



JOYCE DÓRIA RODRIGUES SOARES

**MODELAGEM MATEMÁTICA NA PREVISÃO
DE COLHEITA DE BANANEIRA: REGRESSÃO
LINEAR MÚLTIPLA x REDES NEURAIAS
ARTIFICIAIS**

**LAVRAS – MG
2011**

JOYCE DÓRIA RODRIGUES SOARES

**MODELAGEM MATEMÁTICA NA PREVISÃO DE COLHEITA DE
BANANEIRA: REGRESSÃO LINEAR MÚLTIPLA x REDES NEURAIAS
ARTIFICIAIS**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Agronomia/Fitotecnia, área de concentração em Produção Vegetal, para a obtenção do título de Doutor.

Orientador
Dr. Moacir Pasqual

Dr. Wilian Soares Lacerda
Coorientador

**LAVRAS – MG
2010**

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca da UFLA**

Soares, Joyce Dória Rodrigues.

Modelagem matemática na previsão de colheita de bananeira:
regressão linear múltipla x redes neurais artificiais / Joyce Dória
Rodrigues Soares. – Lavras: UFLA, 2010.

115 p.: il.

Tese (doutorado) – Universidade Federal de Lavras, 2010.

Orientador: Moacir Pasqual.

Bibliografia.

1. Banana. 2. Regressão. 3. Redes neurais. I. Universidade
Federal de Lavras. II. Título.

CDD – 634.7725

SOARES, J. D. R. **Modelagem matemática na previsão de colheita de bananeira**: regressão linear múltipla x redes neurais artificiais (Doutorado em Fitotecnia) – Universidade Federal de Lavras, Lavras, 2010.

ERRATA

Folha	Linha	Onde se lê	Leia-se
50	12	Utilizou-se, para instalação do experimento, mudas micropropagadas, cedidas pela Embrapa Mandioca e Fruticultura, multiplicadas pela Campo Biotecnologia na Biofábrica em Cruz das Almas, BA, aclimatizadas em tubetes, e transportadas para o local do experimento enroladas em embalagem tipo rocambole, onde foram transplantadas diretamente para o campo.	Os dados levantados no experimento de campo que possibilitaram a elaboração deste trabalho é do original de Donato et al. (2008) e as descrições apresentadas a seguir, referentes a condução e implantação da cultura e as avaliações fitotécnicas, são derivadas de Donato (2007). Utilizou-se, para instalação do experimento, mudas micropropagadas, cedidas pela Embrapa Mandioca e Fruticultura, multiplicadas pela Campo Biotecnologia na Biofábrica em Cruz das Almas, BA, aclimatizadas em tubetes, e transportadas

para o local do experimento enroladas em embalagem tipo rocambole, onde foram transplantadas diretamente para o campo.

73	12	idem
98	12	idem

idem
Idem

Folha	Inserir
65	DONATO, S.L.R. Estimativas de tamanho e forma de parcelas experimentais para avaliação de descritores fenotípicos em bananeira (<i>Musa spp.</i>). 2007. 188f . Tese (Doutorado) – Departamento de Fitotecnia, Universidade Federal de Viçosa, Viçosa, 2007.
	DONATO, S.L.R.; SIQUEIRA, D.L.; SILVA, S.O.; CECON, P.R.; SILVA, J.A.; SALOMÃO, L.C.C. Estimativas de tamanho de parcelas para avaliação de descritores fenotípicos em bananeira. <i>Pesquisa Agropecuária Brasileira</i>, Brasília, v.43, n.8, p.957-969, ago. 2008.
90	idem
112	idem

JOYCE DÓRIA RODRIGUES SOARES

MODELAGEM MATEMÁTICA NA PREVISÃO DE COLHEITA DE BANANEIRA: REGRESSÃO LINEAR MÚLTIPLA x REDES NEURAIAS ARTIFICIAIS

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Agronomia/Fitotecnia, área de concentração em Produção Vegetal, para a obtenção do título de Doutor.

APROVADA em 07 de dezembro de 2010.

Dra. Leila Aparecida Salles Pio	PNPD/CAPES/EMBRAPA
Dra. Aparecida Gomes de Araujo	PNPD/CAPES/UFLA
Dr. Sebastião de Oliveira e Silva	UFRB
Dra. Franciane Tavares Braga	PNPD/CAPES/ EPAMIG

Dr Moacir Pasqual
Orientador

Dr. Wilian Soares Lacerda
Coorientador

LAVRAS - MG
2010

A Jesus, filho do Deus vivo, que me proporcionou uma paz interior e
felicidade que levarei comigo eternamente.

A meu marido Charles e meus filhos Maria Eduarda e Arthur.

A minha mãe Miriam, meu irmão Rafael e meus avós João e Rita.

DEDICO

BIOGRAFIA

JOYCE DÓRIA RODRIGUES SOARES, filha de André Luiz Rodrigues (falecido) e Miriam Dória Sousa Rodrigues, nasceu em 29 de abril de 1982, no município de Araçuaí, estado de Minas Gerais. Coursou até a oitava série na cidade de Coronel Murta, mudou-se para Betim (MG), onde concluiu o ensino médio, e, na cidade de Lavras, obteve o título de Engenheira Agrônoma, pela Universidade Federal de Lavras, em julho de 2007.

Sempre estive envolvida em atividades de pesquisa desde o primeiro período de sua graduação, dedicando tempo aos estudos e à descoberta de inovações científicas. Suas atividades se iniciaram no Laboratório de Cultura de Tecidos de Plantas onde, no terceiro período, foi agraciada com bolsa de iniciação científica, pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), durante 36 meses, sendo sua monografia de conclusão de curso realizada com espécies de orquídeas.

Em agosto de 2007, iniciou o Mestrado em Agronomia (área de concentração Fisiologia Vegetal), na Universidade Federal de Lavras, sendo aprovada no Doutorado pela mesma Instituição em junho de 2009 e tendo defendido a dissertação em 18 de julho de 2009. Iniciou o doutorado em Fitotecnia (área de concentração Produção Vegetal), na Universidade Federal de Lavras, onde concluiu o curso em 18 meses.

AGRADECIMENTOS

Primeiramente a Deus, pois hoje enxergo que Ele sempre teve um plano na minha vida. Ele me trouxe aqui pra essa cidade onde vivi e vivo os melhores momentos de minha vida. Conheci-O de verdade e também pessoas que eu sempre lembrarei com muito carinho e saudade.

Ao meu marido Charles e meus filhos Maria Eduarda e Arthur, que sempre estiveram ao meu lado com carinho e amor. Vocês são os amores da minha vida.

À minha mãe que veio de Portugal para morar comigo para me auxiliar no término do doutorado. Meu muito obrigada e amo você (aproveita heim!!!!)

Ao meu orientador Moacir Pasqual, pessoa sem igual, de uma alegria admirável e, principalmente, pela oportunidade, confiança, ensinamentos e presteza, incondicionais, demonstrados e repassados, meus sinceros agradecimentos.

Ao meu coorientador Wilian Soares Lacerda, sou grata por tudo que fez e admiro sua competência e profissionalismo. Agradeço por não ter medido esforços em me ajudar e me direcionar por qual caminho seguir, por ter me atendido várias vezes e sempre com prestatividade e muita educação. Meu sinceros agradecimentos.

Ao Dr. Sebastião de Oliveria e Silva por ter possibilitado, juntamente com meu orientador, a realização e concretização do presente trabalho. Obrigada pela confiança e prestatividade pois sei que não mediu esforços para me ajudar e sempre esteve disposto a me atender. Meu sincero muito obrigada.

A minha avó Rita e meu avô João pela proteção constante, por todas as grandes e pequenas oportunidades, pelo carinho, dedicação e confiança repassados durante toda a minha vida, meu eterno amor e gratidão.

À FAPEMIG pela concessão de bolsas de estudos para o doutorado.

À Universidade Federal de Lavras, instituição na qual obtive a base acadêmica que será parte integrante e fundamental para minha vida.

Aos amigos do laboratório, doutorandos, mestrandos e iniciação científica, nunca me esquecerei do cuidado que tiveram comigo no período de gravidez, obrigada pela amizade e carinho, amo todos vocês.

À 'família' do Laboratório de Cultura de Tecidos Vegetais da UFLA, na pessoa de Vantuil, Claret e Antônio Carlos, profissionais extremamente capazes, dedicados e prestativos, meu muito obrigado.

A minha segunda mãe, líder e amiga, Rosane Freitas Schwan, que nem imagina o quanto é importante pra mim. Eu te amo.

MUITO OBRIGADA!

RESUMO

Um dos entraves relevantes à bananicultura brasileira é a falta de variedades comerciais produtivas com porte adequado, resistentes às principais pragas e doenças e adaptadas a diferentes ecossistemas. O desenvolvimento de cultivares constitui estratégia para a solução deste problema, mediante programas de melhoramento genético, bem como, sua avaliação e caracterização em áreas de produção quando são comparadas às cultivares tradicionais. Os caracteres observados em áreas experimentais têm natureza fenotípica e na maioria das vezes baseia-se apenas na experiência do produtor, mensuração do peso do cacho e as correlações das variáveis envolvidas são estimadas visando mensurar alterações em um caracter quando se altera outro relacionado. O presente trabalho teve como objetivos: 1) Desenvolver um modelo matemático baseado em regressão linear simples; 2) Estudar a relação entre o peso do cacho e as variáveis que compõem o modelo; 3) Criar um modelo de previsão de colheita baseado em redes neurais artificiais; 4) Comparar os sistemas de predição de peso de cacho: regressão e redes neurais. O experimento constituiu-se de um ensaio de uniformidade, conduzido em Guanambi, BA, com a cultivar Tropical (YB42-21), híbrido tetraplóide AAAB, plantado no espaçamento de 3 m x 2 m, formado de 11 fileiras de 52 plantas cada e consideradas como útil as 9 fileiras centrais com 40 plantas por fileira, num total de 360 plantas e área de 2.160 m². Avaliaram-se os caracteres vegetativos, altura da planta, perímetro do pseudocaule, número de filhos emitidos e número de folhas vivas no florescimento e na colheita e os caracteres de rendimento, peso do cacho, número de pencas e frutos, peso da segunda penca, comprimento e diâmetro do fruto em dois ciclos de produção. Nas avaliações, cada planta foi considerada como uma unidade básica (ub), área de 6 m², perfazendo assim, 360 unidades básicas (ub). As principais conclusões são: não é indicado a utilização de modelos lineares para previsão do peso do cacho; a RNA implementada obteve uma eficiente previsão de produção, sendo que a rede que obteve o melhor resultado possui a arquitetura 10:10:1.

Palavras-chave: Banana. Inteligência computacional. Previsão de colheita.

ABSTRACT

One of the barriers relevant to the banana crop in Brazil is the lack of productive commercial varieties with adequate size, resistant to major pests and diseases and adapted to different ecosystems. The development of cultivars is the strategy for solving this problem through breeding programs, as well as its characterization and evaluation in areas of production when compared to traditional cultivars. The characters observed in the experimental areas have phenotypic nature and most often based solely on the experience of the producer, the measurement of bunch weight and the correlations of the variables involved are estimated in order to measure changes in a character when he altered another. This study aimed to: 1) Develop a model based in mathematic simple linear regression, 2) study the relationship between bunch weight and the variables that make up the model, 3) Create a prediction model based on neural networks harvest artificial; 4) Compare systems of predicting bunch weight: regression and neural networks. The experiment consisted of a uniformity trial, conducted in Guanambi, BA, with cultivar Tropical (YB42-21), AAAB tetraploid hybrid, planted at a spacing of 3 mx 2 m, consisting of 11 rows of 52 plants each and considered how useful the nine central rows with 40 plants per row, totaling 360 plants and an area of 2,160 m². We assessed the vegetative characters, plant height, pseudostem circumference, number of children issued and number of green leaves at flowering and harvest, and the characters of yield, bunch weight, number of hands and fruits, weight of the second hand, length and diameter of the fruit in two growing seasons. In the evaluation, each plant was considered as a basic unit (bu), area of 6 m² totaling well, 360 basic units (bu). It is inappropriate to use linear models for predicting the weight of the bunch. The RNA had implemented an efficient production forecast, and the network that obtained the best result has 10:10:1 architecture.

Keywords: Banana. Computational intelligence. Forecasting crop.

SUMÁRIO

	CAPÍTULO 1: Introdução geral.....	12
1	INTRODUÇÃO.....	12
2	REVISÃO DE LITERATURA.....	16
2.1	A cultura da bananeira.....	16
2.2	Modelagem na agricultura.....	17
2.3	Modelo de regressão linear.....	19
2.3.1	Modelo de regressão linear simples.....	20
2.3.2	Modelo de regressão linear múltipla.....	21
2.3.3	Teste de hipótese e intervalo de confiança na regressão linear.....	22
2.3.4	Aplicação da regressão linear: a estimação de valores.....	23
2.3.5	Coefficiente de determinação.....	23
2.4	Redes Neurais Artificiais (RNAs).....	24
2.4.1	Histórico.....	25
2.4.2	Neurônio Artificial.....	27
2.4.3	Função de Ativação.....	28
2.4.4	Redes Neurais Artificiais “<i>Multilayer Perceptron</i>”.....	31
2.4.5	Algoritmo de Aprendizagem “<i>backpropagation</i>”.....	34
2.5	Aplicação da modelagem matemática na agricultura.....	37
	REFERÊNCIAS.....	41
	CAPÍTULO 2: Previsão do peso do cacho da bananeira ‘Tropical’ com base em regressão linear múltipla.....	47
1	INTRODUÇÃO.....	49
2	MATERIAL E MÉTODOS.....	51
2.1	Implantação e condução da cultura.....	51
2.2	Avaliações.....	52
2.3	Caracteres vegetativos.....	52
2.4	Caracteres de rendimento.....	53
2.5	Análise estatística.....	55
3	RESULTADOS E DISCUSSÃO.....	57
4	CONCLUSÃO.....	65
	REFERÊNCIAS.....	66
	CAPÍTULO 3: Utilização de Redes Neurais Artificiais para previsão de peso do cacho da bananeira cv. Tropical.	69
1	INTRODUÇÃO.....	71
2	MATERIAIS E MÉTODOS.....	74
2.1	Implantação e condução da cultura.....	74
2.2	Avaliações.....	75

2.3	Caracteres de rendimento.....	75
2.4	Modelagem de redes neurais artificiais (RNAs)	77
2.4.1	Estrutura do modelo RNA.....	78
2.4.2	Treinamento do modelo.....	78
3	RESULTADOS E DISCUSSÃO.....	81
4	CONCLUSÃO.....	89
	REFERÊNCIAS.....	90
	CAPÍTULO 4: Comparação de técnicas de previsão de rendimento em cacho de bananeira cv. Tropical: Redes Neurais Artificiais x Regressão Linear Múltipla.....	94
1	INTRODUÇÃO.....	96
2	MATERIAIS E MÉTODOS.....	99
2.1	Implantação e condução da cultura.....	99
2.2	Avaliações.....	100
2.3	Desenvolvimento das redes neurais e análise de regressão múltipla.....	102
3	RESULTADOS E DISCUSSÃO.....	105
4	CONCLUSÃO.....	112
	REFERÊNCIAS.....	113

CAPÍTULO 1

INTRODUÇÃO GERAL

1 INTRODUÇÃO

A banana (*Musa* spp.) tem como centro de origem o Sudeste da Ásia e Oeste do Pacífico, sendo cultivada em uma extensa área dos trópicos e subtropicais. A partir dos centros de origem, a bananeira foi introduzida na África, Américas e sul do Pacífico, onde ganhou popularidade e importância econômica (DE LANGHE et al., 2009; VALMAYOR et al., 2001) constituindo-se uma das principais fontes de alimento para milhões de pessoas (BOONRUANGROD; FLUCH; BURG, 2009).

Dentre as frutíferas tropicais a bananeira destaca-se como a mais consumida, tanto pela sua versatilidade em termos de modalidades de consumo (processada, frita, cozida, *in natura*) quanto pelas suas características de sabor, aroma, higiene e facilidade de ser consumida sendo considerada mundialmente um importante alimento em razão da sua composição química e conteúdo em vitaminas e minerais, principalmente potássio. O consumo per capita nacional é em torno de 31 kg hab.⁻¹ ano⁻¹, o que equivale a cerca de 84,9 g hab.⁻¹ dia⁻¹. A banana para consumo *in natura* foi a fruta mais produzida no mundo em 2005 e em conjunto com plátanos (bananas de cozinhar e fritar) foram cultivadas em 134 países, numa área total de 10,25 milhões de hectares, com produção de 125,05 milhões de toneladas. O Brasil é o sexto maior produtor mundial de banana e plátanos, com uma produção de 7,11 milhões de toneladas em uma área cultivada de 513 mil hectares (FOOD AND AGRICULTURAL ORGANIZATION, 2010).

No Brasil a bananeira é a segunda frutífera mais plantada, depois da laranjeira, e a sua produção em 2006 foi 6.996.648 toneladas, numa área cultivada de 505.384 hectares com produtividade de 13,84 toneladas hectare⁻¹. A exportação brasileira de banana tem como principal destino o Mercosul (cerca de 70%) e varia de 1 a 3,5% do total produzido, o que denota uma produção voltada quase que exclusivamente para o mercado interno. Problemas relativos à qualidade e regularidade de oferta do produto, associados à questões específicas de comercialização, boa remuneração do produto no mercado interno, além das variedades cultivadas, concorrem para limitar a participação do país no mercado internacional de bananas (ANUÁRIO BRASILEIRO DA AGRICULTURA, 2007).

O Estado da Bahia apresentou em 2006 a maior área cultivada com bananeiras (74.210 hectares) e a segunda maior produção (1.059.877 toneladas), após o Estado de São Paulo, maior produtor brasileiro (AGRIANUAL, 2007). A produtividade média da cultura na Bahia é baixa, cerca de 14,2 t ha⁻¹, similar a brasileira.

Na região semi-árida brasileira destacam-se os pólos de produção de bananas de Minas Gerais, localizado em Janaúba e Jaíba; Bahia, em Juazeiro, Bom Jesus da Lapa, Santa Maria da Vitória, Barreiras, Livramento de Nossa Senhora, Caraíbas e Guanambi (Ceraima e Estreito); Pernambuco, em Petrolina e Santa Maria da Boa Vista; Rio Grande do Norte, no Vale do Açu; o de Sergipe, no Plato de Neopolis, e Ceará, na Chapada do Apodi e no Baixo Acarau.

A despeito da existência de um número expressivo de cultivares de banana no Brasil, nem todas apresentam potencial para utilização comercial, pois, inexistem cultivares que sejam resistentes a todas as pragas, apresentem alto rendimento, sejam precoces, saborosas, de fácil manejo, com alta longevidade e

tempo de prateleira. Cada cultivar possui vantagens e desvantagens com relação ao que se deseja num ideótipo (DANIELS, 2000). Nesse sentido, é estratégico o desenvolvimento de variedades que possuam características desejáveis, por meio de programas de melhoramento genético, bem como, a execução de experimentos de avaliação de genótipos.

Nesses experimentos normalmente são estudados ciclo da cultura, altura da planta, perímetro do pseudocaule, peso do cacho, número de frutos por cacho, comprimento e diâmetro dos frutos. Esses descritores são relevantes para a identificação e a seleção de indivíduos superiores (SILVA et al., 2000), pois, são quantitativos, fáceis de mensurar, podem estar sobre controle poligênico, sofrem influência ambiental, tem importância econômica direta e indireta (ORTIZ, 1997). E por isso, interessam aos melhoristas e produtores o conhecimento das relações entre os caracteres é de suma importância à bananicultura, possibilitando, ao produtor, estimar a produção do cacho de uma determinada planta a partir de outros atributos. Embora apresente trabalhos relacionando caracteres componentes da produção em diferentes genótipos de bananeira (JARAMILLO, 1982), a literatura carece de informações que permitam uma estimativa do peso do cacho a partir de alguns atributos medidos na fase da colheita.

O número de pesquisas conduzidas contemplando estimativas de produção tem sido feitas, para diversas culturas (LE BAIL et al., 2005). Entretanto, para a bananeira, pouca referência se encontra acerca de trabalhos realizados (JARAMILLO, 1982; DADZIE, 1998) os quais estão defasados, não sendo encontradas referências nacionais a esse respeito.

Condições climáticas, heterogeneidade do solo, práticas culturais e custos da condução do experimento devem ser considerados quando germoplasma melhorado e ou novas tecnologias são testados. Atualmente, os

recursos para pesquisas demandam competitividade acirrada. Adicionalmente, incrementos de produtividade como consequência da elevada evolução tecnológica e de melhoramento genético, tendem ser de pequena magnitude. Por conseguinte, é justificável a necessidade de se procurar identificar as variáveis que explicam parte da variação da produtividade por meio de modelagem matemática para aferição do peso do cacho baseado em caracteres agronômicos coletados ao longo do ciclo de produção, para que o produtor possa utilizar além de sua experiência um modelo que lhe permita realizar a previsão de produção, e de posse disso conseguir financiamentos bancários para a colheita.

2 REVISÃO DE LITERATURA

2.1 A cultura da bananeira

A bananeira de frutos comestíveis (*Musa* spp.) é uma monocotiledônea da classe *Liliopsida*, subclasse *Liliidae*, superordem *Lilinae*, ordem *Zingiberales* (Scitamineae), família *Musaceae*, subfamília *Musoideae*, gênero *Musa*, seção *Eumusa* (SILVA et al., 2002) originada de cruzamentos interespecíficos entre *Musa acuminata* Colla e *M. balbisiana* Colla. Apresenta caracteres das duas espécies e três níveis de ploidia, existindo diplóides (2n) com 22 cromossomos, triplóides (3n) com trinta e três cromossomos e tetraplóides (4n) com 44 cromossomos, sendo que o número básico de cromossomos é 11 ($n = 11$) (SIMMONDS, 1973; SHEPHERD, 1984).

A planta, típica dos trópicos úmidos, é um vegetal herbáceo de grandes dimensões, apresenta pseudocaule aéreo que se origina do rizoma de onde se desenvolve gemas laterais ou filhos (SOTO BALLESTERO, 1992). O centro de origem da maior parte do germoplasma de banana está localizado no continente Asiático, ocorrendo centros secundários na África Oriental, algumas ilhas do Pacífico e uma considerável diversidade genética na África Ocidental (SHEPHERD, 1984), regiões com clima tropical quente e úmido.

O cultivo de bananeira no mundo está situado geograficamente entre latitudes de 30°S e 30°N (SOTO BALLESTERO, 1992). As condições ótimas para o seu cultivo são encontradas entre 15° de latitude Sul e Norte. No Brasil, o cultivo da bananeira ocorre em todos os estados da federação, em diversas condições de solo e clima e nos mais variados ecossistemas.

A banana 'Maçã', uma das cultivares mais apreciadas pelos brasileiros, teve o seu cultivo limitado em todo o País pela alta suscetibilidade ao mal-do-

Panamá. A Embrapa Mandioca e Fruticultura busca soluções através de seu programa de melhoramento genético com o objetivo de obtenção de cultivares tipo Maçã que resultou no lançamento do híbrido Tropical (YB42-21) (SILVA et al., 2002).

2.2 Modelagem na agricultura

Tradicionalmente, as funções de produção usadas para recomendações agrícolas tem sido baseadas em análises de correlação e regressão. O uso desse tipo de análise tem auxiliado no conhecimento qualitativo das variáveis e suas interações envolvidas nos sistemas das culturas e também contribuído para o progresso da ciência da agricultura. No entanto, a informação qualitativa obtida nesse tipo de análise é específica para determinados ambientes, havendo usualmente poucas considerações a respeito dos processos físicos e fisiológicos envolvidos (JAME; CURTFORTH, 1996).

A interação entre as plantas e o ambiente envolve uma complexidade de processos químicos, físicos e biológicos. A fim de se obter melhor conhecimento das respostas da cultura ao ambiente, modelos de simulação são utilizados como ferramenta de grande potencial na área de sistemas cultivados, permitindo o estudo e o entendimento do conjunto, estimando a desempenho da cultura em diferentes áreas e situações.

Visando ao adequado entendimento da definição de modelos de simulação, Costa (1997) menciona a necessidade de se fazer a separação de três termos: sistemas, modelos e simulação. Sistema é uma parte limitada da realidade que contem vários elementos inter-relacionados; modelo é uma representação simplificada de um sistema; simulação é a arte de construir

modelos matemáticos e de estudar suas propriedades em relação às do sistema (DE WIT, 1982).

Haefner (2005) definem modelo como uma série de equações matemáticas, as quais descrevem um sistema físico, no caso, sistema solo-planta-atmosfera. O modelo simula uma cultura pela estimativa do crescimento de seus componentes, como folhas, raízes, caules e grãos. Assim, um modelo de simulação de crescimento da cultura não somente estima a biomassa total ou a produção na colheita, mas também, inclui informações quantitativas sobre a maioria dos processos envolvidos no crescimento e desenvolvimento da planta.

Segundo Monteith (1996), os modelos de simulação de culturas podem ser definidos como um conjunto de equações para estimar o crescimento, desenvolvimento e produção de uma cultura a partir de uma série de coeficientes genéticos e variáveis ambientais. Os modelos permitem analisar detalhadamente os diversos componentes da produção, possibilitando uma visão integrada de sua participação no sistema. Apesar da complexidade envolvida na construção de modelos, os esforços são compensados em função de sua grande aplicabilidade, uma vez que estes auxiliam o agricultor na tomada de decisões, permitindo a organização racional de questões envolvendo distribuição, armazenamento e comercialização do produto agrícola, entre outras (CARDOSO et al., 2004).

Existem três níveis de uso dos modelos de simulação de culturas como ferramenta potencialmente útil na tomada de decisões: em pesquisas, em tecnologia de manejo de cultivos e em políticas de planejamento agrícola (HOOGENBOOM, 2000).

Modelos podem ser utilizados para analisar os efeitos de diferentes estratégias, ajudando na determinação da melhor decisão, e apresentam uma série de vantagens para uma análise econômica, em relação aos dados provenientes de experimentação em campo. Entre essas vantagens destacam-se:

um grande número de estratégias pode ser avaliado rapidamente e com baixo custo; cada estratégia pode ser avaliada em um amplo intervalo de condições incertas, como por exemplo, clima; e dependendo do modelo de crescimento utilizado, pode-se avaliar complexas estratégias de manejo interagindo com diversas decisões como irrigação, fertilização, datas de semeadura, dentre outras (BOOGES; RITCHIE, 1988).

Vale ressaltar que, embora os modelos de simulação tenham grande aplicabilidade e custo inferior em relação aos experimentos convencionais, eles não podem ser considerados substitutos dos experimentos, e sim uma técnica que os complementa.

2.3 Modelo de regressão linear

O termo regressão foi introduzido por Francis Galton, em seu estudo sobre a altura de crianças. Nesse estudo, Galton constatou que a altura média dos filhos de pais de uma determinada altura tendia a “regredir” até a altura média da população. Esse fenômeno foi chamado de lei de regressão universal de Galton e foi confirmado posteriormente por Pearson (GUJARATI, 2000; NETER et al., 2005).

O significado atual da análise de regressão é diferente da abordagem inicial, sendo definida por Montgomery et al. (2001), como “uma técnica estatística para modelar e investigar a relação entre variáveis”. Weisberg (2005) defende que a regressão, assim como a maioria das técnicas estatísticas, tem como meta sumarizar os dados observados da forma mais simples, útil e elegante possível. Para Gujarati (2000) e Neter et al. (2005), a análise de regressão estuda a dependência da característica de interesse, em relação a uma ou mais variáveis, com o objetivo de estimar o valor médio da característica de

interesse em função dos valores conhecidos das variáveis de controle.

A análise de regressão é uma das ferramentas mais utilizadas para a análise de dados, sendo aplicada em quase todas as áreas de conhecimento, incluindo: engenharia, física, química, administração, ciências biológicas e agronomia. A simplicidade e utilidade dessa técnica são resultantes do conceito lógico de utilizar uma equação matemática para expressar a relação entre variáveis, sendo interessante devido à base matemática e bem desenvolvida teoria estatística (MONTGOMERY et al., 2001).

2.3.1 Modelo de regressão linear simples

O modelo de regressão linear simples (MRLS) é representado pela seguinte equação (MONTGOMERY et al., 2001).

$$Y = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (1)$$

O termo ε é chamado de erro aleatório, sendo assumido como normal e independentemente distribuído, com média zero e variância constante e desconhecida σ^2 . Essa suposição é importante para a estimação da reta de regressão e deve ser verificada posteriormente para a validação do modelo estimado (MONTGOMERY et al., 2001; NETER et al., 2005).

Segundo Neter et al. (2005), a equação (1) é dita simples, pois representa a relação entre uma característica de interesse e uma variável de controle, é linear nos parâmetros, pois nenhum dos parâmetros aparece como expoente ou está sendo multiplicado ou dividido por outros parâmetros.

Montgomery et al. (2001) afirmam que a variável x deve ser vista como sendo uma variável de controle e medida com erro desprezível, enquanto a variável Y é vista como uma variável aleatória. Dessa forma, existe uma distribuição de probabilidade para cada Y . A média dessa distribuição é estimada

por:

$$(Y|X) = \beta_0 + \beta_1 X \quad (2)$$

E a variância é:

$$\text{Var}(Y|X) = \text{Var}(\hat{\beta}_0 + \hat{\beta}_1 X_i + \varepsilon_i) = \text{Var}(\varepsilon_i) = \sigma^2 \quad (3)$$

Para o uso de um modelo de regressão linear simples, é necessário utilizar um gráfico de dispersão, que, de acordo com Montgomery; Runger (2003) é uma ferramenta de cada par (X_i, Y_i) , em um ponto plotado em um sistema bidimensional, a qual indica se os pontos estão aleatoriamente dispersos em torno de uma linha reta.

2.3.2 Modelo de regressão linear múltipla

De acordo com Montgomery et al. (2001), o modelo de regressão linear múltipla (MRLM) com k variáveis de controle é representado pela equação (abaixo).

$$Y = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i \quad (4)$$

Os coeficientes de regressão $\beta_0, \beta_1, \dots, \beta_k$ são descritos por Gujarati (2000) e Montgomery et al. (2001) como:

- a) O coeficiente β_0 , chamado de coeficiente de intercepto, o qual corresponde à média de Y_i quando todas as variáveis de controle são iguais a zero;
- b) Os coeficientes β_1, \dots, β_k são chamados de coeficientes de regressão parciais. O coeficiente β_k pode ser interpretado como a derivada parcial de Y_i em relação a

X_{ki} , ou seja, é variação de Y causada por uma variação unitária em x_k dado que as outras variáveis de controle sejam mantidas constantes.

É mais prático trabalhar com modelos de regressão linear múltipla quando a equação (4) é expressa em notação matricial, já que essa notação permite uma apresentação mais compacta do modelo, dados e resultados para qualquer número de variáveis (GUJARATI, 2000; MONTGOMERY et al., 2001). Um modelo de regressão linear múltipla pode ser representado por:

$$y = x\beta + \varepsilon \quad (5)$$

onde:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

2.3.3 Teste de hipótese e intervalo de confiança na regressão linear

A análise de variância (ANOVA) é utilizada para o teste de significância da regressão, ou seja, para verificar a existência de relação linear entre a característica de interesse Y e um subconjunto de variáveis de controle (MONTGOMERY et al.; 2001; NETER et al., 2005). O teste de significância de regressão pode ser representado pelas seguintes hipóteses.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0 \quad (6)$$

$$H_1: \beta_j \neq 0 \text{ para pelo menos um } j$$

A rejeição da hipótese nula $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$, implica que pelo menos uma das variáveis de controle é estatisticamente significativa.

Na ANOVA a estatística de teste F_c apresenta uma distribuição F com k

graus de liberdade no numerador e $n - k$ graus de liberdade no denominador, sob a hipótese de H_0 verdadeira. Dessa forma, rejeita-se a hipótese nula $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$ se:

$$F_0 > F_{\alpha, k, n - k} \quad (7)$$

2.3.4 Aplicação da regressão linear: a estimação de valores

A aplicação mais comum dos modelos de regressão é na estimação da característica de interesse Y para um conjunto de valores das variáveis de controle do modelo. Esta estimativa pode ser utilizada de duas formas: (i) estimação da resposta média e (ii) predição de novas observações. A diferença entre essas duas observações recai no fato de que a estimativa da resposta média é calculada com base em um conjunto de valores das variáveis de controle que foi utilizado para estimar o modelo de regressão, enquanto a previsão de novos valores é estimada através de um conjunto de valores para as variáveis de controle.

2.3.5 Coeficiente de determinação

O coeficiente de determinação é uma medida simples da variabilidade da característica de interesse que pode ser explicada pelo modelo estimado contendo as variáveis de controle X_1, X_2, \dots, X_k (MONTGOMERY et al.; 2001; WEISBERG, 2005):

$$R^2 = \frac{\text{SQReg}}{\text{SQT}} = 1 - \frac{\text{SQR}}{\text{SQT}} \quad (8)$$

em que: $0 \leq R^2 \leq 1$, SQReg é a soma de quadrados da regressão, SQT é a soma de quadrados totais e SQR é a soma de quadrados dos resíduos.

Montgomery et al. (2001) relatam que a estatística R^2 deve ser utilizada com cuidado, pois é possível fazê-la maior simplesmente pela adição de variáveis de controle ao modelo, o que não indica que o modelo seja melhor. Um modelo com mais variáveis não é necessariamente melhor que um modelo com menos variáveis de controle.

2.4 Redes neurais artificiais (RNAs)

As RNAs compõem um dos ramos da Ciência da Computação denominado Inteligência Artificial, que procura desenvolver modelos computacionais inspirados na capacidade natural do cérebro humano de pensar; apresentam ótimos resultados na solução de problemas complexos envolvendo muitas variáveis, por isso, atualmente, as RNAs são aplicadas nas mais diversas áreas, como agronomia, biologia, medicina, negócios, meio ambiente, indústria, engenharia, geologia e física, entre outras.

Suas principais características estão relacionadas à capacidade de aprender por exemplos, de interpolar e/ou extrapolar com base em padrões fornecidos e de selecionar características específicas dentro de um universo amostral. Como já mencionado, as RNAs se baseiam no funcionamento do cérebro humano; simulam, também, sua composição, similar aos neurônios biológicos. São compostas por várias unidades de processamento (neurônios), conectadas por canais de comunicação (sinapses) aos quais são associados pesos que representam a importância de cada sinal transmitido.

As RNAs diferem dos modelos computacionais tradicionais por apresentarem propriedades e características particulares, tais como: adaptabilidade ou aprendizagem, capacidade de generalização, agrupamento, organização de dados, tolerância a falhas e possuem muitas características do

cérebro humano. Certos tipos de redes ensinam a si mesmas, através de exposição repetida de um conjunto de dados, reconhecem características comuns entre eles e agrupa-os ordenadamente.

Outros tipos de redes podem ser programadas para associar um conjunto de entrada com suas respectivas saídas. As RNAs podem generalizar através de exemplos imperfeitos (ruídos) e extrair informações essenciais das entradas contendo tanto dados relevantes como irrelevantes.

A forma pela qual os neurônios estão interconectados constitui a arquitetura da rede. Geralmente essa escolha se faz através do nível de complexidade do problema a ser analisado. Deve-se experimentar várias arquiteturas até se chegar a mais conveniente, pois não existe uma regra a ser seguida, as diversas arquiteturas têm que ser testadas e escolhida a que melhor se comporta para cada problema.

2.4.1 Histórico

O desenvolvimento das RNAs iniciou-se na década de 40, com os trabalhos de McCulloch; Pitts (1943) e Hebb (1949) que discutiam a capacidade de aprendizagem de informações das RNAs. O modelo matemático desenvolvido por McCulloch; Pitts em 1943 tratava-se de um simples neurônio que trabalhava em um processo de decisão lógica (verdadeiro ou falso). O neurônio, nesse modelo, era composto de uma simples unidade de processamento, ativado pela soma ponderada das entradas e das saídas computadas por uma função *threshold* bi-estável (0 ou 1).

Hebb (1949) apresentou uma teoria que explica o processo de aprendizagem que, segundo o autor, quando o axônio de uma célula A está próximo o suficiente para excitar uma célula B e repetida e insistentemente toma

parte na emissão do sinal elétrico da célula B, algum processo de crescimento ou mudança metabólica acontece em uma ou ambas as células tal que a eficiência de A, para fazer a célula B disparar, é aumentada.

Windrow; Hoff (1960) desenvolveram o ADALINE (Adaptive Linear Element) e o MADALINE (*Many ADALINE*) *Perceptron* como dispositivos práticos para resolver tarefas de reconhecimento de padrões. O ADALINE e MADALINE usam saídas analógicas ao invés das binárias originalmente propostas por McCulloch e Pitts. Minsky e Papert fizeram um estudo cuidadoso desses algoritmos e publicaram, em 1969, seu livro *Perceptrons*. Provaram formalmente que uma rede formada de uma única camada de neurônios, independente do algoritmo de aprendizagem, é capaz de resolver o problema de associação de padrões apenas quando os conjuntos são linearmente separáveis. Esses resultados foram devastadores no desenvolvimento das RNAs que ficaram na década de 70 e início da década de 80 em plano secundário.

A deficiência das redes *Perceptron* na resolução de problemas não-lineares foi suprida por Rumelhart et al. (1986). A solução encontrada foi a *Regra Delta Generalizada*, mais conhecida como algoritmo de aprendizagem *backpropagation*, para redes *Perceptron* de várias camadas de neurônios, com entradas e saídas analógicas, denominadas *Multilayer Perceptron*. A partir deste ponto houve um grande crescimento de pesquisas e aplicações das RNAs consolidando o uso das mesmas em diversas áreas.

2.4.2 Neurônio artificial

O Neurônio Artificial é uma unidade fundamental de processamento de uma RNA, o qual recebe uma ou mais entradas, transformando em saídas. Cada entrada tem um peso associado que determina sua intensidade e influencia no dado de saída. A Figura 1 ilustra o esquema do neurônio artificial.

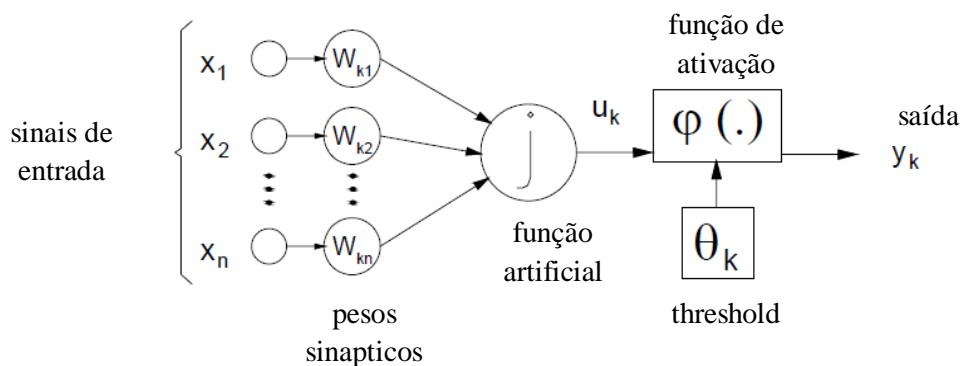


Figura 1 Modelo não linear de um neurônio artificial, segundo Haykin (1994)

Baseando-se na Figura 1 é possível distinguir alguns elementos considerados importantes na estrutura de um neurônio:

- Sinapses: caracterizadas por um peso (w), que pode representar sua intensidade. O papel do peso w_{kj} é multiplicar o sinal x_j na entrada da sinapse j , conectada a um neurônio k . O peso w_{kj} é positivo se a sinapse é excitatória e negativo se a sinapse associada é inibitória;
- Somatório: adiciona as entradas ponderadas pelos seus pesos respectivos, ou seja,

$$u_k = \sum_{j=1}^n w_{kj} x_j \quad (9)$$

- c) Limiar (*threshold*): θ_k , possui um papel determinante na saída do neurônio. Se o valor de u_k for menor que este limiar, então, a saída do neurônio fica inibida. Caso contrário, o neurônio fica ativo;
- d) Função de ativação: funciona como um limitante à amplitude da saída do neurônio, ou seja, a entrada é normalizada dentro de um intervalo fechado, geralmente $[0,1]$ ou $[-1,1]$;
- e) Saída do neurônio: y_k , sendo:

$$y_k = \varphi(u_k - \theta_k) \quad (10)$$

Onde φ é a função de ativação.

2.4.3 Função de ativação

Haykin (2001) afirma que a função de ativação define o valor da ativação de um neurônio conforme o nível de atividade da entrada. As funções de ativação típicas trabalham com adição, comparação ou transformações matemáticas. Algumas das funções de ativação comuns são:

- a) Função linear: $\varphi(v_i(t)) = v_i(t)$

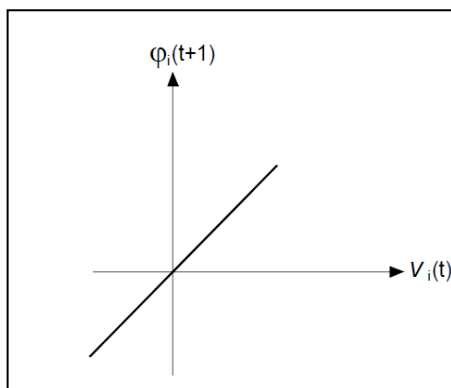


Figura 2 Função de ativação linear

b) Função threshold ou limiar: $\varphi_t(t+1) = \begin{cases} 1, & sev_t(t) \geq \theta \\ 0, & sev_t(t) < \theta \end{cases}$

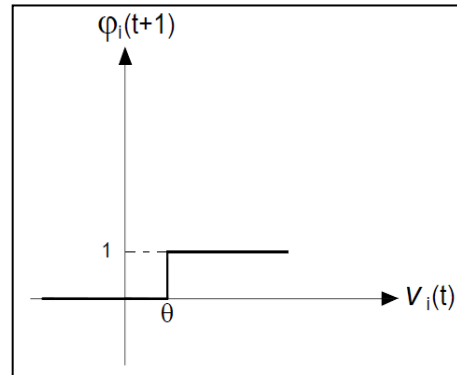


Figura 3 Função de ativação *threshold* ou limiar

c) Função *piecewise* linear: $\varphi_t(t+1) = \begin{cases} 1, & sev_t(t) > \theta \\ 0 \leq v_t(t) \leq \theta \\ -1, & sev_t(t) < -\theta \end{cases}$

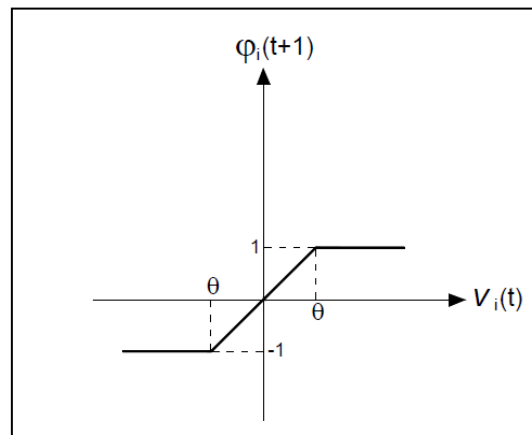


Figura 4 Função de ativação *piecewise* linear

d) Função sigmóide unipolar: $\varphi_i(t+1) = \frac{1}{(1 + e^{-v_i(t)})}$

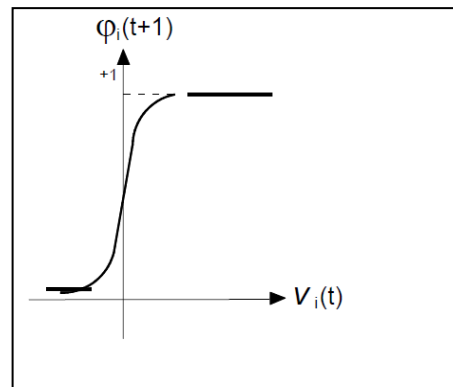


Figura 5 Função de ativação sigmóide unipolar

A função sigmóide é para muitos a forma mais comum de função de ativação usada na construção de RNAs. É definida como uma função de caracter estritamente crescente, que mostra propriedades homogêneas e assintóticas. Enquanto uma função limiar assume o valor de 0 ou 1, uma função sigmóide assume valores em uma faixa contínua entre 0 e 1. Além disso, a função sigmóide é diferenciável, enquanto que a função limiar não.

2.4.4 Redes neurais artificiais “*Multilayer Perceptron*”

As RNAs tipo “*Multilayer Perceptron*” (MLP) são as mais empregadas atualmente, por serem muito versáteis e capazes de resolver problemas desde os

mais simples até os mais complexos. Nas redes tipo MLP, camadas intermediárias são inseridas entre as camadas de entrada e de saída, em um número que pode ser ajustado em função da complexidade do problema e da precisão desejada. Porém, não existe uma arquitetura predefinida de acordo com a complexidade dos problemas. É necessário pesquisar a melhor alternativa, isto é, a melhor arquitetura da rede. A primeira camada de uma rede tipo MLP consiste de unidades de entrada, com as variáveis independentes (Figura 6). A última camada contém as unidades de saída, associadas às variáveis dependentes. Todas as outras unidades no modelo são denominadas unidades escondidas e constituem as camadas intermediárias (BRAGA, 2007).

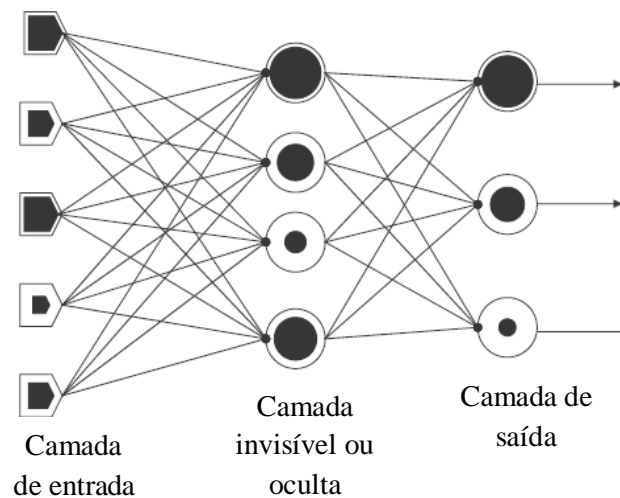


Figura 6 Ilustração de rede “*Multilayer Perceptron*” (MLP)

O algoritmo de treinamento utilizado em uma rede tipo MLP é o de *backpropagation*, que permite a minimização do erro por meio de ajuste dos pesos das sinapses. O desenvolvimento desse algoritmo é o principal

responsável pela recente retomada de interesse pelas redes neurais. Uma das razões para isso é que uma RNA tipo MLP, com um número suficiente de unidades escondidas, pode aproximar qualquer função contínua com qualquer grau de precisão, empregando função de ativação tipo sigmoideal. Esta função tem por finalidade modificar um estado de ativação (ativo ou inativo) de uma unidade em um sinal de saída, comparando sinais antigos e atuais (HAYKIN, 2001).

Escolhida a arquitetura de uma rede tipo MLP e dado o conjunto de entrada e de saída, procede-se ao seu treinamento, empregando o algoritmo *backpropagation*, para estimar as relações funcionais entre as entradas e as saídas. A rede neural pode ser empregada, então, para modelar ou prever a resposta correspondente a um novo padrão. Esse procedimento é semelhante a um problema de regressão, na estatística convencional, em que se deseja encontrar uma relação entre variáveis independentes (entradas) e dependentes (saídas) (HAYKIN, 2001).

Este conhecimento é adquirido através de um processo de aprendizado ou treinamento; durante o treinamento, um conjunto de exemplos (ensaios, por exemplo), com as respostas já conhecidas é apresentado à rede. Os dados de entrada são fornecidos ao sistema e é obtido o conjunto de respostas (saídas). As saídas da rede são comparadas com as respostas previamente conhecidas e, da diferença entre elas, é obtido o erro. Com o valor do erro, os pesos sinápticos são ajustados e o processo é repetido iterativamente, parando quando o erro atingir patamares satisfatórios (BRAGA, 2007).

O conjunto de treinamento deve possuir um número razoável de dados que englobe toda a gama de valores, para que a capacidade de generalização não seja comprometida. Após o treinamento da rede, é testada sua capacidade de generalização. Para tal, apresenta-se à rede dados que não fizeram parte do

conjunto de treinamento e analisam-se as respostas fornecidas por ela, verificando-se seu desempenho (BRAGA, 2007).

Assim, a modelagem através de redes neurais apresenta algumas vantagens, como: flexibilidade da rede (a rede pode ser adaptada para um novo ambiente através da troca de padrões de entrada → saída), facilidade de capturar novas informações do meio (dados de novos ensaios, por exemplo) e treinamento possível mesmo na presença de ruídos, o que é comum em dados experimentais (BRAGA, 2007).

2.4.5 Algoritmo de aprendizagem “*backpropagation*”

Segundo Bishop (2008) algoritmo de aprendizagem *backpropagation* do *Multilayer Perceptron* também é conhecido como Regra Delta Generalizada. Este algoritmo permite minimizar o erro, entre o valor real e o valor previsto pela rede, modificando os pesos das conexões entre os neurônios, por meio de uma retropropagação. O algoritmo *backpropagation* opera basicamente em duas fases, conforme mostra a Figura 7, cada uma percorrendo a rede em um sentido, a saber:

- a) Fase funcional (*forward*): ocorre quando os sinais de entrada são fornecidos à rede e esta produz uma saída. Esta saída é comparada com a saída desejada e o erro para cada neurônio da camada de saída é calculado;
- b) Fase erro (*backward*): se processa a partir da última camada, é quando cada neurônio procura ajustar seus pesos de maneira a reduzir seu erro. Nesta fase a rede procura ajustar seus pesos baseando-se no erro admissível.

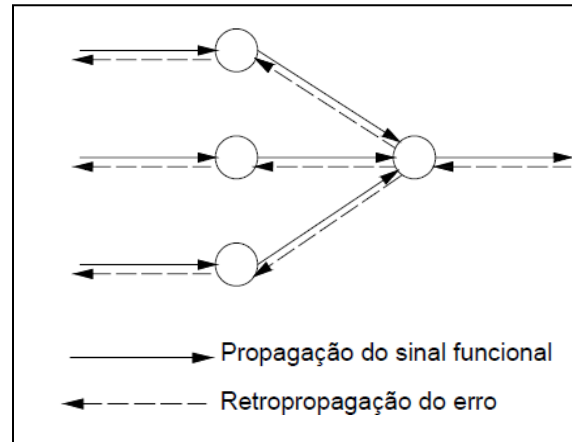


Figura 7 Ilustração das direções das fases de propagação do sinal funcional e retropropagação do erro

O treinamento usando esta regra consiste em fornecer à rede um conjunto de pares de entradas e saídas, onde a cada entrada do treinamento tem-se uma saída desejada. Este algoritmo é um método de gradiente descendente, que não garante chegar ao mínimo erro global, e que pode ser dividido em cinco passos segundo Bishop (2008).

- a) Passo 1: Apresente um padrão de entrada e a saída desejada. Utilizando uma determinada estratégia de apresentação coloque um dos padrões de entrada na camada de entrada e saída desejada nas suas respectivas camadas, então ative o passo *calcule saída*.
- b) Passo 2: Calcule saída. A partir da primeira camada, permita que cada camada produza os valores de saída até atingir a camada de saída e, então ative o passo *ajuste dos pesos da camada de saída*.
- c) Passo 3: Ajuste dos pesos da camada de saída. Para cada neurônio j da camada de saída, atualize todos os pesos w_{ij} conforme Equação 3.3, então ative o passo *ajuste dos pesos das camadas escondidas*.

$$\Delta w_{ij} = \eta \delta_j o_i \quad (11)$$

Onde:

η : taxa de aprendizado;

o_j : saída do neurônio;

\bar{o}_j : diferença (erro) entre a saída computada e a saída desejada do neurônio j , que pode ser calculada segundo a Equação 12:

$$\delta_j = i_j (d_j - o_j)(1 - o_j) \quad (12)$$

Onde:

d_j : saída desejada em n_j .

d) Passo 4: Ajuste de pesos das camadas escondidas - Para ajustar os pesos de todas as camadas escondidas, atualiza-se o peso w_{ki} de um neurônio n_i de uma camada escondida que está ligado a outro neurônio n_k na camada anterior conforme Equação 13.

$$\Delta w_{ki} = \eta \delta_i o_k \quad (13)$$

Onde:

η : taxa de aprendizagem

\bar{o}_i : erro relativo do neurônio n_i , dado pela Equação 14.

$$\delta_i = o_i (1 - o_i) \sum w_{ij} \delta_j \quad (14)$$

o_k : saída do n_k , que estimula n_i via w_{ki}

Depois que todas as conexões tenham sido ajustadas ative o passo *cheque magnitude do erro*.

e) Passo 5: Cheque magnitude do erro - Para se checar a magnitude do erro, pode-se adotar várias estratégias diferenciadas quanto à forma de aceitar o erro como desprezível. Uma regra muito comum é verificar se o erro global de saída da rede em relação à saída desejada é menor que um dado erro predefinido. Se a

condição adotada é satisfeita, então a rede aprendeu o conjunto de treinamento, caso contrário volte a ativar o passo apresente um padrão de entrada e a saída desejada.

f) Fim do algoritmo - Este algoritmo pode ser utilizado em uma arquitetura com qualquer número de camadas. O passo 4 deve ser ativado recursivamente até atingir a camada da entrada. Para se aumentar a velocidade de convergência é possível introduzir na Equação 12 um multiplicador chamado de *momentum*, onde este valor representa o efeito dos ajustes anteriores no ajuste atual.

2.5 Aplicação da modelagem matemática na agricultura

Atualmente as Redes Neurais Artificiais estão sendo aplicadas, com êxito, nas mais variadas áreas da engenharia. Os primeiros trabalhos científicos neste sentido começaram a se destacar no início da década de 1990, crescendo nos últimos anos. Estes estudos envolvem aplicações em classificação de solos, planejamento dos transportes, prospecção, estabilidade de taludes, fundações e estudos dos mais diversos tipos de previsões. Destacam-se a seguir alguns trabalhos nas mais variadas áreas da Engenharia Agrônômica.

Giordano et al. (2010) compararam o desempenho das Redes Neurais Artificiais e da Metodologia de Superfície de Resposta para otimização da recombinação da proteína não simbiótica hemoglobina 1 (OsHb1) em *Oryza sativa* produzida por *Escherichia coli*. Os resultados obtidos pela aplicação de RNAs foram mais confiáveis, já que os melhores parâmetros estatísticos de proteínas recombinantes ($3,50 \text{ g.L}^{-1}$) e um mínimo de concentração de biomassa ($18,48 \text{ g.L}^{-1}$) foram obtidos, demonstrando a possibilidade dessa nova técnica de detectar e quantificar a otimização da recombinação protéica da proteína OsHb1.

Ram et al. (2010) confrontaram a performance das RNAs com modelos de regressão linear para previsão da época ideal de colheita de azeitona (*Olea europaea*) com intuito de otimização do teor de óleo com base em características de qualidade. Os autores concluíram que as RNAs apresentaram um procedimento mais preciso do que a regressão linear, apresentando altas correlações (0,81 e 0,87) entre os fatores estudados para as duas variedades de oliveira, respectivamente.

Rahman; Bala (2010) utilizaram uma série de modelos RNA, com base em uma estrutura comum, para prever o desenvolvimento de diferentes órgãos (expresso em peso seco) de juta (*Corchorus capsularis*). Os modelos previram o crescimento dos diferentes componentes da planta de juta com precisão, sendo que os resultados preditos apresentaram boa concordância com os resultados observados para as partes da planta (folha, raiz, casca e caule) em matéria seca. Este estudo demonstrou que um modelo de RNA poderia ser uma ferramenta útil para prever o rendimento de fibra de juta e que a estrutura utilizada no trabalho pode ser usada para prever o rendimento de outras culturas, se devidamente treinados.

Higgins et al. (2010) propuseram um sistema neural artificial para prever a maturidade de ervilhas (*Pisum sativum*). Segundo os autores as RNAs são ferramentas promissoras para tarefas orientadas pelo homem e pelo computador. O uso de RNAs para a automatização do sistema de previsão de índice de maturidade, que atualmente é realizado pelo homem, torna o sistema apto de corrigir a variância de julgamento do inspetor, assim como acelera o processo completo para uma grande quantidade de classificações. O êxito desse sistema conduz a várias aplicações, como por exemplo, a possibilidade não só de colheita de ervilhas mais perto de seu índice de maturidade ideal, mas também o

plano de colheita e logística de transporte podendo ser realizado em mais tempo de antecedência.

Jiménez et al. (2009) utilizou as RNAs para determinar modelos de produção de amora-preta (*Rubus glaucus*) com base em dados fornecidos por produtores em um banco de dados meteorológicos. A abordagem de análise do trabalho centra-se em identificar as variáveis que explicam a maior parte da variação da produtividade por meio de redes neurais artificiais (perceptron de multicamadas). Os autores concluíram que as RNAs são uma eficaz ferramenta de gestão das informações fornecidas pelos agricultores para verificação de causas de variação de produtividade.

Alvarez (2009) desenvolveu uma rede neural para obter modelos adequados para a estimativa de rendimento e de previsão da produção regional de grãos de trigo (*Triticum aestivum*) com base em amostras de solo e dados climáticos. Neste trabalho foi feita uma comparação entre RNAs (Multilayer Perceptrons) e modelo de regressão simples (linear e quadrática), sendo que o autor verificou que a superfície de resposta do modelo de regressão desenvolvido representou 64% da variância rendimento espacial e inter anual, mas não realizou uma melhor previsão de rendimento em comparação à técnica de RNA, que por sua vez representou 76% da variabilidade do rendimento, e menor erro quadrado médio quando confrontado o rendimento previsto com o observado. Além do fato de que, as variáveis utilizadas para o desenvolvimento das RNAs podem estar disponíveis cerca de 40-60 dias antes da colheita do trigo, o que viabiliza a utilização para a previsão de produção de trigo.

Na busca de um modelo que indicasse respostas confiáveis frente a processos de decisão na agricultura, Ji et al. (2007) propuseram-se a: investigar se modelos de rede neural artificial (RNA) podem prever efetivamente a produtividade do arroz de Fujian na China para as condições climáticas típicas

da região montanhosa, avaliar o desempenho do modelo de RNA em relação às variações dos parâmetros de desenvolvimento e comparar a eficácia de vários modelos de regressão linear com modelos de RNA. Apesar de mais demorado para desenvolver do que modelos de regressão linear múltipla, modelo RNA mostrou-se superior para prever com precisão os rendimentos do arroz em Fujian, obtendo-se índices de aprendizagem da rede de 0,71 a 0,90.

Kaul et al. (2005) fizeram estudo de casos com RNAs para simular produção de milho e soja em condições climáticas típicas da região. Especialistas da área em gestão de nutrientes necessitam de técnicas de estimativa simples e precisa para relacionar a produção agrícola e utilização de nutrientes no processo de planejamento. Os objetivos deste estudo foram: investigar se uma rede neural artificial pode prever efetivamente a produção de milho e soja; comparar a capacidade de predição de modelos no estado, regional e local; avaliar o desempenho do modelo RNA relativo às variações dos parâmetros de desenvolvimento, e comparar a eficácia de vários modelos de regressão linear para os modelos RNA. Foi concluído que a rede apresentou bons resultados para os ensaios empregados tanto no treinamento como nos testes.

REFERÊNCIAS

AGRIANUAL 2007- **Anuário da Agricultura Brasileira**. São Paulo: FNP Consultoria e Comércio, 2007. 516p.

ALVAREZ, R. Predicting average regional yield and production of wheat in the Argentine Pampas by an artificial neural network approach. **European Journal Agronomy**, New York, v. 30, n. 2, p. 70–77, Jan. 2009.

BISHOP, C. M. **Neural networks for pattern recognition**. 2nd. ed. Oxford: University Press, 2008.

BOOGES, W. G.; RITCHIE, J. T. Economic and risk analysis irrigation decisions in humid regions. **Journal Production Agricultural**, v. 2, p. 116-122, 1998.

BOONRUANGROD, R.; FLUCH, S.; BURG, K. Elucidation of origin of the present day hybrid banana cultivars using the 5ETS rDNA sequence information. **MolecularBreeding**, Dordrecht, v. 24, n. 1, p. 24-77, 2009.

BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. 2. ed. Rio de Janeiro: LTC, 2007.

CARDOSO, C.; FARIA, R. T.; FOLEGATTI, M. V. Aplicação do modelo CERES-maize na análise de estratégias de irrigação para milho "safrinha" em Londrina-PR. **Engenharia Agrícola**, Jaboticabal, v. 24 n. 1, jan./abr. 2004.

COSTA, L. C. Modelagem e simulação em agrometeorologia. In: CONGRESSO BRASILEIRO EM AGROMETEOROLOGIA, 10., Piracicaba, 1997. **Anais...** Piracicaba: SBA, 1997. p. 3-6.

DADZIE, B. K. **Post-harvest characteristics of black Sigatoka resistant banana, cooking banana and plantain hybrids**. Montpellier: Inibap, 1998. 75 p. (Inibap Technical Guidelines, 4).

De LANGHE, E. et al. Why bananas matter: an introduction to the history of banana domestication. **Ethnobotany Research and Applications**, Montpellier, v. 7, n. 1, p. 165-177, 2009.

DANIELS, J. Que variedade de banana devo cultivar? **Infomusa**, Montepelier, v. 9, n. 1, p. 31-33, jun. 2000.

DONATO, S. L. R. et al. Comportamento de variedades e híbridos de bananeira (*Musa* spp.), em dois ciclos de produção no sudoeste da Bahia. **Revista Brasileira Fruticultura**, Jaboticabal, v. 28, n. 1, p. 139-144, abr. 2006.

FOOD AND AGRICULTURAL ORGANIZATION. Disponível em: <<http://faostat.fao.org/site/567/default.aspx#ancor>>. Acesso em: 13 ago. 2010.

GIORDANO, P. C. et al. Application of response surface methodology and artificial neural networks for optimization of recombinant *Oryza sativa* non-symbiotic hemoglobin 1 production by *Escherichia coli* in medium containing byproduct glycerol. **Bioresource Technology**, New York, v. 101, p. 7537–7544, Oct. 2010.

GUJARATI, D. N. **Econometria básica**. 3. ed. São Paulo: Makron Books, 2000. 846 p.

HAEFNER, J. W. **Modeling biological systems: principles and applications**. New York: Springer, 2005. 475 p.

HAYKIN, S. **Redes neurais: princípios e prática**. 2 ed. Porto Alegre: Bookman, 2001.

HEBB, D. **The organization of behavior**. New York: J. Wiley, 1949.

HIGGINS, A. et al. Forecasting maturity of green peas: an application of neural networks. **Computers and Electronics in Agriculture**, New York, v.70, n. 1, p. 151–156, Jan. 2010.

HOOGENBOOM, G. Contribution of agrometeorology to the simulation of crop production and its applications. **Agricultural and Forest Meteorology**, New York, v. 103, n. 1-2, p. 137-157, Jun. 2000.

JAME, Y. W.; CURTFORTH, H. W. Crop Growth models for decision support systems. **Canadian Journal of Plant Science**, Ottawa, v. 76, n. 1, p. 9-19, Jan. 1996

JARAMILLO, R. C. **Las principales características morfológicas del fruto de banano, variedad cavendish gigante (Musa AAA) em Costa Rica**. Upeb: Impretex, 1982. 42 p.

J I, B.; SUN, Y.; YANG, S.; WAN, J. Artificial neural networks for rice yield prediction in mountainous regions. **Journal of Agricultural Science**, London, v. 145, n. 2, p.249–261, Jul. 2007.

JIMÉNEZ, J. et al. Analysis of Andean blackberry (*Rubus glaucus*) production models obtained by means of artificial neural networks exploiting information collected by small-scale growers in Colombia and publicly available meteorological data. **Computers and Electronics in Agriculture**, New York, v.69, n. 2, p.198–208, Dec. 2009.

KAUL, M.; HILL, R.L.; WALTHALL, C. Artificial neural networks for corn and soybean yield prediction. **Agricultural System**, New York, v. 85, n. 2, p. 1–18, July 2005.

LE BAIL, M. et al. Is it possible to forecast the grain quality and yield of different varieties of winter wheat from Minolta SPAD meter measurements? **European Journal Agronomy**, v. 23, n. 4, p. 379–391, Dec. 2005.

McCULLOCH, W.; PITTS, W. A Logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, Chicago, 1994, 520p.

MINKY, M.L.; PAPERT, S.A. **Perceptrons**: an introduction to computational geometry. Cambridge: MIT Press, 1969. 258 p.

MONTEITH, J. The quest balance in crop modeling. **Agronomy Journal**, Madison, v. 88, n. 5, p. 695-697, 1996.

MONTGOMERY; D. C. **Design and analysis os experiments**. 6th New York: J. Wiley, 2005. 643p.

MONTGOMERY; D. C.; MASTRANGELO, C. M. Some statistical process control methods for autocorrelated data. **Journal of Quality Technology**, Milwaukee, v. 23, n. 3, p. 179-193, July 1991.

MONTGOMERY; D. C.; RUNGER, G. C. **Estatística aplicada e probabilidade para engenheiros**. 2. ed. Rio de Janeiro: LTC, 2003. 463 p.

MONTGOMERY; D. C.; VINING, G. C.; PECK, E. A. **Introduction to linear regression analysis**. 3rd ed. New York: J. Wiley, 2001. 641 p.

NETER, J. et al. **Applied linear statistical models**. 5th ed. New York: Mc Graw-Hill/Irwin, 2005. 1396 p

ORTIZ, R. Morphological variation in musa germplasm. **Genetic Resources and Crop Evolution**, New York, v. 44, n. 5, p. 393-404, Oct. 1997.

RAHMAN, M. M.; BALA, B. K. Modelling of jute production using artificial neural networks. **Biosystems Engineering**, v. 105, n. 3, p. 350–356, Mar. 2010.

RUMELHART, D.; HILTON, G.; WILLIAMS, R. Parallel distributed processing: explorations in the microstructure of cognition. In: _____. **Learning interval representation by error propagation**. Cambridge: MIT Press, 1986. v. 1, p. 218-362.

SHEPHERD, K. Banana: taxonomia e morfologia. In: SIMPOSIO BRASILEIRO SOBRE BANANICULTURA, 1., 1984, Jaboticabal. **Anais...** Jaboticabal: FCAV/UNESP, 1984. p. 50-74.

SILVA, S. de O. et al. Bananeira. In: BRUCKNER C. H. (Ed.). **Melhoramento de fruteiras tropicais**. Viçosa, MG: UFV, 2002. p. 101-157.

SILVA, S. de O. et al. Caracterizacao morfologica e avaliacao de cultivares e hibridos de bananeira. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 22, n. 2, p. 161-169, ago. 2000.

SIMMONDS, N. W. **Los plátanos**. Barcelona: Blume, 1973. 539 p.

SOTO BALLESTERO, M. **Bananos**: cultivo e comercializacion. 2. ed. San Jose: LIL, 1992. 674 p.

VALMAYOR, R. V. Classification and characterization of *Musa exotica*, *M. alinsanaya* and *M. acuminata* ssp. *Errans*. **Infomusa**, Montpellier, v.10, n.2, p.35-39, 2001.

WIT, C. T. Simulation of living systems. In: PENNING DE VRIES, F. W. T.; VAN LAAR, H. H. (Ed.). **Simulation of plant growth and crop production**. Wageningen: Pudoc, 1982. p. 3-8.

WEISBERG, S. **Applied linear regression**. 3rd ed. New York: J. Wiley, 2005. 310 p.

WINDROW, B.; HOFF, M. E. **Adaptative switching circuits**. San Francisco: Institute of Radio Engineers, Westers Eletronic Show and Convection, 1960. p. 96-104.

WU, B. et al. Harvest evaluation model and system of fast-growing and high-yield poplar plantation. **Mathematical and Computer Modelling**, New York, v. 51, n. 11-12, p. 1444-1452, Jun. 2010.

CAPÍTULO 2

PREVISÃO DO PESO DO CACHO DA BANANEIRA 'TROPICAL' COM BASE EM REGRESSÃO LINEAR MÚLTIPLA

RESUMO

O conhecimento das relações entre os caracteres é de suma importância à bananicultura, possibilitando, ao produtor, estimar a produção do cacho de uma determinada planta a partir de outros atributos. O presente trabalho foi desenvolvido com o objetivo de avaliar o efeito de caracteres mensurados no florescimento e na colheita em relação ao peso do cacho. O experimento constituiu-se de um ensaio de uniformidade, conduzido em Guanambi, BA, com a cultivar Tropical (YB42-21), híbrido tetraplóide AAAB, plantado no espaçamento de 3 m x 2 m, formado de 11 fileiras de 52 plantas cada e consideradas como útil as 9 fileiras centrais com 40 plantas por fileira, em um total de 360 plantas e área de 2.160 m². Avaliaram-se os caracteres vegetativos, altura da planta, perímetro do pseudocaule, número de filhos emitidos e número de folhas vivas no florescimento e na colheita e os caracteres de rendimento, peso do cacho, número de pencas e frutos, peso da segunda penca, comprimento e diâmetro do fruto em dois ciclos de produção. Nas avaliações, cada planta foi considerada como uma unidade básica (ub), área de 6 m², perfazendo assim, 360 unidades básicas (ub). Pela metodologia de regressão linear múltipla foi estimado o peso do cacho, onde as variáveis que mais influenciaram nessa mensuração foram: número de folhas vivas na colheita, número de frutos, peso médio do fruto, comprimento do engaço, peso da ráquis e comprimento do engaço com a seguinte equação de previsão: $PCA = -5,249 + 0,11NFC + 0,066NFR + 0,046PMF + 0,183CEF + 2,039PRA - 0,011CEN$.

Palavras-chave: Produção. Previsão. Banana. Modelo de regressão.

ABSTRACT

The knowledge of the relationships between the characters is of paramount importance to banana plantations, enabling the producer to estimate the output of the cluster of a given plant from other attributes. The present work was to evaluate the effect of traits measured at flowering and harvest in relation to the weight of the bunch. The experiment consisted of a uniformity trial, conducted in Guanambi, BA, with cultivar Tropical (YB42-21), AAAB tetraploid hybrid, planted at a spacing of 3 m x 2 m, consisting of 11 rows of 52 plants each and considered how useful the nine central rows with 40 plants per row, totaling 360 plants and an area of 2,160 m². We assessed the vegetative characters, plant height, pseudostem circumference, number of children issued and number of green leaves at flowering and harvest, and the characters of yield, bunch weight, number of hands and fruits, weight of the second hand, length and diameter of the fruit in two growing seasons. In the evaluation, each plant was considered as a basic unit (bu), area of 6 m² totaling well, 360 basic units (bu). The methodology of multiple linear regression was estimated weight of the bunch, where the most significant variables that were measured: number of leaves at harvest, fruit number, average fruit weight, length of rachis, rachis weight and length of stem with the following prediction equation: $PCA = -5,249 + 0,11NFC + 0,066NFR + 0,046PMF + 0,183CEF + 2,039PRA - 0,011CEN$.

Keywords: Production. Prediction. Banana. Regression model.

1 INTRODUÇÃO

A banana é uma das frutas tropicais mais importantes e constitui alimento básico tanto para a população das áreas rurais quanto urbanas. O seu cultivo é bastante expressivo nos sistemas agrícolas das zonas agroecológicas dos trópicos, sendo considerada um importante alimento em razão da sua composição química e conteúdo em vitaminas e minerais, principalmente potássio, tendo destaque como a frutífera mais consumida, tanto pela sua versatilidade em termos de modalidades de consumo (processada, frita, cozida, *in natura*) quanto pelas suas características de sabor, aroma, higiene e facilidade de ser consumida *in natura* (SILVA et al., 2002).

Com o intuito de se obter melhor conhecimento das respostas da cultura ao ambiente, modelos de simulação são utilizados como ferramenta de grande potencial na área de sistemas cultivados, permitindo o estudo e o entendimento do conjunto, estimando o desempenho da cultura em diferentes áreas e situações, tomando como base o comportamento dos mais diversos tipos de caracteres em relação ao produto de interesse (SILVA et al., 2002).

Assim, identificar, descrever e prever as relações entre eventos que envolvem o desenvolvimento da bananeira cultivada é de fundamental importância. Uma das formas é encontrar modelos que relacionem variáveis que descrevam a realidade, avaliando a possível relação entre uma variável dependente com uma ou mais variáveis independentes. Isto pode ser conseguido por meio da utilização de modelos de regressão lineares.

Apesar de diversas pesquisas estarem sendo realizadas com a cultura da bananeira em grande parte do Brasil nos últimos anos, falta informação técnica para entender o impacto dos caracteres vegetativos e de rendimento na produtividade final (RAHMAN; BALA, 2010). Nesse sentido, os modelos de

simulação são uma ferramenta bastante útil, podendo prever a variabilidade do rendimento de acordo com as variáveis estudadas (GUNGULA et al., 2003).

Prever o desenvolvimento de uma determinada cultura com base em caracteres de rendimento é amplamente utilizado (JAME; CUTFORTH, 1996). A modelagem matemática é usada para estimar a duração de estádios de desenvolvimento vegetal, para escolher a época de plantio, determinar prováveis datas de colheita, para prever uma produção irregular e assim utilizar esses dados em programas de melhoramento genético (ROBERTO et al., 2005; STENZEL et al., 2006; BÍSCARO, 2007).

O conceito de regressão múltipla tem sido largamente difundido aplicado à várias espécies vegetais, como trigo (LE BAIL et al., 2005), café (WYZYKOWSKI, 2009), cana-de-açúcar (SCARPARI; BEAUCLAIR, 2004; 2009), milho (SOLER et al., 2007), morango (ANTUNES et al., 2006), laranja (STENZEL et al., 2006) e arroz (STRECK et al., 2007) com o objetivo de estimar a colheita e a produção. Em bananeira alguns estudos têm focado em caracterizar e avaliar o comportamento de genótipos (variedades e híbridos) mediante o uso de descritores fenotípicos relevantes para a identificação e seleção de indivíduos superiores, porém nenhum trabalho de modelagem matemática de peso de cacho foi realizado.

Dessa forma, objetivou-se mensurar e identificar quais variáveis tem maior influência no peso do cacho e determinar um modelo matemático para previsão de colheita para bananeira cv. Tropical.

2 MATERIAL E MÉTODOS

2.1 Implantação e condução da cultura

O experimento foi implantado num LATOSSOLO VERMELHO-AMARELO distrófico típico A fraco textura média fase caatinga hipoxerofila, relevo plano a suave ondulado na área experimental da Escola Agrotécnica Federal Antônio José Teixeira, localizada no Distrito de Ceraima, Município de Guanambi, Micro Região da Serra Geral, Sudoeste da Bahia, distando 108 km da margem direita do Rio São Francisco e cerca de 90 Km da divisa Minas-Bahia, com latitude de 14°13'30'' sul, longitude de 42°46'53'' oeste de Greenwich, altitude de 525 m, com a média anual de precipitação de 663,69 mm e temperatura média de 26°C.

Utilizou-se, para instalação do experimento, mudas micropropagadas, cedidas pela Embrapa Mandioca e Fruticultura, multiplicadas pela Campo Biotecnologia na Biofábrica em Cruz das Almas, BA, aclimatizadas em tubetes, e transportadas para o local do experimento enroladas em embalagem tipo rocambole, onde foram transplantadas diretamente para o campo.

O experimento constituiu-se de um ensaio de uniformidade com a cultivar Tropical (YB42-21), híbrido tetraplóide derivado da Yangambi número 02, grupo genômico AAAB, resistente a Sigatoka-amarela e tolerante ao mal-do-Panamá, com frutos tipo Maçã, gerado e selecionado pela Embrapa Mandioca e Fruticultura. A área é formada por 11 fileiras de 52 plantas cada, no espaçamento de 3 m x 2 m, perfazendo um total de 572 plantas e 3.432 m², e considerada como área útil as 9 fileiras centrais com 40 plantas cada, num total de 360 plantas e 2.160 m². Nas avaliações procedidas em dois ciclos de produção, considerou-se cada planta como uma unidade básica (ub), área de 6

m², constituindo assim, 360 unidades básicas (ub). O sistema de irrigação utilizado foi aspersão convencional fixo com aspersores de sub-copa. O manejo de irrigação foi calculado pelo método do turno de rega pré-fixado, baseado, nas médias de evaporação de tanque Classe A medidas no Posto Meteorológico da Codevasf, nos coeficientes de cultivos da cultura (COELHO et al., 2001), nas características físico-hídricas do solo e nas características do sistema de irrigação.

A implantação e os tratos culturais e fitossanitários adotados na cultura basearam-se em recomendações técnicas (ALVES, 1997; MOREIRA, 1999;) e as adubações em análises de solo e de folha (MOREIRA, 1999).

2.2 Avaliações

As avaliações foram realizadas na fase de florescimento e de colheita dos cachos nos ciclos de produção da planta-mãe e do filho. Mensuraram-se descritores fenotípicos vegetativos e de rendimento, seguindo a metodologia proposta por (SILVA et al., 1999), apresentada sequencialmente.

2.3 Caracteres vegetativos

a) Altura da planta (APL)

Foi avaliada na época do florescimento com uma trena de cinco metros, medindo-se a distância em centímetros, da base do pseudocaule até a roseta foliar, na altura da inserção do engaço no pseudocaule. Roseta foliar e a região delimitada no pseudocaule entre a folha mais velha e a mais nova, considerando a posição da inserção do pecíolo no pseudocaule (limite pecíolo-bainha).

b) Perímetro do pseudocaule (PPS)

A medida foi feita na época do florescimento com uma fita métrica, medindo-se a circunferência do pseudocaule em centímetros a uma altura de 30 cm do solo.

c) Número de folhas vivas no florescimento (NFF)

O número de folhas vivas presentes nas plantas na época do florescimento foi contado, considerando como viva ou funcional a folha que possuía mais de 50% do limbo verde, ainda que rasgado.

d) Número de filhos emitidos até o florescimento (NFI)

Contou-se o número de filhos emitidos pela planta até o florescimento.

e) Número de folhas vivas na colheita (NFC)

Foi contado o número de folhas vivas presentes nas plantas na época da colheita do cacho.

2.4 Caracteres de rendimento

Os caracteres de rendimento foram avaliados na época da colheita dos cachos. O ponto de colheita baseou-se no diâmetro ou calibre do fruto da fileira externa de frutos da segunda penca, considerando a particularidade do híbrido (tipo de fruto).

a) Peso do cacho (PCA)

A massa do cacho (massa das pencas, engaco e ráquis), em quilogramas, foi avaliada em balança de plataforma com capacidade para 200 kg.

b) Número total de pencas por cacho (NPC)

Anotou-se o número total de pencas por cacho.

c) Número de frutos por cacho (NFR)

Foram contados e anotado o número de frutos do cacho.

d) Peso da segunda penca (PPS)

A segunda penca de cada cacho, considerada de referência, foi pesada em balança digital (precisão de três casas decimais), e o valor expresso em quilogramas.

e) Peso do fruto (PMF)

A massa do fruto ou dedo central da fileira externa de frutos da segunda penca (penca de referência) foi obtida individualmente de cada cacho, em balança digital com precisão de três casas decimais, expresso em gramas.

f) Comprimento do fruto (CEF)

Tomou-se a medida da curvatura externa do fruto ou dedo central da fileira externa de frutos da segunda penca, em centímetros, utilizando fita métrica, da base ao ápice (desconsiderando o pedicelo e o ápice do fruto).

g) Diâmetro do fruto (DFR)

Mediu-se em milímetros, na parte mediana, no sentido do comprimento do fruto central da fileira externa de frutos da segunda penca, utilizando-se paquímetro, posicionando-o nas laterais do fruto. Este diâmetro ou calibração lateral do fruto foi utilizado como critério para a colheita e classificação de frutos.

h) Peso da ráquis (PRA)

A massa da ráquis foi obtida individualmente de cada cacho, em balança digital com precisão de três casas decimais, expresso em gramas.

i) Comprimento do engaço (CEN)

Tomou-se a medida da curvatura do engaço desde a inserção do mesmo no pseudocaule até a inflorescência masculina.

j) Diâmetro do engaço (DEN)

Mediu-se em milímetros, na parte mediana do cacho utilizando-se paquímetro.

2.5 Análise estatística

Para se avaliar a importância das variáveis referentes a caracteres vegetativos e de rendimento e influência desses no peso do cacho, estimou-se uma equação de regressão utilizando-se o procedimento de seleção das variáveis denominado *stepwise* (DRAPER; SMITH, 1981), no software R (R DEVELOPMENT CORE TEAM, 2010). O modelo de regressão linear múltipla (RLM) ajustado pode ser dado por exemplo:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i \quad (1)$$

em que:

Y_i refere-se ao valor do peso do cacho, determinado em função das variáveis de carácter vegetativo e de rendimento ($X_1; \dots; X_k$), , sendo respectivamente, APL, PPS, NFF, NFI, NFC e NPC, NFR, PSP, PMF, CEF, DFR, PRA, CEN, DEN, ε_i é o erro associado à i -ésima observação, assumido normal e independentemente

distribuído, β_0 constante inerente ao modelo e β_1, \dots, β_k coeficientes do modelo. O coeficiente de determinação do modelo foi determinado pela fórmula:

$$R^2 = \frac{SQReg}{SQT} = 1 - \frac{SOR}{SQT} \quad (2)$$

onde: $0 \leq R^2 \leq 1$

Para elucidar o relacionamento de cada variável, que foi significativa na análise de regressão com o peso final do cacho, foram estimadas as correlações de Pearson (DRAPER; SMITH, 1981).

3 RESULTADOS E DISCUSSÃO

Foram realizados os procedimentos estatísticos para estimar a equação de predição dos valores de peso do cacho para bananeira cv. Tropical (onde foram selecionadas apenas as variáveis significativas) levando-se em consideração o coeficiente de determinação (R^2 equação 2) para as equações de predições ajustadas, em cada caso em particular, e a análise de correlação entre as variáveis.

A equação de predição obtida pela regressão linear múltipla, que teve o melhor ajuste para o PCA com variáveis mensuradas no florescimento, apresentou coeficiente de determinação (R^2) de 13%. Na Tabela 1, está apresentada a constante e as estimativas dos coeficientes das variáveis formando a equação que utiliza a composição em APL, PPS, NFF e NFI.

Tabela 1 Componentes da equação de predição de peso de cacho de bananeira cv. Tropical em função de variáveis mensuradas do florescimento

Constante	Coeficientes				R^2	CV(%)
	APL ¹	PPS ¹	NFF ¹	NFI ¹		
+3,783	-0,009	+0,131	0,231	---	0,13	16

¹APL= altura da planta (cm); PPS= perímetro do pseudocaule (cm); NFF= número de folhas vivas no florescimento; NFI= número de filhos emitidos até o florescimento; ---= variáveis não significativas pelo procedimento stepwise.

O ajuste desta equação, com R^2 de 13%, pode ter sido em função da pouca influência que cada variável exerce sobre o peso final ou devido ao reduzido número de variáveis que compõem este modelo (caracteres

vegetativos), o que denota que uma série de outros fatores não considerados nesse trabalho possa estar influenciando o peso do cacho. O baixo valor de R^2 ocorreu apesar do coeficiente de variação ter sido relativamente baixo (CV=16%). Donato (2006a) observou que os caracteres mensurados na época do florescimento, altura da planta e perímetro do pseudocaule, apresentaram significância em relação à predição do peso do cacho dos genótipos Prata anã e Pacovan (AAB); *Cavendish*: Grande naine e Nanicão (AAA); os híbridos tipo Prata: PA42-44 (AAAB), Preciosa, Japira, Pacovan-Ken e ST12-31 (AAAB), gerados e selecionados pela *Embrapa Mandioca e Fruticultura*; o tipo *Gros Michel*: Ambrosia, Calipso e Bucaneiro (AAAA), e o híbrido tipo *Cavendish* FHIA 02 (AAAA). Porém neste estudo não foi estabelecido a equação satisfatória de previsão do peso e somente a relação entre os caracteres.

No caso das variáveis mensuradas na colheita, a melhor equação ajustada para determinar o peso do cacho na colheita foi: $PCA = -5,249 + 0,11NFC + 0,066NFR + 0,046PMF + 0,183CEF + 2,039PRA - 0,011CEN$, com R^2 de 71%, sendo que as demais variáveis não foram significativas (Tabela 2). O fato do coeficiente de determinação deste modelo ser maior é um indício de que as variáveis mensuradas em fase de produção podem representar com maior precisão o peso do cacho, aliado a ocorrência de coeficiente de variação baixo (9%). Pinto (2002) ao estudar a previsão da incidência de ferrugem em cafeeiro também verificaram no modelo ajustado pela análise de regressão um baixo coeficiente de determinação (0,55 a 0,60) apesar das variáveis serem significativas.

Tabela 2 Componentes da equação de predição de peso de cacho de bananeira cv. Tropical em função de variáveis mensuradas na colheita

Constante	Coeficientes											R ²	CV(%)
	NFC ¹	NPC ¹	NFR ¹	PPS ¹	PMF ¹	CEF ¹	DFR ¹	PRA ¹	CEN ¹	DEN ¹			
-5,249	+0,11	---	+0,066	---	+0,04	+0,183	---	+2,039	-0,011	---	0,71	9	

¹NFC= número de folhas vivas na colheita; NPC= número total de pencas por cacho; NFR= número de frutos por cacho; PPS= número de pseudo caule (cm); PMF= peso médio do fruto (g); CEF= comprimento do fruto (cm); DFR= diâmetro do fruto (cm); PRA= peso da ráquis (g); CEN= comprimento do engão (cm); DEN= diâmetro do engão; ---= variáveis não significativas pelo procedimento stepwise.

Os descritores peso do cacho e das pencas, embora expressem diretamente a produtividade, não podem ser considerados isoladamente na escolha de uma cultivar, pois outros caracteres relacionados aos frutos, como peso, comprimento, diâmetro e sabor devem ser considerados neste processo (SILVA et al., 2002). Assim o estudo de previsão associado ao de correlação podem discriminar quais variáveis influenciam na produção ao determinar a magnitude e a significância das associações entre descritores fenotípicos.

As correlações entre o caracter peso do cacho e os caracteres observados na época do florescimento, altura da planta, perímetro do pseudocaule, número de folhas vivas no florescimento e número de filhos emitidos até o florescimento, bem como as correlações entre si, encontram-se na Tabela 3. Foi realizada a análise de correlação apenas nas variáveis significativas pela RLM para predição do peso do cacho.

Tabela 3 Coeficientes de correlação entre variáveis mensuradas na época de florescimento e peso do cacho

	PCA ¹	APL ¹	PPS ¹	NFF ¹
PCA	1,00			
APL	0,857	1,00		
PPS	0,025	0,153	1,00	
NFF	-0,317	-0,178	0,118	1,00

¹PCA= peso do cacho (Kg); APL= altura da planta (cm); PPS= perímetro do pseudocaule (cm); NFF= número de folhas vivas no florescimento.

A associação entre peso do cacho e altura da planta foi significativa e positiva, o que significa que o peso variou de forma direta com o porte da planta. A altura da planta é um descritor importante do ponto de vista fitotécnico e de

melhoramento, pois influi nos aspectos de densidade de plantio e manejo da cultura, interferindo na produção (DONATO et al., 2006b). De forma geral, em bananeira, independente do grupo genômico, os genótipos de porte mais baixo são mais produtivos que os de porte mais alto (DONATO et al., 2003). A cultivar Tropical é considerada de porte médio (3,0 a 4,5m) e de acordo com a análise de correlação existe relação direta do crescimento em altura com produção, mesmo esta não sendo de porte baixo.

A estimativa entre peso do cacho e perímetro do pseudocaule foi não significativa e positiva, o que pode indicar que independente da espessura do pseudocaule o peso do cacho pode ser de valor satisfatório. Este resultado concorda com Donato (2006 a) que obteve a mesma conclusão para todos os genótipos testados, tanto no primeiro quanto no segundo ciclo.

A associação entre peso do cacho e número de folhas vivas no florescimento mostrou-se significativa e negativa o que demonstra que o peso variou de forma inversa ao número de folhas vivas. Como essa correlação apesar de significativa foi baixa, pode ser que o grau de relacionamento entre essas variáveis não seja forte o suficiente para refletir no peso, uma vez que as folhas funcionam como fonte de esqueletos de carbono para formação dos frutos, e conseqüentemente quanto maior a área foliar maior a área para realização da fotossíntese. Relação significativa positiva foi encontrada por Donato (2006a) entre PCA e NFF para PV42-68 e Grande naine (segundo ciclo), PV42-142, PV42-85, Calipso e PA42-44 (primeiro ciclo), sendo que os demais genótipos mostraram-se não significativos.

As correlações entre o caracter peso do cacho e os caracteres observados na época da colheita, comprimento do fruto, peso médio do fruto, número de frutos, comprimento do engaço, peso da ráquis, número folhas vivas na colheita, que foram significativas na análise de regressão encontram-se na Tabela 4.

Tabela 4 Coeficientes de correlação entre variáveis mensuradas na época de colheita e peso do cacho

	PCA ¹	CEF ¹	PMF ¹	NFR ¹	CEN ¹	PRA ¹	NFC ¹
PCA	1,00						
CEF	0,340	1,00					
PMF	0,626	0,250	1,00				
NFR	0,471	-0,255	-0,58	1,00			
CEN	0,060	0,242	0,265	-0,249	1,00		
PRA	0,576	-0,031	-0,216	0,585	0,201	1,00	
NFC	0,199	0,112	0,123	-0,017	0,086	0,124	1,00

¹PCA= peso do cacho; CEF= comprimento do fruto (cm); PMF= peso médio do fruto (g); NFR= número de frutos por cacho; CEN= comprimento do engaço (cm); PRA= peso da ráquis (g); NFC= número de folhas vivas na colheita.

O peso do cacho em relação ao comprimento externo do fruto apresentou uma correlação significativa e positiva, ou seja, o peso variou de forma direta em relação ao comprimento do fruto (0,340). A análise do comprimento externo com o interno dá idéia da curvatura do fruto, porém como neste estudo foi realizado apenas o CEF, o relacionamento deste com o peso pode ser tanto pelo comprimento quanto pelo formato do fruto. A prevalência de valores positivos e significativos indica a existência de uma associação genética entre peso do cacho e comprimento do fruto (DONATO et al., 2006 a; JARAMILO et al., 1982).

A correlação entre peso do cacho e dos frutos foi estatisticamente significativa, positiva e com maior valor em relação às outras variáveis (0,626), como já era previsto, pois as pencas são o próprio cacho, apenas sem a ráquis.

Resultados semelhantes foram encontrados para os genótipos testado por Donato (2006a). A tendência geral encontrada para esta correlação assemelha-se a obtida por Lima Neto et al. (2003). Assim, pode-se inferir que os caracteres peso do cacho e peso médio dos frutos estão diretamente correlacionados.

Analogamente à associação entre PCA e PMF foram encontrados valores significativos e positivos (0,471) para a correlação entre peso do cacho e número de frutos. Estimativas negativas foram encontradas por Donato (2006b) para os genótipos Calipso e PV 42-85, porém entre o peso médio dos frutos foi de correlação positiva o que indica que o cacho de maior peso pode ter menor número de frutos, todavia são de maior tamanho unitário. Lima Neto et al. (2003), encontraram para esta correlação, estimativas positivas e significativas para quase todos os genótipos estudados.

A associação entre peso do cacho e comprimento do engaço foi não significativa e positiva (0,060), o que indica que o peso do cacho independe do comprimento do engaço. Este resultado concorda com Donato (2006) que obteve o mesmo resultado para todos os genótipos testados, tanto no primeiro quanto no segundo ciclo.

O peso da ráquis apresentou correlação significativa, positiva e relativamente alta (0,576), o que denota um padrão geral de variação direta do peso do cacho com o peso da ráquis, um fator já esperado, pois, os cachos são compostos de pencas (frutos) e ráquis, e os cachos grandes têm maiores ráquis. Para cachos com o mesmo número de pencas, a participação percentual da ráquis no peso do cacho oscilou entre 6,7 e 7,3% para cultivares tipo *Cavendish* (JARAMILLO, 1982).

Valores estatisticamente significativos para a associação peso do cacho e número de folhas vivas na colheita foram encontrados (0,199) isso indica que o peso do cacho varia de forma direta em relação ao número de folhas vivas,

porém essa influência não é expressiva (correlação baixa), o que não prejudica o PCA, uma vez que como o cacho já está formado, as folhas já cumpriram sua função de fonte para os frutos, não sendo essencial sua função, até porque essa planta (mãe) em um plantio comercial será cortada para crescimento e posterior produção do “filho”.

Com o modelo de previsão de colheita e o conhecimento de correlação das variáveis com o peso do cacho, o produtor pode executar um planejamento de colheita efetivo, no que diz respeito à predição de produtividade e inclusão em sistemas de financiamento bancário para colheita, pois o estudo científico de previsão de peso de cacho dá garantia de retorno ao agente financiador.

4 CONCLUSÃO

Os caracteres vegetativos não são bons indicadores para previsão do peso do cacho. O modelo matemático que expressa a produção da cv. Tropical é formada pelos caracteres de rendimento: $PCA = -5,249 + 0,11NFC + 0,066NFR + 0,046PMF + 0,183CEF + 2,039PRA - 0,011CEN$.

REFERÊNCIAS

ALVES, E. J. (Org.). **A cultura da banana: aspectos técnicos socioeconômicos e agroindustriais**. Brasília: SPI / Cruz das Almas: Embrapa-CNPMPF, 1997. 585 p.

ANTUNES, O. T. et al. Floração, frutificação e maturação de frutos de morangueiro cultivados em ambiente protegido. **Horticultura Brasileira**, Campinas, v. 24, n. 4, p. 426-430, out. 2006.

BÍSCARO, G. A. **Meteorologia agrícola básica**. Cassilândia: UNIGRAF, 2007. 86 p.

COELHO, E. F.; OLIVEIRA, S. L. de; COSTA, E. L. da. Irrigação da bananeira. In: SIMPOSIO NORTE MINEIRO SOBRE A CULTURA DA BANANA, 1., 2001, Nova Porteirinha. **Anais...** Montes Claros: Unimontes, 2001. p. 91-101.

DONATO, S. L. R. et al. Correlações entre caracteres da planta e do cacho em bananeira (*Musa* spp). **Ciência & Agrotecnologia**, Lavras, v. 30, n. 1, p. 21-30, set./out. 2006a.

DONATO, S. L. R. et al. Comportamento de variedades e híbridos de bananeira (*Musa* spp.), em dois ciclos de produção no sudoeste da Bahia. **Revista Brasileira Fruticultura**, Jaboticabal, v. 28, n. 1, p. 139-144, abr. 2006b.

DRAPER, N. R.; SMITH, H. **Applied regression analysis**. 2. ed. New York: J. Wiley, 1981. 709 p.

GUNGULA, D. T. et al. CERES - maize predictions of maize phenology under nitrogen-stressed conditions in Nigeria. **Agronomy Journal**, Madison, v. 95, n. 4, p. 892-899, 2003.

JAME, Y. W.; CUTFORTH, H.W. Crop growth models for decision support systems. **Canadian Journal of Plant Science**, Ottawa, v. 76, n. 1, p. 9–19, 1996.

JARAMILLO, R. C. **Las principales características morfológicas del fruto de banano, variedad Cavendish Gigante (Musa AAA) em Costa Rica**. Panamá: Upeb-Impretex, 1982. 42 p.

LE BAIL, M. et al. Is it possible to forecast the grain quality and yield of different varieties of winter wheat from Minolta SPAD meter measurements? **European Journal Agronomy**, Ottawa, v. 23, n. 2, p. 379–391, abr. 2005.

LIMA NETO, F. P. et al. Relação entre caracteres de rendimento e desenvolvimento em genótipos de bananeira. **Magistra**, Cruz das Almas, v. 15, n. 2, p. 275-281, fev. 2003.

MOREIRA, R. S. **Banana: teoria e prática de cultivo**. 2. ed. São Paulo: FUNDACAO CARGILL, 1999. 1 CD-ROM.

PINTO, A. C. S. et al. Descrição da epidemia da ferrugem do cafeeiro com redes neuronais. **Fitopatologia Brasileira**, Brasília, v. 27, n. 5, set./out. 2002.

ROBERTO, S. R. et al. Caracterização da fenologia e exigência térmica (graus-dia) para a uva 'Cabernet Sauvignon' em zona subtropical. **Acta Scientiarum Agronomy**, Maringá, v. 27, n. 1, p. 183-187, jan./mar. 2005.

SCARPARI, M. S.; BEAUCLAIR, E. G. F. Sugarcane maturity estimation through Edaphic-climatic parameters. **Scientia Agrícola**, Piracicaba, v. 61, n. 5, p. 486-491, Sept./Oct. 2004.

SILVA, S. de O. et al. Bananeira. In: BRUCKNER, C. H. (Ed.). **Melhoramento de fruteiras tropicais**. Viçosa, MG: UFV, 2002. p. 101-157.

SILVA, S. de O. et al. **Catálogo de germoplasma de bananeira (*Musa spp.*)**. Cruz das Almas: Embrapa-CNPMPF, 1999. 100 p. (Embrapa-CNPMPF. Documentos, 90).

SOLER, C. M. T.; SENTELHAS, P. C.; HOOGENBOOM, G. Application of the CSM-CERES-Maize model for planting date evaluation and yield forecasting for maize grown off-season in a subtropical environment. **European Journal of Agronomy**, New York, v. 27, n. 2-4, p. 165-177, Sept. 2007.

STENZEL, N. M. C. et al. Maturation curves and degree-days accumulation for fruits of 'Folha Murcha' orange trees. **Scientia Agrícola**, Piracicaba, v. 63, n. 3 p. 219-225, May/June 2006.

STRECK, N. A. et al. Soma térmica de algumas fases do ciclo de desenvolvimento da escala de counce para cultivares sul-brasilerias de arroz irrigado. **Bragantia**, Campinas, v. 66, n. 2, p. 357-364, abr. 2007.

WYZYKOWSKI, J. **Modelos de regressão para a descrição do crescimento do cafeeiro irrigado e não irrigado após recepa**. 2009. 80 p. Dissertação (Mestrado em Estatística e Experimentação em Agropecuária), Universidade Federal de Lavras, Lavras, 2009.

CAPÍTULO 3

UTILIZAÇÃO DE REDES NEURAIIS ARTIFICIAIS PARA PREVISÃO DE PESO DO CACHO DA BANANEIRA CV. TROPICAL

RESUMO

Os caracteres observados em áreas experimentais têm natureza fenotípica e na maioria das vezes baseia-se apenas na experiência do observador. Com a mensuração das correlações entre variáveis é possível estimar alterações em um caracter com base nas variações em outros caracteres. O presente trabalho foi desenvolvido com o objetivo de estimar o efeito de caracteres agrônômicos mensurados na colheita da bananeira em relação ao peso do cacho. O experimento constituiu-se de um ensaio de uniformidade, conduzido em Guanambi, BA, com a cultivar Tropical, híbrido tetraplóide AAAB, plantado no espaçamento de 3 m x 2 m, em 9 fileiras com 40 plantas por fileira, num total de 360 plantas e área de 2.160 m². Avaliaram-se os caracteres vegetativos, altura da planta, perímetro do pseudocaule, número de filhos emitidos e número de folhas vivas no florescimento e na colheita e os caracteres de rendimento, peso do cacho, número de pencas e frutos, peso da segunda penca, comprimento e diâmetro do fruto em dois ciclos de produção. Nas avaliações, cada planta foi considerada como uma unidade básica (ub), área de 6 m², perfazendo assim, 360 unidades básicas (ub). Utilizou-se o modelo computacional baseado em redes neurais artificiais (RNA) que obteve com eficiente precisão as estimativas (R² de 91%, EMP de 1,40 e QMD de 2,29) a predição do peso de cacho da bananeira cv. Tropical.

Palavras chave: Inteligência computacional. Rendimento. Produção. *Musa* spp.

ABSTRACT

The characters observed in the experimental areas have phenotypic nature and most often based solely on the experience of the producer, the measurement of bunch weight and the correlations of the variables involved are estimated in order to measure changes in a character when he altered another. This work was carried out to estimate the effect of traits measured at flowering and harvest in relation to the weight of the bunch. The experiment consisted of a uniformity trial, conducted in Guanambi, BA, with cultivar Tropical (YB42-21), AAAB tetraploid hybrid, planted at a spacing of 3 mx 2 m, consisting of 11 rows of 52 plants each and considered how useful the nine central rows with 40 plants per row, totaling 360 plants and an area of 2,160 m². We assessed the vegetative characters, plant height, pseudostem circumference, number of children issued and number of green leaves at flowering and harvest, and the characters of yield, bunch weight, number of hands and fruits, weight of the second hand, length and diameter of the fruit in two growing seasons. In the evaluation, each plant was considered as a basic unit (bu), area of 6 m² totaling well, 360 basic units (bu). We used the computational model based on artificial neural networks (ANN) that obtained with efficient precision (R^2 91%, EMP of 1.40 and QMD of 2.29) the prediction of bunch weight in banana cv. Tropical.

Keywords: Computational intelligence. Yield. Production. Banana cv. tropical.

1 INTRODUÇÃO

Um dos entraves relevantes da bananicultura brasileira é a falta de variedades comerciais produtivas com porte adequado, resistentes às principais pragas e doenças e adaptadas a diferentes ecossistemas. O desenvolvimento de cultivares constitui estratégia para a solução deste problema, mediante programas de melhoramento genético, cuja última etapa consiste na avaliação de novos genótipos em áreas de produção quando são comparados às cultivares tradicionais (SILVA et al., 2000, 2002).

Embora exista um número expressivo de variedades de banana no Brasil, quando se considera preferência dos consumidores, produtividade, tolerância a pragas, porte adequado, resistência à seca e ao frio, restam poucas variedades com potencial agrônomo para utilização comercial. As cultivares mais difundidas no País são as bananas tipo prata (Prata, Pacovan e Prata-Anã), responsáveis por 60% da área cultivada; a Maçã, a Mysore, as bananas tipo Cavendish (Nanica, Nanicão e Grande Naine), preferidas pelo mercado internacional, e as bananas tipo Terra (Terra e D'Angola) (Silva et al., 2000; 2002), existindo ainda, outras variedades em menor proporção, como as do tipo Figo ou Bluggoe, as do tipo Caru e a Ouro (MOREIRA, 1999).

O programa de melhoramento genético da bananeira no Brasil teve início em novembro de 1982, em Cruz das Almas-BA, baseando-se no subgrupo Prata. Desde o início, os genótipos selecionados foram avaliados em diversas regiões do Brasil (SILVA et al., 2002). Como resultados práticos destas avaliações foram recomendados os híbridos PA12-03 (Pioneira), FHIA-18, SH36-40 (Prata Graúda), PV42-68 (Pacovan Ken), PV42-85 (Preciosa), PV42-142 (Japira), YB42-21 (Tropical), FHIA-01 (Maravilha), ST42-08 (Garantida), PC42-01 (Caprichosa) e as variedades Caipira, Nam (Prata Baby) e Thap Maeo.

Para avaliação de genótipos em diferentes regiões estuda-se o ciclo da cultura, a altura da planta, o perímetro do pseudocaule, o peso do cacho, o número de frutos por cacho, o comprimento e o diâmetro dos frutos, caracteres considerados relevantes para a seleção de indivíduos superiores, que possam vir a ser recomendados para incorporação aos sistemas de produção do agricultor da região (FLORES, 2000; SILVA et al., 2000). O conhecimento dessas características e sua relação com o peso final do cacho, podem facilitar a tomada de decisão do melhorista e ou possibilitar ao produtor planejamento a nível de financiamento de safra antes da colheita efetiva do cacho.

Na elaboração de planos de gestão a modelagem matemática tem sido utilizada como um processo de previsão de rendimentos de diversas culturas. Isto requer o desenvolvimento de equações para o processo produtivo e a estimativa dos parâmetros envolvidos. As últimas décadas trouxeram uma nova ferramenta para resolver problemas complexos, como as técnicas neurais (Redes Neurais Artificiais). Esta técnica permite simulações muito rápidas e simples para serem realizadas e não requer nenhuma descrição analítica a ser formulada.

Muitos estudos têm sido relatados sobre a aplicação de redes neurais artificiais (RNAs) na agricultura (JIANG et al., 2004; BALA et al., 2005; DIAMANTOPOULOU, 2005; UNO et al., 2005; MOVAGHARNEJAD; NIKZAD, 2007; SAVIN et al., 2007; ZHANG et al., 2007). A maioria desses estudos tem se dedicado a previsões de rendimento (JIANG et al., 2004; UNO et al., 2005; SAVIN et al., 2007). Jiang et al. (2004) relataram o modelo de RNA com algoritmo de treinamento *back* RNA com algoritmo de treinamento *back-propagation* na previsão de safra de trigo de inverno usando sensoriamento remoto de informações com o modelo RNA usando o algoritmo de treinamento *back-propagation*. Uno et al. (2005) desenvolveram modelos de predição de rendimento para o milho, utilizando métodos estatísticos e RNA com vários

índices de vegetação, sendo que maior precisão da previsão foi obtido com o modelo RNA do que com qualquer um dos três modelos convencionais empíricos. Savin et al. (2007) estudou uso combinado de redes neurais e teoria dos conjuntos difusos ou fuzzy de redes neurais (FNN) ou redes neurais granulares (GNN) para prever o rendimento das culturas em diferentes locais na Rússia.

Embora apresente trabalhos relacionando caracteres componentes da produção em diferentes genótipos de bananeira (JARAMILLO, 1982; DADZIE, 1998), a literatura carece de informações que permitam uma estimativa do peso do cacho a partir de alguns atributos medidos na fase da colheita.

Assim, o objetivo do trabalho foi estabelecer um procedimento que possibilitasse, com base em estudos com redes neurais artificiais, a previsão do peso do cacho da planta de bananeira cv. Tropical.

2 MATERIAIS E MÉTODOS

2.1 Implantação e condução da cultura

O experimento foi implantado num LATOSSOLO VERMELHO-AMARELO distrófico típico A-fraco textura média fase caatinga hipoxerofila, relevo plano a suave ondulado na área experimental da Escola Agrotécnica Federal Antonio José Teixeira, localizada no Distrito de Ceraima, Município de Guanambi, Micro Região da Serra Geral, Sudoeste da Bahia, distando 108 km da margem direita do Rio São Francisco e cerca de 90 km da divisa Minas-Bahia, com latitude de 14°13'30'' sul, longitude de 42°46'53'' oeste de Greenwich, altitude de 525 m, com a média anual de precipitação de 663,69 mm e temperatura média de 26°C.

Utilizou-se, para instalação do experimento, mudas micropropagadas, cedidas pela Embrapa Mandioca e Fruticultura, multiplicadas pela Campo Biotecnologia na Biofábrica em Cruz das Almas, BA, aclimatizadas em tubetes, e transportadas para o local do experimento enroladas em embalagem tipo rocambole, onde foram transplantadas diretamente para o campo.

O experimento constituiu-se de um ensaio de uniformidade com a cultivar Tropical (YB42-21), híbrido tetraplóide derivado da Yangambi número 02, grupo genômico AAAB, resistente à Sigatoka-amarela e tolerante ao mal-do-Panamá, com frutos tipo Maçã, lançada pela Embrapa Mandioca e Fruticultura. A área foi formada por 11 fileiras de 52 plantas cada, no espaçamento de 3 m x 2 m, perfazendo um total de 572 plantas e 3.432 m², e considerada como área útil as nove fileiras centrais com 40 plantas cada, num total de 360 plantas e 2.160 m². Nas avaliações procedidas em dois ciclos de produção, considerou-se cada planta como uma unidade básica (ub), área de 6 m², constituindo assim, 360 unidades básicas (ub). O sistema de irrigação utilizado foi aspersão

convencional fixo com aspersores de sub-copa. O manejo de irrigação foi calculado pelo método do turno de rega pré-fixado, baseado, nas médias de evaporação de tanque Classe A medidas no Posto Meteorológico da Codevasf, nos coeficientes de cultivos da cultura (COELHO et al., 2001), nas características físico-hídricas do solo e nas características do sistema de irrigação.

A implantação e os tratos culturais e fitossanitários adotados na cultura basearam-se em recomendações técnicas (ALVES, 1997; MOREIRA, 1999;) e as adubações em análises de solo e de folha (MOREIRA, 1999).

2.2 Avaliações

As avaliações foram procedidas na fase da colheita dos cachos nos ciclos de produção da planta-mãe e do filho. Mensuraram-se descritores fenotípicos vegetativos e de rendimento, seguindo a metodologia proposta (SILVA et al., 1999), apresentada sequencialmente.

1 Caracter vegetativo

1.1 Número de folhas vivas na colheita (NFC): Foi contado o número de folhas vivas presentes nas plantas na época da colheita do cacho.

2.3 Caracteres de rendimento

Os caracteres de rendimento foram avaliados na época da colheita dos cachos. O ponto de colheita baseou-se no diâmetro ou calibre do fruto da fileira externa de frutos da segunda penca, considerando a particularidade do subgrupo (tipo de fruto).

a) Peso do cacho (PCA)

A massa do cacho (massa das pencas, engaco e ráquis), em quilogramas, foi avaliada em balança de plataforma com capacidade para 200 kg.

b) Número total de pencas por cacho (NPC)

Anotou-se o número total de pencas por cacho.

c) Número de frutos por cacho (NFR)

Foram contados e anotado o número de frutos do cacho.

d) Peso da segunda penca (PSP)

A segunda penca de cada cacho, considerada de referência, foi pesada em balança digital (precisão de três casas decimais), e o valor expresso em quilogramas.

e) Peso do fruto (PMF)

A massa do fruto ou dedo central da fileira externa de frutos da segunda penca (penca de referência) foi obtida individualmente de cada cacho, em balança digital com precisão de três casas decimais, expresso em gramas.

f) Comprimento do fruto (CEF)

Tomou-se a medida da curvatura externa do fruto ou dedo central da fileira externa de frutos da segunda penca, em centímetros, utilizando fita métrica, da base ao ápice (desconsiderando o pedicelo e o ápice do fruto).

g) Diâmetro do fruto (DFR)

Mediu-se em milímetros, na parte mediana, no sentido do comprimento do fruto central da fileira externa de frutos da segunda penca, utilizando-se

paquímetro, posicionando-o nas laterais do fruto. Este diâmetro ou calibração lateral do fruto foi utilizado como critério para a colheita.

h) Peso da ráquis (PRA)

A massa da ráquis foi obtida individualmente de cada cacho, em balança digital com precisão de três casas decimais, expresso em gramas.

i) Comprimento do engaço (CEN)

Tomou-se a medida da curvatura do engaço desde a sua inserção no pseudocaule até a inflorescência masculina.

j) Diâmetro do engaço (DEN)

Mediu-se em milímetros, na parte mediana do cacho utilizando-se paquímetro.

2.4 Modelagem de redes neurais artificiais (RNAs)

Modelos de RNA são projetados para aproximar funções matemáticas. Eles permitem simulações rápidas e simples a serem realizadas. As RNAs são formadas por interconexões de processamento simples entre elementos chamados neurônios que têm a capacidade de identificar a relação entre respostas de entrada e saída dos padrões dados, assim como nas atividades neurais humanas (HAYKIN, 1994; 1999).

2.4.1 Estrutura do modelo RNA

Foram construídos vários modelos de RNA multicamadas para produção de banana, afim de determinar qual melhor representa o peso do cacho, com o uso do ‘software’ gratuito SCILAB (CAMPBELL, 2006), versão 5.3 beta-4. O modelo da abordagem constou de três camadas, sendo a primeira composta de dez neurônios que correspondem as dez variáveis de entrada (peso da ráquis, comprimento do engaço, diâmetro do engaço, peso da segunda penca, número total de pencas por cacho, número de frutos por cacho, peso do fruto, comprimento externo do fruto, diâmetro ou calibração lateral do fruto, número de folhas vivas na colheita), a camada oculta com dez neurônios e a camada de saída com um neurônio, que fornece o peso do cacho.

O número de camadas ocultas na RNA depende do grau da complexidade do problema (CHEN; RAMASWAMY, 2002; MOVAGHARNEJAD; NIKZAD, 2007; IZADIFAR; JAHROMI, 2007; ERZIN et al., 2008) e sobre a aplicação da rede (HUANG; MUJUMDAR, 1993). A seleção do número de neurônios para cada camada oculta é empírico. Entretanto, um número adequado de neurônios para resolver um problema pode acarretar em resultados ruins. A dificuldade está justamente na determinação do número adequado de neurônios em cada camada. Uma ou duas camadas ocultas são na maioria dos casos muito útil para grande parte dos problemas (ERZIN et al., 2008). Para o presente estudo, foi usada uma camada oculta.

2.4.2 Treinamento do modelo

RNAs podem modificar seu comportamento em resposta ao ambiente. Esse fator, mais que qualquer outro, é responsável pela credibilidade recebida do modelo. Ao contrário de outros modelos matemáticos, os modelos de RNA não podem representar o comportamento do sistema, a menos que sejam

devidamente treinados. O objetivo do treinamento da rede é ajustar os pesos da interconexão dos neurônios da rede, de modo que a aplicação de um conjunto de entradas produza o desejado conjunto de saídas. Inicialmente, são atribuídos valores aleatórios aos pesos. Os modelos de RNAs usados aqui foram treinados usando um algoritmo de treinamento *back-propagation*. As etapas do processo de treinamento podem ser resumidas como segue:

- a) Aplica-se um vetor de entrada;
- b) Calcula-se a saída da rede que é comparada com o vetor alvo correspondente;
- c) A diferença (erro) é alimentado de volta através da rede;
- d) Os pesos são alterados de acordo com um algoritmo, chamado de *back-propagation* (WASSERMAN, 1989) que tende a minimizar o erro. Os vetores do conjunto de formação foram aplicados sequencialmente.

Este procedimento foi repetido utilizando o treinamento de todo o conjunto, muitas vezes até que o erro ficasse dentro de um critério aceitável ou até que as saídas não mudem significativamente. Para acelerar o treinamento e convergência dos valores de peso, a taxa de aprendizagem pode ser ajustada. É necessário ter uma taxa de aprendizagem que seja pequena o suficiente para proporcionar a convergência grande o bastante para tornar o tempo de computação razoável. Todas as entradas foram padronizadas entre os valores de 0,00 e 1,00. Como progresso do treinamento, os pesos foram modificados para minimizar o erro na saída. Foram empregados o erro médio de previsão (EMP), o quadrado médio dos desvios (QMD) e coeficiente de determinação (R^2) como auxílio na escolha da melhor rede.

$$EPM(\%) = \sum_i^n \frac{(X_{obs} - X_{calc}) * 100}{(X_{obs})} \quad (1)$$

$$QMD = \sum_i^n \frac{(X_{obs} - X_{calc})^2}{n} \quad (2)$$

em que \sum_i^n é o somatório de i até n ; X_{obs} é o valor do peso do cacho obtido à campo; X_{calc} o valor do peso do cacho obtido calculado pela RNA e “ n ” o número de observações.

$$R^2 = \frac{SQReg}{SQT} = 1 - \frac{SOR}{SQT} \quad (3)$$

onde: $0 \leq R^2 \leq 1$

3 RESULTADOS E DISCUSSÃO

A Figura 1 mostra a arquitetura de redes neurais desenvolvida utilizando o algoritmo back-propagation para a previsão do peso do cacho da bananeira cv. Tropical. As redes que foram treinadas com dados experimentais representam as características do crescimento e desenvolvimento da planta a fim de se obter a produção que poderá alcançar maior capacidade de generalização desde que mais dados sejam fornecidos à rede para treinamento. Assim, os dados do primeiro ciclo foram usados para construção do modelo de predição e os do segundo ciclo estavam reservados para testar a capacidade preditiva dos modelos. Os resultados deste processo são descritos a seguir.

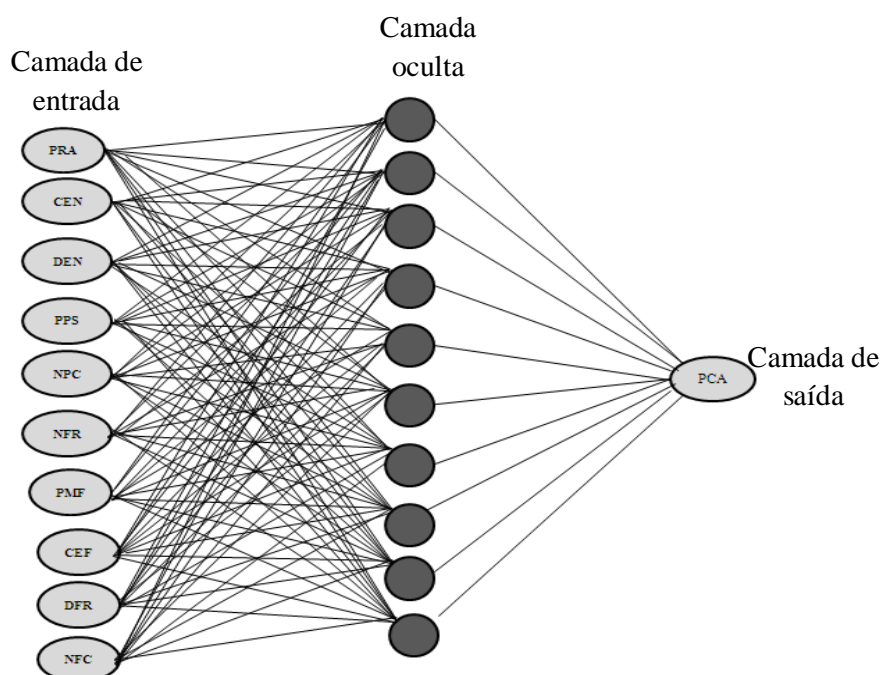


Figura 1 Estrutura do modelo de RNA para previsão de peso do cacho (PCA) de bananeira cv. Tropical

Cada uma das secções de treinamento, até que fosse obtido o menor erro médio, foi realizada com diferentes pesos iniciais, sendo que a rede com o menor erro padrão médio foi selecionada. O melhor modelo de predição de peso de cacho foi composto de uma rede com arquitetura de 10:10:1 para dados da colheita (Tabela 1) e função de transferência sigmóide em ambas as camadas, feitas numa proporção de 80 (288 plantas):20 (72 plantas) entre treinamento (treinar a rede) e validação (feito após o treinamento para verificar o seu desempenho) ambos dados do 1º ciclo e 100% (360 plantas) para testes (feito para verificar a generalização da rede depois do treinamento) com dados do 2º ciclo.

A Tabela 1 resume as características e os parâmetros das RNAs testadas: número de nós em três camadas, a tolerância de erro que foi assumida para a análise, o número de épocas que os menores erros foram alcançados, taxa de aprendizagem (η), coeficiente momento (μ), coeficiente de determinação, erro médio de previsão e o quadrado médio do erro. Os parâmetros η e μ foram fixados em 0,5 porque a menores valores, o processo de aprendizagem seria muito lento, e em valores mais elevados, oscilações podem ocorrer causando a não-convergência da RNA (MARTIN DEL BRIO; SANZ MOLINA, 2002). Os valores previstos pelas RNAs contra os dados reais foram utilizados para calcular os coeficientes de determinação (R^2), erro médio de previsão (EPM) e o quadrado médio dos desvios (QMD) para as respostas.

Tabela 1 Arquiteturas e parâmetros estatísticos para RNA testadas

Arquitetura da rede	RNA-Dados colheita ^a						
	10-1-1	10-5-1	10-10-1	10-15-1	10-20-1	10-30-1	
Tolerância de erro	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	
Número de épocas	500	500	500	500	500	500	
Taxa de aprendizado	0,5	0,5	0,5	0,5	0,5	0,5	
Taxa de momento	0,5	0,5	0,5	0,5	0,5	0,5	
R ²	0,63	0,73	0,91	0,69	0,71	0,71	
EMP (%)	3,83	2,38	1,40	0,97	1,35	1,95	
QMD	2,60	2,50	2,29	2,46	2,49	2,41	

^aRNA-Dados colheita: treinado e validado com dados do 1º ciclo.

De acordo com a Tabela 1 a rede que apresentou os melhores parâmetros estatísticos foi a rede 10-10-1, que contém 10 dados de entrada, 10 neurônios na camada oculta e 1 neurônio de saída (peso do cacho). Nesta RNA 91% dos resultados do peso do cacho são explicados pelas variáveis independentes contidas no modelo e a baixos valores de erro. Quando o número de neurônios na camada oculta é reduzido (1 e 5) ou aumentado (15, 20 e 30), o coeficiente de determinação é reduzido e os valores de erro são aumentados, ou seja, quando o número de neurônios na camada oculta são variados, reduz-se a precisão da previsão e o erro desta aumenta, gerando valores não representativos. A adição da camada oculta torna a rede capaz de extrair estatísticas de ordem elevada, ou seja, a rede adquire uma perspectiva global (BISHOP, 2008), porém o número de nós por camada só é obtido através de várias tentativas e verificação dos parâmetros estatísticos como no presente trabalho.

Na Figura 2, é apresentado o comportamento do erro médio quadrático (EMQ) de treinamento e de validação durante o treinamento da RNA. Pode-se observar que, a partir da época 300, o erro no conjunto de validação apresenta um comportamento constante, enquanto que o erro no conjunto de treinamento decresce. A partir desse momento, a capacidade de generalização da RNA não melhora significativamente, apesar de melhorar os acertos dos dados de treinamento. Oliveira (2010) trabalhando com previsão de produção de álcool através da utilização de RNA obtiveram comportamento constante de erro de quadrado médio a partir da utilização de 10000 épocas, e apesar de serem problemas diferentes, a comparação dos resultados mostra que a RNA para previsão de peso de cacho foi obtida em menor tempo de trabalho da rede.

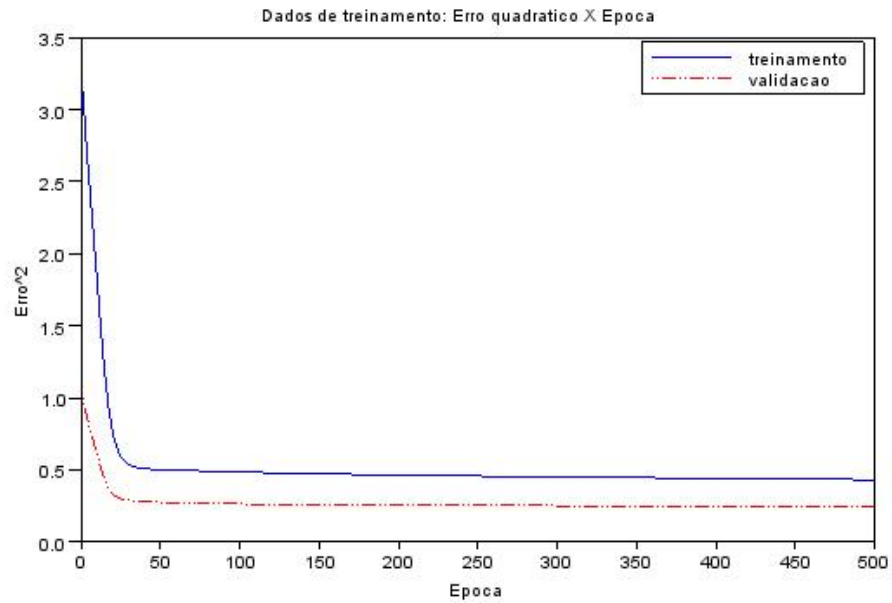


Figura 2 Evolução dos erros de treinamento (azul) e de teste/validação (vermelho) durante o treinamento da RNA com 10 neurônios na camada intermediária

A Figura 3a-b mostra o comportamento entre o real e os valores previstos para as respostas. Pode-se ver que as curvas geradas pela RNA têm, aproximadamente, a mesma forma que as curvas reais, demonstrando que a RNA modela de forma satisfatória o processo gerador dos dados, gerando um modelo confiável com 10 pontos experimentais.

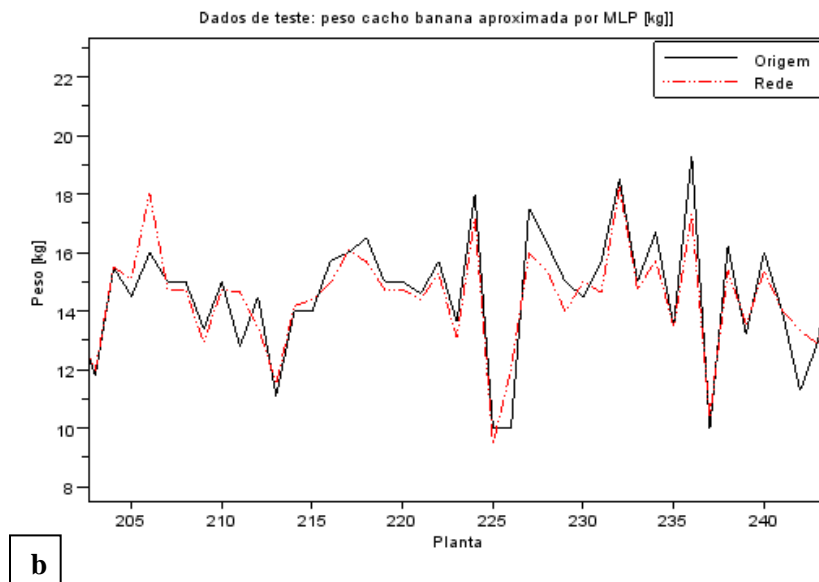
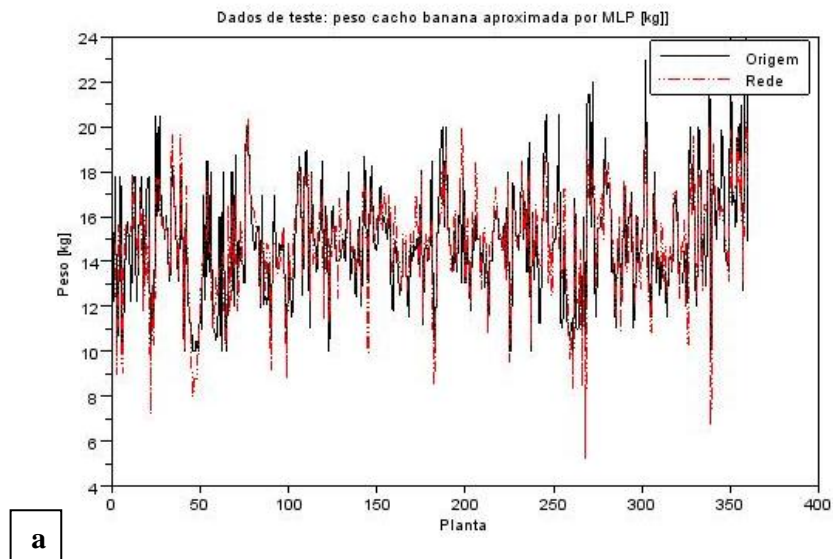


Figura 3 Representação gráfica da série original (preto) e da prevista pela rede neural (vermelho) de peso de cacho da bananeira cv. Tropical: (a) série original de 1 a 360; (b) zoom nas plantas de nº 205 a 240

Os resultados obtidos no presente estudo mostram a capacidade de predição de peso do cacho de bananeira cv. Tropical. Sua principal utilização na prática será a previsão de rentabilidade da colheita, com certa antecedência, o que possibilitará ao produtor a aquisição de financiamentos bancários com garantia científica de retorno ao agente financiador. Os trabalhos atuais em bananeira (DONATO et al., 2006a; 2006b) têm focado em caracterizar e avaliar o comportamento de genótipos (variedades e híbridos) mediante o uso de descritores fenotípicos relevantes para a identificação e seleção de indivíduos superiores, baseados em estudos de correlação mostram qual variável está mais relacionada ao peso do cacho, porém sem que a previsão de produtividade seja feita. Esses estudos mostram que as correlações entre o peso do cacho e os caracteres da planta, de forma geral são não-significativos, e entre o peso do cacho e os caracteres do cacho significativos e positivos. As RNAs serão complementares aos estudos tradicionais de correlação, para estimar o peso do cacho antes da colheita propriamente dita.

Existem várias oportunidades para melhorar as RNAs (RAM et al., 2010), sendo que o fornecimento dos dados de entrada adequados é o principal veículo para melhorias, tanto em número de dados suficientes para treinamento, validação e teste (BISHOP, 2008) quanto a qualidade das variáveis de entrada. Em um estudo feito com oliveira a fim de determinar o tempo ideal de colheita foi verificado que dentre as variáveis analisadas apenas a cor, textura, tamanho e forma são fatores que influenciam no teor final de azeite (RAM et al., 2010).

Uma vez que a RNA estiver totalmente aprovada na prática, mais pesquisas podem ser realizadas com o intuito de promover a ligação do sistema de modelagem de colheita à logística de transportes. Por ter uma boa estimativa de tempo ideal de colheita em que a vida de prateleira seja aumentada (HIGGINS et al., 2010), a RNA dá ao produtor melhor capacidade de

planejamento e transporte de carga para maximizar a eficiência do processo comercial sem perdas. Além disso, caso haja uma fábrica de processamento na propriedade, esta pode desenvolver planos de consumo de multi-colheita, acomodando a capacidade de processamento da fábrica, para maximizar a rentabilidade em toda a sua carteira de produtos. Outras pesquisas também envolvem a tentativa de estender a RNA para previsão de colheita em logística de transporte em culturas como feijão, batata e brócolis, que juntamente com a oliveira, são processadas normalmente pelas mesmas empresas na Austrália que são responsáveis pela distribuição (HIGGINS et al., 2010).

Perceptrons multicamadas foram utilizados como modelos computacionais na identificação e visualização das variáveis mais importantes para modelar a produção de amoreira preta. Redes neurais artificiais foram treinadas com informação de 20 locais na Colômbia, onde a amoreira preta é cultivada. A rede previu com razoável precisão a produção da cultura ($R^2=0,89$). A profundidade do solo, a temperatura média, a drenagem externa e a precipitação acumulada do primeiro mês antes da colheita foram fatores determinantes da produtividade (JIMÉNEZ et al., 2009). Assim como as informações obtidas nos Andes podem ser utilizadas para determinar os locais que são apropriados para a produção de amora, a informação de predição de produção da bananeira pode possibilitar ao produtor a noção de rentabilidade de sua cultura, e incentivo à continuidade dos estudos para verificação de quais variáveis mais influenciam na produção final.

4 CONCLUSÃO

Este trabalho propõe uma metodologia para previsão do peso do cacho da bananeira cv. Tropical utilizando redes neurais artificiais. A RNA implementada obteve uma eficiente previsão de produção, sendo que a rede que obteve o melhor resultado possui a arquitetura 10:10:1.

REFERÊNCIAS

ALVES, E. J. (Org.). **A cultura da banana: aspectos técnicos socioeconômicos e agroindustriais**. Brasília: SPI; Cruz das Almas: Embrapa-CNPMF, 1997. 585 p.

BALA, B. K. et al. Experimental and neural network prediction of the performance of a solar tunnel drier for a solar drying jackfruit bulbs and leather. **Journal of Food Process Engineering**, Hoboken, v. 28, n. 6, p. 552–566, Dec. 2005.

BISHOP, C. M. **Neural networks for pattern recognition**. 2nd. ed. Oxford: University Press, 2008.

CHEN, C. R.; RAMASWAMY, H. S. Modeling and optimization of variable retort temperature (VRT) thermal processing using coupled neural networks and genetic algorithms. **Journal of Food Engineering**, New York, v. 53, n. 3, p. 209–220, July 2002.

COELHO, E. F.; OLIVEIRA, S. L. de; COSTA, E. L. da. Irrigação da bananeira. In: SIMPOSIO NORTE MINEIRO SOBRE A CULTURA DA BANANA, 1., 2001, Nova Porteirinha. **Anais...** Montes Claros: Unimontes, 2001. p. 91–101.

DADZIE, B. K. **Post-harvest characteristics of black Sigatoka resistant banana, cooking banana and plantain hybrids**. Montpellier: Inibap, 1998. 75 p. (Inibap Technical Guidelines, 4).

DIAMANTOPOULOU, M. J. Artificial neural networks as an alternative tool in pine bark volume estimation. **Computers and Electronics in Agriculture**, v. 48, n. 3, p. 235–244, Sept. 2005.

DONATO, S. L. R. et al. Correlações entre caracteres da planta e do cacho Em bananeira (*Musa spp.*). **Ciência & Agrotecnologia**, Lavras, v. 30, n. 1, p. 21-30, set./out. 2006a.

DONATO, S. L. R. et al. Comportamento de variedades e híbridos de bananeira (*Musa spp.*), em dois ciclos de produção no sudoeste da Bahia. **Revista Brasileira Fruticultura**, Jaboticabal, v. 28, n. 1, p. 139-144, abr. 2006b.

ERZIN, Y.; RAO, H. B.; SINGH, D. N. Artificial neural network models for predicting soil thermal resistivity. **International Journal of Thermal Sciences**, v. 47, n. 10, p. 1347–1358, Oct. 2008.

FLORES, J. C. de O. **Avaliação de cultivares e híbridos de bananeira (*Musa spp.*) em quatro ciclos de produção em Cruz das Almas-BA**. 2000. 109 f. Dissertação (Mestrado em Fruticultura Tropical) – Escola de Agronomia, Universidade Federal da Bahia, Cruz das Almas, 2000.

HAYKIN, S. **Neural computing**. 2nd ed. Princeton: Prentice Hall, 1999.

HAYKIN, S. **Neural networks: a comprehensive foundation**. Princeton: Prentice Hall, 1994. p. 696–723.

HIGGINS, A. et al. Forecasting maturity of green peas: an application of neural networks. **Computers and Electronics in Agriculture**, New York, v. 70, n. 1, p. 151–156, Jan. 2010.

HUANG, B.; MUJUMDAR, A. S. Use of neural network to predict industrial dryer performance. **Drying Technology**, London, v. 11, n. 3, p. 525–541, Jan. 1993.

IZADIFAR, M.; JAHROMI, M. Z. Application of genetic algorithm for optimization of vegetable oil hydrogenation process. **Journal of Food Engineering**, New York, v. 78, n. 1, p. 1–8, Jan. 2007.

JARAMILLO, R. C. **Las principales características morfológicas del fruto de banano, variedade Cavendish Gigante (Musa AAA) em Costa Rica**. Panamá: Upeb-Impretex, 1982. 42 p.

JIANG, S. D. et al. An artificial neural network model for estimating crop yields using remotely sensed information. **International Journal of Remote Sensing**, London, v. 25, n. 9, p. 1723–1732, 2004.

JIMÉNEZ, J. et al. Analysis of Andean blackberry (*Rubus glaucus*) production models obtained by means of artificial neural networks exploiting information collected by small-scale growers in Colombia and publicly available meteorological data. **Computers and Electronics in Agriculture**, New York, v.69, n. 2, p.198–208, Dec. 2009.

MARTIN DEL BRÍO, B.; SANZ MOLINA, A. **Redes neuronales y sistemas difusos**. 2.ed. São Paulo: Alfaomega, 2002.

MOREIRA, R. S. **Banana: teoria e prática de cultivo**. 2.ed. São Paulo: Fundação Cargill, 1999. 1 CD-ROM.

MOVAGHARNEJAD, K.; NIKZAD, M. Modelling of tomato drying using artificial neural network. **Computers and Electronics in Agriculture**, New York, v. 59, n. 1-2, p. 78–85, Nov. 2007

OLIVEIRA, A. C. S. et al. Aplicação de redes neurais artificiais na previsão da produção de álcool. **Ciência & Agrotecnologia**, Lavras, v. 34, n. 2, p. 279-284, mar./abr. 2010.

RAM, T. et al. Olive oil content prediction models based on image processing. **Biosystems Engineering**, v. 105, n. 2, p. 221–232, Feb. 2010.

SAVIN, I. Y. et al. Prediction of crop yields with the use of neural networks. **Russian Agricultural Sciences**, Moscow, v. 33, n. 9, p. 361–363, 2007.

SILVA, S. de O. et al. Bananeira. In: BRUCKNER C. H. (Ed.). **Melhoramento de fruteiras tropicais**. Vicosa, MG: UFV, 2002. p .101-157.

SILVA, S. de O. et al. Caracterizacao morfologica e avaliacao de cultivares e hibridos de bananeira. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 22, n. 2, p. 161-169, ago. 2000.

UNO, Y. et al. Artificial neural networks to predict corn yield from compact airborne spectrographic imager data. **Computers and Electronics in Agriculture**, 47, 149–161, 2005.

WASSERMAN, P. D. **Neural Computing: theory and practice**. New York: Van Nostrand Reinhold, 1989. p. 230.

ZHANG, W.; BAI, X. C.; LIU, G. Neural network modeling of ecosystems: a case study on cabbage growth system. **Ecological Modelling**, New York, v. 201, n. 3-4, 317–325, Mar. 2007.

CAPÍTULO 4

COMPARAÇÃO DE TÉCNICAS DE PREVISÃO DE RENDIMENTO EM CACHO DE BANANEIRA CV. TROPICAL: REDES NEURAIAS ARTIFICIAIS X REGRESSÃO LINEAR MÚLTIPLA

RESUMO

Processos de decisão na agricultura, geralmente requerem modelos de resposta para avaliar a rentabilidade da cultura. Este estudo investiga o potencial da cultura em prever respostas de produção pela aplicação de duas técnicas: redes neurais artificiais (RNA) e regressão linear múltipla (RLM) na bananeira cv. Tropical. O experimento constituiu-se de um ensaio de uniformidade, conduzido em Guanambi, BA, com a cultivar Tropical (YB42-21), híbrido tetraplóide AAAB, plantado no espaçamento de 3 m x 2 m, formado de 11 fileiras de 52 plantas cada e consideradas como útil as 9 fileiras centrais com 40 plantas por fileira, num total de 360 plantas e área de 2.160 m². Avaliaram-se os caracteres de rendimento, peso do cacho, número de pencas e frutos, peso da segunda penca, comprimento, diâmetro do fruto e número de folhas vivas na colheita, em dois ciclos de produção. Nas avaliações, cada planta foi considerada como uma unidade básica (ub), área de 6 m², perfazendo assim, 360 unidades básicas (ub). De acordo com as análises realizadas, a RN mostrou-se mais precisa na previsão do peso do cacho em termos de R², EMP e QMD em relação à regressão linear múltipla.

Palavras-chave: Inteligência computacional. Rendimento. Produção. Banana cv. tropical.

ABSTRACT

The evaluation of vegetative characters and yield and its relation to the final weight of the bunch can provide to producers for crop management strategies. Thus, the objective was to compare two methods that allow estimating the effect of traits at flowering and harvest in relation to the weight of the bunch. The experiment consisted of a uniformity trial, conducted in Guanambi, BA, with cultivar Tropical (YB42-21), AAAB tetraploid hybrid, planted at a spacing of 3 mx 2 m, consisting of 11 rows of 52 plants each and considered how useful the nine central rows with 40 plants per row, totaling 360 plants and an area of 2,160 m². We assessed the vegetative characters, plant height, pseudostem circumference, number of children issued and number of green leaves at flowering and harvest, and the characters of yield, bunch weight, number of hands and fruits, weight of the second hand, length and diameter of the fruit in two growing seasons. In the evaluation, each plant was considered as a basic unit (bu), area of 6 m² totaling well, 360 basic units (bu). We used the computational model based on artificial neural networks (ANN) and generalized linear models. According to our analysis, the ANN was more accurate in predicting the weight of the bunch in terms of R², EMP and EMQ in relation to multiple linear regression.

Keywords: Computational intelligence. Yield. Production. Banana cv. tropical.

1 INTRODUÇÃO

O cultivo da bananeira é uma das principais atividades do agronegócio mundial, com plantios em mais de 120 países tropicais e subtropicais (BAKRY et al., 2009). Além de constituir-se em um essencial recurso alimentício, a banana é a fruta mais popular no mundo, e sua produção ocorre praticamente em todos os países em desenvolvimento. O Brasil é quarto produtor mundial de banana, com produção estimada de 6.972.408 toneladas em 2007, em uma área colhida de 508.845 hectares (ANUÁRIO BRASILEIRO DE AGRICULTURA, 2009).

No Brasil, ao contrário da bananicultura latino americana de exportação, baseada nas cultivares do subgrupo Cavendish, Grande Naine e Williams, utiliza-se maior diversidade de cultivares, com predomínio de bananas 'Prata' (SILVA et al., 2000). Porém, não existe uma cultivar que seja resistente a todas as pragas e doenças, apresente elevado rendimento, seja precoce, saborosa, de fácil manejo, com longevidade e tempo de prateleira adequados (DANIELS, 2000). Assim, o programa de melhoramento genético da bananeira no Brasil tem selecionado diversos genótipos promissores, e dentre eles está a cv. YB42-21 (Tropical) (SILVA et al., 2000).

O melhoramento genético assume papel crucial no alcance do ideótipo varietal e, para assegurar a recomendação de genótipos superiores, a etapa final de avaliação em experimentos de campo requer precisão. Nessa etapa, são estudados caracteres fenotípicos que refletem crescimento, precocidade e produtividade, que são importantes para a identificação e seleção de indivíduos superiores (SILVA et al., 1999, 2000), pois, são quantitativos, fáceis de se mensurar, podem estar sob controle poligênico, sofrem influência ambiental e têm importância econômica (ORTIZ, 1997).

Assim o estudo desses fatores possibilita ao produtor uma programação a nível de colheita, uma vez que saber quanto dará o peso do cacho pode viabilizar requisições como programas de financiamento. A literatura carece de estudos recentes sobre previsão de peso de cacho em bananeira, sendo os últimos trabalhos realizados por Jaramillo (1982). Os trabalhos atuais têm focado em caracterizar e avaliar o comportamento de genótipos (variedades e híbridos) mediante o uso de descritores fenotípicos relevantes para a identificação e seleção de indivíduos superiores. Os modelos de simulação são uma ferramenta bastante útil, podendo prever o rendimento de acordo com as variáveis estudadas (GUNGULA et al., 2003). O presente trabalho faz uma abordagem acerca de estudos com regressão múltipla e redes neurais artificiais (RNA).

O conceito de regressão múltipla tem sido largamente difundido e aplicado à várias espécies vegetais, como trigo (LE BAIL et al., 2005), café (WYZYKOWSKI, 2009), cana-de-açúcar (SCARPARI, BEAUCLAIR, 2004; 2009), milho (SOLER et al., 2007), morango (ANTUNES et al., 2006), laranja (STENZEL et al., 2006) e arroz (STRECK et al., 2007) com o objetivo de estimar a colheita e a produção. Em relação às redes neurais artificiais, muitos estudos têm sido relatados na agricultura (JIANG et al., 2004; BALA et al., 2005; DIAMANTOPOULOU, 2005; UNO et al., 2005; MOVAGHARNEJAD, NIKZAD, 2007; SAVIN et al., 2007; ZHANG et al., 2007). A maioria desses estudos tem se dedicado a previsões de rendimento (JIANG et al., 2004; UNO et al., 2005; SAVIN et al., 2007). Embora apresente trabalhos relacionando caracteres componentes da produção em diferentes genótipos de bananeira (JARAMILLO, 1982; DADZIE, 1998), a literatura carece de informações que permitam uma estimativa do peso do cacho a partir de alguns atributos medidos na fase da colheita.

Assim, o objetivo do trabalho foi estabelecer um procedimento que possibilitasse, com base em estudos com regressão múltipla e redes neurais artificiais, a previsão do peso do cacho da bananeira Tropical.

2 MATERIAIS E MÉTODOS

2.1 Implantação e condução da cultura

O experimento foi implantado num LATOSSOLO VERMELHO-AMARELO distrófico típico A fraco textura media fase caatinga hipoxerófila, relevo plano a suave ondulado na área experimental da Escola Agrotécnica Federal Antonio José Teixeira, localizada no Distrito de Ceraima, Município de Guanambi, Micro Região da Serra Geral, Sudoeste da Bahia, distando 108 km da margem direita do Rio Sao Francisco e cerca de 90 km da divisa Minas-Bahia, com latitude de 14°13'30'' sul, longitude de 42°46'53'' oeste de Greenwich, altitude de 525 m, com a média anual de precipitação de 663,69 mm e temperatura média de 26°C.

Utilizou-se, para instalação do experimento, mudas micropropagadas, cedidas pela Embrapa Mandioca e Fruticultura, multiplicadas pela Campo Biotecnologia na Biofábrica em Cruz das Almas, BA, aclimatizadas em tubetes, e transportadas para o local do experimento enroladas em embalagem tipo rocambole, onde foram transplantadas diretamente para o campo.

O experimento constituiu-se de um ensaio de uniformidade com a cultivar Tropical (YB42-21), híbrido tetraplóide derivado da Yangambi número 02, grupo genômico AAAB, resistente a Sigatoka-amarela e tolerante ao mal-do-Panamá, com frutos tipo Maçã, gerado e selecionado pela Embrapa Mandioca e Fruticultura. A área é formada por 11 fileiras de 52 plantas cada, no espaçamento de 3 m x 2 m, perfazendo um total de 572 plantas e 3.432 m², e considerada como área útil as 9 fileiras centrais com 40 plantas cada, num total de 360 plantas e 2.160 m². Nas avaliações procedidas em dois ciclos de produção, considerou-se cada planta como uma unidade básica (ub), área de 6 m², constituindo assim, 360 unidades básicas (ub). O sistema de irrigação

utilizado foi aspersão convencional fixo com aspersores de sub-copa. O manejo de irrigação foi calculado pelo método do turno de rega pré-fixado, baseado, nas médias de evaporação de tanque Classe A medidas no Posto Meteorológico da Codevasf, nos coeficientes de cultivos da cultura (COELHO et al., 2001), nas características físico-hídricas do solo e nas características do sistema de irrigação.

A implantação e os tratos culturais e fitossanitários adotados na cultura basearam-se em recomendações técnicas (ALVES, 1997; MOREIRA, 1999;) e as adubações em análises de solo e de folha (MOREIRA, 1999).

2.2 Avaliações

As avaliações foram procedidas na fase de colheita dos cachos nos ciclos de produção da planta-mãe e do filho. Mensuraram-se descritores fenotípicos vegetativos e de rendimento, seguindo a metodologia proposta (SILVA et al., 1999), apresentada sequencialmente.

a) Caráter vegetativo

Número de folhas vivas na colheita (NFC): Foi contado o número de folhas vivas presentes nas plantas na época da colheita do cacho.

b) Caracteres de rendimento

Os caracteres de rendimento foram avaliados na época da colheita dos cachos. O ponto de colheita baseou-se no diâmetro ou calibre do fruto da fileira externa de frutos da segunda penca, considerando a particularidade do subgrupo (tipo de fruto).

c) Peso do cacho (PCA)

A massa do cacho (massa das pencas, engaço e ráquis), em quilogramas, foi avaliada em balança de plataforma com capacidade para 200 kg.

d) Número total de pencas por cacho (NPC)

Anotou-se o número total de pencas por cacho.

e) Número de frutos por cacho (NFR)

Foram contados e anotado o número de frutos do cacho.

f) Peso da segunda penca (PSP)

A segunda penca de cada cacho, considerada de referência, foi pesada em balança digital (precisão de três casas decimais), e o valor expresso em quilogramas.

g) Peso do fruto (PMF)

A massa do fruto ou dedo central da fileira externa de frutos da segunda penca (penca de referência) foi obtida individualmente de cada cacho, em balança digital com precisão de três casas decimais, expresso em gramas.

h) Comprimento do fruto (CEF)

Tomou-se a medida da curvatura externa do fruto ou dedo central da fileira externa de frutos da segunda penca, em centímetros, utilizando fita métrica, da base ao ápice (desconsiderando o pedicelo e o ápice do fruto).

i) Diâmetro do fruto (DFR)

Mediu-se em milímetros, na parte mediana, no sentido do comprimento do fruto central da fileira externa de frutos da segunda penca, utilizando-se

paquímetro, posicionando-o nas laterais do fruto. Este diâmetro ou calibração lateral do fruto foi utilizado como critério para a colheita e classificação de frutos.

j) Peso da ráquis (PRA)

A massa da ráquis foi obtida individualmente de cada cacho, em balança digital com precisão de três casas decimais, expresso em gramas.

k) Comprimento do engaço (CEN)

Tomou-se a medida da curvatura do engaço desde a inserção do mesmo no pseudocaule até a inflorescência masculina.

l) Diâmetro do engaço (DEN)

Mediu-se em milímetros, na parte mediana do cacho utilizando-se paquímetro.

2.3 Desenvolvimento das redes neurais e análise de regressão múltipla

A RN foi desenvolvida com o ‘software’ gratuito SCILAB (CAMPBELL, 2006), versão 5.3 beta-4. O erro de previsão da rede foi estipulado em 1% e o algoritmo utilizado no treinamento da rede foi o ‘*back-propagation*’. Com o objetivo de comparar a RN com modelos explicativos foi desenvolvida equação de regressão múltipla. Na análise de regressão linear múltipla as variáveis foram selecionadas no procedimento ‘stepwise’ (DRAPER; SMITH, 1981) no software R (VELABLES; SMITH, 2005). As variáveis selecionadas deveriam ser significativas a 5% de probabilidade pelo teste t. Para comparar a RN com a equação de regressão foram empregados o erro médio de

previsão (EMP), o quadrado médio dos desvios (QMD) e o coeficiente de determinação (R^2).

$$EPM(\%) = \sum_i^n \frac{(X_{obs} - X_{calc}) * 100}{X_{obs}} \quad (1)$$

$$QMD = \sum_i^n \frac{(X_{obs} - X_{calc})^2}{n} \quad (2)$$

em que \sum_i^n é o somatório de i até n ; X_{obs} é o valor do peso do cacho obtido à campo; X_{calc} o valor do peso do cacho obtido calculado pela RNA ou pela RLM e “ n ” o número de observações.

$$R^2 = \frac{SQReg}{SQT} = 1 - \frac{SQR}{SQT} \quad (3)$$

onde: $0 \leq R^2 \leq 1$

Foi desenvolvida RNA e construído o modelo de regressão para predição de peso de cacho. Os dados do primeiro ciclo foram utilizados para treinar e validar e os dados do segundo ciclo para testar a rede e a regressão. Na rede não houve distinção de qual variável teria maior influência no peso do cacho, pois seria necessário o desenvolvimento de um algoritmo específico para esse fim (STEVENSON et al., 1990), assim, a camada de entrada foi formada pelas variáveis de rendimento e a de saída pela variável dependente peso do cacho.

Na regressão linear múltipla foram selecionadas as variáveis independentes pelo procedimento ‘stepwise’, que seleciona somente as variáveis significativas, sendo utilizado o modelo:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (4)$$

em que:

Y_i refere-se ao valor do peso do cacho, determinado em função das variáveis de rendimento ($x_i; \dots; x_k$), , sendo respectivamente NPC, NFR, PSP, PMF, CEF, DFR, PRA, CEN, DEN e NFC. ε é o erro associado à i -ésima observação, assumido normal e independentemente distribuído. β_0 constante atribuída ao modelo e β_1, \dots, β_k coeficientes.

3 RESULTADOS E DISCUSSÃO

No modelo ajustado pela análise de regressão linear múltipla, as variáveis de rendimento: número total de pencas por cacho, perímetro do pseudocaule, diâmetro do fruto e diâmetro do engaço não foram incluídas, uma vez que este método seleciona apenas as variáveis significativas. Essa ausência pode estar relacionada à pouca influência que essas variáveis tem sobre o peso do cacho. Foram selecionadas apenas as variáveis número de folhas vivas na colheita, número de frutos, peso médio do fruto, comprimento do engaço, peso da ráquis e comprimento do engaço, de acordo com equação linear múltipla abaixo (Tabela 1).

Tabela 1 Componentes da equação de predição de peso de cacho de bananeira cv. Tropical em função de variáveis mensuradas na colheita

Constante	Coeficientes										R ²	CV(%)
	NFC ¹	NPC ¹	NFR ¹	PPS ¹	PMF ¹	CEF ¹	DFR ¹	PRA ¹	CEN ¹	DEN ¹		
-5,249	+0,11	---	+0,066	---	+0,046	+0,183	---	+2,039	-0,011	---	0,71	9

¹NFC= número de folhas vivas na colheita; NPC= número total de pencas por cacho; NFR= número de frutos por cacho; PPS= arímetro do pseudo caule (cm); PMF= peso médio do fruto (g); CEF= comprimento do fruto (cm); DFR= diâmetro do fruto (cm); RA= peso da ráquis (g); CEN= comprimento do engão (cm); DEN= diâmetro do engão; ---= variáveis não significativas pelo procedimento stepwise.

Ao determinar a magnitude e a significância das associações entre descritores fenotípicos, utilizados para seleção de indivíduos em trabalhos de avaliação, pode-se discriminar quais fatores influenciam na produção. O estudo das variáveis que afetam o peso do cacho foi feito com o propósito futuro de mensurar se a alteração em um caracter pode provocar alteração em outro, porém no presente trabalho, o coeficiente de determinação da RLM foi considerado relativamente baixo (71%), uma vez que as variáveis independentes explicam apenas 71% da variação total do peso do cacho.

Após inúmeras tentativas para descobrir qual a rede que melhor representa o peso do cacho baseado no coeficiente de determinação, erro médio de previsão e quadrado médio do erro foi verificado que a melhor previsão foi obtida pela rede 10-10-1, ou seja, 10 variáveis de entrada, 10 neurônios na camada oculta e 1 neurônio de saída, que corresponde ao PCA. A todas as redes testadas a tolerância do erro assumida na análise foi 0,0001, o número de épocas em que os menores erros foram alcançados foi 500 épocas e a taxa de aprendizagem e o coeficiente momento fixados em 0,5, pois a menores valores o processo de aprendizagem seria muito lento e a valores mais elevados, oscilações poderiam ocorrer causando a não convergência da RNA (MARTIM DEL BRIO; SANZ MOLINA, 2002). A Tabela 2 mostra que os parâmetros estatísticos obtidos foram de alta qualidade com R^2 de 91%, EMP de 1,4% e QMD de 2,29 para a rede selecionada, o que gera alta precisão de previsão de dados. Ram (2010) em estudos com oliveira verificou que dentre as arquiteturas de redes analisadas para previsão do conteúdo de azeite foram obtidos melhores parâmetros estatísticos no formato 80-10-10 com erro padrão de 0,137 para a variedade Pictual.

Tabela 2 Arquiteturas e parâmetros estatísticos para RNA testadas

Arquitetura da rede	RNA-Dados colheita					
	10-1-1	10-5-1	10-10-1	10-15-1	10-20-1	10-30-1
R ²	0,63	0,73	0,91	0,69	0,71	0,71
EMP (%)	3,83	2,38	1,40	0,97	1,35	1,95
QMD	2,60	2,50	2,29	2,46	2,49	2,41

R²= coeficiente de determinação; EPM= erro de previsão médio (%); QMD= quadrado médio dos desvios.

Na Figura 1 é comparada a precisão da modelagem RLM e RNA frente aos dados reais de previsão de colheita, sendo que quanto mais se aproximam as linhas azuis e vermelhas da preta mais próximos estão os dados previstos dos reais. Apesar da não distinção das variáveis significativas na rede neural, a curva ajustada por valores obtidos pela RNA aproximou-se mais dos valores mensurados à campo do peso do cacho, enquanto os valores estimados pela equação de regressão tiveram menor ajuste em relação aos dados observados o que pode ser inferido pelo coeficiente de determinação que mede a proporção da variação total da variável dependente PCA, que é explicada pela variação das variáveis independentes, e no presente trabalho a RN explica 91% da variação do peso do cacho e a RLM explica 71% do peso final (Tabela 3).

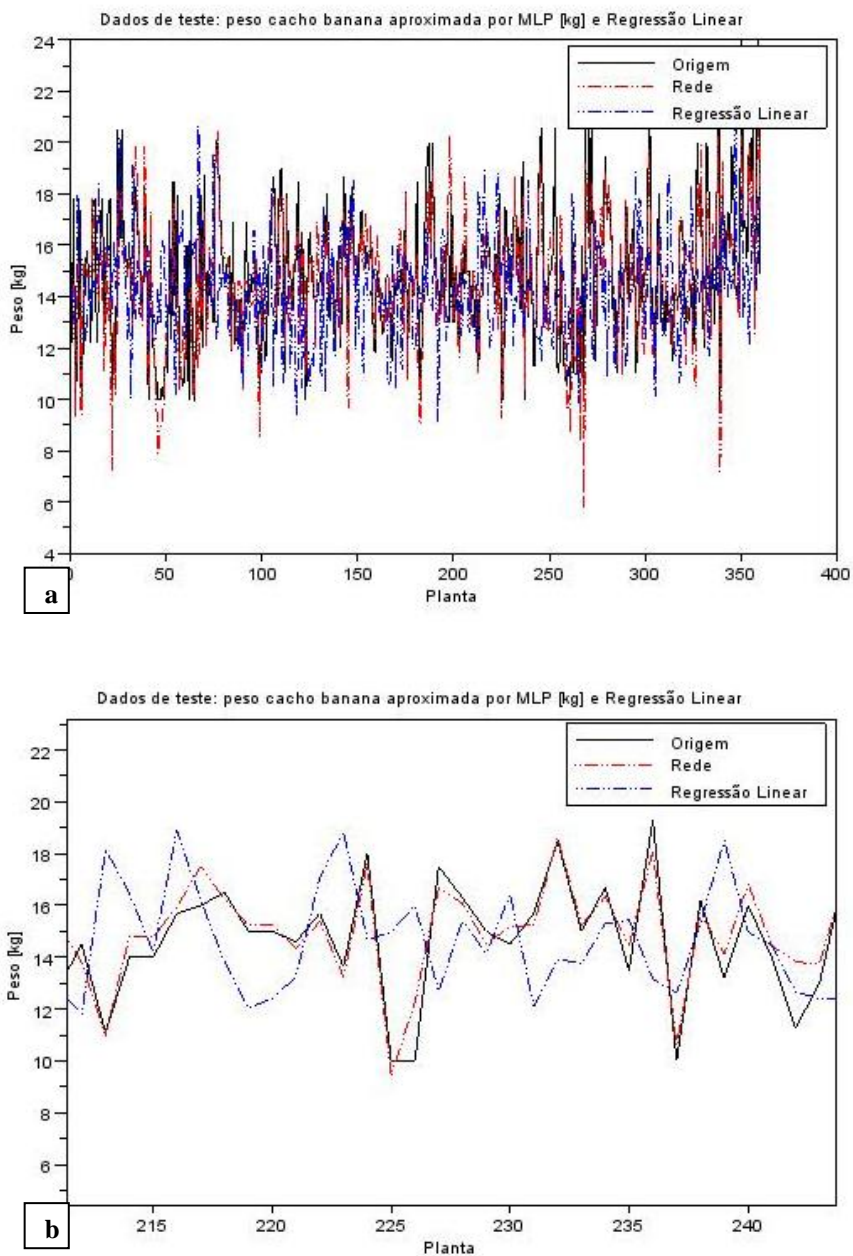


Figura 1 Representação gráfica da série original (preto), da prevista pela rede neural (vermelho) e da prevista pela RLM (azul) de peso de cacho da bananeira cv. Tropical: a- série original de 1 a 360; zoom nas plantas de nº 215 a 240

Os valores previstos pelas RNAs e pela RLM juntamente com os dados reais foram utilizados para calcular os coeficientes de determinação (R^2), erro médio de previsão (EMP) e quadrado médio dos desvios (QMD). Com base na precisão da previsão geral (Figura 1) e na comparação entre os modelos (Tabela 3), a RNA por detectar com maior eficiência a mudança na taxa de progresso ou as inflexões da curva de progresso, demonstrou ser adequada para estabelecer relações entre as variáveis estudadas e o peso do cacho, de maneira a proporcionar melhor ajuste entre os dados observados e estimados.

Estes resultados concordam com Pinto (2002) que avaliaram o potencial das redes neuronais como método alternativo aos sistemas clássicos para descrever a epidemia da ferrugem do cafeeiro (*Coffea arabica*) e perceberam menores valores de EPM (4,72%) e QMD (3,95) para RN em comparação à regressão que apresentou EPM (6,58%) e QMD (4,36), que significa que a RNA são uma ferramenta útil para descrever epidemias de ferrugem.

Tabela 3 Comparação entre redes neurais e análise de regressão, elaboradas a partir de variáveis de rendimento obtidas da bananeira cv. Tropical

	Regressão múltipla			Redes Neurais		
	R^2	EPM(%)	QMD	R^2	EPM(%)	QMD
PCA	0,71	6,52	11,08	0,91	1,40	2,29

¹PCA= peso do cacho (Kg); R^2 = coeficiente de determinação; EPM= erro de previsão médio (%); QMD= quadrado médio dos desvios.

Uma das várias aplicações das RNA é a tarefa de previsão (BISHOP, 2008), e baseado nesse princípio alguns estudos têm sido realizados com o intuito de testar a eficiência das redes neurais em

comparação com modelos clássicos de regressão múltipla. Em um estudo feito com arroz (*Oryza sativa*) foi verificado que o modelo de RN foi mais eficiente para predição de recombinação da proteína hemoglobina 1 quando comparado ao modelo de RLM com 17 variáveis de entrada, obtendo-se R^2 de 85 e 65%, respectivamente (GIORDANO et al., 2010).

Os requisitos obrigatórios para o planejamento da gestão de nutrientes em todos os terrenos agrícolas levou a uma necessidade de técnicas precisas para estimar a produtividade das culturas sendo que as exigências de utilização de nutrientes possam ser baseadas na gestão de processo de planejamento sustentável (KAUL et al., 2005). Assim, com base maior R^2 e menor EMP foi observado que as RN conseguem estimar com maior precisão a previsão de produção de milho e soja em comparação à regressão linear múltipla.

Com o implemento das redes neurais artificiais na previsão de colheita de bananeira cv. Tropical o produtor poderá ter melhor gestão de rendimentos em sua propriedade, o que poderá aumentar o interesse em pesquisas futuras acerca da influência que cada variável exerce sobre o peso final, uma vez que com essa informação, estudos de melhoramento genético podem ser realizados com maior foco na variável que expressa maior rendimento.

4 CONCLUSÃO

A regressão linear múltipla não é adequada para a predição do peso de cacho da bananeira cv. Tropical. A RNA mostrou superioridade de previsão de produção em detrimento a RLM. O produtor pode usufruir da rede para prever sua produção de forma a planejar seu lucro.

REFERÊNCIAS

- ANTUNES, O. T. et al. Floração, frutificação e maturação de frutos de morangueiro cultivados em ambiente protegido. **Horticultura Brasileira**, Campinas, v. 24, n. 4, p. 426-430, out. 2006.
- BALA, B. K. et al. Experimental and neural network prediction of the performance of a solar tunnel drier for a solar drying jackfruit bulbs and leather. **Journal of Food Process Engineering**, Hoboken, v. 28, n. 6, p. 552–566, Dec. 2005.
- BISHOP, C. M. **Neural networks for pattern recognition**. 2nd. Oxford: University Press, 2008.
- CAMPBELL, S. **Modeling and Simulation in Scilab/Scicos**. New York: Springer, 2006.
- DADZIE, B. K. **Post-harvest characteristics of black Sigatoka resistant banana, cooking banana and plantain hybrids**. Montpellier: Inibap, 1998. 75 p. (Inibap Technical Guidelines, 4).
- DANIELS, J. Que variedade de banana debo cultivar? **Infomusa**, Montepelier, v. 9, n. 1, p. 31-33, jun. 2000.
- DIAMANTOPOULOU, M. J. Artificial neural networks as an alternative tool in pine bark volume estimation. **Computers and Electronics in Agriculture**, v. 48, n. 3, p. 235–244, Sept. 2005.
- DRAPER, N. R.; SMITH, H. **Applied regression analysis**. 2nd ed. New York: J. Wiley, 1981. 709 p.
- GIORDANO, P. C. et al. Application of response surface methodology and artificial neural networks for optimization of recombinant *Oryza sativa* non-symbiotic hemoglobin 1 production by *Escherichia coli* in medium containing byproduct glycerol. **Bioresource Technology**, New York, v. 101, p. 7537–7544, Oct. 2010.

GUNGULA, D. T. et al. CERES - Maize predictions of maize phenology under nitrogen-stressed conditions in Nigeria. **Agronomy Journal**, Madison, v. 95, n. 4, p. 892-899, 2003.

JARAMILLO, R. C. **Las principales características morfológicas del fruto de banano, variedad cavendish gigante (Musa AAA) em Costa Rica**. Upeb: Impretex, 1982. 42 p.

JIANG, S. D. et al. An artificial neural network model for estimating crop yields using remotely sensed information. **International Journal of Remote Sensing**, London, v. 25, n. 9, p. 1723–1732, 2004.

KAUL, M.; HILL, R. L.; WALTHALL, C. Artificial neural networks for corn and soybean yield prediction. **Agricultural System**, New York, v. 85, n. 2, p. 1–18, July 2005.

LE BAIL, M. et al. Is it possible to forecast the grain quality and yield of different varieties of winter wheat from Minolta SPAD meter measurements? **European Journal Agronomy**, v. 23, n. 4, p. 379–391, Dec. 2005.

MARTIN DEL BRÍO, B.; SANZ MOLINA, A. **Redes neuronales y sistemas difusos**. 2.ed. São Paulo: Alfaomega, 2002

MOVAGHARNEJAD, K.; NIKZAD, M. Modelling of tomato drying using artificial neural network. **Computers and Electronics in Agriculture**, New York, v. 59, n. 1-2, p. 78–85, Nov. 2007

ORTIZ, R. Morphological variation in musa germplasm. **Genetic Resources and Crop Evolution**, New York, v. 44, n. 5, p. 393-404, Oct. 1997.

PINTO, A. C. S. et al. Descrição da epidemia da ferrugem do cafeeiro com redes neuronais. **Fitopatologia Brasileira**, Brasília, v. 27, n. 5, set./out. 2002.

RAM, T. et al. Olive oil content prediction models based on image processing. **Biosystems Engineering**, v. 105, n. 2, p. 221–232, Feb. 2010.

SAVIN, I. Y. et al. Prediction of crop yields with the use of neural networks. **Russian Agricultural Sciences**, Moscow, v. 33, n. 9, p. 361–363, 2007.

SCARPARI, M. S.; BEAUCLAIR, E. G. F. Physiological model to estimate the maturity of sugarcane. **Scientia Agricola**, Piracicaba, v. 66, n. 5, p. 622-628, Sept./Oct. 2009.

SILVA, S. de O. et al. Caracterizaco morfolgica e avalicao de cultivares e hbridos de bananeira. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 22, n. 2, p. 161-169, ago. 2000.

SILVA, S. de O. et al. **Catlogo de germoplasma de bananeira (*Musa spp.*)**. Cruz das Almas: Embrapa-CNPMF, 1999. 100 p. (Embrapa-CNPMF. Documentos, 90).

SOLER, C. M. T.; SENTELHAS, P. C.; HOOGENBOOM, G. Application of the CSM-CERES-Maize model for planting date evaluation and yield forecasting for maize grown off-season in a subtropical environment. **European Journal of Agronomy**, New York, v. 27, n. 2-4, p. 165-177, Sept. 2007.

STENZEL, N. M. C. et al. Maturation curves and degree-days accumulation for fruits of 'Folha Murcha' orange trees. **Scientia Agrcola**, Piracicaba, v. 63, n. 3 p. 219-225, May/June 2006.

STRECK, N. A. et al. Soma trmica de algumas fases do ciclo de desenvolvimento da escala de counce para cultivares sul-brasilerias de arroz irrigado. **Bragantia**, Campinas, v. 66, n. 2, p. 357-364, abr. 2007.

UNO, Y. et al. Artificial neural networks to predict corn yield from compact airborne spectrographic imager data. **Computers and Electronics in Agriculture**, 47, 149–161, 2005.

WYZYKOWSKI, J. **Modelos de regressão para a descrição do crescimento do cafeeiro irrigado e não irrigado após recepa**. 2009. 80 p. Dissertação (Mestrado em Estatística e Experimentação em Agropecuária), Universidade Federal de Lavras, Lavras, 2009.

ZHANG, W.; BAI, X. C.; LIU, G. Neural network modeling of ecosystems: a case study on cabbage growth system. **Ecological Modelling**, New York, v. 201, n. 3-4, 317–325, Mar. 2007.