,24	339,66	169,79	126,74
-0,19	1,14	254,58	122,21	-123,06
0,74	0,78	-892,42	5,68	180,56
-0,92	-1,26	633,39	450,24	402,07
-1,32	0,81	-108,17	91,06	128,42
-1,60	0,49	286,10	435,58	531,77
-0,24	0,64	-196,12	34,33	-82,25
0,69	1,18	91,90	51,07	140,02
0,62	0,05	205,29	11,18	170,63
0,48	-0,06	-338,75	8,49	129,29
-0,16	0,06	-201,33	0,62	-15,06
0,97	-1,08	6,61	414,65	337,35
-0,17	-0,49	9,08	28,01	40,25
-0,86	-0,52	229,96	161,17	177,68
-0,41	0,34	-166,36	3,78	-59,39
-0,68	-0,33	-37,13	63,96	70,22
-0,05	-0,10	-605,95	1,52	13,75
0,10	-1,29	-124,54	169,23	133,07
1,22	1,00	-228,65	24,60	381,27
1,20	1,25	27,46	3,81	361,36
-0,83	0,35	107,82	13,24	-8,79
-0,84	-1,62	571,68	551,46	439,45
-1,13	-0,97	1196,81	506,72	502,66
0,73	0,45	461,04	0,59	191,07
0,77	-0,04	405,91	40,92	228,92
-0,58	1,63	-836,42	169,52	-222,55
-1,17	1,44	-276,25	3,48	-108,62
0,54	0,31	-195,57	0,03	131,70
-0,44	0,82	-355,80	40,46	-118,70
0,63	-1,34	-236,00	303,52	235,17
1,28	0,35	1066,35	170,73	431,39
0,84	-0,39	421,62	121,74	268,02
-0,33	0,56	474,21	21,37	-81,56
-0,32	0,21	52,77	2,03	-42,87
1,19	-0,30	297,36	301,28	407,57
1,70	2,12	291,69	60,98	574,26
-0,09	0,45	392,50	20,18	-43,01
0,73	0,31	-123,97	5,09	198,36
-0,24	0,07	-29,88	0,66	-23,03
-1,27	0,68	174,44	89,76	126,66
-0,16	-0,06	2,16	1,22	-2,96
-0,36	0,78	64,16	43,97	-109,32
-0,62	-0,89	-236,66	164,64	152,19
0,24	-0,51	-87,48	32,00	94,72
0,71	0,77	253,78	7,23	169,84
0,49	1,39	263,92	134,82	51,05
1,04	-0,94	1141,04	413,85	359,93
-1,36	-0,96	988,94	798,60	806,64
1,19	0,67	289,34	58,67	381,03
-0,18	1,06	186,89	106,95	-114,41
0,29	0,46	-631,53	13,95	44,65
0,36	1,59	279,40	215,30	-7,19
-0,74	0,23	-132,14	11,80	-10,29
0,39	-1,86	-31,98	407,32	209,64
-0,42	0,92	-231,01	56,86	-128,61
0,31	-1,01	246,95	124,57	143,95
1,22	0,71	-222,32	60,84	391,22
0,95	2,36	354,30	214,78	192,80
-0,46	-1,30	366,65	232,09	174,49
-2,11	-0,77	3113,64	2747,17	2869,30
-1,38	1,48	83,99	21,63	-13,61

<i>Dados Experimentais</i>		<i>Dados Modelados</i>		
<i>x1</i>	<i>x2</i>	<i>yreal</i>	<i>yRNL</i>	<i>yRNA</i>
0,54	-0,63	262,64	85,29	177,84
-1,40	-0,08	-77,54	417,80	492,26
-0,23	0,34	-180,13	8,93	-49,79
-0,73	0,48	-355,79	1,87	-54,56
1,16	0,60	-418,95	58,02	368,85
0,46	0,34	-2,85	1,58	102,81
-0,47	1,31	168,17	121,32	-177,49
-0,88	1,27	-630,78	27,65	-157,65
-0,36	-1,63	334,62	312,22	200,02
-0,18	0,87	-20,16	70,66	-96,26
-1,32	-0,35	649,11	451,36	509,42
-0,05	1,04	-388,40	108,02	-85,20
-0,32	0,21	-412,25	1,89	-42,43
1,59	-0,24	534,99	769,48	586,96
0,53	0,59	-88,27	9,12	113,56
0,01	-0,32	366,88	10,88	40,99
0,51	-0,18	-576,24	19,33	146,44
-0,11	0,05	-111,33	0,60	-8,30
0,66	-0,20	-439,94	40,44	194,48
-0,63	-0,74	44,57	131,72	129,06
-0,97	-0,77	186,52	293,67	302,60
1,55	-0,53	841,23	867,06	572,48
-1,51	1,19	456,17	122,28	168,36
0,88	0,58	205,73	3,69	245,18
-0,69	-0,24	70,34	53,52	58,62
-3,13	-2,86	16332,87	16109,66	15330,19
-1,70	-0,58	1663,87	1218,14	1278,63
-1,43	-1,89	1634,10	1569,08	1459,77
-2,39	-1,48	4873,77	5204,98	5710,76
0,35	2,30	449,40	475,75	-55,84
0,78	-0,47	335,58	118,16	249,87
-0,01	0,80	140,46	64,03	-57,16
-0,32	-0,66	-246,23	59,36	60,69
1,45	0,08	239,23	415,15	516,34
-1,84	-2,82	4176,70	3867,12	4200,81
-0,90	-1,20	1212,77	406,16	366,07
-2,26	1,48	1305,66	1328,01	1333,16
-2,14	-0,67	2518,39	2790,95	2911,07
-0,71	-0,88	106,06	193,84	183,33
0,16	-1,11	316,27	128,99	124,68
0,38	2,10	1022,79	383,63	-30,71
-0,29	0,34	-413,03	7,39	-55,53
-0,80	0,22	159,82	20,09	7,94
-0,54	-0,27	-331,18	33,33	31,62
1,08	0,69	-505,17	22,22	327,23
-0,79	-0,14	-187,78	60,59	70,70
0,40	-0,35	197,73	26,29	124,83
-1,08	0,19	185,16	98,83	132,99
2,45	0,39	2732,05	3177,95	1024,76
0,28	-1,25	380,53	178,72	153,98
-0,04	2,21	502,73	492,54	-180,91
0,47	-0,86	340,65	116,38	169,33
0,45	-0,38	-426,36	33,96	139,83
-0,18	-0,35	446,19	15,71	25,13
0,83	-0,70	-216,05	192,71	273,12
0,76	-0,07	46,79	41,84	225,81
0,56	-0,57	-36,70	79,13	180,84
-0,90	0,06	143,28	59,70	73,90

APÊNDICE B – Resolução pelo *software* Minitab

A modelagem dos dados através do método de regressão não linear foi feita no *software* Minitab a partir da estatística regressão não linear como mostra a Figura B.1. Primeiramente é necessário selecionar a resposta a ser modelada. Essa variável corresponde aos valores da resposta observada (*y_{real}*). Em seguida, deve-se escrever na caixa Função expectativa a função utilizada para gerar o *y*, essa função foi determinada por cada um dos experimentos do DOE (*Rosenbrock* ou *Ackley N.4*), conforme mostra a Figura B.1.

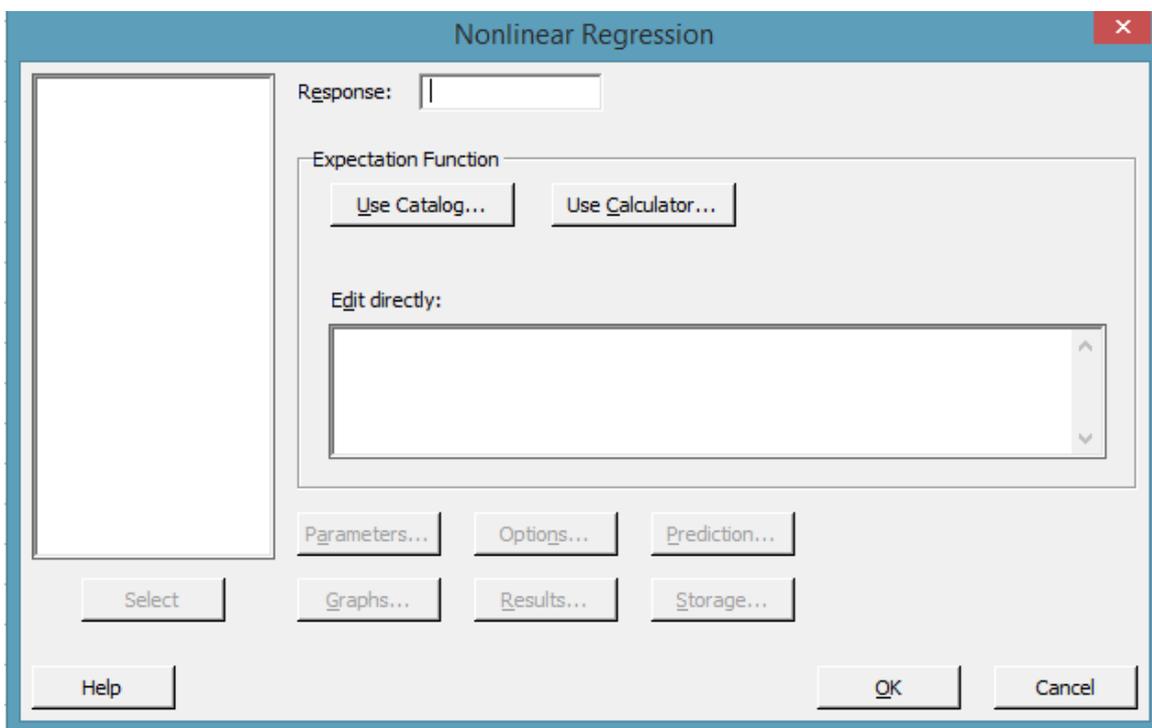


Figura B.1 - Tela do Minitab para definir a resposta a ser modelada pela RNL e a função utilizada

O próximo passo, como mostra a Figura B.2, é escolher os valores iniciais para os parâmetros da equação. Conforme foi explicado no item 2.2.3.1, o algoritmo de Gauss-Newton precisa partir de um ponto de partida para desenvolver seus cálculos. Dessa forma, assumimos o valor zero para todos os parâmetros como ponto de partida para esse algoritmo.

Nonlinear Regression: Parameters ✕

Function

```
exp(-theta1)*(( x1 **2 + x2 **2)**(1/2))+theta2*(cos(theta3 * x1) + sin(theta3* x2
))
```

Required starting values:

Parameter	Values	Locked
theta1	0	<input type="checkbox"/>
theta2	0	<input type="checkbox"/>
theta3	0	<input type="checkbox"/>

Optional constraints:

Parameter	Lower Bound	Upper Bound
theta1		
theta2		
theta3		

Figura B.2 - Tela do Minitab para definir os valores iniciais dos parâmetros

Em seguida, é necessário determinar o tipo de algoritmo iterativo utilizado e conforme dito anteriormente, será o algoritmo de Gauss-Newton. Optou-se por utilizar os valores padrões do *software* para o nível de confiança para todos os intervalos, máximo número iterações e tolerância de convergência (Figura B.3).

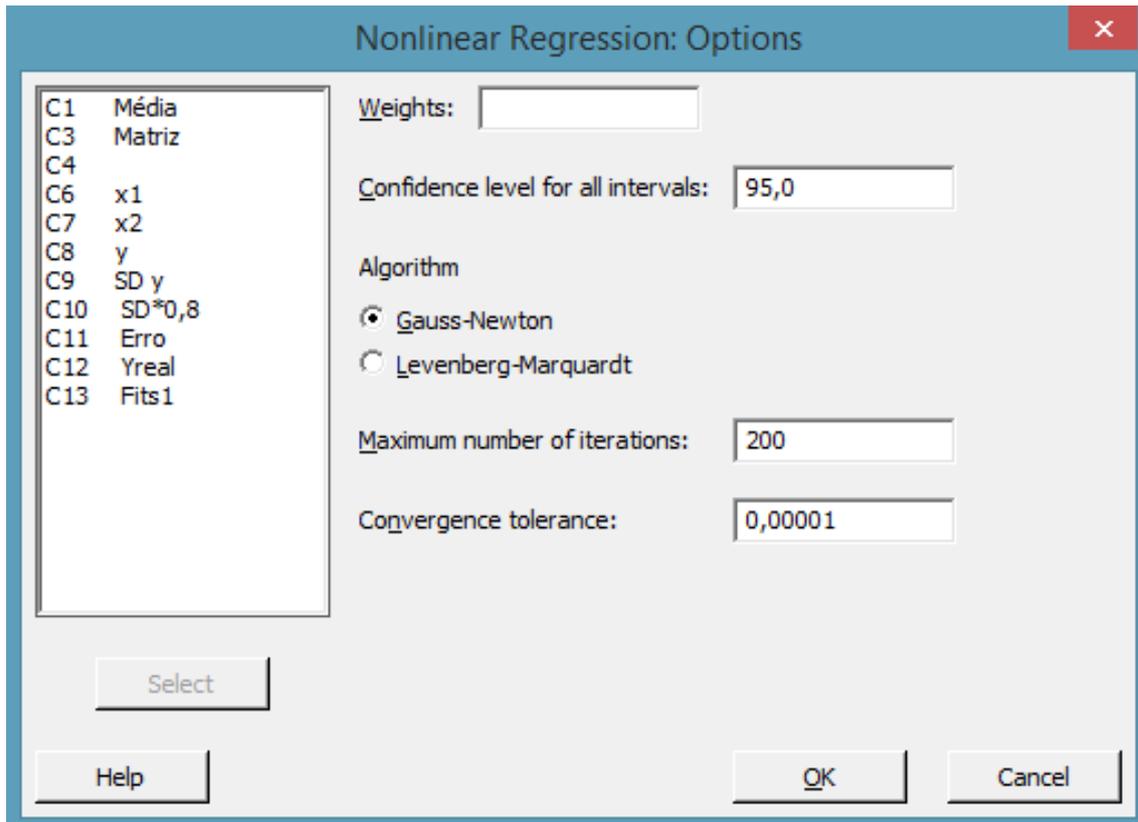


Figura B.3 - Tela do Minitab para definir os parâmetros do método de Gauss-Newton

Por fim, deve-se selecionar o item “*Fits*” da aba “*Storage*”, para obter os valores ajustados da resposta modelados pela RNL (y_{RNL}), conforme mostra a Figura B. 4.

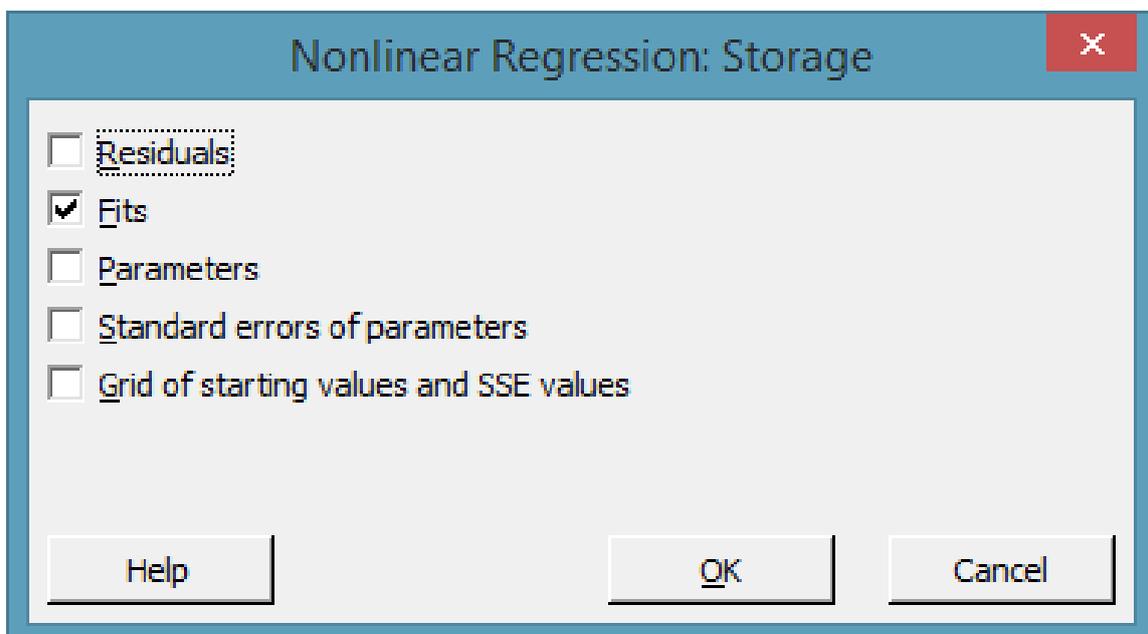


Figura B. 4 - Tela do Minitab para armazenar os valores modelados pela RNL

APÊNDICE C – Resolução pelo *software* *Statistica -Intelligent Problem Solver (IPS)*

A resolução da modelagem dos dados através do método de RNAs foi feita no *software* *Statistica* na sua versão 7.0. Primeiramente, o usuário deve escolher qual o critério de parada do algoritmo, que pode ser através da quantidade de redes testadas ou da quantidade de horas/minutos que o algoritmo será executado, e também a quantidade de redes a serem retidas no final do processo de treinamento. O tempo de treinamento definido nesse trabalho foi de 1 minuto e optou-se por reter as 5 melhores redes, conforme mostra a Figura C.1.

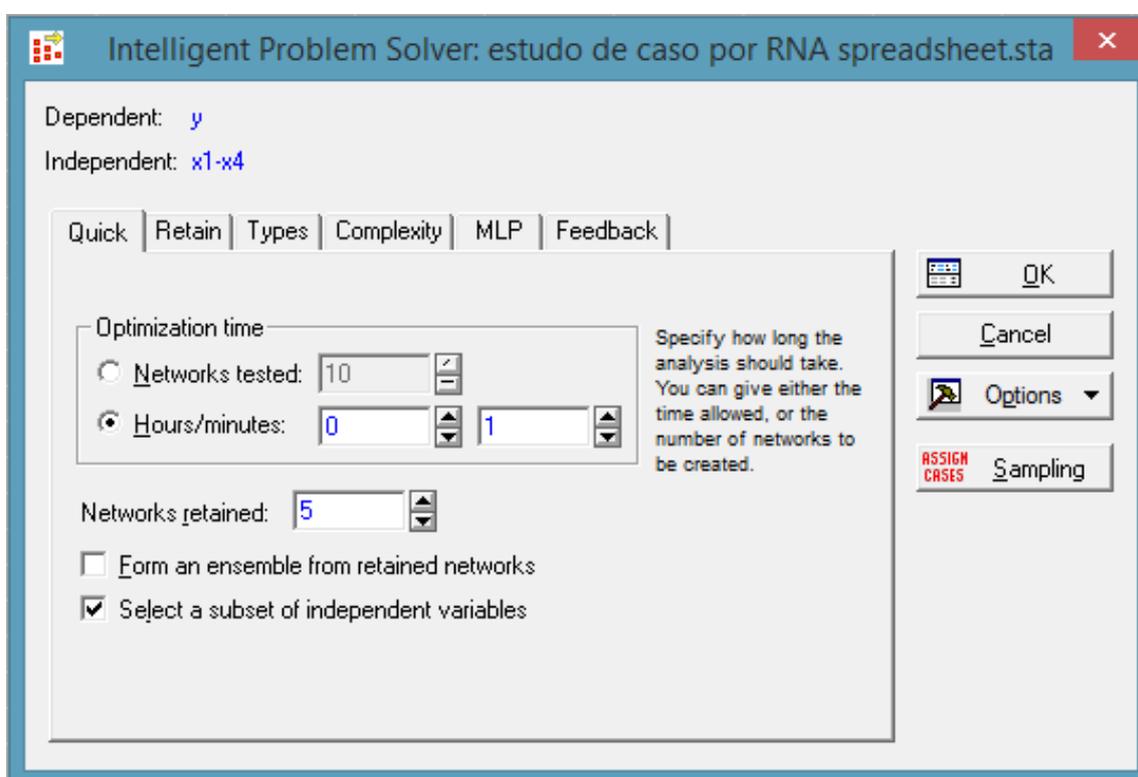


Figura C.1 - Tela do IPS para definir critério de parada e quantidade de redes neurais retidas

O próximo passo é escolher o critério de seleção das redes neurais. O presente estudo escolheu como melhores redes as que possuem menores valores de erro, conforme mostra a Figura C.2. O cálculo do erro no IPS é feito através da soma do quadrado da diferença entre a resposta esperada e a resposta ajustada.

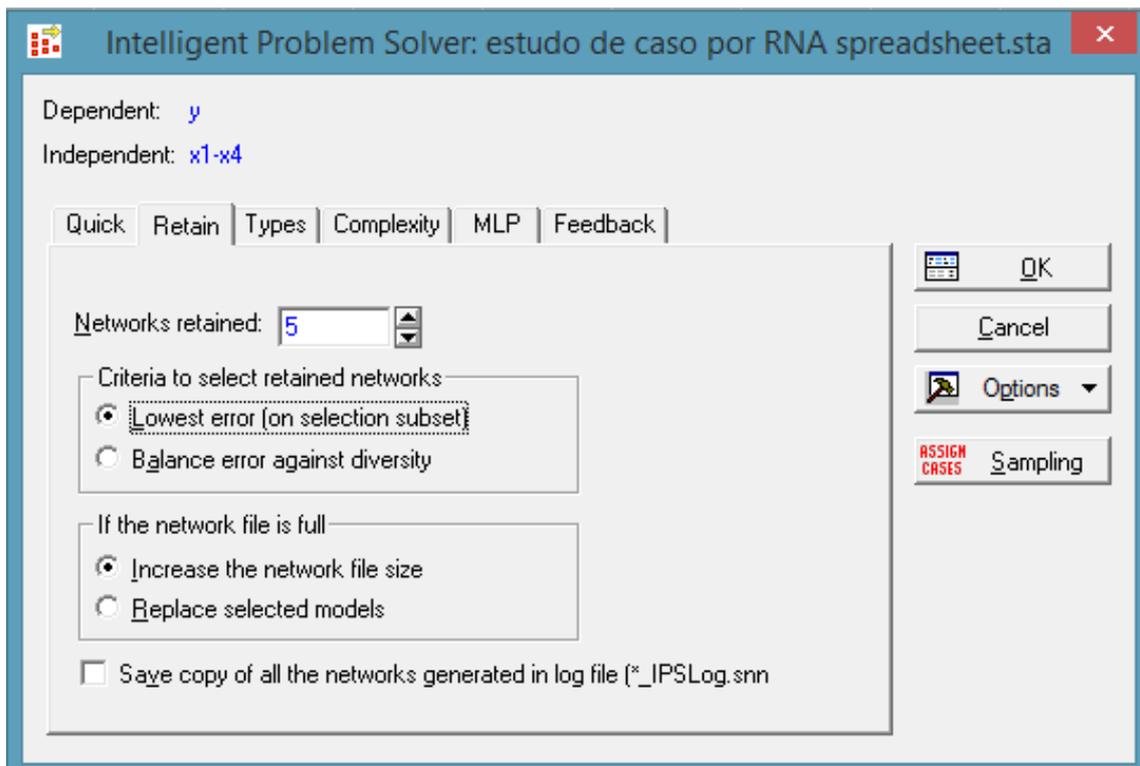


Figura C.2 - Tela do IPS para definir o critério de escolha da melhor rede neural

Conforme dito no item 2.3.3, o tipo de rede utilizado neste trabalho foi a *Three Layer Perceptron* (MLP), pois esse tipo de rede apresenta um desempenho excelente na modelagem de dados não lineares, conforme mostra a Figura C.3.

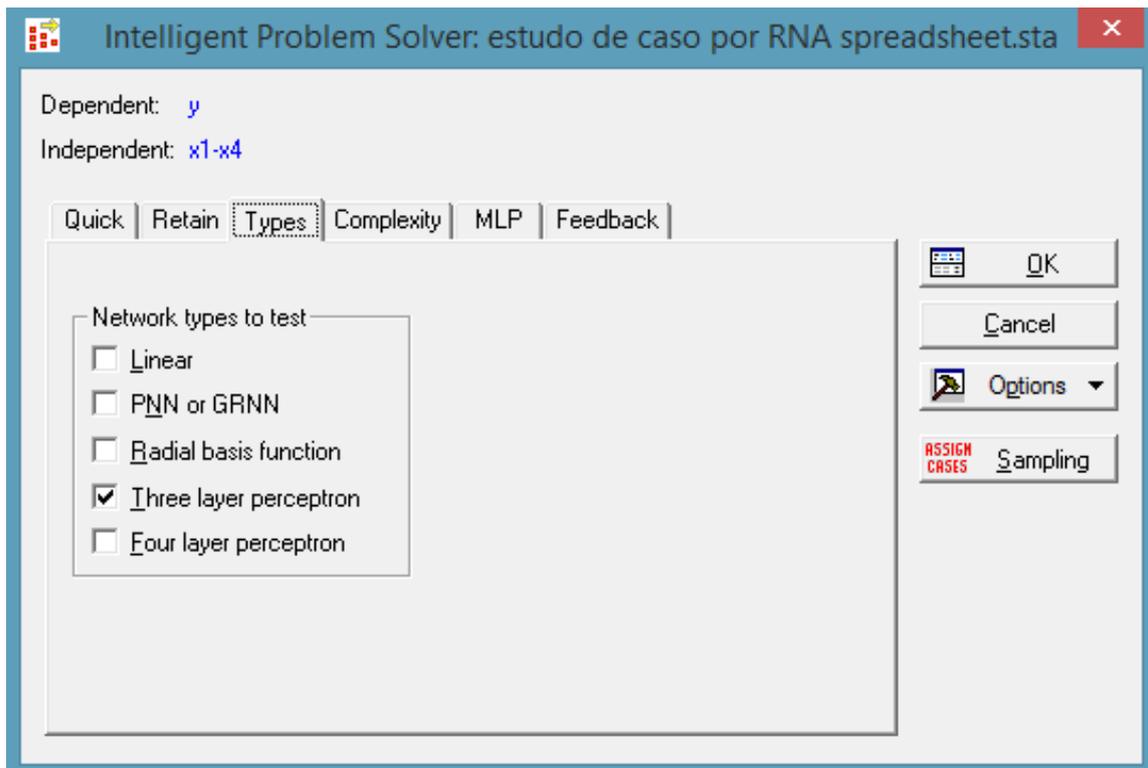


Figura C.3 - Tela do IPS para definição dos tipos de rede testadas

Em seguida, é necessário determinar o intervalo relacionado à quantidade de neurônios da camada oculta da rede *Three Layer Perceptron*. Quanto maior o número dessas unidades maior o tempo de processamento para realizar o treinamento da rede e maior é a probabilidade de ocorrência de *overfitting*, ou seja, a rede neural começa a decorar as informações e vai perdendo completamente sua capacidade de previsão. Nesse caso, optou-se por utilizar o intervalo padrão do *software*, ou seja, de 1 a 10 unidades, conforme pode ser observado na Figura C.4.

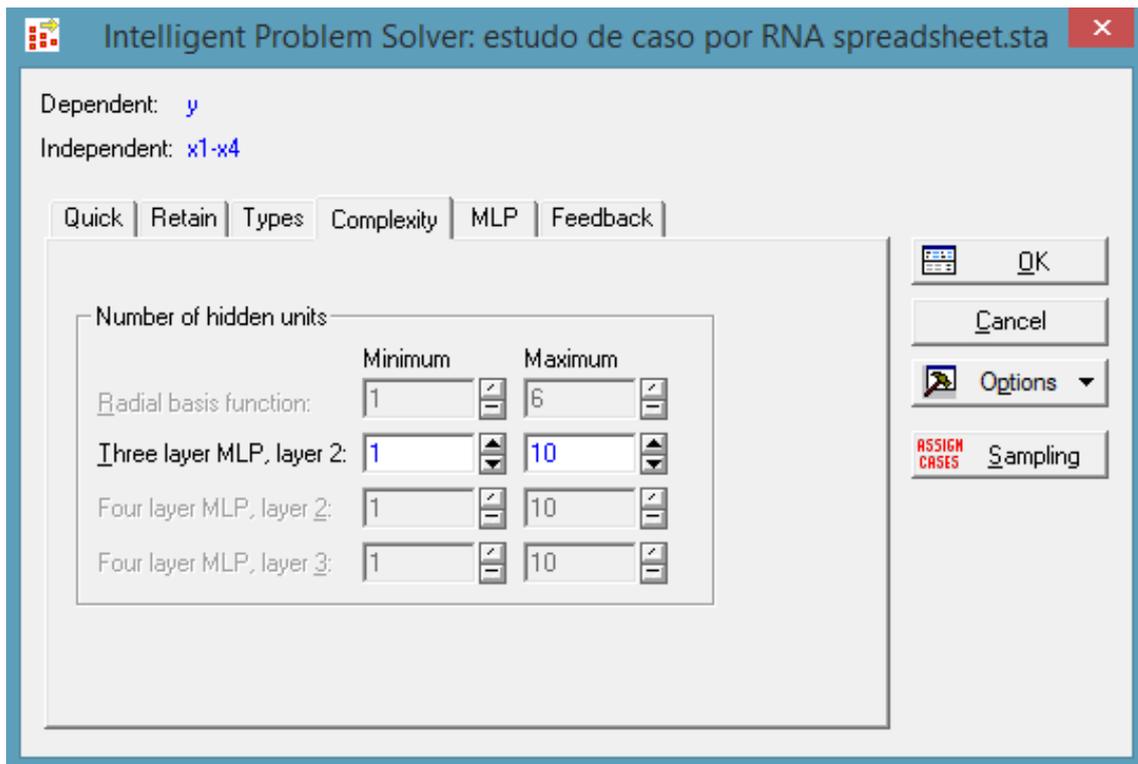


Figura C.4 - Tela do IPS para definição da quantidade de neurônios na camada oculta

Por fim, pode-se escolher duas formas para codificação da resposta, linear ou logística, conforme mostra a Figura C.5. A linear utiliza a função de ativação identidade, enquanto que a logística utiliza a função de ativação logística. Nesse trabalho optou-se por utilizar as duas funções.

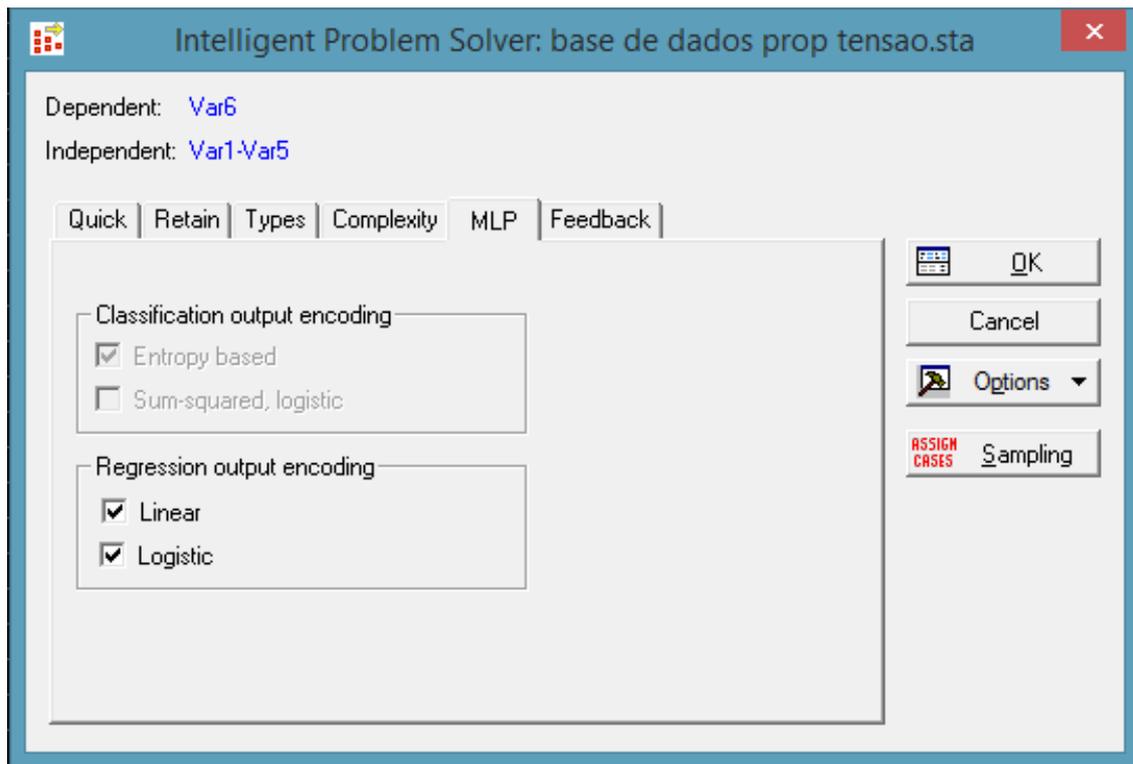


Figura C.5 - Tela do IPS para definição do tipo de função de ativação