

DIRCEU ZEFERINO RODRIGUES

**REDES NEURAIS, IDENTIDADE DE MODELOS E RESPOSTA DA CEBOLA À
ADUBAÇÃO NITROGENADA**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS - BRASIL
2013

**Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV**

T

R696r
2013

Rodrigues, Dirceu Zeferino, 1969-

Redes neurais, identidade de modelos e resposta da cebola à adubação nitrogenada / Dirceu Zeferino Rodrigues. – Viçosa, MG, 2013.

82f. : il. (algumas color.) ; 29cm.

Orientador: Cosme Damião Cruz

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f. 81-82

1. Redes neurais (Computação). 2. Análise de regressão.
3. Cebola - Efeito do nitrogênio. 4. Nitrogênio na nutrição animal. I. Universidade Federal de Viçosa. Departamento de Estatística. Programa de Pós-Graduação em Estatística Aplicada e Biometria. II. Título.

CDD 22. ed. 006.32

DIRCEU ZEFERINO RODRIGUES

**REDES NEURAIS, IDENTIDADE DE MODELOS E RESPOSTA DA CEBOLA À
ADUBAÇÃO NITROGENADA**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Magister Scientiae*.

APROVADA: 21 de março de 2013.

Prof. Moysés Nascimento
(Coorientador)

Prof. Sanzio Mollica Vidigal

Prof. Cosme Damião Cruz
(Orientador)

DEDICATÓRIA

À minha mãe Maria da Conceição,

Aos meus filhos Isabela e Igor, e

À minha amiga e namorada Vanessa

Sem vocês nenhuma conquista valeria a pena...

“Os dias que estes homens passam nas montanhas são os dias em que realmente vivem. Quando a mente se limpa das teias de aranha e o sangue corre com força pelas veias. Quando os cinco sentidos recobram a vitalidade e o homem completo se torna mais sensível, e então já pode ouvir as vozes da natureza, e ver as belezas que só estão ao alcance dos mais ousados.”

Reinhold Messner

SUMÁRIO

RESUMO	
ABSTRACT	
1. INTRODUÇÃO GERAL	1
2. REVISÃO BIBLIOGRÁFICA	3
2.1 A cebola e os efeitos das doses da adubação nitrogenada	3
2.2 Análise de regressão	5
2.3 Teste de identidade de modelos	8
2.4 A bioinformática	10
2.5 Redes neurais artificiais	11
2.5.1 Histórico	11
2.5.2 O cérebro humano e os neurônios biológicos	12
2.5.3 Os neurônios artificiais e as RNAs	14
2.5.4 Função de ativação e capacidade de aprendizagem	16
2.5.5 Classificação das redes neurais e a rede MLP	18
3. CAPÍTULO 1: Identidade de modelos e a resposta da cebola à adubação nitrogenada	22
3.1 Introdução	22
3.2 Material e Métodos	24
3.3 Resultados e discussão	26
3.4 Conclusão	36
4. CAPÍTULO 2: Redes neurais artificiais e a predição da resposta da cebola à adubação nitrogenada	37
4.1 Introdução	37
4.2 Material e Métodos	40
4.3 Resultados e discussão	43
4.3.1 Analisando o caso da interpolação de valores	43
4.3.2 Analisando o caso da extrapolação de valores	62
4.4 Conclusão	80
5. PERSPECTIVAS FUTURAS	80
6. REFERÊNCIAS BIBLIOGRÁFICAS	81

RESUMO

RODRIGUES, Dirceu Zeferino, M. Sc. Universidade Federal de Viçosa, março de 2013. **Redes neurais, identidade de modelos e resposta da cebola à adubação nitrogenada.** Orientador: Cosme Damião Cruz. Coorientadores: Moysés Nascimento e Paulo Roberto Cecon.

O estudo das curvas de produtividade comparadas com a quantidade de nitrogênio absorvido pela cultura da cebola é de fundamental importância para a formulação de um plano de adubação que seja mais eficiente tanto em termos técnicos quanto econômicos. Diversas técnicas estatísticas têm sido propostas, testadas e aprimoradas com o intuito de contribuir para alavancar pesquisas nesta direção. A justificativa para este trabalho de pesquisa está na necessidade de avaliar e aprimorar novas técnicas estatísticas que ajudem na obtenção de informações precisas com a finalidade de auxiliar na tomada de decisão visando melhorar a produtividade. Para isso, este estudo teve como objetivo empregar e avaliar duas metodologias de auxílio à estatística, mas com objetivos específicos distintos com respeito à avaliação da aplicação de nitrogênio na produção dos cultivares da cebola. Na primeira avaliação, objetivou-se utilizar técnicas estatísticas baseadas em modelos de regressão e ajustar curvas para alguns níveis de doses de nitrogênio, relacionadas à produtividade, para uma pesquisa realizada com quatro cultivares em locais distintos de cebola e, em seguida, avaliar a possibilidade de agrupamento desses modelos estatísticos obtidos, utilizando o teste de identidade de modelos. Nesta etapa, procurou-se estimar uma curva que representasse, em conjunto, o padrão de resposta à adubação em todos os quatro locais avaliados. No segundo estudo, a meta era verificar a eficiência de técnicas baseadas em redes neurais. Assim, a proposta foi constatar se já é possível utilizar, com segurança, esse novo conceito baseado em redes neurais artificiais em pesquisas relacionadas à resposta de cultivares de cebola à adubação nitrogenada. De uma maneira geral, o trabalho descreve o êxito da utilização de novas técnicas estatísticas com ênfase em redes neurais que ajudem melhorar a produtividade da cebola para, a partir daí, permitir aplicar e difundir técnicas baseadas em inteligência computacional para fins de estudos de predição e modelagem.

ABSTRACT

RODRIGUES, Dirceu Zeferino, M. Sc., Universidade Federal de Viçosa, March, 2013. **Neural networks, model identity and onions' response to nitrogen fertilization.** Advisor: Cosme Damião Cruz. Co-advisors: Moysés Nascimento and Paulo Roberto Cecon.

The study of the productivity curves compared with the amount of nitrogen absorbed by the onion crop is fundamentally important for the elaboration of a more efficient fertilization plan in technical terms as well as in economic terms. Many statistical techniques have been proposed, tested, and improved in order to help boost research in this direction. The justification for this research is the need to assess and improve new statistical techniques that help in obtaining accurate information in order to assist in decision making for improving productivity. For this case, this study aimed to use and evaluate two statistical methods with different specific objectives with respect to the evaluation of nitrogen application in the production of onion cultivars. In the first evaluation, statistical techniques based on regression models were used for adjusting curves for some nitrogen levels related to productivity, performing a survey with four onion cultivars in different locations, and then to carry out the evaluation of the grouping possibility of these statistical models using the models' identity test. In this step, it was tried to estimate a curve that could represent together the fertilization response pattern in all four evaluated sites. In the second study, the goal was to verify the techniques efficiency based on neural networks. So, the proposal was to see the possibility of using safely this new concept based on artificial neural networks in research related to the onion cultivars response to nitrogen fertilization. In general, this study describes the successful use of new statistical techniques with emphasis on neural networks that help improve the onion productivity and thereafter to implement and disseminate techniques based on computational intelligence for purposes of study prediction and modeling.

1. INTRODUÇÃO GERAL

O estudo das curvas de produtividade comparadas com a quantidade de nitrogênio absorvido pela cultura da cebola é de fundamental importância para a formulação de um plano de adubação que seja mais eficiente tanto em termos técnicos quanto econômicos. Vidigal et al. (2010), ressaltaram que a adubação nitrogenada na produção de cebola pode ser mais eficiente, sendo que a quantidade recomendada é de 148 kg ha^{-1} de nitrogênio para o solo da região pesquisada, no entanto mais estudos devem ser desenvolvidos para uma melhor recomendação do nitrogênio para solos da região considerando o teor de matéria orgânica desses solos.

Diversas técnicas estatísticas têm sido propostas, testadas e aprimoradas com o intuito de contribuir para alavancar pesquisas nesta direção. Dentre elas, a que está em destaque no momento é a área das redes neurais artificiais (RNAs). Segundo Braga (2000), a solução de problemas por meio de RNAs está bastante atrativa atualmente, já que a forma como os neurônios são representados internamente pela rede e o paralelismo natural inerente à arquitetura das redes neurais artificiais criam a possibilidade de um desempenho superior ao dos modelos convencionais de apoio à decisão. Em RNAs, o procedimento usual na solução de problemas passa inicialmente por uma fase de aprendizagem, em que um conjunto de exemplos é apresentado para a rede, a qual extrai as características necessárias para representar a informação fornecida. Essas características são utilizadas posteriormente para gerar respostas para o problema. A capacidade de aprender por meio de exemplos e de generalizar a informação aprendida é, sem dúvida, o atrativo principal da solução de problemas pelas RNAs.

Diante do exposto, a justificativa para este trabalho de pesquisa está pautada na necessidade que existe atualmente de avaliar e aprimorar novas técnicas estatísticas que ajudem na obtenção de informações úteis e precisas para tomada de decisão. Essas novas metodologias, depois de testadas, validadas e utilizadas nas pesquisas de campo com cebola, certamente contribuirão para melhorar a produtividade e, futuramente, será viável aplicar e divulgar essas técnicas para fins de estudos de predição e modelagem no melhoramento genético.

Este estudo teve como finalidade avaliar dois métodos de auxílio à tomada de decisão na estatística, mas com objetivos específicos distintos com respeito à avaliação da aplicação de nitrogênio na produção dos cultivares da cebola. Ou seja, avaliou-se, utilizando a mesma pesquisa de campo, duas propostas com métodos de auxílios à tomada de decisão não só de origens, mas também escolas de pensamentos diferentes. Agrega-se a isto, a contribuição em aplicar e difundir técnicas alternativas baseadas em inteligência computacional para fins de estudos de predição de valores.

Na primeira avaliação, objetivou-se utilizar técnicas estatísticas baseadas em modelos de regressão e ajustar curvas para alguns níveis de doses de nitrogênio, relacionadas à produtividade, para uma pesquisa realizada com quatro cultivares distintos de cebola e, então, avaliar a possibilidade de agrupamento desses modelos estatísticos obtidos pela análise de regressão, com base no teste de identidade de modelos. Procurou-se, com isso, estimar uma curva que representasse, em conjunto, o padrão de resposta à adubação em todos os locais estudados. Este estudo será detalhado no primeiro capítulo.

Na segunda avaliação, verificou-se a eficiência e precisão de técnicas baseadas em redes neurais, comparando os resultados obtidos após a utilização desta metodologia com os resultados provenientes dos cálculos realizados aplicando-se a análise de regressão polinomial, de uso rotineiro na estatística experimental. Assim, objetivou-se constatar se já é possível utilizar, com segurança, esse novo conceito para auxílio à tomada de decisão baseado em redes neurais artificiais em pesquisas que exigem resposta para esse tipo de questionamento referente à adubação na agricultura. Uma explanação bem mais detalhada será assunto do segundo capítulo.

Deste modo e considerando os avanços tecnológicos atuais no campo da inteligência artificial verificou-se a possibilidade de recomendar e difundir os modelos baseados em redes neurais para serem aplicados a estudos de pesquisas experimentais agrícolas, tendo como referência ensaios com cebolas, que visem aperfeiçoar a produção e, conseqüentemente, obter ganhos em quantidade e qualidade do produto comercial.

2. REVISÃO BIBLIOGRÁFICA

2.1 A cebola e os efeitos das doses da adubação nitrogenada

A cebola é uma hortaliça de grande importância econômica, sendo a terceira mais cultivada no mundo, ultrapassada apenas pelo tomate e a batata. A área de produção no Brasil é de 70.000 ha ano⁻¹, com rendimento, nos últimos anos, próximo de 19 t ha⁻¹, mas cultivos bem conduzidos têm rendimentos entre 40 e 60 t ha⁻¹ ou até mesmo superior (VIDIGAL et al., 2007).

A cebola que é uma das espécies de hortaliças mais antigas, sendo, inclusive, mencionada na Bíblia e no Corão islâmico. O registro mais antigo sobre o cultivo da cebola data de cerca de 3.200 anos A.C., sendo provavelmente a região da antiga Pérsia (atualmente Paquistão e Irã) um dos primeiros centros de domesticação segundo o geneticista russo Vavilov. Cultivada na China e Índia desde a antiguidade, foi consumida na Grécia, Roma e Egito desde tempos remotos, passando pela África e todo o continente Europeu até vir a ser difundida na América, pelos nossos colonizadores. No Brasil, o consumo da cebola se deu, inicialmente, no Rio Grande do Sul e, somente tempos mais tarde, se espalhou para o restante dos estados do país. Atualmente, o país é o maior produtor da América do Sul e está entre os 10 maiores do mundo.

Nesta espécie, os experimentos de seleção visam, de maneira geral, modificar o desempenho ou aspecto em características como: o formato, a coloração, a retenção de escamas e o tamanho de bulbos. Tem-se também objetivo de aumentar a produtividade, melhorar a conservação pós-colheita e o nível de resistência a pragas e doenças, e a adaptação a diferentes condições edafoclimáticas. Como resultado marcante, pode-se ressaltar a adaptação da cebola a diferentes latitudes em relação ao seu centro de origem, considerando-se que o fotoperíodo é fator limitante no processo de bulbificação. Hoje, a cebola está sendo cultivada em regiões distintas, dentro de uma grande amplitude geográfica, estendendo-se do equador até regiões mais próximas aos círculos polares (COSTA et al., 1978).

A importância econômica desta cultura é evidenciada por Boiteux (2004) quando disse que a cebola (*Allium cepa* L.) é uma das plantas cultivadas de mais ampla difusão no mundo, sendo a segunda hortaliça em importância econômica,

com valor da produção estimado em cerca de US\$ 6 bilhões anuais. A produção mundial apresentou aumento de cerca de 25% na última década, o que coloca a cebola como uma das três hortaliças mais importantes ao lado do tomate e da batata. Somado a isto, o valor social da cultura de cebola é inestimável, sendo consumida por quase todos os povos do planeta, independente da origem étnica e cultural, constituindo-se em um importante elemento de ocupação de mão-de-obra familiar.

Outro aspecto singular, segundo Machado (2006), refere-se à variedade de tipos de cebolas que podem ser encontradas mundo afora. Essa variação é percebida no tamanho, na cor e no aroma. As cebolas podem ser brancas, amarelas, roxas ou rosadas; são encontradas em tamanho grande, médio e pequeno; estão disponíveis nos mercados in natura (frescas), em pasta, em pó ou ainda em flocos.

Dentre as práticas culturais para a melhoria do manejo da adubação nitrogenada da cebola, inclui-se o parcelamento das doses aplicadas, que pode diminuir as perdas por lixiviação, principalmente em solos de textura arenosa e, também, reduzir custos de produção e promover aumentos na produtividade. Mas, para isto, a melhoria do manejo da adubação nitrogenada se faz necessário, para que sua absorção pela planta seja eficiente (VIDIGAL et al., 2010).

Oliveira et al. (2003) relatam que o nitrogênio é um macronutriente primário essencial para as plantas, por participar da formação de proteínas, aminoácidos e de outros compostos importantes no metabolismo das plantas. Sua ausência bloqueia a síntese de citocinina, hormônio responsável pelo crescimento das plantas, causando redução do seu tamanho e conseqüentemente redução da produção econômica das sementes.

O suprimento adequado de nitrogênio é necessário para a máxima produtividade e desenvolvimento da cebola, sendo que a excessiva aplicação, entretanto, pode limitar a produtividade e, assim sendo, aumentar as perdas no armazenamento (SOUZA; RESENDE, 2002). Além disso, o nitrogênio contribui marcadamente para a melhoria da produção de cebola, sendo absorvido em grandes quantidades e superado somente pelo potássio (VIDIGAL et al., 2007).

2.2 Análise de regressão

Muitos problemas encontrados na agricultura envolvem explorar as relações entre duas ou mais variáveis. A análise de regressão é uma técnica de grande importância na estatística e é utilizada para investigar e modelar o relacionamento existente entre as diversas variáveis de um processo na análise de dados (MONTGOMERY, 2009). Sua utilização vem se ampliando a cada dia, principalmente devido ao fato de a análise de regressão ser baseada na ideia relativamente simples de se empregar uma equação para expressar o relacionamento entre as variáveis de interesse. A partir do uso da equação obtida, fica mais fácil a identificação dos fatores que estão agindo sobre os resultados de um processo e a determinação do foco de um problema. Além disso, a própria equação de regressão pode fornecer indicações sobre como deve ser a atuação no processo, de modo que o problema seja solucionado (WERKEMA, 1996).

A primeira notícia que se tem sobre o modelo estatístico chamado de modelo clássico de regressão teve origem nos trabalhos sobre astronomia realizados por Gauss de 1809 a 1821. Mas o termo “regressão” somente foi referenciado pela primeira vez mais de meio século depois por Francis Galton (1822-1911), época em que esse realizava trabalhos sobre hereditariedade de alturas entre pais e filhos (BARBETTA, 2011). Nestas pesquisas antropométricas, ele verificou que os filhos herdavam as mesmas características de seus pais, ou seja, pais com estatura baixa teriam filhos pequenos também, porém com estatura maior em média. Já pais com estatura alta teriam filhos maiores, porém com estatura menor em média. Por verificar que a altura média de filhos de pais de uma dada estatura tendia a se deslocar ou a “regredir” para a altura média da população como um todo, começou-se a fazer uso deste termo. A terminologia e a técnica passaram, então, a serem utilizadas e, hoje em dia, a análise de regressão é uma das técnicas mais importantes da estatística e é aplicada em problemas estatísticos de uma infinidade de áreas do conhecimento como, por exemplo: a área de biometria (agronomia, biologia, genética e melhoramento, medicina, psicologia, etc.), engenharia e economia.

Este relacionamento pode ser modelado tanto por uma equação linear quanto por uma função não linear, dependendo do caso. Dentre as finalidades

que valem a pena mencionar estão a descrição, a predição, o controle e a estimação.

Apesar da ampla utilidade da análise de regressão, é preciso atentar para o fato de que os valores obtidos através de uma variável em termos de outra, podem ou não verificar, na prática, uma relação de causa e efeito entre estas variáveis. Mas neste caso, a lógica de uma relação causal provém de teorias externas ao âmbito da estatística, caso em que será necessário usar o raciocínio da área em questão para a conclusão da análise.

Para se obter uma equação para um determinado modelo estatístico que explique o comportamento de uma variável de interesse para uma dada pesquisa é imprescindível calcular a tendência desta variável chamada dependente de se alterar como consequência da alteração dos valores das variáveis explicativas, e também, da dispersão em torno da média desta tendência. Para isso, primeiramente é necessário coletar uma amostra aleatória com os valores das características que se deseja avaliar sua influência na variável resposta. Em seguida, determinar os valores dos parâmetros de uma função que melhor se adapte a esse conjunto de valores observados.

O método mais conhecido e usado para se obter as estimativas de uma equação de regressão é chamado de método dos mínimos quadrados. Segundo Barbetta (2011), a idéia básica da construção da parte estrutural do modelo, é encontrar a curva que passe o mais próximo possível dos pontos observados, ou seja, consiste em fazer com que a soma quadrática dos efeitos aleatórios, seja a menor possível.

De autoria do matemático alemão Carl Friedrich Gauss, este método apresenta algumas propriedades estatísticas importantes para obter uma equação de regressão de forma mais eficiente, como por exemplo, imparcialidade e consistência. A curva de regressão obtida pelo método dos mínimos quadrados passa pela média da amostra dos valores das variáveis dependente (resposta) e independente (explicativa), mostrando se elas realmente mantêm alguma relação estatística, assim como se são positivamente ou negativamente correlacionadas e igualando a soma de seus desvios a zero.

A regressão polinomial pode ser vista como uma generalização da regressão linear e é definida como sendo um tipo de regressão onde o relacionamento entre a variável independente (x) e a variável dependente (y) é

modelado utilizando uma expressão polinomial de ordem n . É considerado um caso especial da regressão linear múltipla e o termo regressão polinomial é somente usado para expressões com n maior ou igual a 2. A técnica é utilizada para representar um relacionamento não linear entre o valor da variável explicativa e a esperança condicional correspondente da variável resposta, denotada por $E(y|x)$, e tem sido usado para descrever uma variedade de fenômenos de diversas áreas do conhecimento. Modelos de regressão polinomial são mais comumente obtidos utilizando o método dos mínimos quadrados cuja finalidade é minimizar a variância dos estimadores não viesados. Em geral, pode-se modelar o valor esperado de y como um polinômio de ordem n , obtendo o modelo de regressão polinomial geral, expresso pela seguinte fórmula:

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots + a_mx^m + \xi$$

em que:

- Y é o valor da variável dependente;
- a_i é o parâmetro associado ao efeito de x^i ;
- x^i é o valor da variável independente de grau i associado a y ;
- ξ é o erro aleatório não observável no modelo.

O coeficiente de determinação (R^2) foi utilizado para averiguar quão bem a equação de regressão se ajustou aos dados amostrais estudados. O coeficiente de determinação tem como função nos fornecer uma informação auxiliar para que possamos verificar se o modelo proposto é adequado ou não para descrever o fenômeno em estudo. Apesar de ser muito influenciado pela área sob a qual o problema está sendo estudado, quanto maior for o valor calculado para o R^2 , melhor será o ajuste da regressão pelo modelo proposto aos dados observados.

A fórmula que determina o coeficiente de determinação é a seguinte:

$$R^2 = \frac{SQ_{reg}}{SQ_{total}} = \frac{\text{variação explicada}}{\text{variação total}}$$

em que SQ_{reg} é a soma de quadrados da regressão e SQ_{total} é a soma de quadrados total e pode ser interpretado como uma medida descritiva da proporção da variação de y que poder ser explicada por x , segundo o modelo especificado (BARBETTA, 2011).

A aplicação da análise de regressão exige conhecimento teórico e experiência com análise de dados, e necessita extensivamente de cálculos matemáticos, tanto que atualmente, é impossível se pensar em análise de regressão nos estudos experimentais mais complexos sem o uso de computadores para realizar inversão de matrizes, métodos numéricos, grandes bases de dados, etc.

2.3 Teste de identidade de modelos

Na análise de regressão e em experimentos relacionados a uma mesma finalidade, pode acontecer de serem geradas mais de uma curva de equações com parâmetros diferentes numericamente. Muitas vezes é verificado que estas equações são iguais estatisticamente e uma única curva pode ser obtida para representar todas elas sem perda de informação para a inferência.

De acordo com Regazzi (2003), alguns problemas têm aplicações importantes, como determinar se um conjunto de curvas é idêntico, se tem um intercepto comum ou se alguns dos parâmetros do modelo são os mesmos de modelo para modelo. Em muitos casos, o interesse maior está em saber se um conjunto de equações pode ser representado por uma equação comum.

Neter et al. (1985) testaram se duas equações de regressão linear simples eram idênticas utilizando o teste F. Os pesquisadores concluíram que o teste pode ser aplicado para verificar a igualdade de duas equações de regressão polinomial ou de duas equações de regressão múltipla, desde que sejam feitas as modificações adequadas, e, ainda, que o teste pode ser estendido, em caso de três ou mais equações.

Para exemplificar o método, considerou-se verificar a igualdade de duas equações de regressão lineares, apesar de o método ser bastante geral podendo, com isso, ser generalizado para qualquer número de equações polinomiais de grau maior que dois. Assim, dadas as equações:

- $y_{1i} = a_1 + b_1x_{1i} + e_{1i}$, com $i = 1, \dots, n_1$
- $y_{2i} = a_2 + b_2x_{2i} + e_{2i}$, com $i = 1, \dots, n_2$

em que:

- y_{ij} é o valor observado na i -ésima unidade experimental do i -ésimo grupo;
- x_{ij} é o valor da variável independente associado a y_{ij} ;
- e_{ij} é o erro aleatório não observável no modelo.

A metodologia que está sendo exemplificada pressupõe $e_{ij} \sim NID(0, \sigma^2)$. Outra importante pressuposição é a da homocedasticidade entre os grupos cuja verificação pode ser feita utilizando-se um teste de homogeneidade de variâncias apropriado, como por exemplo, o teste de Bartlett. As hipóteses a serem testadas são as seguintes:

- $H_0^1: a_1 = a_2 = a$ vs H_a^1 : nem todos os a_i são iguais.
- $H_0^2: b_1 = b_2 = b$ vs H_a^2 : nem todos os b_i são iguais.
- $H_0^3: a_1 = a_2 = a$ e $b_1 = b_2 = b$ vs H_a^3 : pelo menos uma igualdade não é verdade.

Inicialmente, obtém-se a soma de quadrados do resíduo para as duas equações em separado, ou seja, SQ_{res1} e SQ_{res2} , com os graus de liberdade $n_1 - p$ e $n_2 - p$, respectivamente, onde n_i é o número de observações. As somas de quadrados calculadas individualmente são somadas para se obter $SQ_{res3} = SQ_{res1} + SQ_{res2}$. E somam-se também os graus de liberdade $n_1 + n_2 - 2p$. Em seguida, combinam-se as $n_1 + n_2$ observações e calcula-se, através do método dos mínimos quadrados, as estimativas para os parâmetros a e b da equação de regressão combinada:

$$y = a + bx + e$$

Desta equação combinada, obtém-se a soma de quadrados do resíduo SQ_{resc} , com grau de liberdade igual a $n_1 + n_2 - p$, em que p é o número de parâmetros a ser estimado. Para este exemplo, p é igual a 2. É necessário subtrair também da soma de quadrados combinada, as somas de quadrado individuais:

$$SQ_{res4} = SQ_{resc} - SQ_{res3}$$

A estatística F é calculada da seguinte forma:

$$F_{calc.} = \frac{SQ_{res4}/p}{SQ_{res3}/(n_1 + n_2 - 2p)}$$

com (p) e $(n_1 + n_2 - 2p)$ graus de liberdade. Se F_{calc} for maior que F_{tabelado} , para um determinado nível de significância, rejeita-se a hipótese de que os parâmetros a'_s e b'_s são os mesmos para os dois conjuntos de observações. Caso contrário, não se rejeita a hipótese e conclui-se que a equação combinada pode ser usada para representar as equações estudadas.

Em seu trabalho, Regazzi (2003) concluiu que a identidade de modelos de regressão e igualdade de qualquer subconjunto de parâmetros pode ser verificada por meio do teste F, e que a metodologia apresentada é geral e pode ser usada em modelos polinomiais de qualquer grau e também em modelos de regressão múltipla.

2.4 A bioinformática

A bioinformática é um campo das ciências biológicas que, ultimamente, vem crescendo rapidamente e está sendo desenvolvida com o objetivo de atender à necessidade de manipular e analisar grandes quantidades de dados genéticos. Para compreender a relação entre estas informações provenientes das diversas subáreas, a bioinformática usa o poder computacional para catalogar, organizar e estruturar estas informações em uma entidade compreensiva e extremamente importante para, posteriormente, analisar e tirar conclusões sobre esses dados.

Segundo Attwood (2011), Bioinformática é o estudo da aplicação de técnicas computacionais e matemáticas à geração e gerenciamento de bioinformação.

Devido ao fato de a bioinformática ser ainda uma ciência recente, ela é referida como tendo a tarefa de organizar e analisar dados normalmente complexos resultantes de modernas técnicas de biologia molecular. Para muitos a bioinformática é uma importante ferramenta para a compreensão de como as informações coletadas podem ser armazenadas, ou seja, é a ciência da criação e manutenção de base de dados complexa e, eventualmente, a simulação de fenômenos relacionados ao melhoramento genético.

A bioinformática utiliza ferramentas metodológicas da matemática aplicada e da ciência da computação. A Biologia molecular de hoje, por exemplo, seria impossível sem os recursos de bioinformática, tais como o armazenamento,

distribuição e atualização das informações, as análises estatísticas, a modelagem de dados e a simulação de fenômenos biológicos em computador.

A neuroinformática também é um novo e emergente ramo da bioinformática. Ela surgiu através do esforço colaborativo entre neurologistas e psicólogos. O cérebro e os neurônios estão sendo encarados como um sistema complexo que serve de modelo para o desenvolvimento de redes neurais de computadores, devido à forma com que os neurônios trabalham. Algoritmos genéticos e raciocínio não linear atualmente têm sido utilizados para o desenvolvimento de inteligência artificial e evolução computacional.

2.5 Redes neurais artificiais

Com os avanços tecnológicos, os cientistas estão cada vez mais empenhados em compreender os segredos e as características biológicas do cérebro humano para ser capaz de imitá-lo, ou seja, tentar reproduzir computacionalmente sua complexa capacidade de realizar julgamentos baseados em experiências passadas e relacioná-las com situações que jamais foram encontradas antes. Portanto o grande desafio da computação, atualmente, é tentar juntar a velocidade de processamento computacional com estas características do cérebro humano com o objetivo de permitir resolver uma infinidade de problemas ainda sem soluções. Dentre suas principais aplicações pode-se citar: reconhecimento de fala, reconhecimento automático de caracteres escritos à mão, localização de fontes de radar e otimização de processos químicos e biológicos.

2.5.1 Histórico

Segundo Braga (2000), o primeiro modelo artificial de um neurônio biológico foi fruto do trabalho pioneiro de Warren McCulloch e Walter Pitts em 1943. O trabalho inicial de McCulloch e Pitts não teve como meta apresentar técnicas de aprendizado e focou esforços em mostrar as capacidades computacionais de um neurônio e definir um modelo artificial. Em 1949, Donald Hebb apresentou seu trabalho sobre aprendizagem de redes neurais, obtida variando-se os pesos de entrada dos nodos, técnica que é usada até hoje em algoritmos de aprendizado.

Idealizado por Frank Rosenblatt, Charles Wightman e outros, Mark I Perceptron surgiu em 1957 e 1958, como o primeiro neuro computador a obter sucesso. O perceptron simples possuía três camadas: uma de entrada, a segunda cuja eficiência de transmissão era ajustável e, a terceira, de resposta. Na época, ele tinha como objetivo inicial o reconhecimento de padrões e como característica principal as sinapses ajustáveis na segunda camada. Pela dedicação e pelos esforços despendidos nesta área, Rosenblatt é visto até hoje como o precursor da neuro computação.

Em 1962, o Adaline que foi projetado bem diferente do Perceptron e composto com um potente sistema de aprendizado, foi desenvolvido por Widrow auxiliado por estudantes. Foi Widrow que também fundou a primeira companhia de hardware de neurocomputadores e componentes. Algum tempo depois, em 1969, começaram as críticas sobre o perceptron. Minsk e Papert mostraram que o perceptron somente solucionava problemas linearmente separáveis e não era capaz de detectar paridade, conectividade e simetria.

Depois disso, a falta de algoritmos para treinar redes com camadas intermediárias foi um dos fatores que fez com que as pesquisas sobre redes neurais ficassem um bom tempo paradas até que, em 1982, Hopfield escreveu artigos mostrando a relação entre redes recorrentes auto-associativas e sistemas físicos. Apesar desse impulso, foi somente em 1986 que as pesquisas nesta área ganharam força com a publicação do livro "Parallel Distributed Processing" (Processamento Distribuído Paralelo) de autoria de David Rumelhart e James McClelland. Desde então muitos institutos de pesquisa foram criados com a finalidade de incentivar estudos na área de neuro computação.

2.5.2 O cérebro humano e os neurônios biológicos

O cérebro humano é responsável pelo que chamamos de emoção, pensamento, percepção e cognição, assim como pela execução de funções sensoriomotoras e autônomas. Além disso, sua rede de nodos tem a capacidade de reconhecer padrões e relacioná-los, usar e armazenar conhecimento por experiência, além de interpretar observações (BRAGA, 2000).

Assim, plasticidade neuronal é a propriedade do sistema nervoso que permite o desenvolvimento de alterações estruturais em resposta à experiência, e como adaptação a condições mutantes e a estímulos repetidos. A aprendizagem

pode levar a alterações estruturais no cérebro. A cada nova experiência do indivíduo, portanto, redes de neurônios são rearranjadas, outras tantas sinapses são reforçadas e múltiplas possibilidades de respostas ao ambiente tornam-se possíveis.

Segundo Haykin (2002), o sistema nervoso humano, cujo centro é o cérebro, pode ser visto como um sistema de três estágios. Os receptores convertem estímulos do corpo humano ou do ambiente externo em impulsos elétricos que transmitem informação para a rede neural (cérebro). Os atuadores convertem impulsos elétricos gerados pela rede neural em respostas discerníveis como saídas do sistema. Uma representação do diagrama em blocos do sistema nervoso pode ser visto na Figura 1 a seguir.

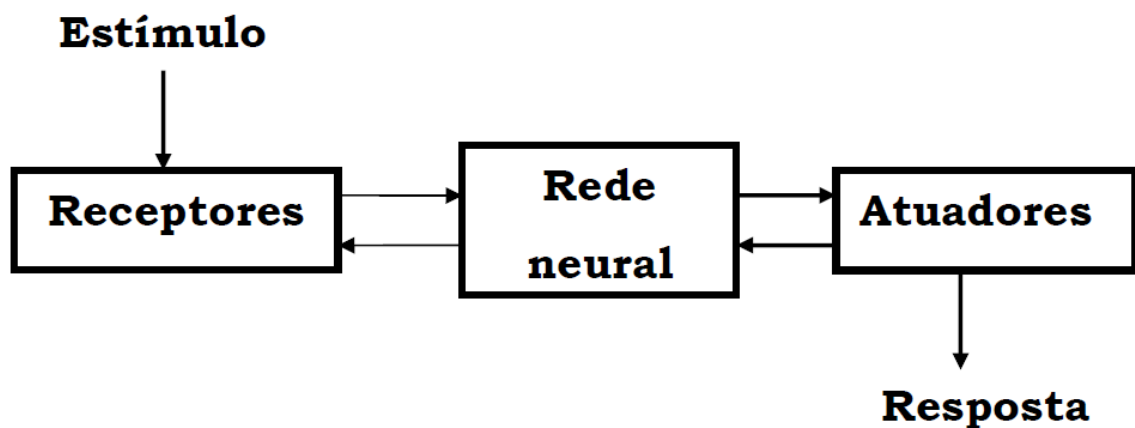


Figura 1. Representação em diagrama de blocos do sistema nervoso. Fonte: Larissa (2004).

De uma maneira geral, os neurônios podem ser divididos em três partes, dependendo da função que executa para a condução do impulso nervoso dentro do cérebro (Figura 2): o corpo celular chamado soma; as ramificações chamadas dendritos que são receptores de mensagem para conduzir até o corpo da célula onde a informação é processada; e o axônio, prolongamento que conduz os impulsos nervosos vindos do corpo celular para suas extremidades (é a saída do sinal). As extremidades do axônio são conectadas com dendritos de outros neurônios pelas sinapses, formando grandes redes. Sinapse é, então, o espaço de contato entre os neurônios e os sinais são transmitidos através dela por neurotransmissores.

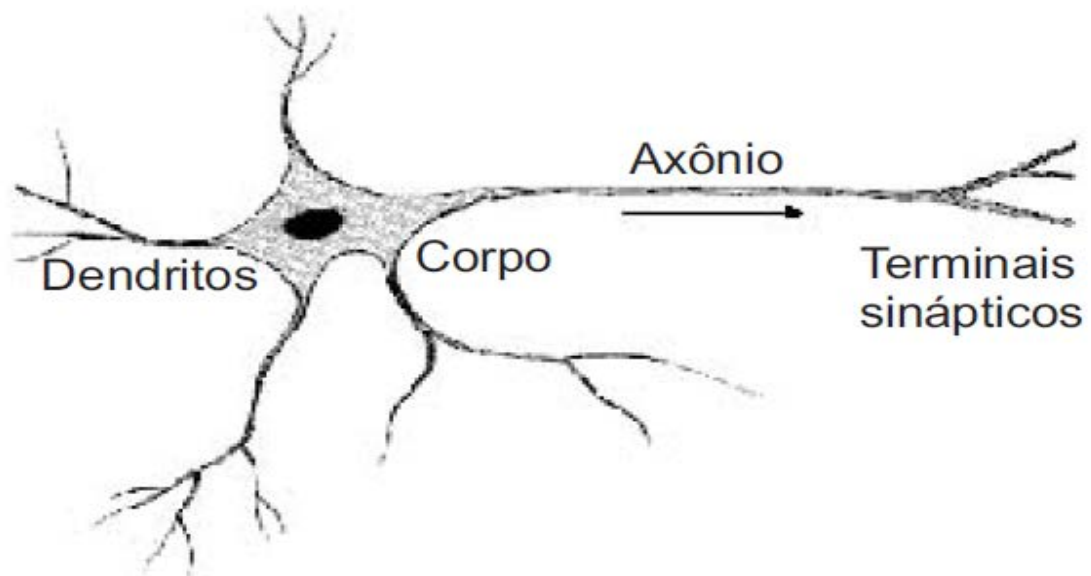


Figura 2. Ilustração simplificada de um neurônio biológico. Fonte: Braga (2000)

2.5.3 Os neurônios artificiais e as RNAs

Para Haykin (2002), uma rede neural artificial é um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso. Ele se assemelha ao cérebro humano em dois aspectos:

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

Com a difícil missão de tentar imitar o sistema nervoso humano, que é composto por bilhões de células nervosas, os cientistas da computação foram levados a projetar uma rede neural artificial que também deve ser formada por diversas unidades modulares que simulam o funcionamento de um neurônio. Para isso, esses módulos computacionais devem funcionar de acordo com os elementos em que foram inspirados, recebendo e retransmitindo informações. Com o avanço da tecnologia computacional surgiu o neurônio artificial que é uma

unidade de processamento de informação de fundamental importância para o funcionamento de uma RNA.

Segundo Haykin (2002), um neurônio artificial é composto de três partes distintas: um conjunto de sinapses ou elos de conexão, cada uma delas com um peso ou força própria, um somador para somar os sinais de entrada e uma função de ativação para restringir a amplitude de saída de um neurônio. Em analogia, os dendritos são substituídos por entradas, onde as ligações com o corpo celular artificial são realizadas através de elementos chamados de peso, simulando as sinapses. Os estímulos captados pelas entradas são processados pela função de somatório, e o limiar de disparo do neurônio biológico foi substituído por uma função de transferência.

Apresentada de uma forma mais simplificada, uma rede neural artificial pode ser visualizada como um grafo onde os nós representam os neurônios e as ligações funcionam como as sinapses, como mostrado na Figura 3.

O modelo de neurônio proposto inicialmente por McCulloch e Pitts é uma simplificação do que se sabia até então a respeito do neurônio biológico. Sua descrição matemática resultou em um modelo com n terminais de entrada x_1, x_2, \dots, x_n (que representam os dendritos) e apenas um terminal de saída y (representando o axônio). Para emular o comportamento das sinapses, os terminais de entrada do neurônio têm pesos acoplados p_1, p_2, \dots, p_n cujos valores podem ser positivos ou negativos, dependendo de as sinapses correspondentes serem inibitórias ou não, respectivamente. O efeito de uma sinapse particular i no neurônio pós-sináptico é dado por $x_i p_i$ (BRAGA, 2000).

O diagrama em blocos da Figura 3 mostra o modelo simplificado de um neurônio, que forma a base para o projeto de redes neurais artificiais. Resumindo, este modelo proposto compreende três elementos básicos:

- um conjunto de n conexões de entrada (x_1, x_2, \dots, x_n), caracterizadas por pesos (p_1, p_2, \dots, p_n);
- um somador (Σ) para acumular os sinais de entrada;
- uma função de ativação (Φ) que limita o intervalo permissível de amplitude do sinal de saída (y) a um valor fixo.

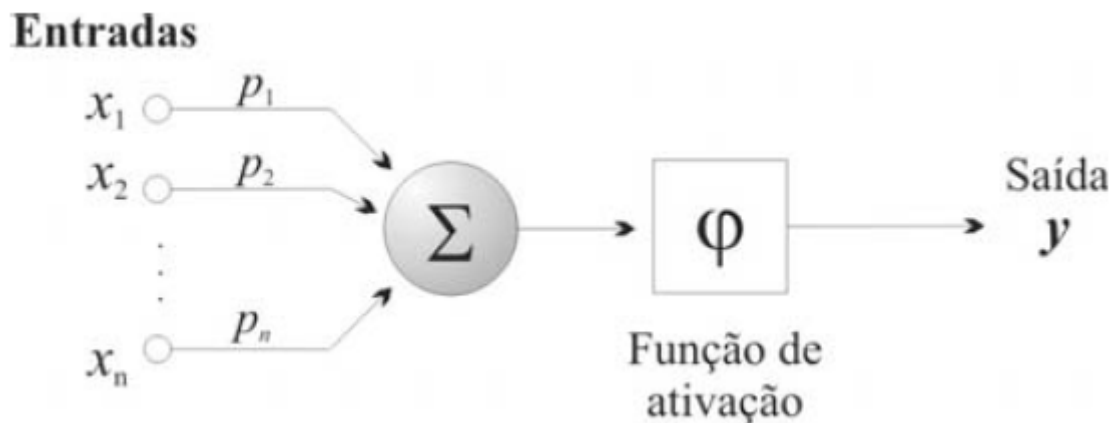


Figura 3: Representação de um modelo de um neurônio parte de um projeto de uma rede neural artificial. Fonte: Haykin (2002)

2.5.4 Função de ativação e capacidade de aprendizagem

As funções de ativação controlam o sinal de saída de um neurônio artificial. Analogamente a um neurônio biológico, a ativação no modelo de McCulloch e Pitts (MCP), é fiscalizada por uma função de ativação que decide se libera ou não o sinal de saída, de acordo com o valor da soma das suas entradas ponderada pelos seus respectivos pesos. No modelo original MCP, a função de ativação é dada pela função de limiar, descrita a seguir (BRAGA, 2000):

- $\sum x_i w_i \geq \Theta$,

onde n é o número de entradas do neurônio, w_i é o peso associado à entrada x_i e Θ é o limiar de ativação (threshold) do neurônio. Esse modelo por ser simples, considera que os nodos em cada camada da rede disparam ao mesmo tempo, o que não é verificado em sistemas biológicos.

Diferentemente de outras técnicas computacionais, o principal propósito de uma rede neural artificial é sua capacidade de aprender com exemplos e melhorar, com isso, seu aprendizado. Ela se baseia nos dados externos e na forma como se relaciona com o meio ambiente para definir um modelo geral de aprendizado. Para isso a rede vai ajustando seus pesos através de um processo iterativo chamado de treinamento da rede que, de uma forma simplificada, é o processo que faz uma rede neural adquirir a habilidade de responder corretamente a estímulos, ajustando uma série de parâmetros internos de acordo com um esquema de aprendizado pré-estabelecido.

Haykin (2002) diz que uma rede neural aprende acerca do seu ambiente através de um processo de ajuste, aplicados a seus pesos sinápticos e níveis de bias. Idealmente, a rede se torna mais instruída sobre o seu ambiente após cada interação do processo de aprendizagem.

A cada iteração do processo de aprendizado de uma rede neural, os pesos são determinados ou modificados de acordo com a seguinte equação:

- $w(t+1) = w(t) + \Delta w(t)$

Em que $w(t)$ e $w(t+1)$ são os valores dos pesos nos instantes t e $t+1$, respectivamente; $\Delta w(t)$ é o ajuste sináptico atualizado depois de cada iteração. Analisando estatisticamente, o aprendizado corresponde a estimar os parâmetros do modelo através de um conjunto de dados informado.

O tipo de aprendizagem é determinado pela maneira na qual a modificação dos parâmetros ocorre. Essa definição do processo de aprendizagem implica a seguinte seqüência de eventos, de acordo Haykin (2002):

- A rede neural é estimulada por um ambiente.
- A rede neural sofre modificações nos seus parâmetros livres como resultado desta estimulação.
- A rede neural responde de uma maneira nova ao ambiente, devido às modificações ocorridas na sua estrutura interna.

Pensando em uma classificação mais generalizada, o tipo de aprendizado pode ser dividido em dois grandes grupos: associativo e não associativo e serão descritos a seguir.

No tipo associativo, o aprendizado se dá sobre o relacionamento existente entre pares de estímulos. Este tipo de aprendizado é amplamente utilizado em modelos de redes neurais artificiais supervisionadas ou aprendizagem com professor. No treinamento da rede neural, pares de entradas e saídas são apresentados à rede que recebe cada entrada, processa e produz uma única resposta na saída. Esta resposta é comparada e, se a resposta real for diferente da resposta desejada, a rede gera um sinal de erro. Caso seja gerado um erro, os pesos sinápticos da rede são ajustados ou reajustados até atingir o resultado desejado, ou seja, os erros são minimizados com a supervisão ou ajuda do professor.

Já no tipo não-associativo que é utilizado em modelos de redes neurais artificiais não-supervisionadas ou aprendizagem sem professor, não requer uma saída desejada para cada vetor de entrada. Ao invés disso, as saídas são organizadas em classes ou categorias. Quando a rede recebe uma entrada, ela gera uma resposta de saída indicando uma determinada classe a qual a entrada pertence. Neste tipo de aprendizado, utilizado em sistemas classificadores, uma nova classe é gerada se o padrão de entrada não encontrar a classe obtida. Neste caso, normalmente várias aproximações são necessárias para minimizar a função erro.

2.5.5 Classificação das redes neurais e a rede MLP

Existem diversos tipos de redes e para se classificar corretamente estas redes tem-se que verificar, pelo menos, a organização dos neurônios e o tipo de conexões estabelecidas entre eles. Dependendo da necessidade, a rede pode possuir uma ou mais camadas intermediárias. A primeira camada é chamada de camada de entrada, as camadas intermediárias são conhecidas como camadas escondidas (ou ocultas) e a última camada é a camada de saída. A complexidade do problema define o total de camadas e o número de neurônios em cada uma delas, principalmente em relação às intermediárias.

A definição da arquitetura de uma RNA é um parâmetro importante na sua concepção, uma vez que ela restringe o tipo de problema que pode ser trabalhado pela rede. Redes com uma camada única de nós MCP, por exemplo, só conseguem resolver problemas linearmente separáveis. Redes recorrentes, por sua vez, são mais apropriadas para resolver problemas que envolvem processamento temporal. Fazem parte da definição da arquitetura os seguintes parâmetros: número de camadas da rede, números de nós em cada camada, tipo de conexão entre os nós e topologia da rede. (BRAGA, 2000)

As redes podem ser classificadas quanto ao número de camadas em:

- Redes de camada única quando somente há um nó entre as entradas e as saídas da rede.
- Redes de múltiplas camadas quando há mais de um nó entre as entradas e as saídas da rede.

Quanto ao tipo de conexão dos nós, as redes podem ser classificadas em:

- Redes alimentadas adiante (feedforward) quando a saída de um neurônio na *i-ésima* camada da rede não pode ser usada para alimentar nodos em camadas de índice menor ou igual a *i*, ou seja, um nodo no mesmo nível ou atrás dele.
- Redes recorrentes (feedback) quando a saída de um neurônio na *i-ésima* camada da rede é usada para alimentar nodos em camadas de índice menor ou igual a *i*, ou seja, não há restrição na alimentação dos nodos.

Juntando estas classificações, podemos definir algumas classes de arquiteturas de redes mais usadas atualmente como: as redes alimentadas adiante com camada única, as redes alimentadas diretamente com múltiplas camadas e as redes recorrentes.

Segundo Haykin (2002), na forma mais simples de uma rede em camadas, tem-se uma camada de entrada de nós de fonte que se projeta sobre uma camada de saída de neurônios (nós computacionais), mas não vice-versa. Esta rede é estritamente do tipo alimentada adiante ou acíclica. A designação camada única se refere à camada de saída, pois a camada de entrada de nós de fonte não é contada, porque lá não é realizada qualquer computação.

A segunda classe de uma rede neural alimentada adiante se distingue pela presença de uma ou mais camadas ocultas, cujos nós computacionais são chamados correspondentemente de neurônios ocultos ou unidades ocultas. A função dos neurônios ocultos é intervir entre a entrada externa e a saída da rede de uma maneira útil. Adicionando-se uma ou mais camadas ocultas, a rede torna capaz de extrair estatísticas de ordem elevada. E isso é importante quando o tamanho da camada de entrada é grande. Em um sentido bastante livre, a rede adquire uma perspectiva global apesar de sua conectividade local, devido ao conjunto extra de conexões sinápticas e da dimensão extra de interações neurais (CHURCHLAND; SEJNOWSKI, 1992).

Ainda de acordo com Haykin (2002), uma rede neural recorrente se distingue de uma rede neural alimentada adiante por ter pelo menos um laço de realimentação. Uma rede recorrente pode consistir, por exemplo, de uma única camada de neurônios com cada neurônio alimentando seu sinal de saída de volta para as entradas de todos os outros neurônios. A presença de laços de realimentação tem um impacto profundo na capacidade de aprendizagem da rede e no seu desempenho. Além disso, os laços de realimentação envolvem o uso de

ramos particulares compostos de elementos de atraso unitário, o que resulta em um comportamento dinâmico não-linear, admitindo-se que a rede neural contenha unidades não-lineares.

As primeiras redes neurais possuíam apenas uma camada e, portanto, eram limitadas em relação à solução de problemas resolvendo somente os linearmente separáveis. Para sanar esse problema foi preciso projetar redes com camadas intermediárias, também conhecidas como camadas escondidas. Na arquitetura de rede chamada de perceptrons de múltiplas camadas (MLP, multilayer perceptron), além de ser possível tratar problemas não linearmente separáveis, ainda possui características computacionais muito mais vantajosas.

De acordo com Cybenko (1989), uma rede com uma camada intermediária pode implementar qualquer função contínua. A utilização de duas camadas intermediárias permite a aproximação de qualquer função. No entanto, existe uma desvantagem no uso de camadas escondidas que é o fato de o aprendizado da rede se tornar muito mais difícil.

Mas daí surgiu um novo problema, o de como treinar as redes com múltiplas camadas. Os perceptrons de múltiplas camadas têm sido aplicados com sucesso para resolver diversos problemas desta natureza, através do seu treinamento de forma supervisionada com um algoritmo muito popular conhecido como algoritmo de retropropagação de erro. Basicamente, a aprendizagem da rede consiste de dois passos através das diferentes camadas da rede: um passo para frente, a propagação, e um passo para trás a retropropagação (HAYKIN, 2002). A propagação das respostas que são recebidas pela camada posterior gera a saída da rede para um padrão de entrada pré-definido e o erro é calculado. Na retropropagação, o erro é propagado em sentido contrário até a camada de entrada, atualizando os pesos, comparando a saída desejada com a saída gerada pela rede. A Figura 4 ilustra a fase de propagação, onde depois de apresentado o padrão de entrada, a resposta de uma unidade é propagada como entrada para as unidades na camada seguinte, até a camada de saída, onde é obtida a resposta da rede e o erro é calculado. Ademais a Figura 5 exemplifica a fase de retropropagação, onde a partir da camada de saída até a camada de entrada, são feitas alterações nos pesos sinápticos.

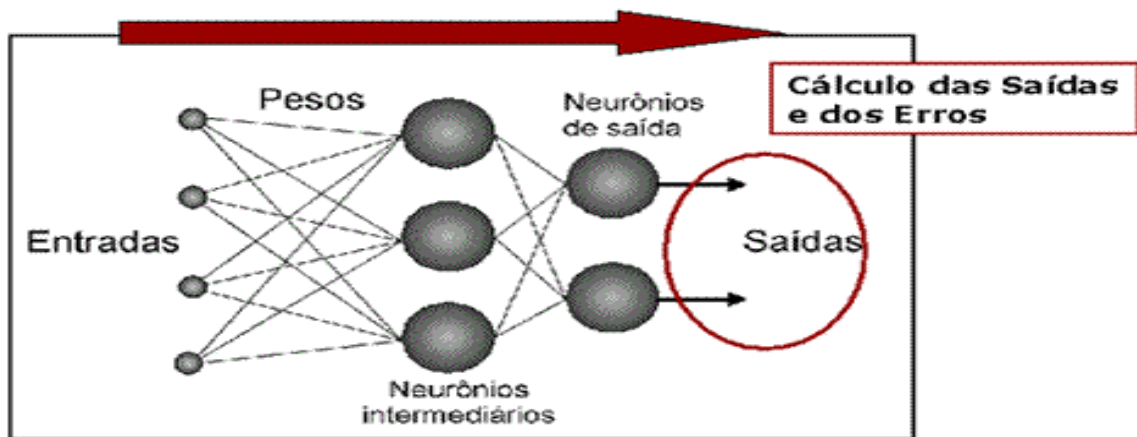


Figura 4: Diagrama representando a fase de propagação do algoritmo de retropropagação de erro.

O algoritmo back-propagation é resumido a seguir (BRAGA, 2000):

1. Inicializar pesos e parâmetros da rede;
2. Repetir até o erro ser mínimo ou até a realização de um número de ciclos:
 - 2.1 Para cada padrão de treinamento X, faça:
 - 2.1.1 Definir saída da rede através do passo propagação.
 - 2.1.2 Comparar saídas produzidas com as saídas desejadas.
 - 2.1.3 Atualizar pesos do nodo através do passo retropropagação.

Ultimamente, as pesquisas tem se concentrado no desenvolvimento de redes neurais construtivas, nas quais a estrutura da rede é determinada dinamicamente, ou seja, a quantidade de nodos ou conexões pode ser aumentada ou diminuída no período de treinamento da rede. A rede Cascade Correlation é um exemplo deste novo paradigma de RNAs.

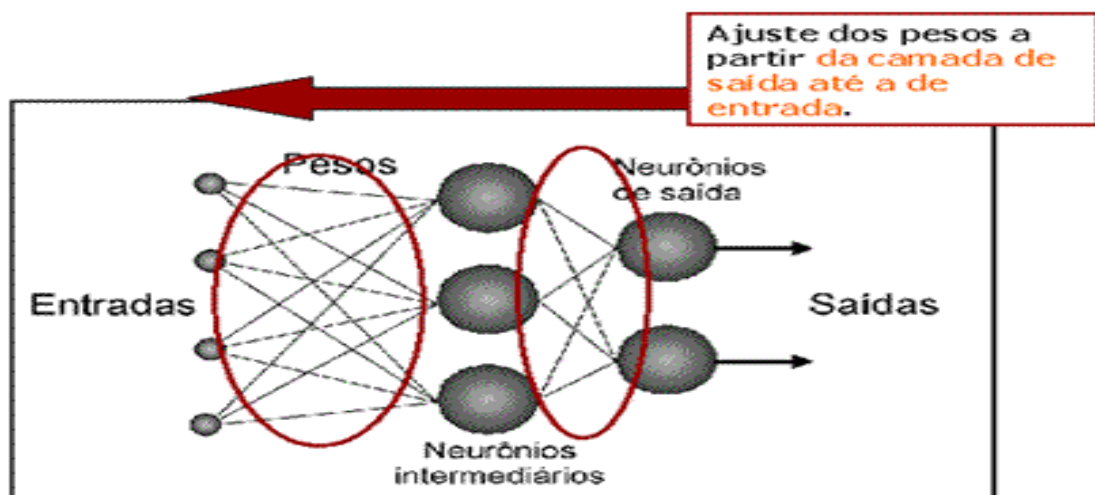


Figura 5: Diagrama representando a fase de retropropagação do algoritmo de retropropagação de erro.

3. CAPÍTULO 1: Identidade de modelos e a resposta da cebola à adubação nitrogenada

3.1 Introdução

A expressão fenotípica é resultado do efeito do genótipo sob a influência do ambiente no qual o mesmo é cultivado. Entretanto, quando o genótipo é avaliado em diversos ambientes, deve-se considerar um efeito adicional, que é dado pela interação entre o genótipo e o ambiente em questão (CRUZ, 2004). Então, quando for verificada a presença deste efeito adicional, pode-se deduzir que genótipos com produtividade alta em certas regiões, nem sempre tenham essa mesma boa produtividade em outras regiões. Por isso, fica evidente a necessidade de estudos que visem mensurar a interação entre genótipos x ambientes para que seja possível a recomendação de genótipos no melhoramento genético de uma maneira geral.

Segundo Madeira et al. (2004), a cebola é a terceira hortaliça mais importante economicamente no Brasil e compreende área total média anual (1999-2001) de cerca de 65,9 mil ha (1,1 milhões t). A área plantada concentra-se em oito estados: 23,4 mil ha em Santa Catarina (393 mil t), 16,3 mil ha no Rio Grande do Sul (179 mil t), 10,5 mil ha em São Paulo (234 mil t), 5,2 mil ha no Paraná (58 mil t), 4,4 mil ha na Bahia (78 mil t), 3,8 mil há em Pernambuco (60 mil t), 2,1 mil ha em Minas Gerais (55 mil t), e 1 mil ha em Goiás (40 mil t). A totalidade da produção é destinada ao mercado interno, basicamente para consumo in natura, como condimento e salada, considerando que a produção de bulbos para industrialização, nas formas de pasta, desidratada e pickles é incipiente. Socialmente se caracteriza como típica de propriedades pequenas e médias e de natureza familiar, principalmente no Sul e no Nordeste brasileiro.

Oliveira et al. (2003) relatam que o nitrogênio é um macronutriente primário essencial para as plantas, por participar da formação de proteínas, aminoácidos e de outros compostos importantes no metabolismo das plantas. Sua ausência bloqueia a síntese de citocinina, hormônio responsável pelo crescimento das plantas, causando redução do seu tamanho e conseqüentemente redução da produção econômica das sementes.

Neste ponto é interessante mencionar também que o suprimento adequado de nitrogênio é necessário para a máxima produtividade e

desenvolvimento da cebola, sendo que a excessiva aplicação, entretanto, pode limitar a produtividade e, assim sendo, aumentar as perdas no armazenamento (SOUZA; RESENDE, 2002). Além disso, o nitrogênio contribui marcadamente para a melhoria da produção de cebola, sendo absorvido em grandes quantidades e superado somente pelo potássio (VIDIGAL et al., 2007).

Do exposto acima, entende-se a necessidade de se estudar as curvas de produtividades provenientes da análise de regressão e compará-las com a quantidade de nitrogênio absorvido pelos cultivares da cebola que, evidentemente, é de suma importância para a formulação de um plano de adubação que seja mais eficiente em termos de ganhos na produtividade.

Para se estudar o efeito de um fator quantitativo e sua relação funcional com as variáveis respostas, utiliza-se a técnica da análise de regressão. Usualmente, procede-se a análise de variância considerando o efeito do fator quantitativo como se fosse a fonte de variação “tratamentos” com o intuito de mensurar a variância residual. Em seguida, o efeito de tratamentos é desmembrado em duas partes: uma para o ajuste de um modelo de regressão e outra para a falta de ajuste deste modelo. A escolha do modelo é realizada selecionando-se aquele que mais se aproxima dos pontos médios observados para cada nível da variável independente.

Finalmente, para avaliar a possibilidade de uma única equação representar o comportamento de variáveis em diferentes locais, aplica-se o teste de identidade entre modelos, descrito por Graybill (1976). Em estudos realizados com características da Candeia (*Eremanthus erythropappus*), em Minas Gerais, também foi empregado por Fabio Camolesi et al. (2010) o teste de identidade entre modelos na tentativa de averiguar agrupamentos possíveis entre municípios.

A justificativa para o estudo que resultou neste capítulo está na necessidade de aplicar e aprimorar o uso de técnicas estatísticas pouco empregadas nas pesquisas com cebolas, mas que ajudem melhorar a produtividade para que, futuramente, seja viável aplicar e divulgar essas técnicas para fins de estudos de predição e modelagem no melhoramento genético.

Para tanto, este estudo teve como objetivo utilizar técnicas estatísticas baseadas em modelos de regressão e ajustar curvas para alguns níveis de doses de nitrogênio, relacionadas à produtividade, para uma pesquisa realizada com

quatro cultivares distintos de cebola e, então, avaliar a possibilidade de agrupamento desses modelos estatísticos obtidos, com base no teste de identidade de modelos. Procurou-se estimar uma curva que representasse, em conjunto, o padrão de resposta à adubação em todos os locais estudados. Em síntese, o objetivo geral é permitir modelagem e previsões mais eficientes no futuro.

3.2 Material e Métodos

Para o desenvolvimento desta pesquisa, foram utilizados dados coletados de dois experimentos que foram conduzidos no período de maio a outubro de 2007 com cebola híbrida Superex.

O primeiro experimento foi realizado na horta de pesquisa da Fazenda Experimental Vale do Piranga pertencente à EPAMIG, localizada no Município de Oratórios, Zona da Mata de Minas Gerais (20,43°S; 42,80°W; 430 m). O solo, Argissolo Vermelho Amarelo, apresentou na camada de 0 a 20 cm de profundidade, as seguintes características: pH (água) 4,5; Ca, Mg, Al, e H+Al respectivamente 1,00; 0,30; 0,50 e 3,96; em $\text{cmol}_c \text{dm}^{-3}$; P e K, 21,80 e 43,00 mg dm^{-3} e matéria orgânica de 11,0 g kg^{-1} . Neste experimento, a área experimental recebeu calagem com antecedência de dois meses sendo aplicados 4,0 t ha^{-1} de calcário dolomítico. Os tratamentos foram distribuídos no delineamento experimental em blocos casualizados com quatro repetições (Cultivares 1 e 2).

O segundo experimento foi realizado na Fazenda Experimental de Mocambinho, pertencente à EPAMIG, localizada no Município de Jaíba, Norte de Minas Gerais (15,08°S; 44,02°W; 452 m). O solo, Neossolo Quartzarênico, apresentou na camada de 0 a 40 cm de profundidade, as seguintes características: pH (água) 5,8; Ca, Mg, Al, e H+Al respectivamente 2,10; 0,25; 0,00 e 0,90; em $\text{cmol}_c \text{dm}^{-3}$; P e K, 70,65 e 44,00 mg dm^{-3} e matéria orgânica de 5,90 g kg^{-1} . O delineamento experimental utilizado foi o de blocos casualizados com três repetições (Cultivares 3 e 4).

Seis foram as doses de nitrogênio utilizadas: 0, 80, 160, 240, 320 e 480 ($\text{kg} \cdot \text{ha}^{-1}$), aplicadas em três épocas diferentes: aos 50, 64 e 99 dias após a semeadura (DAS), na forma de uréia dissolvida em água e distribuída em toda

área da parcela. A semeadura foi realizada nos dias 09 e 21 de maio de 2007 no primeiro e no segundo experimento, respectivamente, e o transplântio 51 dias após, no espaçamento de 0,07 m x 0,25 m. Dentre todas as características mensuradas no experimento, as selecionadas para este trabalho de pesquisa devido às informações estarem mais completas foram: número total de bulbos (NTOT); peso total dos bulbos (PTOT); peso médio dos bulbos (PMTOT); número comerciável de bulbos (NCOM); peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

Neste trabalho, utilizou-se de técnicas estatísticas tradicionais: a análise de variância (ANOVA), a análise de regressão e o teste de identidade de modelos que, aplicados nessa ordem, têm a finalidade de verificar se existe e qual é o melhor modelo de equação que represente, em conjunto, todos os locais para cada variável estudada.

Primeiramente, procedeu-se a análise de variância cujo objetivo é verificar se mudanças ocorridas nas diferentes doses de nitrogênio (causas) sobre as seis variáveis dependentes citadas acima (característica de qualidade) são relevantes para o problema em estudo com cebolas.

A análise de variância foi utilizada com o objetivo maior de certificar quais das variáveis testadas deram resultados significativos para as doses de nitrogênio estudadas. Se o resultado for significativo, conclui-se que as diversas doses de nitrogênio aplicadas nos cultivares de cebola têm efeitos diferentes na produtividade. Em seguida, somente nas que tiveram o resultado significativo, utilizou-se da técnica da análise de regressão polinomial para auxiliar na busca de respostas de como determinadas variáveis influenciam outra variável. Para isso, foram utilizadas técnicas estatísticas com a finalidade de se estabelecer fórmulas empíricas que levam a interpretar a relação funcional entre as variáveis do problema com boa aproximação.

Em seguida, aplicou-se a técnica da análise de regressão polinomial, separadamente, para todas as variáveis que deram resultados significativos para a análise de variância, com o intuito de verificar qual o grau de polinômio do modelo se adequava mais a uma dada variável. Foram feitas as análises de regressão polinomial nos seguintes modelos: linear, quadrática e cúbica para

verificar qual, dentre os três, era o modelo que melhor se adequava, de forma individual, a cada variável em estudo.

Por fim, encerrou-se a aplicação dos métodos pré-determinados para esta pesquisa com o teste de identidade de modelos cuja finalidade é verificar se há possibilidade de determinar uma equação que represente, em conjunto, todos os locais para cada variável estudada. Se isso for possível, ainda é necessário encontrar qual é a melhor equação para esse propósito. Para isso, foram testadas todas as combinações possíveis de agrupamento entre os quatro cultivares estudados.

Todas as análises estatísticas dos dados coletados no experimento com os cultivares de cebola foram processadas através do uso do aplicativo computacional Genes (2006), versão Windows (CRUZ, 2011).

3.3 Resultados e Discussão

Da análise de variância realizada para todas as seis características selecionadas em cada um dos quatro cultivares para o estudo proposto: número total de bulbos (NTOT); peso total dos bulbos (PTOT); peso médio os bulbos (PMTOT); número comerciável de bulbos (NCOM); peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM), verificou-se que as variáveis (NTOT) e (NCOM) não apresentaram diferenças significativas entre as médias de tratamentos para os cultivares 1, 2 e 4, como mostrados nas Tabelas 1, 2 e 4 a seguir. No caso do cultivar 3, somente a variável (NTOT) não apresentou diferença significativa (Tabela 3).

No caso em que a diferença foi significativa, tanto a 1%, sinalizado com dois asteriscos (**); quanto a 5%, sinalizado com apenas um asterisco (*), concluiu-se que as diferentes doses de nitrogênio tiveram efeitos substanciais em relação à combinação de cada variável em cada local, tomados separadamente. Caso, contrário, ou seja, quando a diferença entre as médias das doses de nitrogênio foi não significativa (*ns*), é entendido que a alteração na dosagem do nutriente estudado não surtiu efeito nas variáveis respostas avaliadas. Por esse motivo e, devido ao fato de o objetivo final deste trabalho ser o de fazer uma avaliação dos cultivares em conjunto, os estudos prosseguiram considerando-se apenas as quatro características que tiveram resultados significativos em todos os cultivares

do experimento. São elas: peso total dos bulbos (PTOT), peso médio dos bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM). Resposta à adubação nitrogenada que foi relatada por outros autores. Assim, segundo Vidigal et al. (2010), a resposta encontrada é esperada pelo fato de a adubação nitrogenada em cobertura influenciar a produção de bulbos comercializáveis e, considerando a economia de fertilizante, respostas positivas ao nitrogênio têm sido observadas em pesquisas realizadas em regiões produtoras de cebola do Brasil, como Pernambuco, Minas Gerais e São Paulo, com diferentes variedades e híbridos, no entanto existe uma variação entre as doses de nitrogênio estimadas para a máxima produção. Esta variação pode ser atribuída aos diferentes tipos de solo, época e local de cultivo.

As Tabelas 1, 2, 3 e 4 mostram, para cada híbrido cultivado, individualmente, os resultados da análise de variância para as seis variáveis estudadas em cada um deles, contendo o valor do quadrado médio de cada variável para: blocos, tratamento e resíduo, e as médias e coeficientes de variação de cada característica estudada para os cultivares 1, 2, 3 e 4, respectivamente. Vale mencionar que o coeficiente de variação (CV) é uma medida de dispersão utilizada para estimar a precisão de experimentos e é obtido pela soma dos quadrados do desvio em relação à média. Quanto menor o CV, mais preciso tende a ser o experimento, apesar de essa idéia de precisão ter uma variabilidade inerente a cada área de pesquisa.

Tabela 1: Resultado da análise de variância do cultivar de cebola 1 contendo o quadrado médio e os respectivos coeficientes de variação das variáveis: número total de bulbos (NTOT), peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), número comerciável de bulbos (NCOM), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

ANOVA do cultivar 1		QUADRADO MÉDIO DA VARIÁVEL					
FV	GL	NTOT	PTOT	PMTOT	NCOM	PCOM	PMCOM
BLOCOS	3	220.79	111.36	419.72	220.79	111.36	419.72
TRATAMENTO	5	69.82ns	1513.49**	10878.78**	69.82ns	1513.49**	10878.78**
RESÍDUO	15	238.04	78.47	588.66	238.04	78.47	588.66
MÉDIA		379.33	71.56	189.11	379.33	71.56	189.11
CV		4.06	12.37	12.82	4.06	12.37	12.82

*, significativo a 5%; **, significativo a 1% e ns, não significativo

A Tabela 1 mostra que as variáveis número total de bulbos (NTOT) e número comerciável de bulbos (NCOM) tiveram resultados não significativos enquanto que as variáveis peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM) foram significativas.

Tabela 2: Resultado da análise de variância do cultivar de cebola 2 contendo o quadrado médio e os respectivos coeficientes de variação das variáveis: número total de bulbos (NTOT), peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), número comerciável de bulbos (NCOM), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

ANOVA do cultivar 2		QUADRADO MÉDIO DA VARIÁVEL					
FV	GL	NTOT	PTOT	PMTOT	NCOM	PCOM	PMCOM
BLOCOS	3	9.09	395.45	5250.46	9.58	385.82	5243.38
TRATAMENTO	5	21.83ns	264.06*	3771.56*	22.71ns	268.31*	3634.97*
RESÍDUO	15	16.89	64.13	861.73	23.23	63.05	865.09
MÉDIA		273.03	49.87	182.84	270.89	49.66	183.31
CV		1.50	16.05	16.05	1.77	15.98	16.04

*, significativo a 5%; **, significativo a 1% e ns, não significativo

As Tabelas 2 e 4 mostram resultados semelhantes com relação a quais variáveis foram ou não significativas, porém a Tabela 3 mostra que apenas a variável número total de bulbos (NTOT) teve resultado não significativo.

Tabela 3: Resultado da análise de variância do cultivar de cebola 3 contendo o quadrado médio e os respectivos coeficientes de variação das variáveis: número total de bulbos (NTOT), peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), número comerciável de bulbos (NCOM), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

ANOVA do cultivar 3		QUADRADO MÉDIO DA VARIÁVEL					
FV	GL	NTOT	PTOT	PMTOT	NCOM	PCOM	PMCOM
BLOCOS	2	517.74	914.29	12666.62	381.16	924.81	11707.22
TRATAMENTO	5	276.44ns	813.87**	11254.86**	4015.15**	873.85**	8646.72**
RESÍDUO	10	215.15	40.59	466.49	346.43	39.94	540.44
MÉDIA		270.89	48.31	179.49	252.06	48.01	183.84
CV		5.41	13.18	12.03	7.38	13.16	12.64

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Tabela 4: Resultado da análise de variância do cultivar de cebola 4 contendo o quadrado médio e os respectivos coeficientes de variação das variáveis: número total de bulbos (NTOT), peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), número comerciável de bulbos (NCOM), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

ANOVA do cultivar 4		QUADRADO MÉDIO DA VARIÁVEL					
FV	GL	NTOT	PTOT	PMTOT	NCOM	PCOM	PMCOM
BLOCOS	2	1162.74	432.49	1651.93	3577.41	413.38	1738.96
TRATAMENTO	5	200.61ns	1635.49**	14184.46**	745.54ns	1424.79**	15405.87**
RESÍDUO	10	1075.09	93.60	282.52	869.38	74.65	258.81
MÉDIA		340.38	80.12	235.26	289.99	71.07	242.53
CV		9.63	12.07	7.14	10.16	12.15	6.63

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Trabalhos anteriores já evidenciaram essa resposta da produtividade da cebola relacionada à adubação nitrogenada. Segundo Vidigal et al. (2010), a fertilização com nitrogênio aumenta a produção de cebola, mas a produção de bulbos pode variar com a cultivar, época de plantio e com o tipo de solo.

Após ser realizada a análise de variância, procedeu-se a análise de regressão para os quatro cultivares de cebola, considerando-se apenas as características em que os resultados foram significativos: (PTOT), (PMTOT), (PCOM) e (PMCOM), comparadas em função da variação das seis doses de nitrogênio. Isso foi feito para que fosse possível selecionar dentre as equações obtidas na análise de regressão polinomial, a que tenha o melhor ajuste aos dados.

Para exemplificar, na Tabela 5, a variável peso total dos bulbos (PTOT) do cultivar 1, teve as três equações polinomiais estudadas: linear, quadrática e cúbica; com o resultado significativo. Depois disso, dentre as significativas, verificou-se a que continha o maior R^2 , ou seja, a equação cúbica com o valor de 98,02%. Por isso, essa foi a equação selecionada para representar o modelo. Assim foi feito para a combinação de todos os cultivares com todas as quatro variáveis que seguiram o mesmo raciocínio de interpretação.

Nas Tabelas 5, 6, 7 e 8 são apresentados, individualmente, os resultados da análise de regressão para as quatro variáveis estudadas contendo o valor do quadrado médio de cada variável para os modelos linear, quadrático e cúbico; e o

respectivo coeficiente de determinação de cada característica estudada para os cultivares 1, 2, 3 e 4, respectivamente.

Tabela 5: Resultado da análise de regressão do cultivar de cebola 1 contendo o quadrado médio e os respectivos coeficientes de determinação das variáveis: peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM) para os modelos linear, quadrático e cúbico.

Cultivar 1		QUADRADO MÉDIO E R ²							
FV	GL	PTOT		PMTOT		PCOM		PMCOM	
		QM	R ²	QM	R ²	QM	R ²	QM	R ²
Regressão	3	2472.58**		17802.11**		2472.58**		17802.11**	
Linear	1	2918.77**	78.47	22363.53**	41.11	2918.77**	38.56	22363.53**	41.11
Quadrática	1	3241.09**	81.39	22422.68**	82.33	3241.09**	81.39	22422.68**	82.33
Cúbica	1	1257.87**	98.02	8620.14**	98.18	1257.87**	98.02	8620.14**	98.18
Desvio	2	74.86		493.78		74.86		493.78	
Resíduo	15	78.47		588.66		78.47		588.66	

*, significativo a 5%; **, significativo a 1% e ns, não significativo

As Tabelas 5, 6, 7 e 8 mostram que, na maioria dos casos, a equação cúbica foi a que melhor representou o modelo para a maioria das variáveis.

Tabela 6: Resultado da análise de regressão do cultivar de cebola 2 contendo o quadrado médio e os respectivos coeficientes de determinação das variáveis: peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM) para os modelos linear, quadrático e cúbico.

Cultivar 2		QUADRADO MÉDIO E R ²							
FV	GL	PTOT		PMTOT		PCOM		PMCOM	
		QM	R ²	QM	R ²	QM	R ²	QM	R ²
Regressão	3	438.52**		6272.60**		445.29**		6044.72**	
Linear	1	426.35*	32.29	5547.42*	29.41	434.83*	32.41	5286.48*	29.08
Quadrática	1	526.48*	72.16	8120.79**	72.48	513.79*	70.70	7917.14**	72.64
Cúbica	1	362.73*	99.64	5149.58*	99.78	387.25*	99.57	4930.53*	99.77
Desvio	2	2.36		20.01		2.84		20.35	
Resíduo	15	64.13		861.73		63.05		865.09	

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Tabela 7: Resultado da análise de regressão do cultivar de cebola 3 contendo o quadrado médio e os respectivos coeficientes de determinação das variáveis: peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM) para os modelos linear, quadrático e cúbico.

Cultivar 3		QUADRADO MÉDIO E R ²							
FV	GL	PTOT		PMTOT		PCOM		PMCOM	
		QM	R ²	QM	R ²	QM	R ²	QM	R ²
Regressão	3	1347.66**		18649.18**		1445.87**		14351.55**	
Linear	1	1033.08**	25.38	18622.05**	33.09	1128.09**	25.81	13885.08**	32.11
Quadrática	1	1916.93**	72.49	22590.78**	73.23	2048.91**	72.71	17118.82**	71.71
Cúbica	1	1092.98**	99.35	14734.70**	99.41	1160.61**	99.27	12050.73**	99.58
Desvio	2	13.18		163.38		15.82		89.49	
Resíduo	11	40.59		466.49		39.94		540.44	

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Tabela 8: Resultado da análise de regressão do cultivar de cebola 4 contendo o quadrado médio e os respectivos coeficientes de determinação das variáveis: peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM) para os modelos linear, quadrático e cúbico.

Cultivar 4		QUADRADO MÉDIO E R ²							
FV	GL	PTOT		PMTOT		PCOM		PMCOM	
		QM	R ²	QM	R ²	QM	R ²	QM	R ²
Regressão	3	2701.94**		23371.57**		2361.12**		25516.14**	
Linear	1	3851.49**	47.09	38659.68**	54.50	4263.15**	59.84	43137.08**	56.00
Quadrática	1	4032.29**	96.40	29195.81**	95.67	2502.98**	94.97	32064.04**	97.62
Cúbica	1	222.03ns	99.12	2259.23*	98.86	317.23ns	99.42	1347.31*	99.37
Desvio	2	35.82		403.79		20.30		240.47	
Resíduo	11	93.60		282.52		74.65		258.81	

*, significativo a 5%; **, significativo a 1% e ns, não significativo

A partir dos resultados da análise de regressão mostrados acima, verificou-se que todos os quatro cultivares de cebola foram mais bem representados por equações cúbicas, quando se tratava do peso médio dos bulbos (PMTOT), ou seja, uma equação de grau polinomial 3 foi o modelo estatístico que melhor explicou ou representou o fenômeno estudado (Tabela 10). A mesma conclusão foi obtida ao ser analisado o peso médio comerciável dos

bulbos (PMCOM) por também ter sido encontrada uma equação polinomial de grau 3 para melhor representar o problema (Tabela 12).

Para o peso total dos bulbos (PTOT) e o peso comerciável de bulbos (PCOM), as respostas dos cultivares 1, 2 e 3 também foram mais bem representadas por uma equação cúbica, porém a resposta do cultivar 4 foi mais bem descrita por uma equação polinomial de grau 2 (Tabelas 9 e 11). Nessas Tabelas 9, 10, 11 e 12, pode-se observar também que, na maioria das equações obtidas, a de terceiro grau foi a que melhor representou o modelo para a maioria dentre as quatro variáveis analisadas.

O que se fez depois disso foi obter as estimativas para os parâmetros, a equação e seu respectivo coeficiente de determinação. Nas Tabelas 9, 10, 11 e 12 são apresentadas, separadamente, a equação e o seu coeficiente de determinação (R^2) de cada característica estudada para os cultivares 1, 2, 3 e 4, respectivamente.

Tabela 9: Resultado da análise do peso total dos bulbos (PTOT) contendo a melhor equação polinomial obtida e o coeficiente de determinação para cada um dos quatro cultivares.

Cultivar	Modelo	R²
1	$33,67 + 0,63X^1 - 24,23 \times 10^{-4}X^2 + 2,69 \times 10^{-6}X^3$	98,02
2	$33,92 + 0,30X^1 - 12,36 \times 10^{-4}X^2 + 1,45 \times 10^{-6}X^3$	99,64
3	$16,27 + 0,61X^1 - 25,16 \times 10^{-4}X^2 + 2,90 \times 10^{-6}X^3$	99,35
4	$40,33 + 0,39X^1 - 6,25 \times 10^{-4}X^2$	96,40

Tabela 10: Resultado da análise do peso médio os bulbos (PMTOT) contendo a melhor equação polinomial obtida e o coeficiente de determinação para cada um dos quatro cultivares.

Cultivar	Modelo	R²
1	$87,51 + 1,68X^1 - 63,4 \times 10^{-4}X^2 + 7,05 \times 10^{-6}X^3$	98,18
2	$122,64 + 1,14X^1 - 46,89 \times 10^{-4}X^2 + 5,45 \times 10^{-6}X^3$	99,78
3	$59,24 + 2,23X^1 - 91,37 \times 10^{-4}X^2 + 10,64 \times 10^{-6}X^3$	99,41
4	$107,26 + 1,61X^1 - 46,80 \times 10^{-4}X^2 + 4,17 \times 10^{-6}X^3$	98,86

Tabela 11: Resultado da análise peso comerciável de bulbos (PCOM) contendo a melhor equação polinomial obtida e o coeficiente de determinação para cada um dos quatro cultivares.

Cultivar	Modelo	R²
1	$33,67 + 0,63X^1 - 24,23 \times 10^{-4}X^2 + 2,69 \times 10^{-6}X^3$	98,02
2	$33,60 + 0,30X^1 - 12,68 \times 10^{-4}X^2 + 1,49 \times 10^{-6}X^3$	99,57
3	$14,80 + 0,63X^1 - 25,94 \times 10^{-4}X^2 + 2,99 \times 10^{-6}X^3$	99,27
4	$34,51 + 0,33X^1 - 4,92 \times 10^{-4}X^2$	94,97

Tabela 12: Resultado da análise do peso médio comerciável de bulbos (PMCOM) contendo a melhor equação polinomial obtida e o coeficiente de determinação para cada um dos quatro cultivares.

Cultivar	Modelo	R²
1	$87,51 + 1,68X^1 - 63,4 \times 10^{-4}X^2 + 7,05 \times 10^{-6}X^3$	98,18
2	$124,24 + 1,11X^1 - 45,95 \times 10^{-4}X^2 + 5,33 \times 10^{-6}X^3$	99,77
3	$78,48 + 1,98X^1 - 82,13 \times 10^{-4}X^2 + 9,62 \times 10^{-6}X^3$	99,58
4	$111,04 + 1,55X^1 - 40,78 \times 10^{-4}X^2 + 3,22 \times 10^{-6}X^3$	99,37

Após a obtenção das equações visualizadas nas quatro Tabelas anteriores, iniciou-se a última etapa do processo que era a análise conjunta das equações obtidas com a expectativa de ajustar uma única equação de regressão que representasse dois ou mais locais estudados. Para tanto, aplicou-se ao teste de identidade entre modelos, comparando-se as equações de regressão dos quatro cultivares, levando-se em conta cada uma das quatro características mensuradas separadamente. Os resultados podem ser visualizados nas Figuras 6, 7, 8 e 9 mostradas nas páginas seguintes.

O teste de identidade de modelos fora anteriormente estudado por outros autores. Neter et al. (1985), testaram se duas equações de regressão linear simples eram idênticas utilizando o teste F. Os pesquisadores concluíram que o teste pode ser aplicado para verificar a igualdade de duas equações de regressão polinomial ou de duas equações de regressão múltipla, desde que sejam feitas as modificações adequadas e, ainda, que o teste pode ser estendido, em caso de três ou mais equações.

Regazzi e Silva (2004) concluíram que a identidade de modelos de regressão não-linear e a igualdade de qualquer subconjunto de parâmetros podem ser verificadas por meio do teste da razão de verossimilhança utilizando as aproximações F ou qui-quadrado, mas a aproximação F deve ser preferida. Os mesmos autores ainda relataram, em outro trabalho, que é importante verificar se não há grandes discrepâncias entre os valores observados e preditos, olhar a magnitude da variância residual para decidir se é suficientemente pequena, utilizar análise de resíduos, dentre outras técnicas de diagnóstico e, assim, optar pelo modelo mais adequado.

Como pode ser observado nas Figuras 6 e 7, ao serem analisadas as características peso total dos bulbos (PTOT) e peso comerciável dos bulbos de cebola (PCOM), foi possível gerar uma curva polinomial cúbica que descrevesse somente o comportamento dos cultivares 1 e 4 conjuntamente. Para as características peso médio do total de bulbos (PMTOT) e peso médio comerciável dos bulbos de cebola (PMCOM), foi possível determinar uma curva capaz de representar conjuntamente os cultivares 1, 2 e 3, como mostram as Figuras 8 e 9.

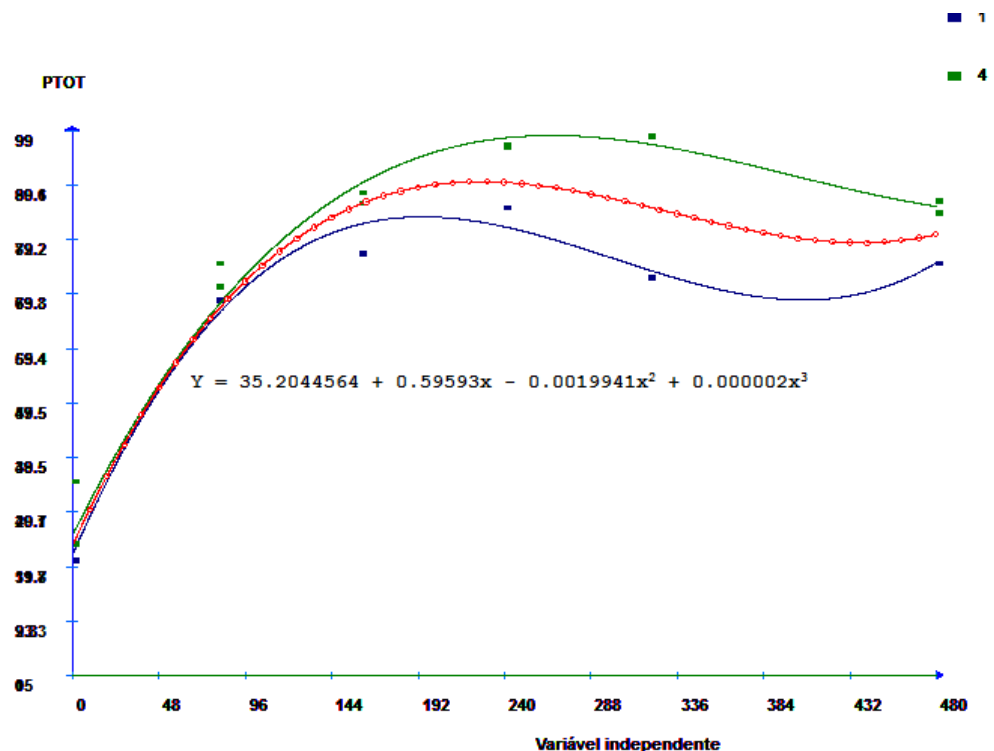


Figura 6: Curva representando os cultivares de cebola 1 e 4 conjuntamente (vermelha pontilhada/círculo) e separadamente (azul e verde) quando considerada a característica peso total dos bulbos (PTOT).

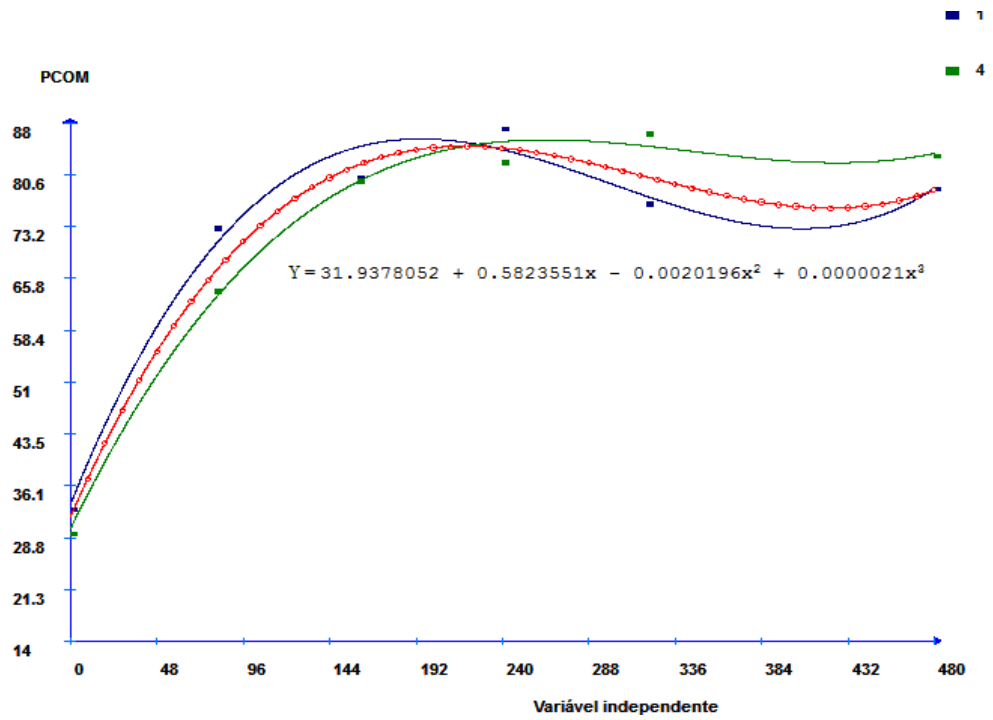


Figura 7: Curva representando os cultivares de cebola 1 e 4 conjuntamente (vermelha pontilhada/círculo) e separadamente (azul e verde) quando considerada a característica peso comercializável dos bulbos de cebola (PCOM).

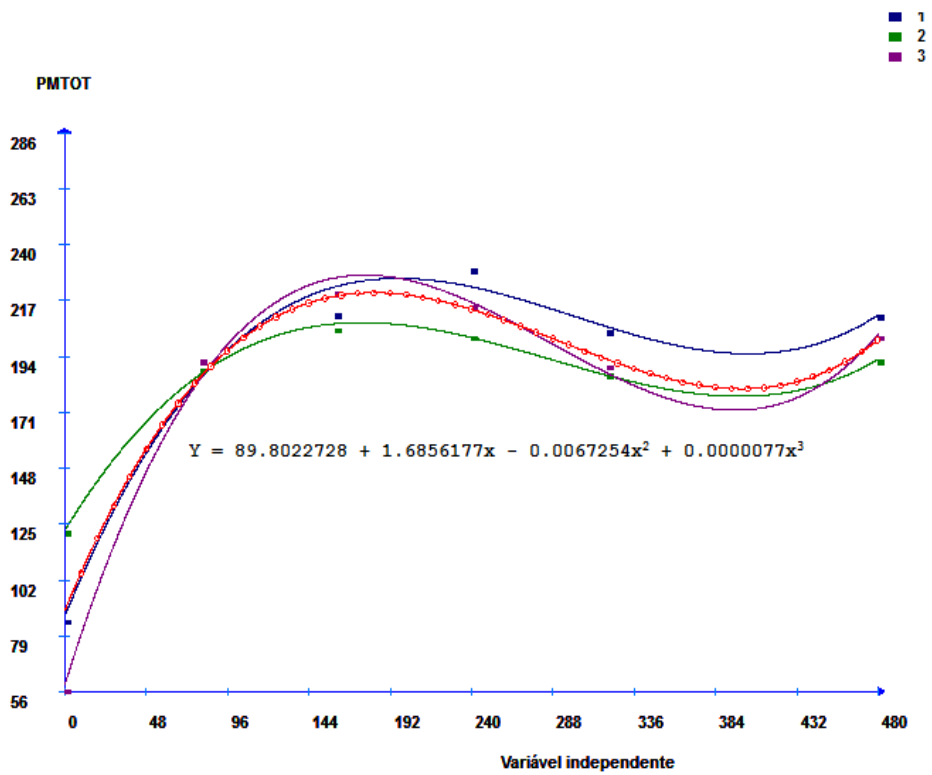


Figura 8: Curva representando os cultivares de cebola 1, 2 e 3 conjuntamente (vermelha pontilhada/círculo) e separadamente (azul, verde e roxa) quando considerada a característica peso médio do total de bulbos (PMTOT).

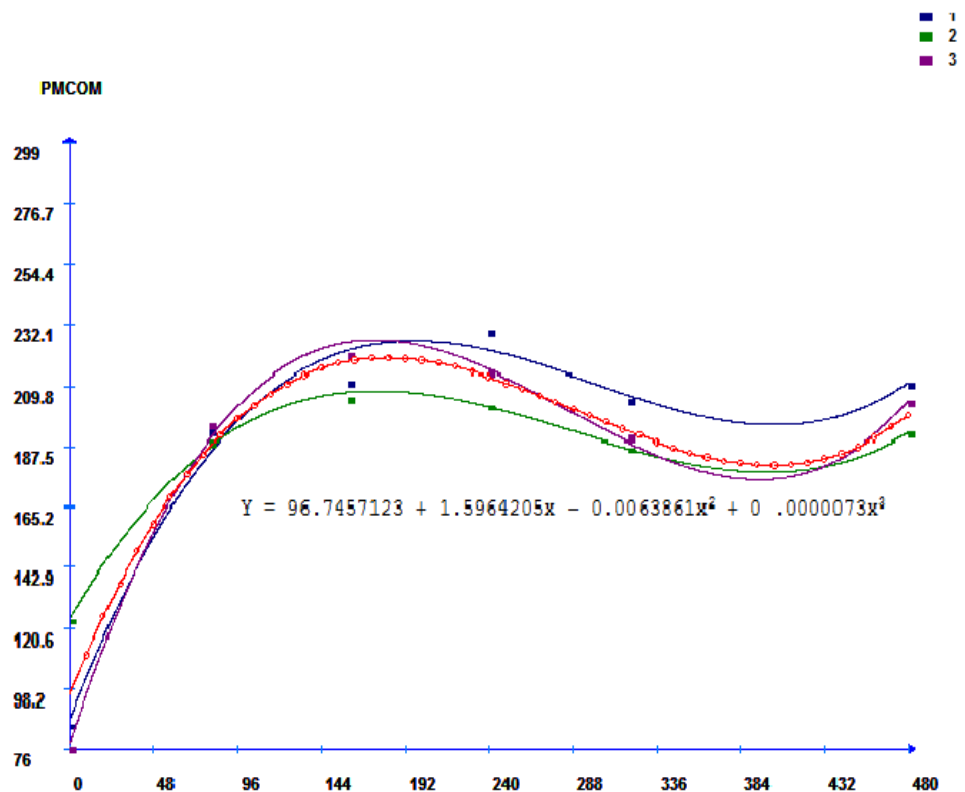


Figura 9: Curva representando os cultivares de cebola 1, 2 e 3 conjuntamente (vermelha pontilhada/círculo) e separadamente (azul, verde e roxa) quando considerada a característica peso médio comerciável dos bulbos de cebola (PMCOM).

3.4 Conclusão

A análise de regressão, seguida de teste de identidade entre modelos, mostrou-se útil para descrever a resposta dos cultivares de cebola à adubação nitrogenada.

Não foi possível estabelecer uma única curva de resposta à adubação nitrogenada, por cultivares de cebola, que fosse capaz de representar os quatro locais estudados conjuntamente tendo em vista a interação genótipos x ambientes.

Foi possível estabelecer uma curva única a partir de determinados pares e de determinados trios de cultivares de cebola, indicando a possibilidade de generalização parcial do comportamento de resposta da cebola à aplicação de nitrogênio.

4. CAPÍTULO 2: Redes neurais artificiais e a predição da resposta da cebola à adubação nitrogenada

4.1 Introdução

Técnicas inovadoras da área da ciência da computação têm surgido, ultimamente, com o propósito de auxiliar análises estatísticas na área de bioinformática. Dados coletados de experimentos agrícolas estão sendo testados para, com isso, verificar se esses dados podem ser analisados utilizando-se técnicas específicas em bioinformática, as denominadas redes neurais artificiais. Essa tecnologia tem sido bastante promissora em diversos problemas relacionados ao melhoramento genético, mas, principalmente, situações em que se deseja prever algo ou para reconhecer padrões.

A inteligência artificial (IA) é conhecida como uma área da ciência da computação que tenta, através dos meios computacionais, construir mecanismos e dispositivos que simulem a capacidade do ser humano de pensar, resolver problemas, ou seja, de agir como um ser inteligente. O interesse no estudo e desenvolvimento desse ramo de pesquisa na área da informática teve início na época da Segunda Guerra Mundial. Dentre os mais importantes idealizadores, podemos citar os seguintes cientistas: Hebert Simon, Allen Newell, John McCarthy e vários outros, que com objetivos em comum tinham a intenção de criar um “ser” que simulasse a vida do ser humano (SCHUTZER, 1987).

Outro fator não menos importante foi a evolução da computação, onde a inteligência artificial ganhou ainda mais força, tendo em vista que o seu desenvolvimento possibilitou grande avanço na análise computacional, podendo a máquina chegar a fazer análise e síntese da voz humana, como exemplo bem impactante. Inicialmente, as pesquisas na área de inteligência artificial buscavam apenas uma forma de reproduzir somente a capacidade do ser humano de pensar, mas assim como todas as pesquisas que evoluem, com essa não foi diferente.

Então, ao perceber que essa subárea da ciência tinha muito mais a ser descoberto, os pesquisadores e cientistas focaram na idéia de fazer, também, com que uma máquina pudesse imitar não só a capacidade de um ser humano pensar como também a capacidade de sentir, de ter criatividade, e de ter auto-aperfeiçoamento e uso da linguagem, ou seja, de agir como um ser humano de

verdade. Mas também não podemos nos esquecer de que os avanços na principal área dessa pesquisa, que é a de fazer uma inteligência similar à do ser humano, é um processo muito lento (SCHUTZER, 1987).

Mesmo assim, os estudos nessa área têm surtido efeito em vários ramos da área agrícola como o diagnóstico e o acompanhamento de doenças, o controle de pragas agrícolas, análise de expressão gênica, alinhamento de seqüências, reconhecimento de genes e outras mais. A bioinformática é um campo das ciências biológicas ligada à ciência da computação que, ultimamente, vem crescendo rapidamente e está sendo desenvolvida com o objetivo de atender à necessidade de manipular e analisar grandes quantidades de dados genéticos. Para compreender a relação entre estas informações provenientes das diversas subáreas, a bioinformática usa o poder computacional para catalogar, organizar e estruturar estas informações em uma entidade compreensiva e extremamente importante para, posteriormente, analisar e tirar conclusões sobre esses dados.

A bioinformática utiliza ferramentas metodológicas da matemática aplicada e da ciência da computação. A Biologia molecular de hoje, por exemplo, seria impossível sem os recursos de bioinformática, tais como o armazenamento, distribuição e atualização das informações, as análises estatísticas, a modelagem de dados e a simulação de fenômenos biológicos em computador.

Segundo Madeira et al. (2004), a cebola é a terceira hortaliça mais importante economicamente no Brasil e compreende área total média anual (1999-2001) de cerca de 65,9 mil ha (1,1 milhões t). A área plantada concentra-se em oito estados: 23,4 mil ha em Santa Catarina (393 mil t), 16,3 mil ha no Rio Grande do Sul (179 mil t), 10,5 mil ha em São Paulo (234 mil t), 5,2 mil ha no Paraná (58 mil t), 4,4 mil ha na Bahia (78 mil t), 3,8 mil há em Pernambuco (60 mil t), 2,1 mil ha em Minas Gerais (55 mil t), e 1 mil ha em Goiás (40 mil t). A totalidade da produção é destinada ao mercado interno, basicamente para consumo in natura, como condimento e salada, considerando que a produção de bulbos para industrialização, nas formas de pasta, desidratada e pickles é incipiente. Socialmente se caracteriza como típica de propriedades pequenas e médias e de natureza familiar, principalmente no Sul e no Nordeste brasileiro.

Oliveira et al. (2003) relatam que o nitrogênio é um macronutriente primário essencial para as plantas, por participar da formação de proteínas,

aminoácidos e de outros compostos importantes no metabolismo das plantas. Sua ausência bloqueia a síntese de citocinina, hormônio responsável pelo crescimento das plantas, causando redução do seu tamanho e conseqüentemente redução da produção econômica das sementes.

Vale a pena mencionar, também, que o suprimento adequado de nitrogênio é necessário para a máxima produtividade e desenvolvimento da cebola, sendo que a excessiva aplicação, entretanto, pode limitar a produtividade e, assim sendo, aumentar as perdas no armazenamento (SOUZA; RESENDE, 2002). Além disso, o nitrogênio contribui marcadamente para a melhoria da produção de cebola, sendo absorvido em grandes quantidades e superado somente pelo potássio (VIDIGAL et al., 2007).

Do exposto acima e com a proposta de aplicar conceitos fundamentados em bioinformática ou, mais precisamente em redes neurais, escolheu-se avaliar algumas variáveis relacionadas à produção da cebola em função da aplicação de doses de nitrogênio.

O presente estudo se justifica pela crescente necessidade de soluções inovadoras e mais eficientes para o desenvolvimento e transferência de conhecimentos entre técnicas estatísticas de computação e de bioinformática que provêm auxílio à tomada de decisão, em benefício da genética e da agricultura como um todo. Por isso, resolveu-se analisar através das redes neurais, algumas características relacionadas com a produtividade de um experimento realizado com cebola, em quatro locais diferentes, para alguns níveis de doses de nitrogênio.

Este trabalho tem o objetivo de apresentar uma aplicação de técnicas específicas provenientes das redes neurais, contendo um modelo computacional baseado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência, em análises de estudos experimentais realizados com híbridos de cebola. Comparações serão feitas com técnicas tradicionais utilizadas atualmente, para que seja possível, em seguida, mensurar a eficiência relativa ao uso dessas técnicas para fins de estudos de predição e modelagem tanto para interpolação quanto para extrapolação de valores auxiliando, com isso, a área de melhoramento genético.

4.2 Material e Métodos

Para o desenvolvimento desta pesquisa, foram utilizados dados coletados de dois experimentos que foram conduzidos no período de maio a outubro de 2007 com cebola híbrida Superex.

O primeiro experimento foi realizado na horta de pesquisa da Fazenda Experimental Vale do Piranga pertencente à EPAMIG, localizada no Município de Oratórios, Zona da Mata de Minas Gerais (20,43°S; 42,80°W; 430 m). O solo, Argissolo Vermelho Amarelo, apresentou na camada de 0 a 20 cm de profundidade, as seguintes características: pH (água) 4,5; Ca, Mg, Al, e H+Al respectivamente 1,00; 0,30; 0,50 e 3,96; em $\text{cmol}_c \text{dm}^{-3}$; P e K, 21,80 e 43,00 mg dm^{-3} e matéria orgânica de 11,0 g kg^{-1} . Neste experimento, a área experimental recebeu calagem com antecedência de dois meses sendo aplicados 4,0 t ha^{-1} de calcário dolomítico. Os tratamentos foram distribuídos no delineamento experimental em blocos casualizados com quatro repetições (Cultivares 1 e 2).

O segundo experimento foi realizado na Fazenda Experimental de Mocaminho, pertencente à EPAMIG, localizada no Município de Jaíba, Norte de Minas Gerais (15,08°S; 44,02°W; 452 m). O solo, Neossolo Quartzarênico, apresentou na camada de 0 a 40 cm de profundidade, as seguintes características: pH (água) 5,8; Ca, Mg, Al, e H+Al respectivamente 2,10; 0,25; 0,00 e 0,90; em $\text{cmol}_c \text{dm}^{-3}$; P e K, 70,65 e 44,00 mg dm^{-3} e matéria orgânica de 5,90 g kg^{-1} . O delineamento experimental utilizado foi o de blocos casualizados com três repetições (Cultivares 3 e 4).

Seis foram as doses de nitrogênio utilizadas: 0, 80, 160, 240, 320 e 480 ($\text{kg} \cdot \text{ha}^{-1}$), aplicadas em três épocas diferentes: aos 50, 64 e 99 dias após a semeadura (DAS), na forma de uréia dissolvida em água e distribuída em toda área da parcela. A semeadura foi realizada nos dias 09 e 21 de maio de 2007 no primeiro e no segundo experimento, respectivamente, e o transplante 51 dias após, no espaçamento de 0,07 m x 0,25 m. Dentre todas as características mensuradas no experimento, as selecionadas para este trabalho de pesquisa devido ao fato de as informações estarem mais completas foram: número total de bulbos (NTOT); peso total dos bulbos (PTOT); peso médio dos bulbos (PMTOT); número comerciável de bulbos (NCOM); peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

Para que fosse possível estudar a metodologia baseada em redes neurais, procedeu-se com a análise de regressão tradicional da estatística, a análise baseada em redes neurais e, finalmente, compará-las em duas situações diferentes: interpolação e extrapolação. Inicialmente, retirou-se um valor central conhecido (240 kg.ha^{-1}) dentre as seis doses de nitrogênio pesquisadas para, então, realizar os cálculos sem esse valor, obter a equação de regressão, estimar o valor correspondente à dose de nitrogênio retirada, comparar os resultados obtidos e verificar a eficiência da predição em ambas as metodologias. Posteriormente, decidiu-se estimar e comparar as duas metodologias retirando-se um valor conhecido extremo (480 kg.ha^{-1}), ou seja, comparar as duas metodologias semelhantes ao caso anterior, mas verificando a predição em relação a uma extrapolação.

Primeiramente, aplicou-se a análise de variância nas variáveis pré-selecionadas. O objetivo era verificar se mudanças ocorridas nas diferentes doses de nitrogênio (causas) sobre as seis variáveis dependentes citadas acima (característica de qualidade) são relevantes para o problema em estudo com cebolas.

Em seguida, aplicou-se a técnica da análise de regressão polinomial, separadamente, para todas as variáveis que deram resultados significativos para a análise de variância. Foram feitas as análises de regressão polinomial para os seguintes modelos: linear, quadrático e cúbico para verificar qual, dentre os três, era o modelo que melhor se adequava, de forma individual, a cada variável em estudo. Depois de obtida a equação, foram feitas as estimativas correspondentes aos valores 240 kg.ha^{-1} e 480 kg.ha^{-1} , retirados previamente.

Após serem determinados todos os valores necessários para a comparação, iniciou-se o processamento do aplicativo das redes neurais. O programa também foi executado individualmente para cada uma das variáveis em todos os quatro cultivares de cebola.

O aplicativo simulador das redes neurais deste estudo foi criado utilizando o software MATLAB R2012a e seu kit de ferramentas *ntoll* que engloba pacotes contendo diversas funções matemáticas e estatísticas específicas para confecção de algoritmos utilizados em redes neurais. Antes de escolher uma rede para predição de valores, foram testadas diferentes arquiteturas com o objetivo de encontrar uma que melhor se adequasse ao problema em questão.

O tipo de treinamento determinado para a rede foi o supervisionado, pois foi utilizada a arquitetura feedforward muito requisitada atualmente. Para isso, foi necessário selecionar um conjunto de dados para o treinamento inicial da rede. Esses dados nada mais são do que um conjunto de pares de entradas e saídas requeridas pela rede, para que ela aprenda a executar seu trabalho reconhecendo padrões em exemplos que são previamente inseridos nela. Outro parâmetro importante na configuração da rede são os pesos que se alteram dinamicamente através das chamadas regras de treinamento. A regra mais utilizada no tipo feedforward é o algoritmo de retro-propagação de erros (Errorbackpropagation) que atualiza os erros pelas derivadas dos erros em relação aos pesos.

A rede neural usada neste trabalho foi a perceptron multicamadas, testando-se um número variado de neurônios em cada uma das três camadas escondidas. O número de neurônios utilizado em cada camada não foi fixado, podendo variar (mínimo de um e máximo de três) a quantidade deles de cenário para cenário. As funções de ativação empregadas nesta configuração foram: a função *purelin* para a camada de entrada, *trainbr* para camada de saída da rede; e a *logsig*, *tansig* ou *purelin* para as camadas escondidas. Convencionou-se escolher esta configuração para manter a não-linearidade da rede.

Já com relação ao número de ciclos de treinamento da rede, este foi fixado em 2000 épocas (cenário 3) e 7000 épocas (cenários 1 e 2). Estipulou-se limitar o número de iterações quando certa precisão fosse alcançada, porém tendo o cuidado para que esse número também não fosse muito pequeno, evitando com isso, que a rede não deixasse de atingir o seu melhor desempenho.

Após definidos todos os parâmetros e o algoritmo estando sem nenhum tipo de erro: sintático ou lógico, considera-se o programa pronto para fazer as previsões. A configuração final da rede que foi utilizada está apresentada na Tabela 13.

Com a rede neural criada e treinada, realizou-se o processamento para a obtenção do valor previsto pela rede, de cada uma das variáveis para cada um dos quatro cultivares. Esse processamento teve a finalidade de obter o valor previsto pela rede tanto para a interpolação como também para a extrapolação. Depois disso, fez-se a comparação com os valores obtidos pela análise de regressão polinomial convencional e, então, ratificar se os resultados obtidos

pelas redes neurais são próximos ou melhores que aqueles previstos na análise estatística.

Tabela 13: Configuração dos parâmetros utilizados pela rede neural para predição dos valores do estudo de resposta dos cultivares de cebolas à adubação nitrogenada.

Parâmetro	Descrição
Número de camadas	3 (entrada, camada de escondida, saída)
Número de neurônios por camadas	10 (entrada), 6 (escondida), 3 (saída)
Função de ativação	purelin (entrada), trainbr (saída) e logsig, tansig ou purelin (escondida)
Função de treinamento	traingdx - backpropagation
Função de performance	mse – erro quadrático médio
Número de épocas de treinamento	2000 (cenário 3) e 7000 (cenários 1 e 2)
Número de dados para treinamento	1050
Número de dados para validação	300
Número de dados para teste	100

4.3 Resultados e Discussão

O tópico resultados e discussão foi dividido em dois subtópicos com o objetivo de facilitar o entendimento dos resultados por parte do leitor, primeiro referente à interpolação e, segundo, à extrapolação de valores. Pela considerável quantidade de cálculos executados para que fosse possível a comparação entre os dois métodos, foi gerado grande número de tabelas e figuras o que exigiu essa divisão para facilitar a visualização e entendimento, apesar de os resultados e conclusões obtidos serem semelhantes.

4.3.1 Analisando o caso da interpolação de valores

Após a retirada do valor 240 kg.ha^{-1} correspondente a uma dentre as doses de nitrogênio do experimento, realizou-se novamente a análise de variância de todas as seis características selecionadas em cada um dos quatro cultivares do estudo proposto: número total de bulbos (NTOT); peso total dos bulbos (PTOT); peso médio os bulbos (PMTOT); número comerciável de bulbos (NCOM); peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM). Os cálculos determinaram que as variáveis (NTOT) e (NCOM) não apresentaram diferenças significativas entre as médias de tratamentos para os

cultivares 1, 2 e 4, como mostrados nas Tabelas 14, 15 e 17 a seguir. No caso do cultivar 3, somente a variável (NTOT) não apresentou diferença significativa (Tabela 16).

No caso em que a diferença foi significativa, tanto a 1%, sinalizado com dois asteriscos (**); quanto a 5%, sinalizado com apenas um asterisco (*), concluiu-se que as diferentes doses de nitrogênio tiveram efeitos substanciais em relação à combinação de cada variável em cada local, tomados separadamente. Caso contrário, quando a diferença entre as médias das doses de nitrogênio foi não significativa (*ns*), é entendido que a alteração na dosagem do nutriente estudado não surtiu efeito nas variáveis respostas avaliadas. Por esse motivo e, devido ao fato de o objetivo final deste trabalho ser o de fazer uma avaliação de todos os cultivares, os estudos prosseguiram considerando-se apenas as quatro características que tiveram resultados significativos em todos os cultivares do experimento. São elas: peso total dos bulbos (PTOT), peso médio dos bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM). Resposta à adubação nitrogenada que foi relatada por outros autores. Assim, segundo Vidigal et al. (2010), a resposta encontrada é esperada pelo fato de a adubação nitrogenada em cobertura influenciar a produção de bulbos comercializáveis e, considerando a economia de fertilizante, respostas positivas ao nitrogênio têm sido observadas em pesquisas realizadas em regiões produtoras de cebola do Brasil, como Pernambuco, Minas Gerais e São Paulo, com diferentes variedades e híbridos, no entanto existe uma variação entre as doses de nitrogênio estimadas para a máxima produção. Esta variação pode ser atribuída aos diferentes tipos de solo, época e local de cultivo.

As Tabelas 14, 15, 16 e 17 mostram, para cada híbrido, os resultados da análise de variância para as seis variáveis estudadas em cada um deles, contendo o valor do quadrado médio de cada variável para: blocos, tratamento e resíduo, e as médias e coeficientes de variação de cada característica estudada para os cultivares 1, 2, 3 e 4, respectivamente. O coeficiente de variação (CV) é uma medida de dispersão utilizada para estimar a precisão de experimentos e é obtido pela soma dos quadrados do desvio em relação à média. Quanto menor o CV, mais preciso tende a ser o experimento, apesar de essa idéia de precisão ter uma variabilidade inerente a cada área de pesquisa.

Tabela 14: Resultado da análise de variância do cultivar de cebola 1 contendo o quadrado médio e os respectivos coeficientes de variação das variáveis: número total de bulbos (NTOT), peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), número comerciável de bulbos (NCOM), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

ANOVA do cultivar 1		QUADRADO MÉDIO DA VARIÁVEL					
FV	GL	NTOT	PTOT	PMTOT	NCOM	PCOM	PMCOM
BLOCOS	3	179.58	94.80	416.53	179.58	94.80	416.53
TRATAMENTO	4	86.06ns	1590.19**	11609.95**	86.06ns	1590.19**	11609.95**
RESÍDUO	12	278.34	97.80	719.34	278.34	97.80	719.34
MÉDIA		379.13	68.39	180.97	379.13	68.39	180.97
CV		4.40	14.46	14.82	4.40	14.46	14.82

*, significativo a 5%; **, significativo a 1% e ns, não significativo

As Tabelas 14, 15 e 17 mostram que as variáveis número total de bulbos (NTOT) e número comerciável de bulbos (NCOM) tiveram resultados não significativos (*ns*) enquanto que as variáveis peso total dos bulbos (PTOT), peso médio dos bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM) foram significativas (**). A Tabela 16 exibiu um resultado um pouco diferente, que pode ser notado observando-se a quinta coluna, se comparado com as outras três Tabelas. Nela pode ser visto que apenas a variável número total de bulbos (NTOT) teve resultado não significativo (*ns*).

Tabela 15: Resultado da análise de variância do cultivar de cebola 2 contendo o quadrado médio e os respectivos coeficientes de variação das variáveis: número total de bulbos (NTOT), peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), número comerciável de bulbos (NCOM), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

ANOVA do cultivar 2		QUADRADO MÉDIO DA VARIÁVEL					
FV	GL	NTOT	PTOT	PMTOT	NCOM	PCOM	PMCOM
BLOCOS	3	6.82	223.11	2926.02	7.01	220.16	2901.14
TRATAMENTO	4	22.51ns	301.11*	4266.86*	25.43ns	308.10*	4100.63*
RESÍDUO	12	19.00	58.18	789.14	26.99	59.28	781.81
MÉDIA		273.42	48.89	178.98	271.20	48.71	179.47
CV		1.59	15.59	15.69	1.91	15.80	15.57

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Tabela 16: Resultado da análise de variância do cultivar de cebola 3 contendo o quadrado médio e os respectivos coeficientes de variação das variáveis: número total de bulbos (NTOT), peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), número comerciável de bulbos (NCOM), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

ANOVA do cultivar		QUADRADO MÉDIO DA VARIÁVEL					
FV	GL	NTOT	PTOT	PMTOT	NCOM	PCOM	PMCOM
BLOCOS	2	420.38	676.41	9615.97	358.51	686.39	8710.17
TRATAMENTO	4	344.69ns	925.53**	12935.63**	4665.88**	994.97**	9936.56**
RESÍDUO	8	257.21	46.73	532.79	402.64	46.05	611.82
MÉDIA		270.69	46.29	172.39	248.10	45.93	177.62
CV		5.92	14.76	13.38	8.08	14.77	13.92

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Tabela 17: Resultado da análise de variância do cultivar de cebola 4 contendo o quadrado médio e os respectivos coeficientes de variação das variáveis: número total de bulbos (NTOT), peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), número comerciável de bulbos (NCOM), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

ANOVA do cultivar		QUADRADO MÉDIO DA VARIÁVEL					
FV	GL	NTOT	PTOT	PMTOT	NCOM	PCOM	PMCOM
BLOCOS	2	1297.45	526.02	2018.17	4942.13	558.31	2082.06
TRATAMENTO	4	248.68ns	1794.01**	15479.64**	821.48ns	1661.70**	16678.51**
RESÍDUO	8	1243.94	74.29	216.61	530.24	32.65	172.02
MÉDIA		340.69	76.78	225.26	292.20	68.77	231.82
CV		10.35	11.22	6.53	7.88	8.30	5.65

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Feita a análise de variância considerando como fatores as cinco doses de nitrogênio do estudo, iniciou-se a análise de regressão para os quatro cultivares de cebola, avaliando apenas as características em que os resultados foram significativos: (PTOT), (PMTOT), (PCOM) e (PMCOM). Isso foi feito para que fosse possível selecionar dentre as equações significativas obtidas na análise de regressão polinomial, a que tivesse o melhor ajuste aos dados.

Para exemplificar, na Tabela 18, a variável peso total dos bulbos (PTOT) do cultivar 1, teve as três equações polinomiais estudadas: linear, quadrática e cúbica; com o resultado significativo a 5% (*). Então, dentre as significativas,

escolheu-se a que continha o maior grau polinomial, ou seja, para este exemplo a selecionada para representar o modelo em estudo foi a equação cúbica. Ainda para o peso total dos bulbos (PTOT), as respostas dos cultivares 2, 3 também foram mais bem representadas por uma equação cúbica (Tabelas 19 e 20), porém a resposta do cultivar 4 foi mais bem descrita por uma equação polinomial quadrática (Tabela 21).

Continuando a análise, verificou-se que todos os quatro cultivares de cebola foram mais bem representados por equações cúbicas, quando se tratava do peso médio dos bulbos (PMTOT), ou seja, uma equação de grau polinomial 3 foi o modelo matemático que melhor explicou ou representou o fenômeno estudado. A mesma conclusão foi obtida ao ser analisado o peso médio comerciável dos bulbos (PCOM) por também ter sido encontrada uma equação polinomial de grau 3 para representar o problema. O mesmo sendo verificado em todos os cultivares para a variável peso comerciável (PCOM).

Assim foi feito para a combinação de todos os cultivares com todas as quatro variáveis. Nas Tabelas 18, 19, 20 e 21 são apresentados, individualmente, os resultados da análise de regressão para as quatro variáveis estudadas contendo o valor do quadrado médio de cada variável para os modelos linear, quadrático e cúbico; e o respectivo coeficiente de determinação de cada característica estudada para os cultivares 1, 2, 3 e 4, respectivamente.

Tabela 18: Resultado da análise de regressão do cultivar de cebola 1 contendo o quadrado médio e os respectivos coeficientes de determinação das variáveis: peso total dos bulbos (PTOT), peso médio dos bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PCOM) para os modelos linear, quadrático e cúbico.

Cultivar 1		QUADRADO MÉDIO E R ²							
FV	GL	PTOT		PMTOT		PCOM		PCOM	
		QM	R ²	QM	R ²	QM	R ²	QM	R ²
Regressã	3	2092.34*		15260.57*		2092.34*		15260.57*	
Linear	1	2657.12*	41.	20509.77*	44.1	2657.12*	41.7	20509.77*	44.1
Quadrátic	1	2311.48*	78.	16373.99*	79.4	2311.48*	78.1	16373.99*	79.4
Cúbica	1	1308.42*	98.	8897.94**	98.5	1308.42*	98.6	8897.94**	98.5
Desvio	2	83.77		658.08		83.77		658.08	
Resíduo	12	97.80		719.34		97.80		719.34	

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Tabela 19: Resultado da análise de regressão do cultivar de cebola 2 contendo o quadrado médio e os respectivos coeficientes de determinação das variáveis: peso total dos bulbos (PTOT), peso médio dos bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM) para os modelos linear, quadrático e cúbico.

Cultivar 2		QUADRADO MÉDIO E R ²							
FV	GL	PTOT		PMTOT		PCOM		PMCOM	
		QM	R ²	QM	R ²	QM	R ²	QM	R ²
Regressão	3	399.92**		5676.19**		408.90**		5454.61**	
Linear	1	395.68*	32.85	5110.39*	29.94	404.82*	32.84	4861.61*	29.63
Quadrática	1	445.39*	69.82	6820.96*	69.90	439.68*	68.52	6615.70*	69.97
Cúbica	1	358.70*	99.61	5097.22*	99.77	382.21*	99.53	4886.52*	99.76
Desvio	2	4.69		38.88		5.69		38.69	
Resíduo	12	58.18		789.14		59.28		781.81	

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Pode-se verificar, observando as Tabelas 18, 19, 20 e 21 que, exceto a variável peso total dos bulbos (PTOT) que foi quadrática, todas as outras variáveis mensuradas em todos os cultivares, tiveram a equação polinomial cúbica como a que melhor representou o modelo. Além disso, nestas mesmas tabelas é possível observar também o coeficiente de determinação (R²) que diz o quanto o modelo conseguiu explicar a variação nas doses de nitrogênio em relação a cada variável estudada. Como todos eles estão próximos a 100%, conclui-se que a precisão do experimento foi satisfatória.

Tabela 20: Resultado da análise de regressão do cultivar de cebola 3 contendo o quadrado médio e os respectivos coeficientes de determinação das variáveis: peso total dos bulbos (PTOT), peso médio dos bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM) para os modelos linear, quadrático e cúbico.

Cultivar 3		QUADRADO MÉDIO E R ²							
FV	GL	PTOT		PMTOT		PCOM		PMCOM	
		QM	R ²	QM	R ²	QM	R ²	QM	R ²
Regressã	3	1225.55*		17141.25*		1316.49*		13189.72*	
Linear	1	947.47**	25.	17358.28*	33.5	1036.04*	26.0	12926.84*	32.5
Quadrátic	1	1658.32*	70.	19611.01*	71.4	1777.41*	70.6	14790.73*	69.7
Cúbica	1	1070.85*	99.	14454.47*	99.3	1136.02*	99.2	11851.60*	99.5
Desvio	2	25.47		318.78		30.41		177.09	
Resíduo	8	46.73		532.79		46.05		611.82	

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Tabela 21: Resultado da análise de regressão do cultivar de cebola 4 contendo o quadrado médio e os respectivos coeficientes de determinação das variáveis: peso total dos bulbos (PTOT), peso médio dos bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM) para os modelos linear, quadrático e cúbico.

Cultivar 4		QUADRADO MÉDIO E R ²							
FV	GL	PTOT		PMTOT		PCOM		PMCOM	
		QM	R ²	QM	R ²	QM	R ²	QM	R ²
Regressã	3	2369.70*		20374.81*		2213.96*		22084.78*	
Linear	1	3580.76*	49.	36096.73*	58.2	4073.53*	61.2	40236.78*	60.3
Quadrátic	1	3316.77*	96.	22758.62*	95.0	2279.63*	95.5	24649.11*	97.2
Cúbica	1	211.58ns	99.	2269.08*	98.7	288.72*	99.9	1368.45*	99.3
Desvio	2	66.94		794.11		4.93		459.72	
Resíduo	8	74.29		216.61		32.65		172.02	

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Os resultados obtidos e descritos nas tabelas anteriores permitiram determinar qual é o grau da equação polinomial que melhor se ajustou aos dados de cada variável analisada em separado. O que se fez depois disso foi obter as estimativas para os parâmetros, a equação, a estimativa para a dosagem que foi retirada antes da realização dos cálculos e o coeficiente de determinação da equação. Nas Tabelas 22, 23, 24 e 25 são apresentadas, separadamente, a equação, a estimativa para a dosagem 240 kg.ha⁻¹ de nitrogênio (Est.₂₄₀) e o coeficiente de determinação (R²) do modelo ajustado de cada característica estudada para os cultivares 1, 2, 3 e 4, respectivamente.

Tabela 22: Resultado da análise do peso total dos bulbos (PTOT) contendo a melhor equação polinomial obtida, a estimativa para a dosagem 240 kg.ha⁻¹ de nitrogênio (Est.₂₄₀) e o coeficiente de determinação para cada um dos quatro cultivares.

Cultivar	Modelo	Est. ₂₄₀	R ²
1	$34,04 + 0,62X^1 - 24,37 \times 10^{-4}X^2 + 2,76 \times 10^{-6}X^3$	82,25	98,68
2	$33,93 + 0,29X^1 - 12,36 \times 10^{-4}X^2 + 1,45 \times 10^{-6}X^3$	54,70	99,61
3	$16,22 + 0,61X^1 - 25,14 \times 10^{-4}X^2 + 2,89 \times 10^{-6}X^3$	59,14	99,31
4	$40,03 + 0,40X^1 - 6,45 \times 10^{-4}X^2$	99,61	96,11

Tabela 23: Resultado da análise do peso médio os bulbos (PMTOT) contendo a melhor equação polinomial obtida, a estimativa para a dosagem 240 kg.ha⁻¹ de nitrogênio (Est.₂₄₀) e o coeficiente de determinação para cada um dos quatro cultivares.

Cultivar	Modelo	Est. ₂₄₀	R ²
1	$88,31 + 1,65X^1 - 63,80 \times 10^{-4}X^2 + 7,21 \times 10^{-6}X^3$	218,41	98,58
2	$122,69 + 1,13X^1 - 46,90 \times 10^{-4}X^2 + 5,46 \times 10^{-6}X^3$	201,55	99,77
3	$59,10 + 2,23X^1 - 91,32 \times 10^{-4}X^2 + 10,61 \times 10^{-6}X^3$	217,06	99,38
4	$107,44 + 1,61X^1 - 46,88 \times 10^{-4}X^2 + 4,20 \times 10^{-6}X^3$	282,56	98,71

Tabela 24: Resultado da análise peso comerciável de bulbos (PCOM) contendo a melhor equação polinomial obtida, a estimativa para a dosagem 240 kg.ha⁻¹ de nitrogênio (Est.₂₄₀) e o coeficiente de determinação para cada um dos quatro cultivares.

Cultivar	Modelo	Est. ₂₄₀	R ²
1	$34,04 + 0,62X^1 - 24,37 \times 10^{-4}X^2 + 2,76 \times 10^{-6}X^3$	82,25	98,68
2	$33,60 + 0,30X^1 - 12,68 \times 10^{-4}X^2 + 1,49 \times 10^{-6}X^3$	54,35	99,53
3	$14,74 + 0,63X^1 - 25,92 \times 10^{-4}X^2 + 2,97 \times 10^{-6}X^3$	59,17	99,23
4	$29,88 + 0,53X^1 - 16,04 \times 10^{-4}X^2 + 1,50 \times 10^{-6}X^3$	86,95	99,92

Tabela 25: Resultado da análise do peso médio comerciável de bulbos (PMCOM) contendo a melhor equação polinomial obtida, a estimativa para a dosagem 240 kg.ha⁻¹ de nitrogênio (Est.₂₄₀) e o coeficiente de determinação para cada um dos quatro cultivares.

Cultivar	Modelo	Est. ₂₄₀	R ²
1	$88,31 + 1,65X^1 - 63,80 \times 10^{-4}X^2 + 7,21 \times 10^{-6}X^3$	218,41	98,58
2	$124,30 + 1,11X^1 - 45,97 \times 10^{-4}X^2 + 5,34 \times 10^{-6}X^3$	201,62	99,76
3	$78,41 + 1,99X^1 - 82,11 \times 10^{-4}X^2 + 9,61 \times 10^{-6}X^3$	216,04	99,55
4	$111,28 + 1,54X^1 - 40,87 \times 10^{-4}X^2 + 3,26 \times 10^{-6}X^3$	292,66	99,31

Com a conclusão dos cálculos através da análise de regressão, iniciaram-se os mesmos cálculos, mas agora utilizando as redes neurais para que fosse possível a comparação entre as metodologias. Para isso, em cada um dos quatro cultivares foi utilizado o algoritmo da rede para prever valores correspondentes ao valor 240 kg.ha⁻¹, para cada uma das seguintes características: peso total dos bulbos (PTOT), peso médio dos bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

Foram plotados dois gráficos, para cada um dos quatro cultivares estudados em cada uma das quatro variáveis citadas acima, como resultado da execução do algoritmo desenvolvido para simular a rede neural artificial. O gráfico superior de cada figura contém os resultados obtidos utilizando como entrada os valores individuais coletados no experimento de campo com os cultivares de cebola. O gráfico inferior da figura mostra o desempenho da rede utilizando-se as médias aritméticas e não os próprios valores coletados em campo. Decidiu-se usar este artifício para que a comparação destes resultados com os obtidos pela análise de regressão fosse mais correta e, dessa forma, permitir uma conclusão justa sobre a eficiência relativa entre as duas metodologias.

Os gráficos seguintes mostram os resultados obtidos após os processamentos dos algoritmos da rede contendo, no gráfico superior, o coeficiente de determinação obtido e o código de seis números gerado pela rede que obteve essa previsão de interpolação. Já no gráfico inferior, também tem o coeficiente de determinação da melhor rede obtida pelo algoritmo, o valor observado na pesquisa de campo e o valor previsto pela seqüência da rede que obteve a melhor previsão. Os círculos verdes em ambos os gráficos (superior e inferior) indicam a posição do valor predito pela rede neural em relação aos valores colhidos para a referida variável no experimento original, ou seja, é uma forma auxiliar de visualizar a eficiência da RNA.

O coeficiente de determinação mede a eficiência da rede que, de acordo com a estrutura do algoritmo desenvolvido, seleciona a melhor configuração dentre todas as combinações possíveis a ser testada pela rede neural. Os três primeiros números do código da rede indicam o número de neurônios das três camadas intermediárias (que podem variar de 1 a 3) usadas para encontrar tal resultado: primeira, segunda e terceira camada, respectivamente. Já os três números seguintes indicam o tipo de função de ativação utilizada pelas camadas intermediárias desta mesma rede neural selecionada pelo algoritmo como a mais eficiente. Os códigos destes tipos podem ser: 1, quando a função for *logsig*; 2, quando for *tansig* ou 3, quando a função de ativação escolhida for *purelin*.

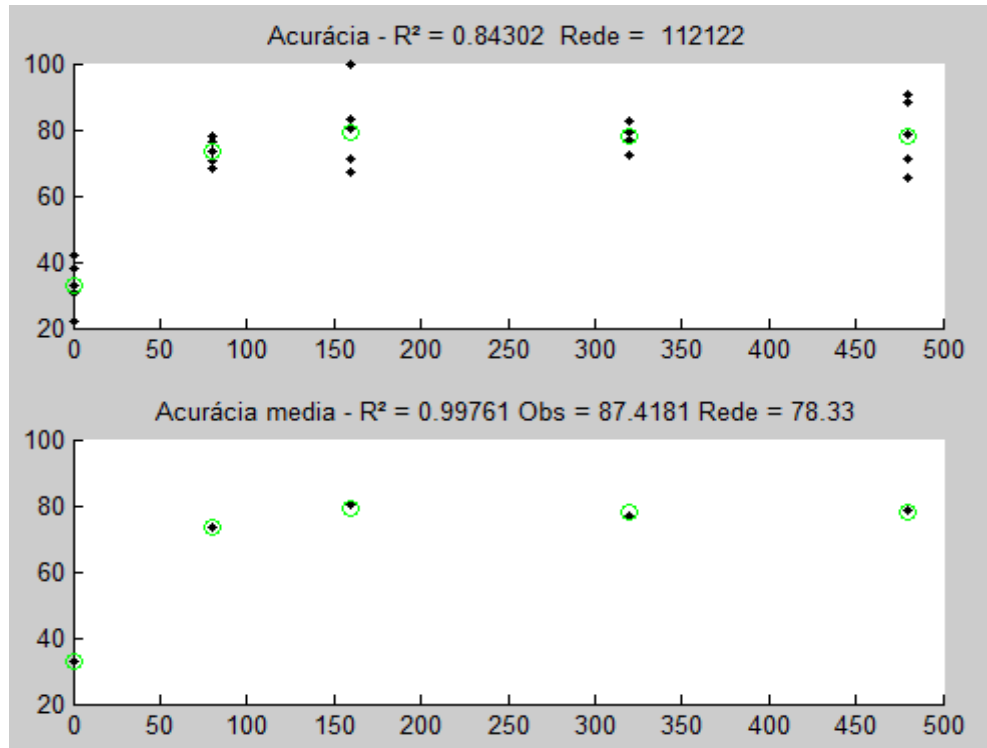


Figura 10: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 1 e da variável peso total dos bulbos (PTOT), com todos os valores obtidos (superior) e com a média (inferior).

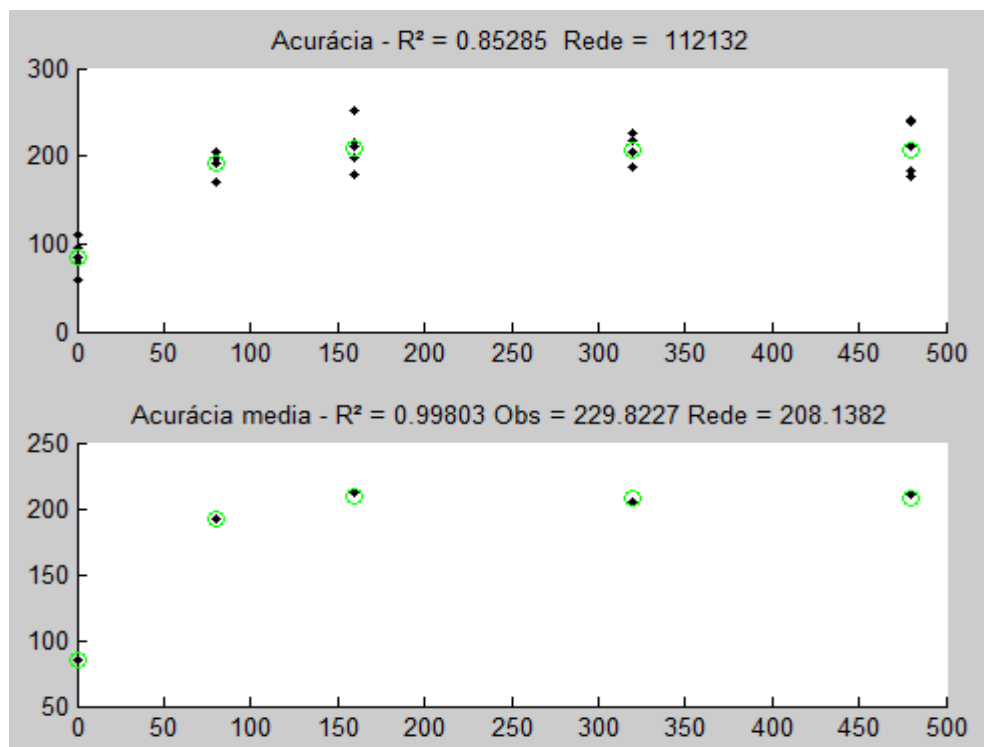


Figura 11: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 1 e da variável peso médio dos bulbos (PMTOT), com todos os valores obtidos (superior) e com a média (inferior).

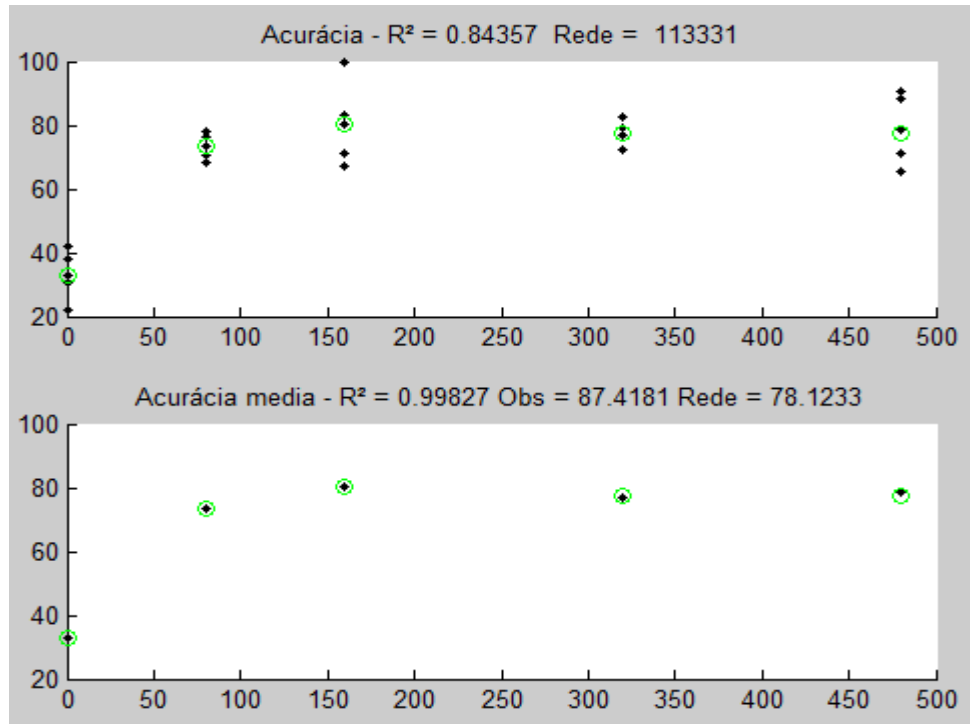


Figura 12: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 1 e da variável peso comerciável de bulbos (PCOM), com todos os valores obtidos (superior) e com a média (inferior).

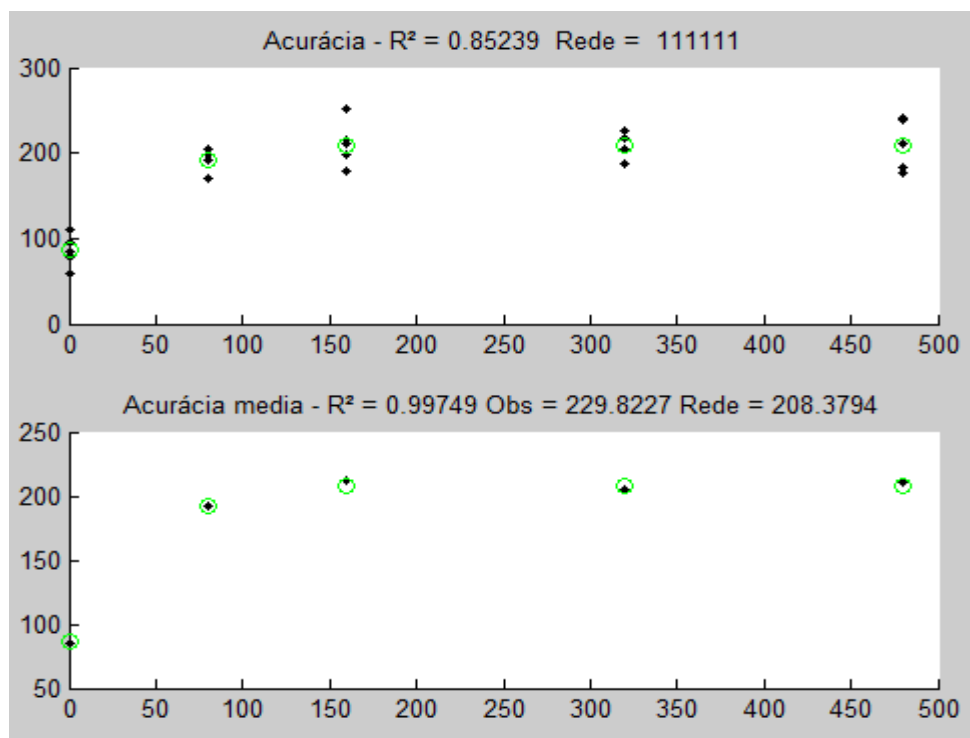


Figura 13: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 1 e da variável peso médio comerciável de bulbos (PMCOM), com todos os valores obtidos (superior) e com a média (inferior).

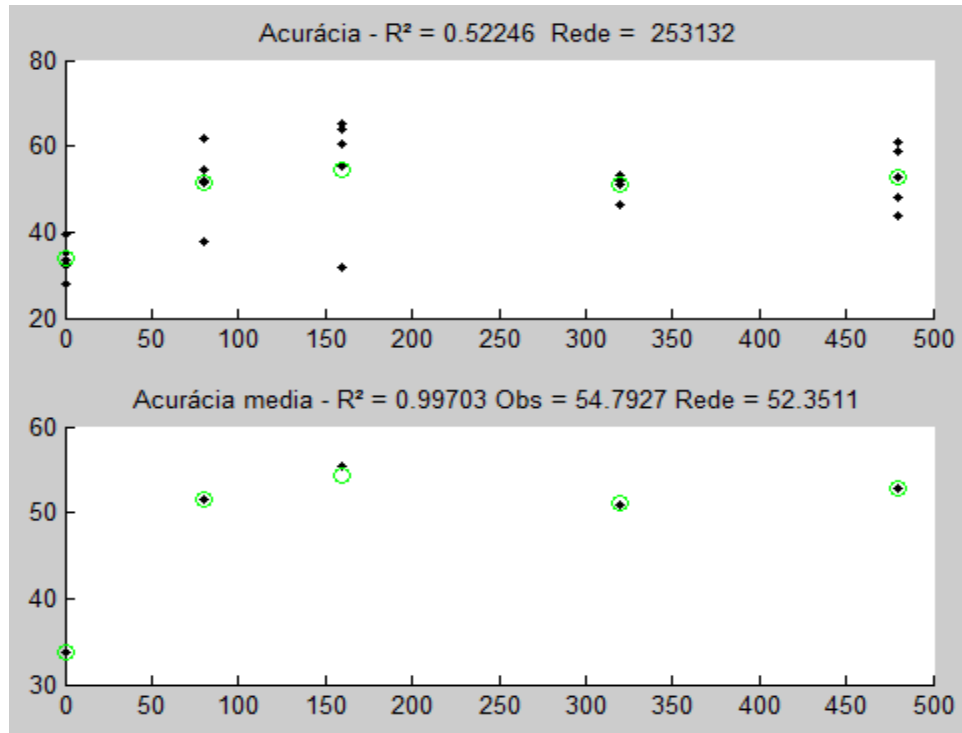


Figura 14: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 2 e da variável peso total dos bulbos (PTOT), com todos os valores obtidos (superior) e com a média (inferior).

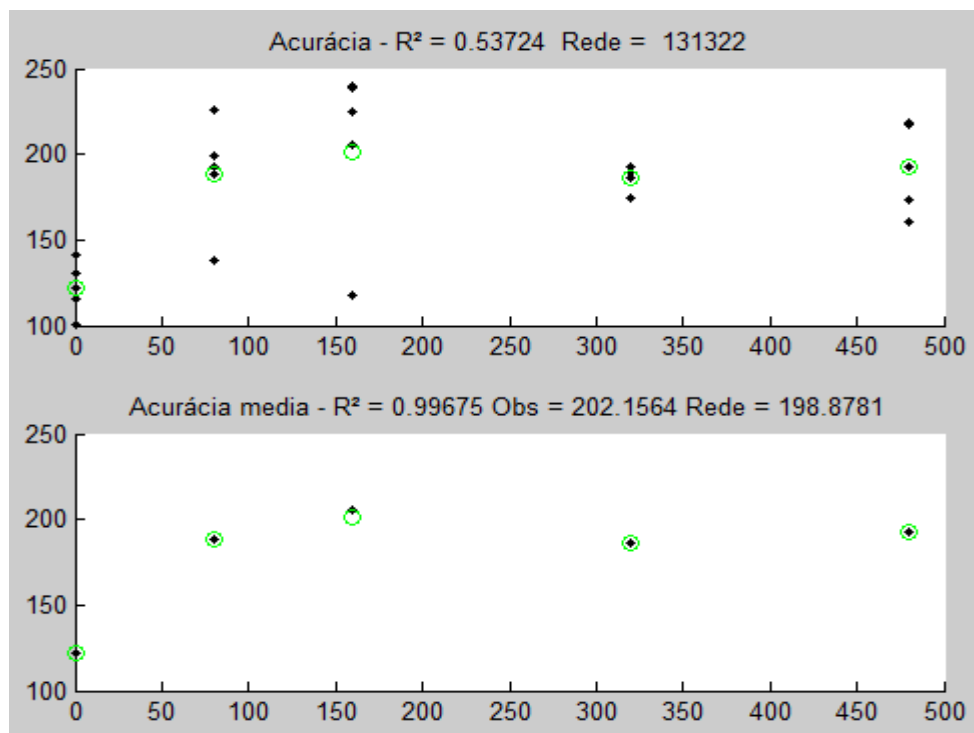


Figura 15: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 2 e da variável peso médio dos bulbos (PMTOT), com todos os valores obtidos (superior) e com a média (inferior).

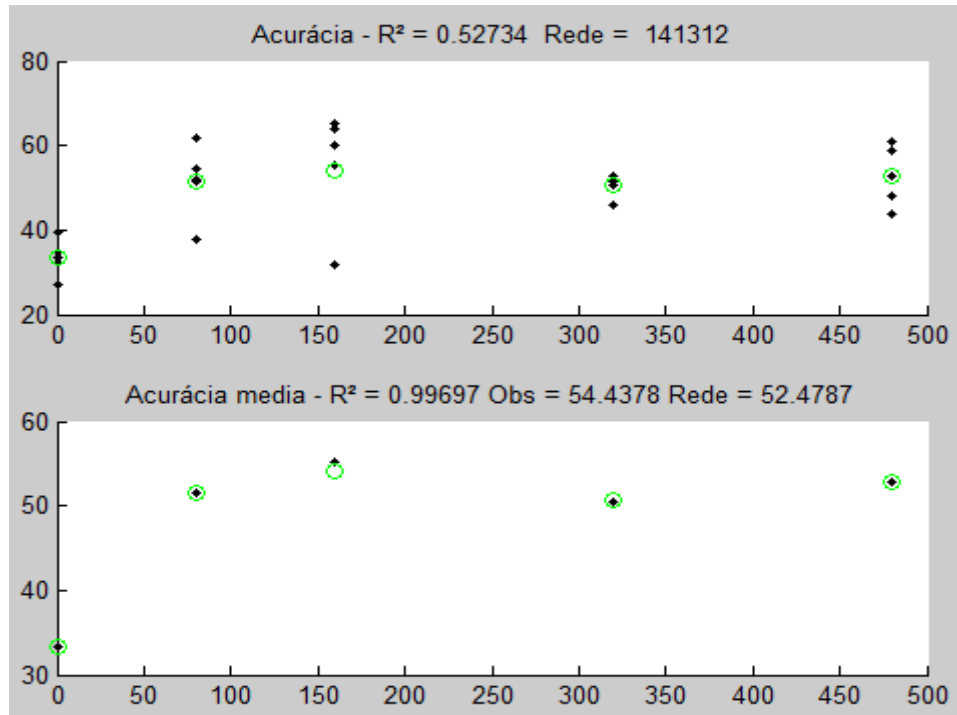


Figura 16: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 2 e da variável peso comerciável de bulbos (PCOM), com todos os valores obtidos (superior) e com a média (inferior).

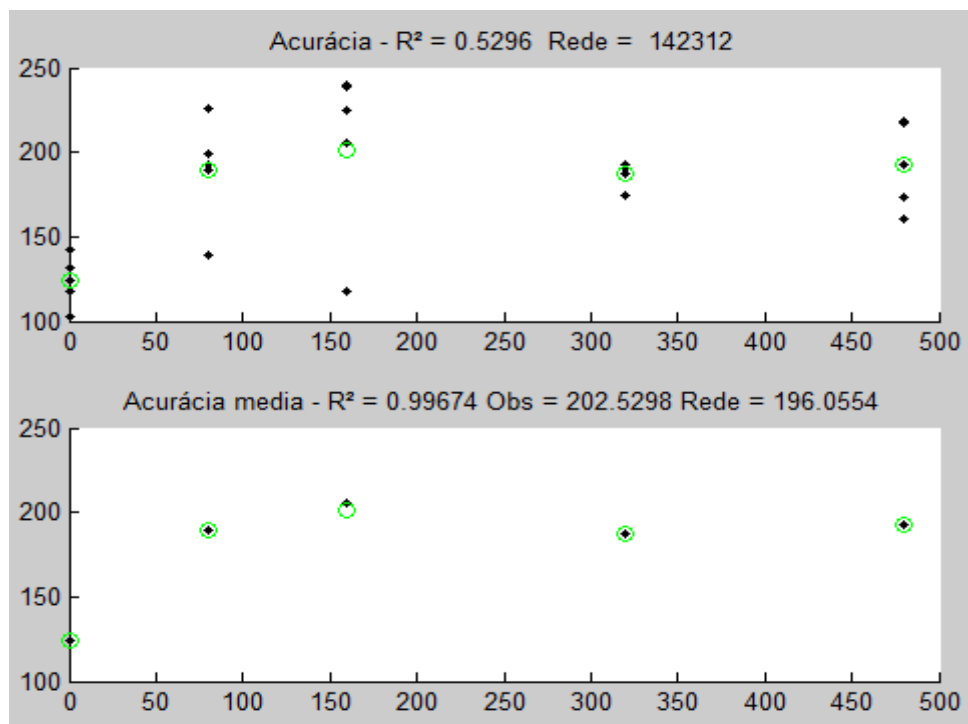


Figura 17: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 2 e da variável peso médio comerciável de bulbos (PMCOM), com todos os valores obtidos (superior) e com a média (inferior).

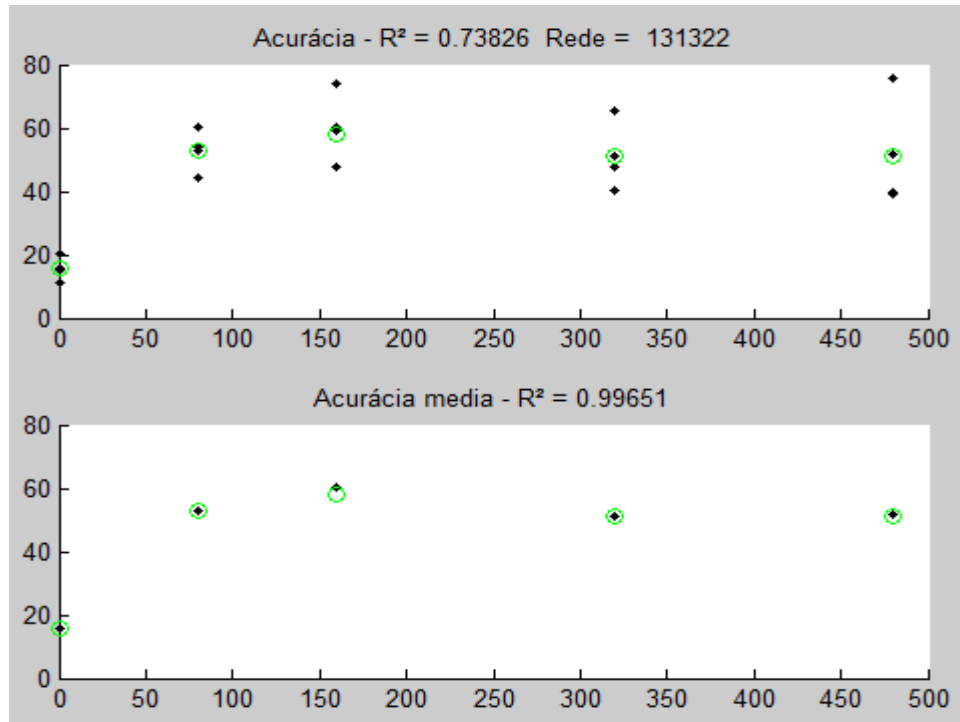


Figura 18: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 3 e da variável peso total dos bulbos (PTOT), com todos os valores obtidos (superior) e com a média (inferior).

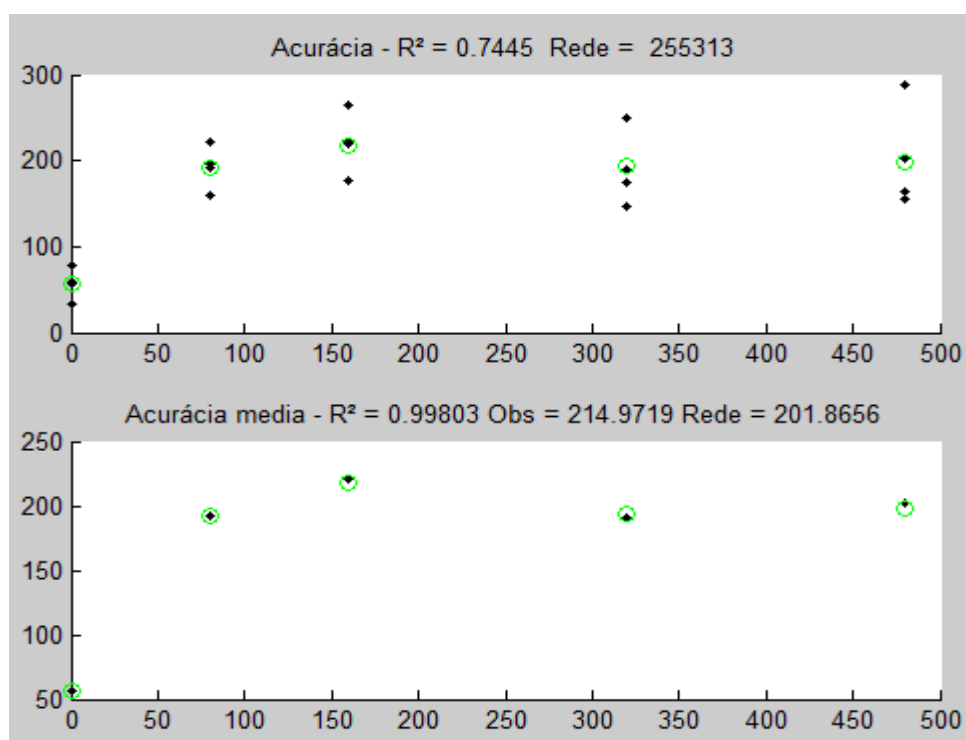


Figura 19: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 3 e da variável peso médio dos bulbos (PMTOT), com todos os valores obtidos (superior) e com a média (inferior).

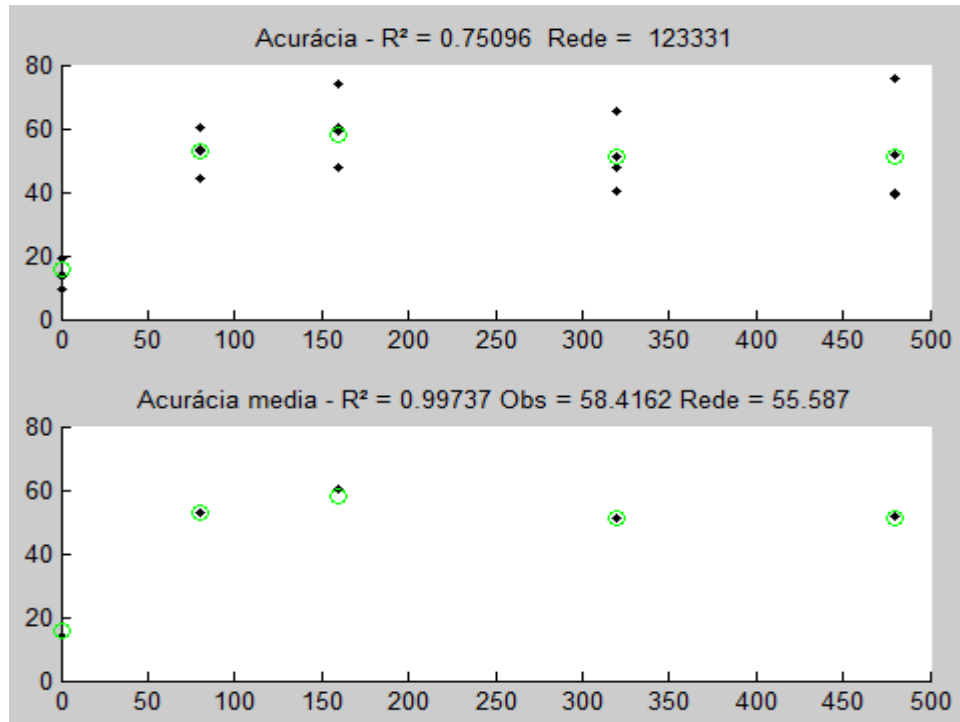


Figura 20: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 3 e da variável peso comerciável de bulbos (PCOM), com todos os valores obtidos (superior) e com a média (inferior).

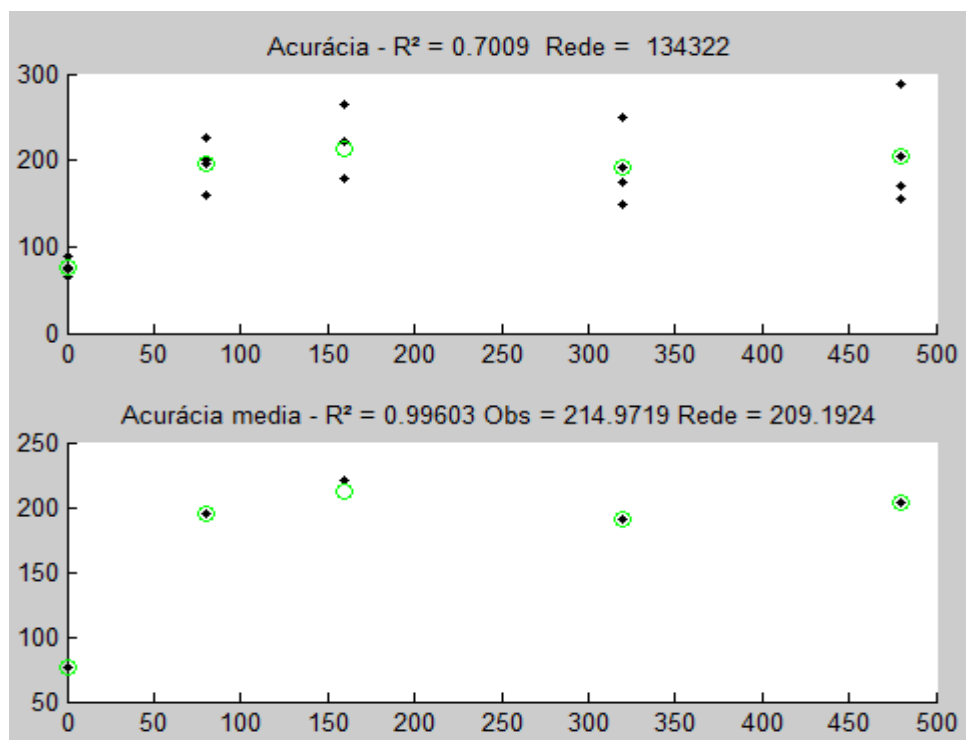


Figura 21: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 3 e da variável peso médio comerciável de bulbos (PMCOM), com todos os valores obtidos (superior) e com a média (inferior).

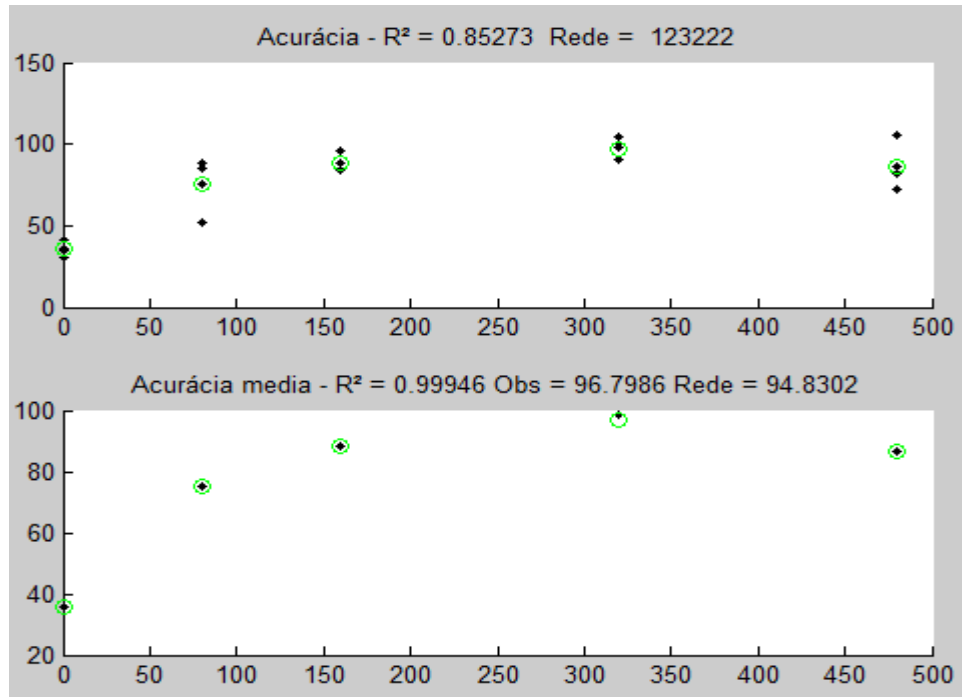


Figura 22: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 4 e da variável peso total dos bulbos (PTOT), com todos os valores obtidos (superior) e com a média (inferior).

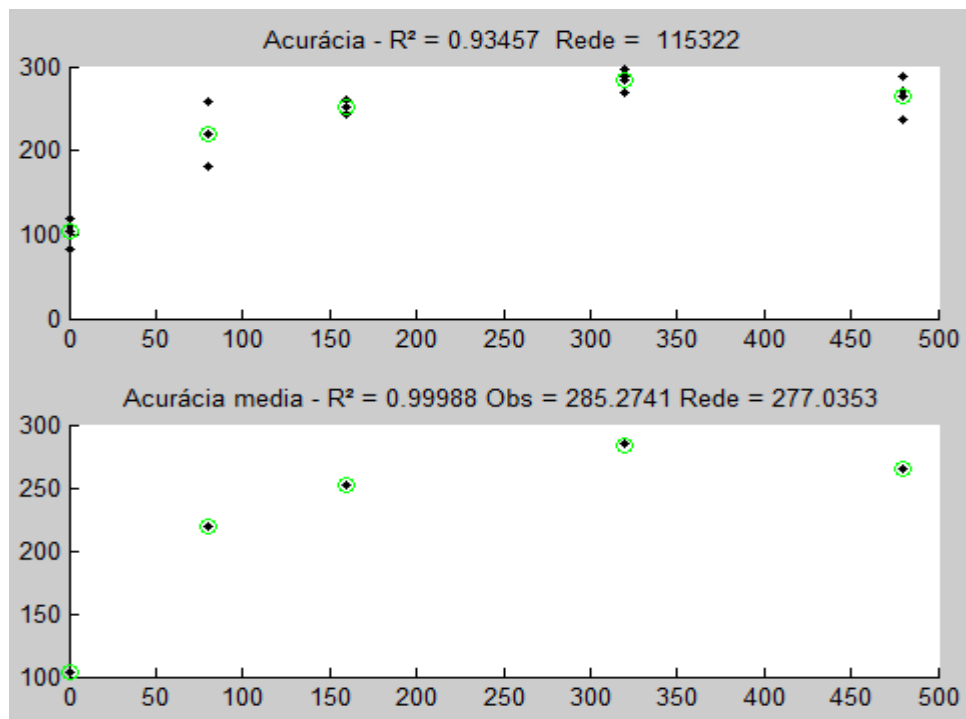


Figura 23: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 4 e da variável peso médio dos bulbos (PMTOT), com todos os valores obtidos (superior) e com a média (inferior).

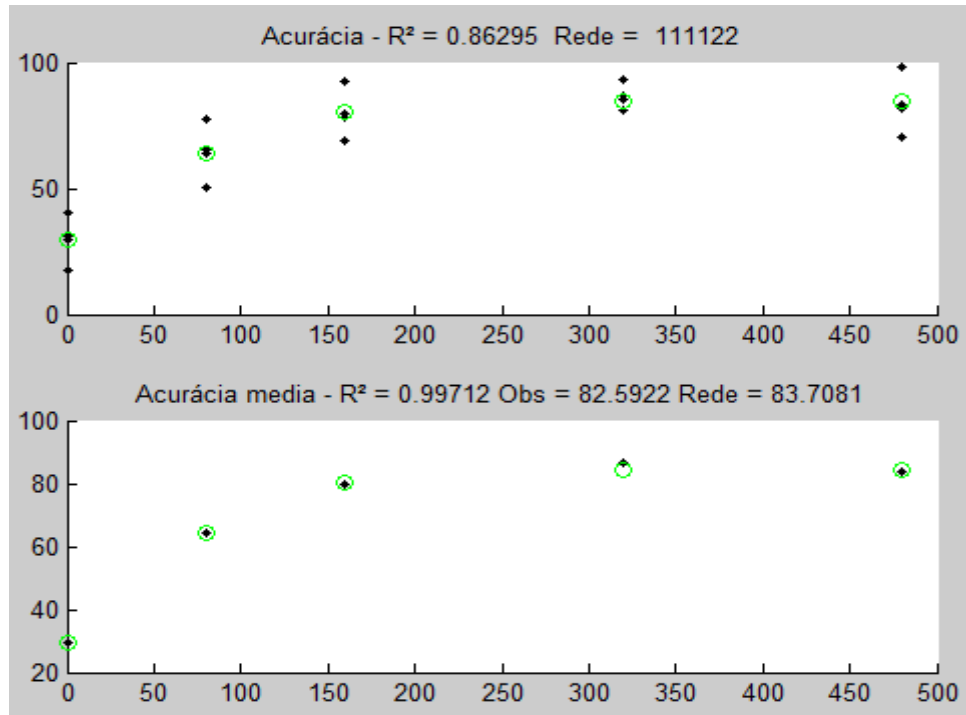


Figura 24: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 4 e da variável peso comerciável de bulbos (PCOM), com todos os valores obtidos (superior) e com a média (inferior).

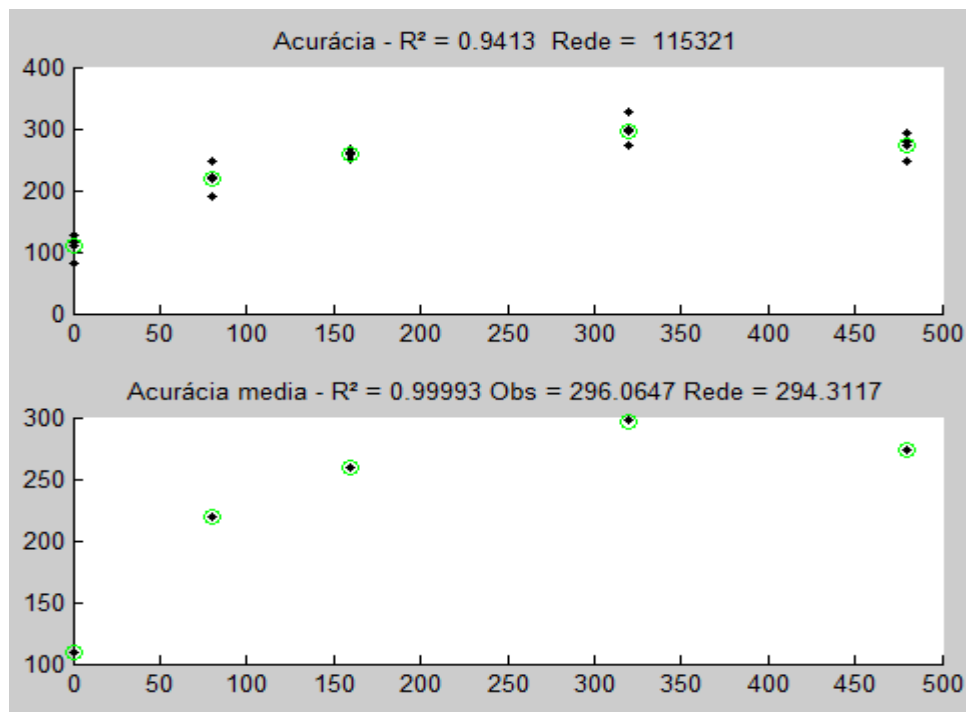


Figura 25: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 4 e da variável peso médio comerciável de bulbos (PMCOM), com todos os valores obtidos (superior) e com a média (inferior).

Como uma forma auxiliar de comparação entre os valores previstos, depois de realizadas as análises de regressão e as previsões provenientes do processamento do algoritmo da rede neural, os resultados obtidos em todos os processamentos realizados foram colocados em colunas no quadro comparativo a seguir. A Tabela 26 mostra a variável estudada, o local de cultivo da cebola e os valores preditos pela análise de regressão e pela rede neural, além dos valores observados.

Tabela 26: Resumo ilustrando os valores previstos pela análise de regressão, pela rede neural, além do valor observado em cada uma dos quatro cultivares de cebola para cada uma das seguintes variáveis: peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

Variável	Local	Regr.	Rede	Obs.
Peso total dos bulbos (PTOT)	1	82,25	78,33	87,42
Peso médio os bulbos (PMTOT)	1	218,41	208,14	229,82
Peso comerciável de bulbos (PCOM)	1	82,25	78,12	87,42
Peso médio comerciável de bulbos (PMCOM)	1	218,41	208,38	229,82
Peso total dos bulbos (PTOT)	2	54,70	52,35	54,79
Peso médio os bulbos (PMTOT)	2	201,55	198,38	202,16
Peso comerciável de bulbos (PCOM)	2	54,35	52,48	54,44
Peso médio comerciável de bulbos (PMCOM)	2	201,62	196,06	202,53
Peso total dos bulbos (PTOT)	3	59,14	57,27	60,21
Peso médio os bulbos (PMTOT)	3	217,06	201,87	214,97
Peso comerciável de bulbos (PCOM)	3	59,17	55,59	58,42
Peso médio comerciável de bulbos (PMCOM)	3	216,04	209,19	214,97
Peso total dos bulbos (PTOT)	4	99,61	94,83	96,80
Peso médio os bulbos (PMTOT)	4	282,56	277,04	285,27
Peso comerciável de bulbos (PCOM)	4	86,95	83,71	82,52
Peso médio comerciável de bulbos (PMCOM)	4	292,66	294,31	296,06

A comparação entre os resultados da análise de regressão e o das redes neurais pode melhor ser visualizada no quadro comparativo mostrado a seguir, que mostra a eficiência relativa entre as metodologias. A Tabela 27 mostra a variável estudada, o local de cultivo do híbrido, o coeficiente de determinação ($R^2_{\text{regressão}}$) da análise de regressão, o coeficiente de determinação medindo a eficiência da rede (R^2_{rede}) e uma última coluna exibindo a eficiência relativa que calcula a porcentagem entre o coeficiente de determinação da análise de regressão e o coeficiente de determinação da rede (E_{relativa}).

Tabela 27: Resumo ilustrando a eficiência relativa entre os coeficientes de determinação da análise de regressão e da rede neural em cada uma dos quatro cultivares de cebola para cada uma das seguintes variáveis: peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

Variável	Loca	$R^2_{\text{regressã}}$	R^2_{rede}	E_{relativa}
Peso total dos bulbos (PTOT)	1	98,68	99,76	101,09%
Peso médio os bulbos (PMTOT)	1	98,58	99,80	101,24%
Peso comerciável de bulbos (PCOM)	1	98,68	99,82	101,16%
Peso médio comerciável de bulbos (PCOM)	1	98,58	99,75	101,19%
Peso total dos bulbos (PTOT)	2	99,61	99,70	100,09%
Peso médio os bulbos (PMTOT)	2	99,77	99,67	99,90%
Peso comerciável de bulbos (PCOM)	2	99,53	99,69	100,16%
Peso médio comerciável de bulbos (PCOM)	2	99,76	99,67	99,91%
Peso total dos bulbos (PTOT)	3	99,31	99,65	100,34%
Peso médio os bulbos (PMTOT)	3	99,38	99,80	100,42%
Peso comerciável de bulbos (PCOM)	3	99,23	99,73	100,50%
Peso médio comerciável de bulbos (PCOM)	3	99,55	99,60	100,05%
Peso total dos bulbos (PTOT)	4	99,11	99,94	100,84%
Peso médio os bulbos (PMTOT)	4	98,71	99,98	101,29%
Peso comerciável de bulbos (PCOM)	4	99,92	99,71	99,79%
Peso médio comerciável de bulbos (PCOM)	4	99,31	99,99	100,68%

4.3.2 Analisando o caso da extrapolação de valores

Apesar de a extrapolação de valores nem sempre ser recomendada nas análises estatísticas, para que fosse possível comparar a eficiência das redes neurais até mesmo a valores fora do intervalo da amplitude dos dados amostrados, decidiu-se retirar os dados correspondentes ao valor de um dos extremos ($480 \text{ kg}\cdot\text{ha}^{-1}$) dentre as doses de nitrogênio do experimento. Este procedimento foi feito para todas as seis características selecionadas em cada um dos quatro cultivares para a pesquisa: número total de bulbos (NTOT); peso total dos bulbos (PTOT); peso médio dos bulbos (PMTOT); número comerciável de bulbos (NCOM); peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM). Após isso, realizou-se novamente a análise de variância. Os cálculos constataram que as variáveis (NTOT) e (NCOM) não apresentaram diferenças significativas entre as médias de tratamentos para os cultivares 1, 2 e 4 como pode ser visto nas Tabelas 28, 29 e 31. Para o cultivar 3, somente a variável número total de bulbos (NTOT) não apresentou diferença significativa (Tabela 30).

No caso em que a diferença foi significativa, tanto a 1%, sinalizado com dois asteriscos (**); quanto a 5%, sinalizado com apenas um asterisco (*), concluiu-se que as diferentes doses de nitrogênio tiveram efeitos substanciais em relação à combinação de cada variável em cada local, tomados separadamente. Caso contrário, quando a diferença entre as médias das doses de nitrogênio foi não significativa (*ns*), é entendido que a alteração na dosagem do nutriente estudado não surtiu efeito nas variáveis respostas avaliadas. Por esse motivo e, devido ao fato de o objetivo final deste trabalho ser o de fazer uma avaliação de todos os cultivares, os estudos prosseguiram considerando-se apenas as quatro características que tiveram resultados significativos em todos os cultivares do experimento. São elas: peso total dos bulbos (PTOT), peso médio dos bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

As Tabelas 28, 29, 30 e 31 mostram, para cada híbrido individualmente, os resultados da análise de variância para as seis variáveis estudadas em cada um deles, contendo o valor do quadrado médio de cada variável para: blocos,

tratamento e resíduo, e as médias e coeficientes de variação de cada característica estudada para os cultivares 1, 2, 3 e 4, respectivamente.

Tabela 28: Resultado da análise de variância do cultivar de cebola 1 contendo o quadrado médio e os respectivos coeficientes de variação das variáveis: número total de bulbos (NTOT), peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), número comerciável de bulbos (NCOM), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

ANOVA do cultivar 1		QUADRADO MÉDIO DA VARIÁVEL					
FV	GL	NTOT	PTOT	PMTOT	NCOM	PCOM	PMCOM
BLOCOS	3	142.29	127.18	497.09	142.29	127.18	497.09
TRATAMENTO	4	61.60ns	1828.26**	13041.74**	61.60ns	1828.26**	13041.74**
RESÍDUO	12	292.82	55.40	412.16	292.82	55.40	412.16
MÉDIA		380.26	70.10	184.80	380.26	70.10	184.80
CV		4.50	10.61	10.98	4.50	10.61	10.98

*, significativo a 5%; **, significativo a 1% e ns, não significativo

As Tabelas 28, 29 e 31 mostram que as variáveis número total de bulbos e número comerciável de bulbos tiveram resultados não significativos (*ns*) enquanto que as variáveis peso total dos bulbos, peso médio dos bulbos, peso comerciável de bulbos e peso médio comerciável de bulbos foram significativas (*) ou (**). Já na Tabela 30 verifica-se que apenas a variável número total de bulbos teve resultado não significativo (*ns*).

Tabela 29: Resultado da análise de variância do cultivar de cebola 2 contendo o quadrado médio e os respectivos coeficientes de variação das variáveis: número total de bulbos (NTOT), peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), número comerciável de bulbos (NCOM), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

ANOVA do cultivar 2		QUADRADO MÉDIO DA VARIÁVEL					
FV	GL	NTOT	PTOT	PMTOT	NCOM	PCOM	PMCOM
BLOCOS	3	14.81	376.21	5014.67	22.41	366.73	5006.62
TRATAMENTO	4	21.05ns	318.95*	4608.93*	4.97ns	322.67*	4448.52*
RESÍDUO	12	14.81	67.72	912.62	20.95	66.34	917.06
MÉDIA		272.57	49.27	180.96	270.00	49.01	181.53
CV		1.41	16.70	16.69	1.69	16.61	16.68

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Tabela 30: Resultado da análise de variância do cultivar de cebola 3 contendo o quadrado médio e os respectivos coeficientes de variação das variáveis: número total de bulbos (NTOT), peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), número comerciável de bulbos (NCOM), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

ANOVA do cultivar 3		QUADRADO MÉDIO DA VARIÁVEL					
FV	GL	NTOT	PTOT	PMTOT	NCOM	PCOM	PMCOM
BLOCOS	2	457.11	596.06	8513.03	271.15	604.49	7661.72
TRATAMENTO	4	64.37ns	1007.49**	13599.70**	5016.59**	1081.07**	10443.60**
RESÍDUO	8	241.05	21.99	257.89	380.98	21.03	368.07
MÉDIA		274.42	47.65	174.92	252.38	47.30	179.82
CV		5.65	9.84	9.18	7.73	9.69	10.66

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Tabela 31: Resultado da análise de variância do cultivar de cebola 4 contendo o quadrado médio e os respectivos coeficientes de variação das variáveis: número total de bulbos (NTOT), peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), número comerciável de bulbos (NCOM), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

ANOVA do cultivar 4		QUADRADO MÉDIO DA VARIÁVEL					
FV	GL	NTOT	PTOT	PMTOT	NCOM	PCOM	PMCOM
BLOCOS	2	204.86	236.19	1253.66	2259.58	265.19	1309.98
TRATAMENTO	4	60.13ns	2006.88**	16902.60**	729.75ns	1642.63**	18397.79**
RESÍDUO	8	855.87	93.99	283.33	999.58	80.83	285.70
MÉDIA		343.29	78.82	229.19	286.99	68.59	236.35
CV		8.52	12.29	7.34	11.01	13.10	7.15

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Finalizada a análise de variância dos dados considerando como fatores apenas as cinco doses de nitrogênio do estudo, ou seja, após a retirada da dose de 480 kg.ha^{-1} , iniciou-se a análise de regressão para os quatro cultivares de cebola, avaliando apenas as características em que os resultados foram significativos: (PTOT), (PMTOT), (PCOM) e (PMCOM). Isso foi feito para que fosse possível selecionar dentre as equações significativas obtidas na análise de regressão polinomial, a que tivesse evidenciado o melhor ajuste aos dados.

Na Tabela 32, por exemplo, a variável peso total dos bulbos (PTOT) do cultivar 1, teve as duas primeiras equações polinomiais estudadas: linear e

quadrática com o resultado significativo a 5% (*). Então, dentre as significativas, escolheu-se a que continha o maior grau polinomial, ou seja, para este exemplo a selecionada para representar o modelo em estudo foi a equação quadrática. Ainda para o peso total dos bulbos (PTOT), as respostas dos cultivares 2 e 4 também foram mais bem representadas por uma equação quadrática (Tabelas 33 e 35), porém a resposta do cultivar 3 foi mais bem descrita por uma equação polinomial de grau 3, ou seja, cúbica (Tabela 34).

Prosseguindo com a interpretação dos resultados, é possível verificar que os cultivares 1, 2 e 4 de cebola foram mais bem representados por equações quadráticas, qualquer que fosse as variáveis analisadas. Porém o cultivar 3 variou de modelo dependendo da variável analisada. Quando se tratava do peso total dos bulbos (PTOT), do peso médio dos bulbos (PMTOT) e do peso comerciável dos bulbos (PCOM) uma equação de grau polinomial 3 foi o modelo matemático que melhor explicou ou representou o fenômeno estudado. Já para a variável peso médio comerciável dos bulbos (PMCOM), o modelo selecionado foi o quadrático.

Nas Tabelas 32, 33, 34 e 35 são apresentados os resultados da análise de regressão para as quatro variáveis estudadas contendo o valor do quadrado médio de cada variável para os modelos linear, quadrático e cúbico; e o respectivo coeficiente de determinação de cada característica estudada para os cultivares 1, 2, 3 e 4, respectivamente.

Tabela 32: Resultado da análise de regressão do cultivar de cebola 1 contendo o quadrado médio e os respectivos coeficientes de determinação das variáveis: peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM) para os modelos linear, quadrático e cúbico.

Cultivar 1		QUADRADO MÉDIO E R ²							
FV	GL	PTOT		PMTOT		PCOM		PMCOM	
		QM	R ²	QM	R ²	QM	R ²	QM	R ²
Regressão	3	2388.08*		17060.11*		2388.08*		17060.11*	
Linear	1	4109.78*	56.	30255.75*	57.9	4109.78*	56.1	30255.75*	57.9
Quadrática	1	2963.54*	96.	20123.20*	96.5	2963.54*	96.7	20123.20*	96.5
Cúbica	1	90.93ns	97.	801.38ns	98.1	90.93ns	97.9	801.38ns	98.1
Desvio	2	148.79		986.62		148.79		986.62	
Resíduo	12	55.40		412.16		55.40		412.16	

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Tabela 33: Resultado da análise de regressão do cultivar de cebola 2 contendo o quadrado médio e os respectivos coeficientes de determinação das variáveis: peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM) para os modelos linear, quadrático e cúbico.

Cultivar 2		QUADRADO MÉDIO E R ²							
FV	GL	PTOT		PMTOT		PCOM		PMCOM	
		QM	R ²	QM	R ²	QM	R ²	QM	R ²
Regressão	3	424.13**		6134.71**		428.98**		5920.07**	
Linear	1	568.29*	44.54	8108.66*	43.98	557.77*	43.21	7837.79*	44.04
Quadrática	1	659.09**	96.20	9722.51**	96.72	678.24**	95.76	9384.02**	96.78
Cúbica	1	44.99ns	99.73	572.95ns	99.82	50.94ns	99.70	538.39ns	99.80
Desvio	2	3.43		31.59		3.74		33.88	
Resíduo	12	67.72		912.62		66.34		917.06	

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Pode-se verificar, observando as Tabelas 32, 33, 34 e 35 que, no caso da extrapolação de valores, a maioria das variáveis teve a equação polinomial de grau 2 (quadrática) como a que melhor representou o modelo. A exceção foi para o cultivar 4 que teve a equação cúbica representando melhor três de suas variáveis.

Tabela 34: Resultado da análise de regressão do cultivar de cebola 3 contendo o quadrado médio e os respectivos coeficientes de determinação das variáveis: peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM) para os modelos linear, quadrático e cúbico.

Cultivar 3		QUADRADO MÉDIO E R ²							
FV	GL	PTOT		PMTOT		PCOM		PMCOM	
		QM	R ²	QM	R ²	QM	R ²	QM	R ²
Regressã	3	1339.34		18080.47**		1436.79*		13892.84*	
Linear	1	1751.52	43.4	25104.39**	46.	1897.55*	43.8	18465.25*	44.2
Quadrátic	1	2087.59	95.2	26810.82**	95.	2214.86*	95.1	21525.64*	95.7
Cúbica	1	178.91*	99.7	2326.21*	99.	197.96*	99.6	1687.63ns	99.7
Desvio	2	11.96		157.38		13.91		95.89	
Resíduo	8	21.99		257.89		21.03		368.07	

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Tabela 35: Resultado da análise de regressão do cultivar de cebola 4 contendo o quadrado médio e os respectivos coeficientes de determinação das variáveis: peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM) para os modelos linear, quadrático e cúbico.

Cultivar 4		QUADRADO MÉDIO E R ²							
FV	GL	PTOT		PMTOT		PCOM		PMCOM	
		QM	R ²	QM	R ²	QM	R ²	QM	R ²
Regressã	3	2666.74*		22333.43*		2189.06*		24398.55*	
Linear	1	6505.97*	81.	54705.01*	80.9	5266.62*	80.1	62275.96*	84.6
Quadrátic	1	1378.53*	98.	11542.11*	97.9	1177.66*	98.0	10533.32*	98.9
Cúbica	1	115.73ns	99.	753.18ns	99.0	122.90ns	99.9	386.38ns	99.4
Desvio	2	27.28		610.10		3.33		395.50	
Resíduo	8	93.99		283.33		80.83		285.70	

*, significativo a 5%; **, significativo a 1% e ns, não significativo

Os resultados obtidos e descritos nas tabelas anteriores permitiram determinar qual é o grau da equação polinomial que melhor se ajustou aos dados de cada variável analisada em separado. O passo seguinte a essas análises foi obter as estimativas para os parâmetros da equação, a equação e o respectivo coeficiente de determinação. Nas Tabelas 36, 37, 38 e 39 são apresentadas a equação, a estimativa para a dosagem 480 kg.ha⁻¹ de nitrogênio (Est.₄₈₀) que foi retirada antes da realização dos cálculos e o coeficiente de determinação (R²) do modelo ajustado de cada característica estudada, individualmente, para os cultivares 1, 2, 3 e 4. Além disso, é possível observar o coeficiente de determinação (R²) da equação em cada situação e, como todos eles estão próximos a 100%, conclui-se que a precisão do experimento foi satisfatória.

Tabela 36: Resultado da análise do peso total dos bulbos (PTOT) contendo a melhor equação polinomial obtida e o coeficiente de determinação para cada um dos quatro cultivares.

Cultivar	Modelo	Est. ₄₈₀	R ²
1	$35,28 + 0,49X^1 - 11,36 \times 10^{-4}X^2$	8,81	96,72
2	$34,87 + 0,21X^1 - 5,36 \times 10^{-4}X^2$	207,64	96,20
3	$15,83 + 0,66X^1 - 30,09 \times 10^{-4}X^2 + 3,97 \times 10^{-6}X^3$	81,55	99,70
4	$37,91 + 0,47X^1 - 8,95 \times 10^{-4}X^2$	57,53	98,21

Tabela 37: Resultado da análise do peso médio os bulbos (PMTOT) contendo a melhor equação polinomial obtida e o coeficiente de determinação para cada um dos quatro cultivares.

Cultivar	Modelo	Est.₄₈₀	R²
1	$91,88 + 1,29X^1 - 29,61 \times 10^{-4}X^2$	29,43	96,57
2	$126,13 + 0,83X^1 - 20,58 \times 10^{-4}X^2$	734,70	96,72
3	$57,73 + 2,41X^1 - 108,27 \times 10^{-4}X^2 + 14,33 \times 10^{-6}X^3$	306,51	99,71
4	$110,63 + 1,36X^1 - 25,90 \times 10^{-4}X^2$	167,92	97,98

Tabela 38: Resultado da análise peso comerciável de bulbos (PCOM) contendo a melhor equação polinomial obtida e o coeficiente de determinação para cada um dos quatro cultivares.

Cultivar	Modelo	Est.₄₈₀	R²
1	$35,28 + 0,49X^1 - 11,36 \times 10^{-4}X^2$	8,81	96,72
2	$34,58 + 0,22X^1 - 5,43 \times 10^{-4}X^2$	218,72	95,76
3	$14,31 + 0,69X^1 - 31,41 \times 10^{-4}X^2 + 4,18 \times 10^{-6}X^3$	85,23	99,67
4	$31,50 + 0,43X^1 - 8,27 \times 10^{-4}X^2$	47,47	98,07

Tabela 39: Resultado da análise do peso médio comerciável de bulbos (PMCOM) contendo a melhor equação polinomial obtida e o coeficiente de determinação para cada um dos quatro cultivares.

Cultivar	Modelo	Est.₄₈₀	R²
1	$91,88 + 1,29X^1 - 29,61 \times 10^{-4}X^2$	29,43	96,57
2	$127,64 + 0,82X^1 - 20,22 \times 10^{-4}X^2$	716,53	96,78
3	$84,92 + 1,44X^1 - 35,37 \times 10^{-4}X^2$	-37,88	95,73
4	$113,55 + 1,36X^1 - 24,74 \times 10^{-4}X^2$	196,89	98,93

Com o encerramento dos cálculos através da análise de regressão, iniciaram-se os mesmos cálculos, mas agora utilizando as redes neurais para que fosse possível a comparação entre as metodologias. Para isso, em cada um dos quatro cultivares foi utilizado o algoritmo da rede para prever valores correspondentes ao valor 480 kg.ha⁻¹, para cada uma das seguintes características: peso total dos bulbos (PTOT), peso médio dos bulbos (PMTOT),

peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

Foram plotados dois gráficos, para cada um dos quatro cultivares estudados em cada uma das quatro variáveis citadas acima, como resultado da execução do algoritmo desenvolvido para simular a rede neural artificial. O gráfico superior de cada figura contém os resultados obtidos utilizando como entrada os valores individuais coletados no experimento de campo com os cultivares de cebola. O gráfico inferior da figura mostra o desempenho da rede utilizando-se as médias aritméticas e não os próprios valores coletados em campo. Decidiu-se usar este artifício para que a comparação dos resultados obtidos pela análise de regressão com os da rede neural fosse mais correta e, dessa forma, permitir uma conclusão justa sobre a eficiência relativa entre as duas metodologias.

Os gráficos seguintes mostram os resultados obtidos após os processamentos dos algoritmos da rede contendo, no gráfico superior, o coeficiente de determinação obtido e o código de seis números gerado pela rede que obteve essa previsão de interpolação. Já no gráfico inferior, também tem o melhor coeficiente de determinação gerado pela rede, o valor observado na pesquisa de campo e o valor previsto pela seqüência da rede que obteve a melhor previsão. Os círculos verdes em ambos os gráficos (superior e inferior) indicam a posição do valor predito pela rede neural em relação aos valores colhidos para a referida variável no experimento original, ou seja, é uma forma auxiliar de visualizar a eficiência da RNA.

De acordo com a estrutura do algoritmo desenvolvido, é selecionada a melhor configuração dentre todas as combinações possíveis a ser testada pela rede neural. Os três primeiros números do código da rede indicam o número de neurônios das três camadas intermediárias usadas para encontrar tal resultado: primeira, segunda e terceira camada, respectivamente. Já os três números seguintes indicam o tipo de função de ativação utilizada por estas mesmas camadas intermediárias da rede neural selecionada pelo algoritmo como a mais eficiente. O valor da rede igual a 113331 (Figura 26) indica que a rede ideal obtida pelo algoritmo teve a primeira e a segunda camada com 1 neurônio cada uma e a terceira com três neurônios. Já o quarto número sendo 1 indica que a primeira camada intermediária teve a função *purelin* como função de ativação; o quinto

número sendo 3, indica que novamente a função *purelin* foi a função obtida; e, o sexto número sendo 1, mostrou *logsig* como a função de ativação.

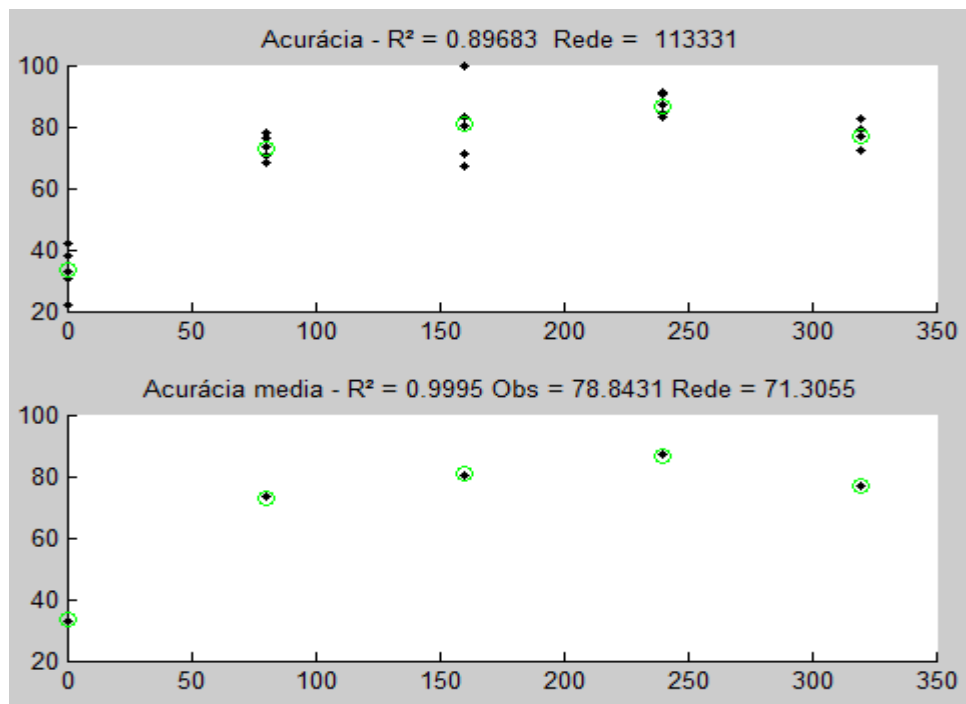


Figura 26: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 1 e da variável peso total dos bulbos (PTOT), com todos os valores obtidos (superior) e com a média (inferior).

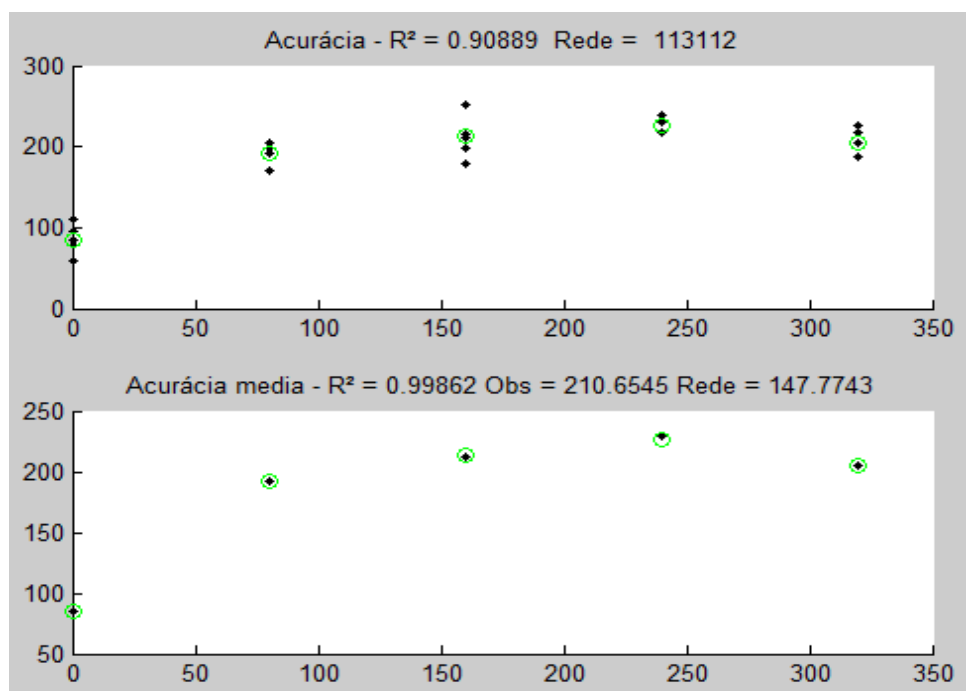


Figura 27: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 1 e da variável peso médio dos bulbos (PMTOT), com todos os valores obtidos (superior) e com a média (inferior).

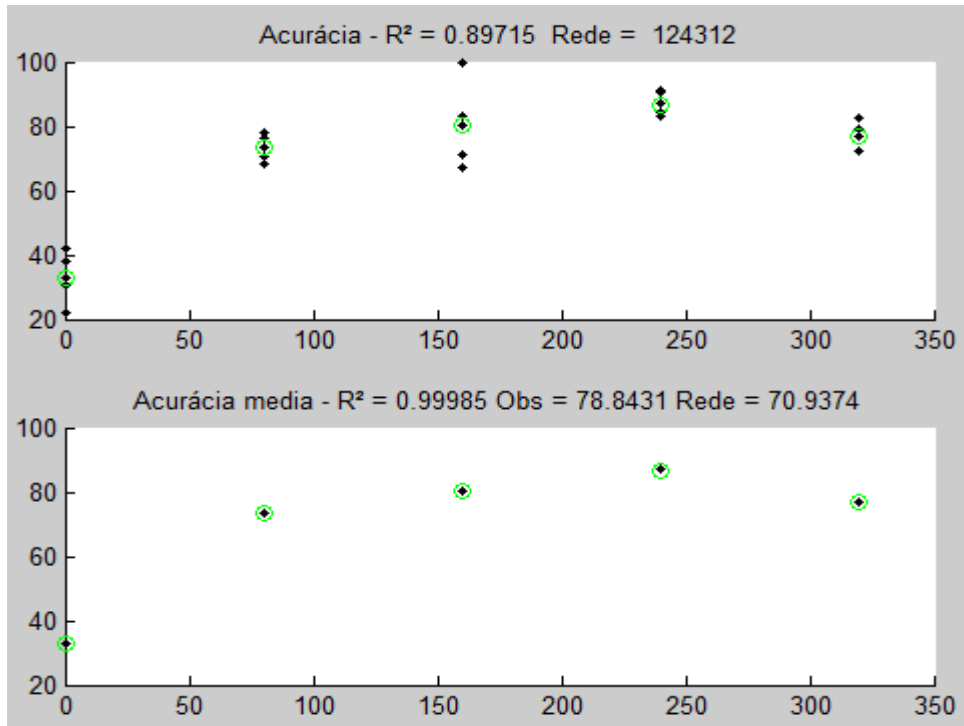


Figura 28: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 1 e da variável peso comerciável de bulbos (PCOM), com todos os valores obtidos (superior) e com a média (inferior).

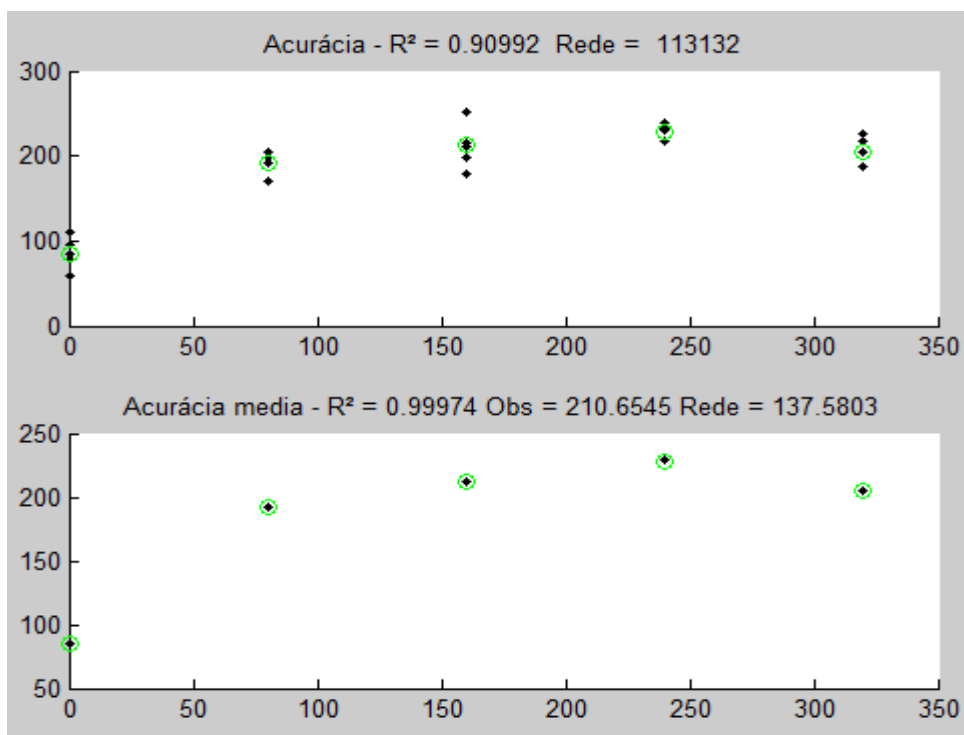


Figura 29: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 1 e da variável peso médio comerciável de bulbos (PMCOM), com todos os valores obtidos (superior) e com a média (inferior).

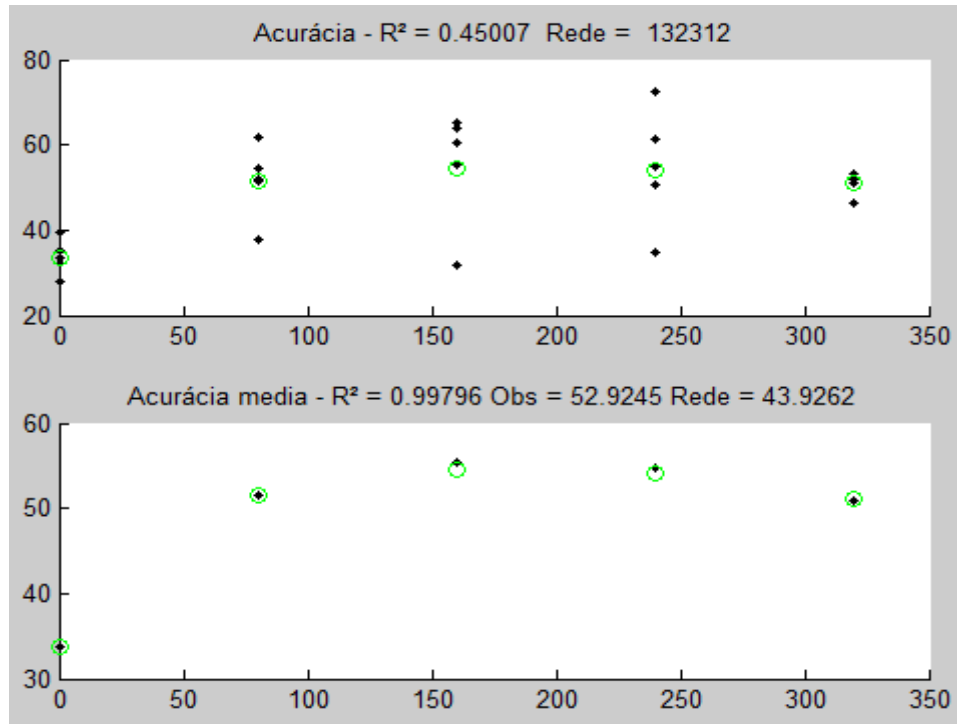


Figura 30: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 2 e da variável peso total dos bulbos (PTOT), com todos os valores obtidos (superior) e com a média (inferior).

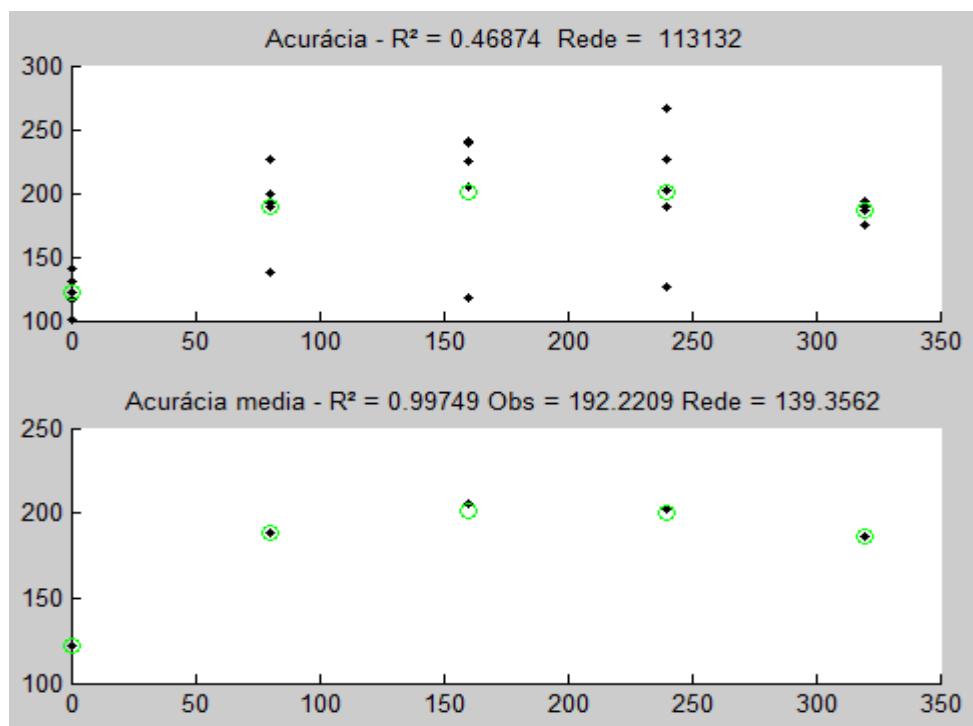


Figura 31: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 2 e da variável peso médio dos bulbos (PMTOT), com todos os valores obtidos (superior) e com a média (inferior).

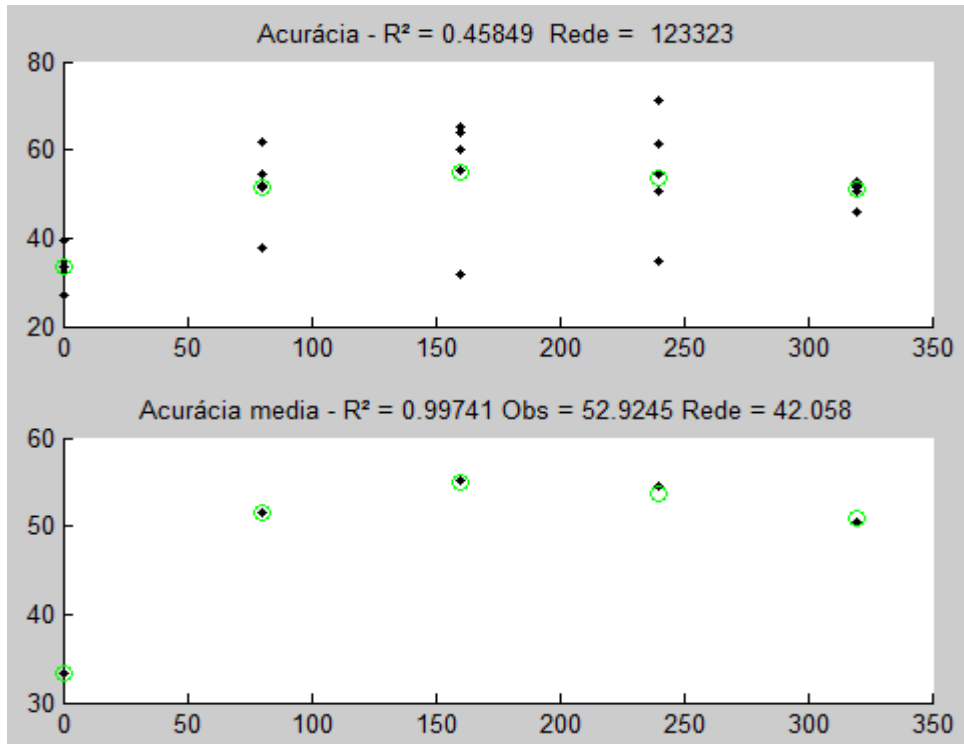


Figura 32: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 2 e da variável peso comerciável de bulbos (PCOM), com todos os valores obtidos (superior) e com a média (inferior).

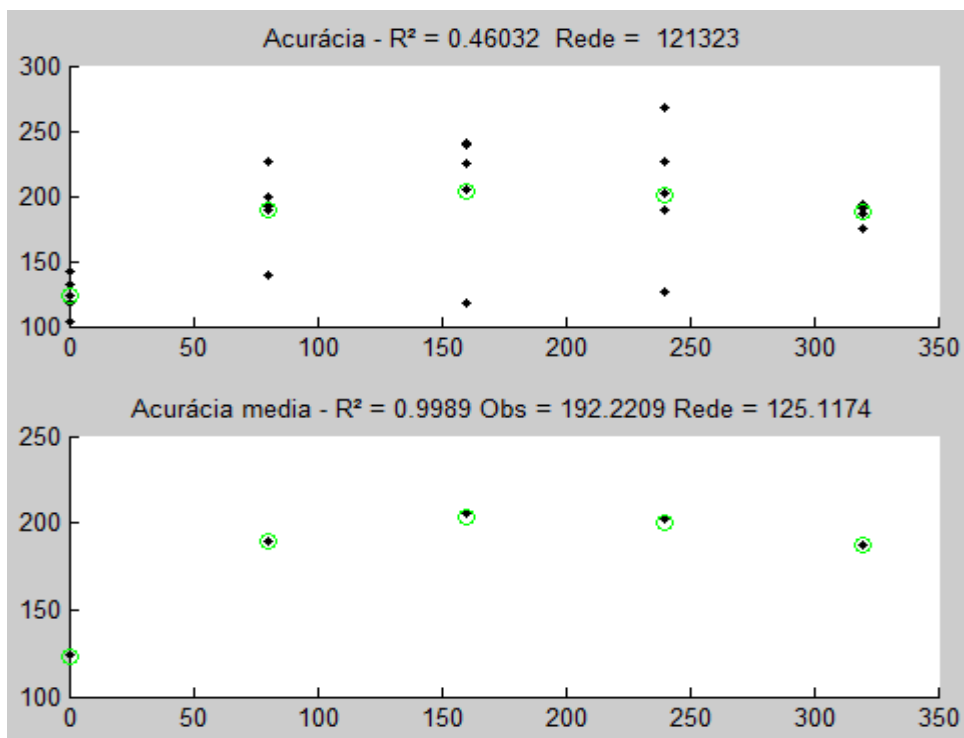


Figura 33: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 2 e da variável peso médio comerciável de bulbos (PMCOM), com todos os valores obtidos (superior) e com a média (inferior).

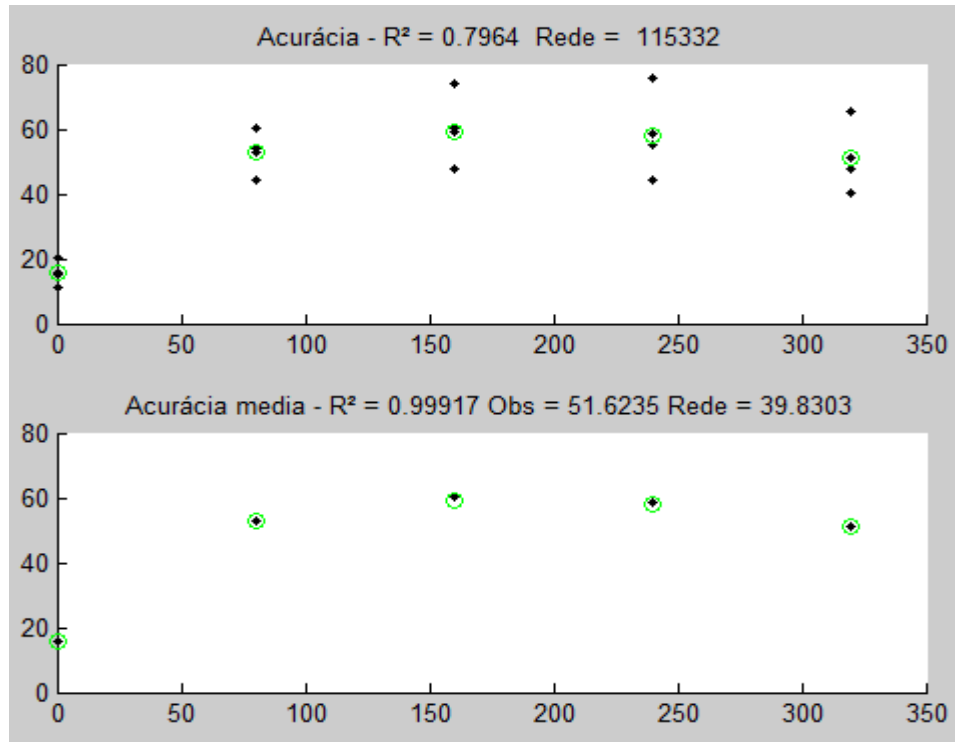


Figura 34: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 3 e da variável peso total dos bulbos (PTOT), com todos os valores obtidos (superior) e com a média (inferior).

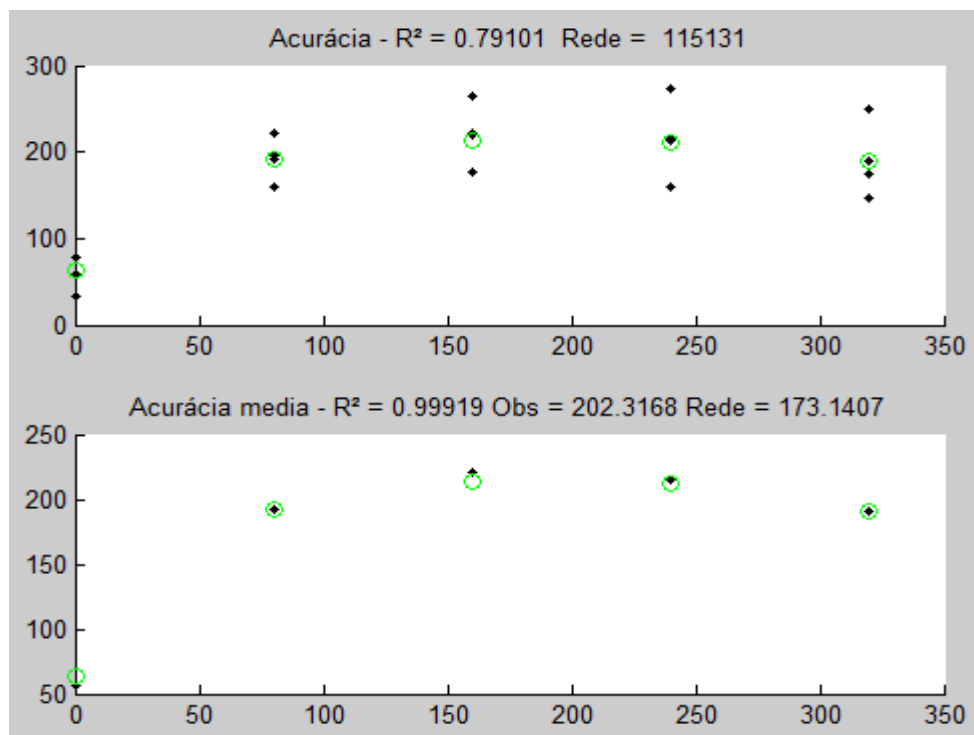


Figura 35: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 3 e da variável peso médio dos bulbos (PMTOT), com todos os valores obtidos (superior) e com a média (inferior).

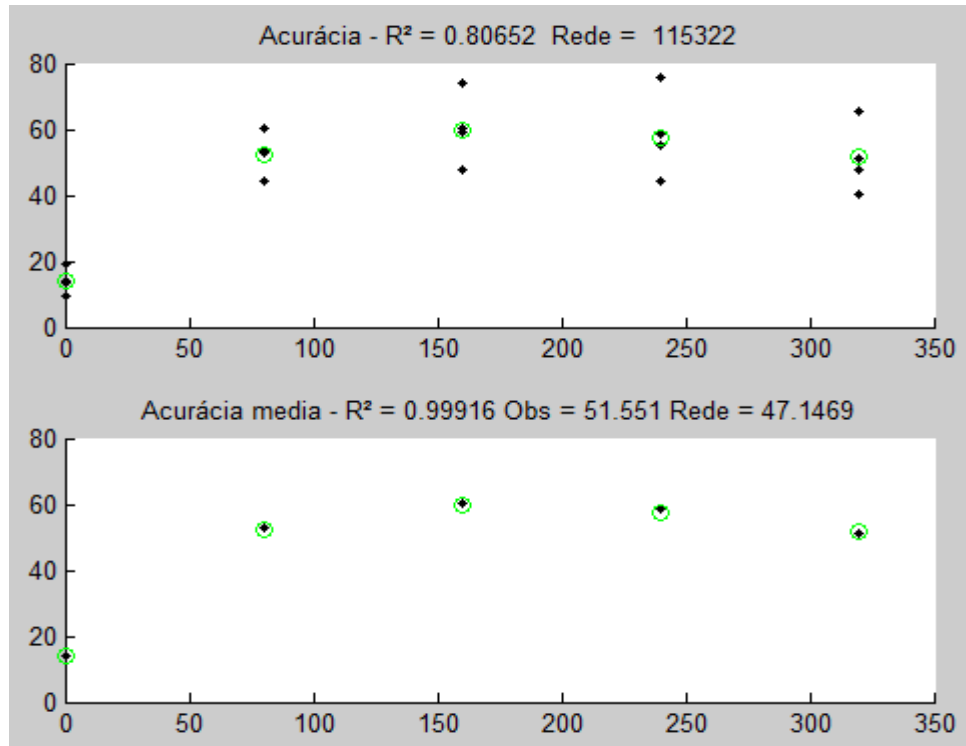


Figura 36: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 3 e da variável peso comerciável de bulbos (PCOM), com todos os valores obtidos (superior) e com a média (inferior).

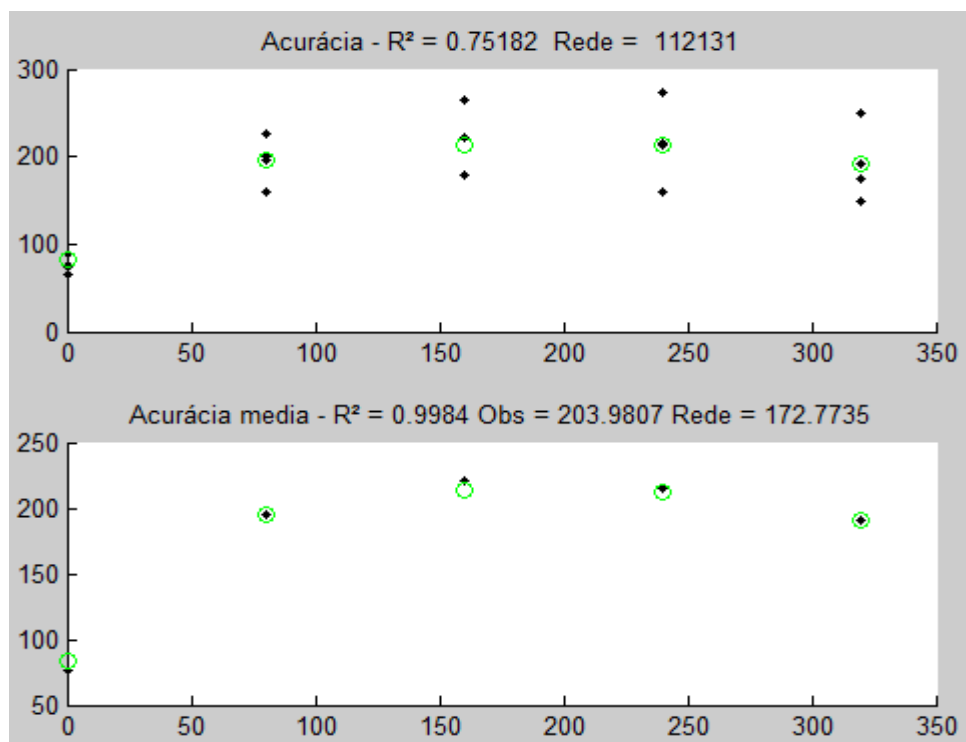


Figura 37: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 3 e da variável peso médio comerciável de bulbos (PMCOM), com todos os valores obtidos (superior) e com a média (inferior).

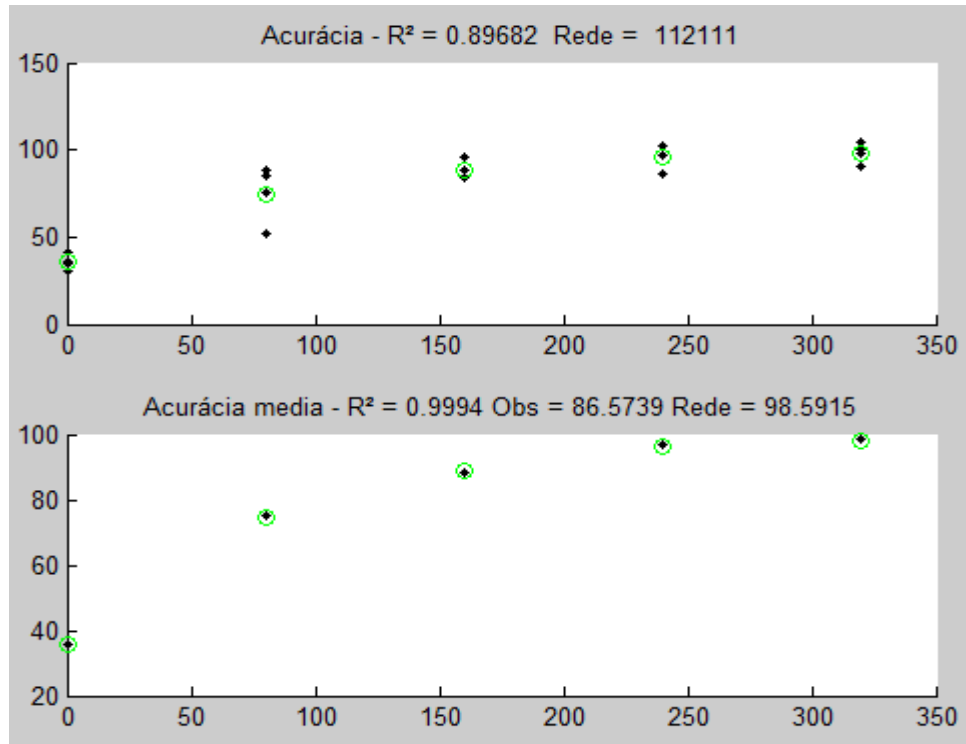


Figura 38: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 4 e da variável peso total dos bulbos (PTOT), com todos os valores obtidos (superior) e com a média (inferior).

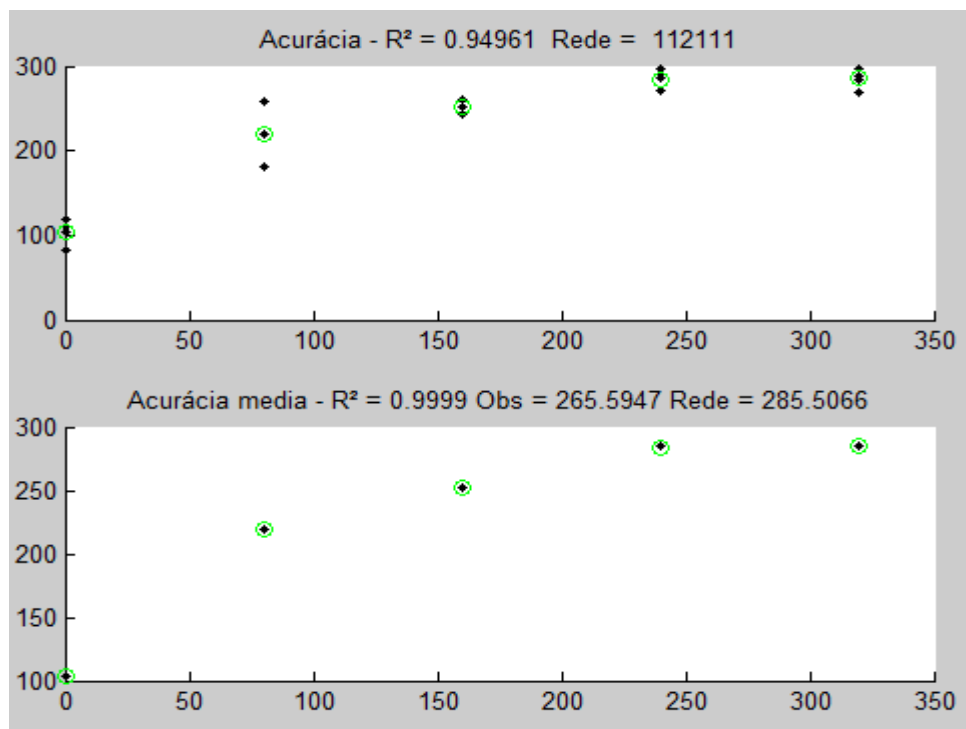


Figura 39: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 4 e da variável peso médio dos bulbos (PMTOT), com todos os valores obtidos (superior) e com a média (inferior).

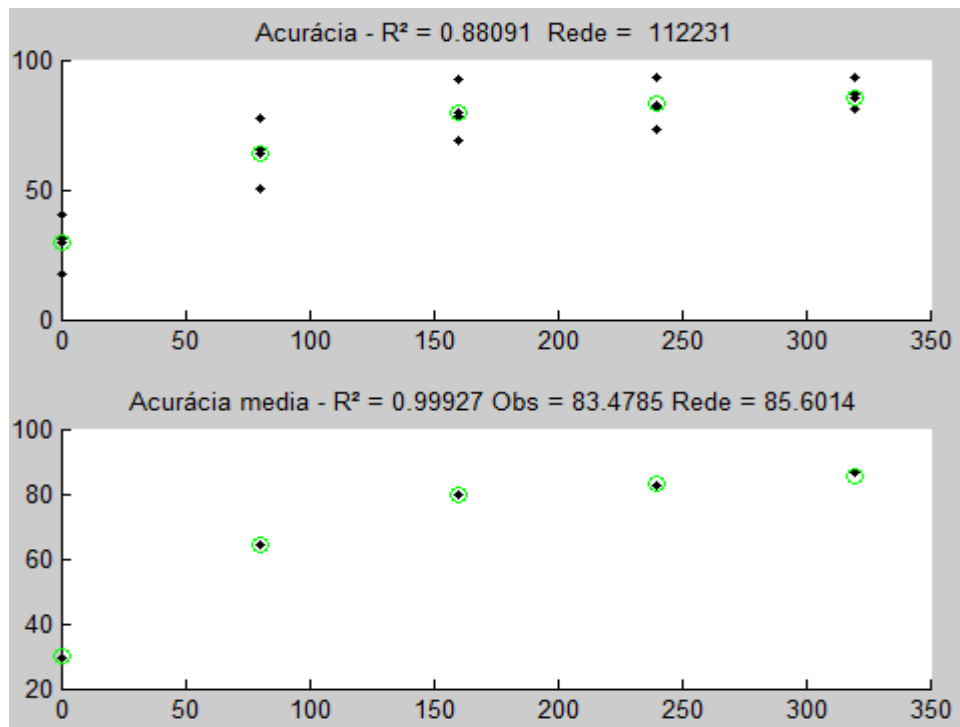


Figura 40: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 4 e da variável peso comerciável de bulbos (PCOM), com todos os valores obtidos (superior) e com a média (inferior).

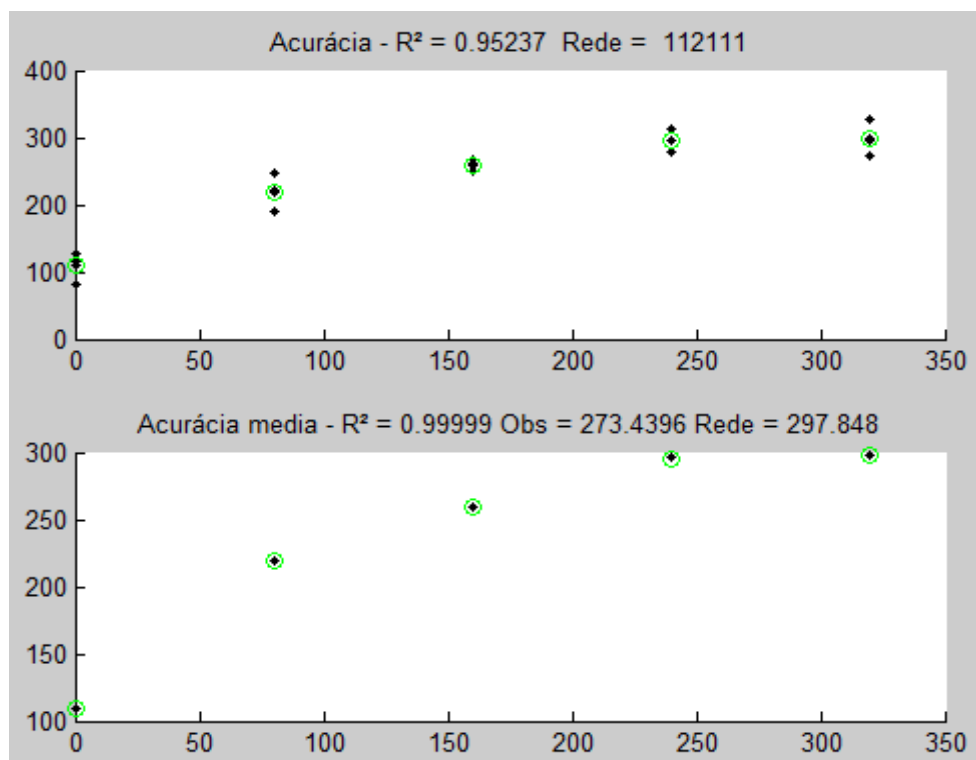


Figura 41: Resultado do processamento do algoritmo computacional projetado para a rede neural do cultivar de cebola 4 e da variável peso médio comerciável de bulbos (PMCOM), com todos os valores obtidos (superior) e com a média (inferior).

Em relação à extrapolação os resultados foram obtidos da mesma forma que os da interpolação. Para que fosse possível a comparação entre os valores previstos depois de realizadas as análises de regressão e as previsões provenientes do processamento do algoritmo da rede neural, os resultados obtidos em todos os processamentos realizados foram colocados em colunas no quadro comparativo a seguir. A Tabela 40 mostra a variável estudada, o local de cultivo da cebola e os valores preditos pela análise de regressão e pela rede neural, além dos valores observados.

Tabela 40: Resumo ilustrando os valores previstos pela análise de regressão, pela rede neural, além do valor observado em cada uma dos quatro cultivares de cebola para cada uma das seguintes variáveis: peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

Variável	Local	Regr.	Rede	Obs.
Peso total dos bulbos (PTOT)	1	8,81	71,31	78,84
Peso médio os bulbos (PMTOT)	1	29,43	147,77	210,65
Peso comerciável de bulbos (PCOM)	1	8,81	70,94	78,84
Peso médio comerciável de bulbos (PMCOM)	1	29,43	137,58	210,65
Peso total dos bulbos (PTOT)	2	207,64	43,93	52,92
Peso médio os bulbos (PMTOT)	2	734,70	139,36	192,22
Peso comerciável de bulbos (PCOM)	2	218,72	42,06	52,92
Peso médio comerciável de bulbos (PMCOM)	2	716,53	125,12	192,22
Peso total dos bulbos (PTOT)	3	81,55	39,83	51,62
Peso médio os bulbos (PMTOT)	3	306,51	173,14	202,32
Peso comerciável de bulbos (PCOM)	3	85,23	47,15	51,55
Peso médio comerciável de bulbos (PMCOM)	3	-37,88	172,77	203,98
Peso total dos bulbos (PTOT)	4	57,53	98,79	86,57
Peso médio os bulbos (PMTOT)	4	167,92	285,51	265,59
Peso comerciável de bulbos (PCOM)	4	47,47	85,60	83,48
Peso médio comerciável de bulbos (PMCOM)	4	196,89	297,85	273,44

Outra forma muito útil de comparação entre os resultados da análise de regressão e o das redes neurais pode melhor ser visualizada no quadro comparativo mostrado a seguir, que mostra a eficiência relativa entre as metodologias. A Tabela 41 mostra a variável estudada, o local de cultivo do híbrido, o coeficiente de determinação ($R^2_{\text{regressão}}$) da análise de regressão, o coeficiente de determinação medindo a eficiência da rede (R^2_{rede}) e uma última coluna exibindo a eficiência relativa que calcula a porcentagem entre o coeficiente de determinação da análise de regressão e o coeficiente de determinação da rede (E_{relativa}).

Tabela 41: Resumo ilustrando a eficiência relativa entre os coeficientes de determinação da análise de regressão e da rede neural em cada uma dos quatro cultivares de cebola para cada uma das seguintes variáveis: peso total dos bulbos (PTOT), peso médio os bulbos (PMTOT), peso comerciável de bulbos (PCOM) e peso médio comerciável de bulbos (PMCOM).

Variável	Loca	$R^2_{\text{regressã}}$	R^2_{rede}	E_{relativa}
Peso total dos bulbos (PTOT)	1	96,72	99,95	103,34%
Peso médio os bulbos (PMTOT)	1	96,57	99,86	103,41%
Peso comerciável de bulbos (PCOM)	1	96,72	99,98	103,37%
Peso médio comerciável de bulbos (PCOM)	1	96,57	99,97	103,52%
Peso total dos bulbos (PTOT)	2	96,20	99,79	103,73%
Peso médio os bulbos (PMTOT)	2	96,72	99,75	103,13%
Peso comerciável de bulbos (PCOM)	2	95,76	99,74	104,16%
Peso médio comerciável de bulbos (PCOM)	2	96,78	99,89	103,21%
Peso total dos bulbos (PTOT)	3	99,70	99,91	100,21%
Peso médio os bulbos (PMTOT)	3	99,71	99,92	100,21%
Peso comerciável de bulbos (PCOM)	3	99,67	99,91	100,24%
Peso médio comerciável de bulbos (PCOM)	3	95,73	99,84	104,29%
Peso total dos bulbos (PTOT)	4	98,21	99,94	101,76%
Peso médio os bulbos (PMTOT)	4	97,98	99,99	102,05%
Peso comerciável de bulbos (PCOM)	4	98,07	99,92	101,89%
Peso médio comerciável de bulbos (PCOM)	4	98,93	99,99	101,07%

4.4 Conclusão

As previsões de valores de interpolação e extrapolação utilizando-se as metodologias baseadas em redes neurais artificiais mostraram-se bastante satisfatórias na previsão de valores se comparadas às análises feitas através da regressão.

Na interpolação de valores, os resultados comparados mostraram que a eficiência da rede em relação à análise de regressão, apesar de ter sido muito próxima, na quase totalidade dos testes, foi um pouco melhor.

Já para os resultados da extrapolação de valores, a rede neural não foi tão eficiente quanto na interpolação, porém, os resultados foram muito melhores quando comparados com os da análise de regressão, resultado que já era esperado.

5. PERSPECTIVAS FUTURAS

Não é difícil de imaginar, pela proporção dos avanços dos acontecimentos tecnológicos, que a grande tendência é a aplicação de redes neurais em grande escala nas mais diversas áreas dentro de muito pouco tempo. O que já existe atualmente será aperfeiçoado e, sendo assim, podemos esperar que esta metodologia de redes neurais mensurando a produtividade em função da quantidade de nitrogênio aplicada aos cultivares de cebola, utilizada neste estudo, venha a contribuir em breve para subsidiar outras pesquisas na área de melhoramento genético.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- ATTWOOD T.K.; GISEL A.; ERIKSSON N-E.; BONGCAM-RUDLOFF E. Europe, 2011. **Concepts, Historical Milestones and the Central Place of Bioinformatics in Modern Biology: A European Perspective. Bioinformatics - Trends and Methodologies.** Versão eletrônica.
Disponível em : <http://www.intechopen.com/books/bioinformatics-trends-and-methodologies/concepts-historical-milestones-and-the-central-place-of-bioinformatics-in-modern-biology-a-european>. Acesso em 8/01/2012.
- BARBETTA, P. A. **Estatística Aplicada às Ciências Sociais.** Florianópolis: Ed. Da UFSC, 2011. 320p.
- BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDEMIR, T. B. **Redes Neurais Artificiais: Teoria e Aplicações.** Rio de Janeiro: LTC, 2000.
- BOITEUX, L.S.; MELO, P.C.T. **Sistema de produção de cebola (*Allium cepa* L.): taxonomia e origem.** Brasil, 2004. Versão eletrônica.
Disponível em: <<http://www.cnph.embrapa.br/sistprod/cebola/index.htm>>.
Acesso em: 21/09/ 2012.
- CHURCHLAND, P.; SEJNOWSKI, T. J. **"The Computational Brain"**, Cambridge: MIT Press, 1992.
- COSTA, C. P. d. **Melhoramento de cebola (*Allium cepa* L.) de curtos para sistemas de cultivo.** Piracicaba: ESALQ, 1978. Tese Livre Docência.
- CRUZ, C. D. **Aplicativo Computacional Genético e Estatística: Análise multivariada e semelhança.** Viçosa: UFV, 2011.
- CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético.** Viçosa: UFV, 2004. v.1. 480p.
- CYBENKO, G. **Approximation by superpositions of a sigmoidal function.** ControlSignalsSyst., New York, 1989.
- FABIO CAMOLESI, J.; SOARES SCOLFORO, J. R.; OLIVEIRA, A. D. de; ACERBI JÚNIOR, F. W.; RUFINI, A. L.; MELLO, J. M. de. **Ajuste, Seleção e Teste de Identidade de Modelo para Volume e Número de Moirões da Candeia (*Eremanthus Erythropappus*).** Cerne, Vol. 16, Núm. 4, Outubro-Diciembre, 2010, p. 431-441. Universidade Federal De Lavras, Lavras, Brasil.
- GRAYBILL, F. A. **Theory and application of the linear model.** Belmont: Duxbury, 1976.
- HAYKIN, S. **Redes Neurais, Princípios e prática.** 2. ed. [S.l.]: Bookman, 2002.
- MACHADO, J. L. A. Brasil, 2006. **Cebola: sabor que nos leva a emocionadas lágrimas.** Versão eletrônica. Disponível em: <www.artigocientifico.uol.com.br>.
Acesso em: 4/01/2013.

MADEIRA, N. R.; RESENDE, F. V.; SOUZA, R. B. de. **Sistema de plantio direto. In: Sistema de produção de cebola (*Allium cepa* L.)**. Versão eletrônica. Brasília: EMBRAPA, 2004. Versão eletrônica.

Disponível em: <http://www.cnph.embrapa.br/sistprod/cebola/plantio_direto.htm>. Acesso em: 19/01/2012.

MONTGOMERY, C. D.; RUNGER, G. C. **Estatística Aplicada e Probabilidade para Engenheiros**. Rio de Janeiro: LTC, 2009. 493p.

NETER, J.; KUTNER, M. H.; NACHTSHEIM, C. J.; WASSERMAN, W. **Applied linear statistical models**. 4. ed. New York: McGraw-Hill, 1985.

OLIVEIRA, A. P. d.; PEREIRA, E. L.; BRUNO, R. d. L. A.; ALVES, E. U.; COSTA, R. F. d.; LEAL, F. R. F. Brasil, 2003. **Rendimento e qualidade de sementes do feijão-vagem em função de fontes e níveis de nitrogênio**. Revista brasileira de Sementes 25 (1). Versão eletrônica.

Disponível em: URL http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-31222003000100009&lng=en&nrm=isso. Acesso em 4/01/2013.

REGAZZI, A. J. Brasil, 2003. **Teste para verificar a igualdade de parâmetros e a identidade de modelos de regressão não-linear**. Revista Ceres, 50:9-26.

REGAZZI, A. J.; SILVA, C. H. O. Brasil, 2004. **Teste para verificar a igualdade de parâmetros e a identidade de modelos de regressão não-linear. I. dados no delineamento inteiramente casualizado**. Revista de Matemática e Estatística, 22:33-45.

SCHUTZER, D. **Artificial intelligence: an applications-oriented approach**. New York: Van Nostrand Reinhold Company, 1987.

SOUZA, R. J. d.; RESENDE, G. M. d., Brasil, 2002. **Cultura de cebolas**. Vol. 1. Lavras: UFLA/FAEPE. 115 p.

VIDIGAL, S. M.; COSTA, E. L.; CIOCIOLA, J. A. I. Brasil, 2007. **Cebola (*Allium cepa* L.). 101 Culturas – Manual de Tecnologias Agrícolas**. Belo Horizonte: EPAMIG. p. 243-252

VIDIGAL, S. M.; PEDROSA, M. W.; FONSECA, M. S.; SANTOS, I. C. Brasil, 2010. **Adubação com nitrogênio em cobertura na produção de cebola**. Horticultura Brasileira 28: S3705-S3711.

WERKEMA, M. C. C. **Análise de regressão: como entender o relacionamento entre as variáveis de um processo**. Belo Horizonte: Fundação Christiano Ottoni; UFMG, 1996. 311p.