

ELVÉCIO GOMES DA SILVA JÚNIOR

SELEÇÃO DE GENÓTIPOS DE ALGODOEIRO UTILIZANDO REDES NEURAIAS
ARTIFICIAIS

Trabalho de Conclusão de Curso apresentado à Universidade Federal de Uberlândia, como parte das exigências do Instituto de Ciências Agrárias, para obtenção do título de “Engenheiro Agrônomo”.

Orientadora

Prof.^a Dr.^a Larissa Barbosa de Sousa

UBERLÂNDIA
MINAS GERAIS – BRASIL
2017

ELVÉCIO GOMES DA SILVA JÚNIOR

SELEÇÃO DE GENÓTIPOS DE ALGODOEIRO UTILIZANDO REDES NEURAIAS
ARTIFICIAIS

Trabalho de Conclusão de Curso apresentado à Universidade
Federal de Uberlândia, como parte das exigências do
Instituto de Ciências Agrárias, para obtenção do título de
“Engenheiro Agrônomo”.

APROVADA em 28 de abril de 2017.

Eng. Agr. Suelen Martins de Oliveira

UFU/ICIAG

Eng. Agr. Daniel Bonifácio Oliveira Cardoso

UFU/ICIAG

Prof.^a Dr.^a Larissa Barbosa de Sousa
ICIAG-UFU
(Orientadora)

UBERLÂNDIA
MINAS GERAIS – BRASIL
2017

A Deus, minha força e
fortaleza, que ilumina, rege e abençoa
meus passos. Aos meus pais Elvécio e
Rosangela, e namorada Ana Flávia
pela dedicação, amor e incentivo para
com o término deste trabalho e a
minha orientadora Prof.^a Larissa pelo
respeito, dedicação e profissionalismo
demonstrados por todo o período de
orientação.

DEDICO.

LISTA DE TABELAS

Tabela 1. Valores referência para as características intrínsecas da fibra avaliadas.	8
Tabela 2. Genótipos selecionados para análises com RNA's e seus respectivos grupos de classificação.	9
Tabela 3. Resumos das análises de variância dos experimentos realizados com 20 genótipos de algodoeiro.	12
Tabela 4. Resumo das análises de variância conjuntas dos caracteres avaliados em 20 genótipos de algodoeiro nas safras 2013/14 e 2014/15.	13
Tabela 5. Taxas de erro aparente calculadas nas análises com redes neurais.	13
Tabela 6. Resumo do percentual de acerto da classificação utilizando redes neurais artificiais.	14
Tabela 7. Topologia das RNA's, quanto ao número de neurônios e função de ativação nas camadas ocultas (O1, O2 e O3), em relação aos caracteres avaliados em 20 genótipos de algodoeiro.	15

SUMÁRIO

1. INTRODUÇÃO.....	6
2. MATERIAL E MÉTODOS.....	8
3. RESULTADOS	12
4. DISCUSSÃO	16
5. CONCLUSÕES	18
REFERÊNCIAS.....	19

1. INTRODUÇÃO

O *Gossypium hirsutum* é a espécie de algodão mais cultivada no mundo. A planta foi domesticada há milhares de anos no sul da Arábia, e como consequência deste processo incessante de domesticação, o algodão é a fibra de origem vegetal mais importante na indústria têxtil do mundo (COTTON INCORPORATED, 2017).

Contudo, para a indústria têxtil é primordial a alta qualidade da fibra, e está influenciando diretamente sua comercialização, pois sua qualidade determinará seu valor de mercado e consequentemente sua aceitação (BONIFACIO; MUNDIM; SOUSA, 2015). Como consequência desta importância, os programas de melhoramento visam maiormente, a qualidade de fibra, assim como a melhoria na produtividade de pluma (MORELLO; FREIRE, 2005; DE ARAÚJO et al., 2013).

A qualidade da fibra do algodão é determinada principalmente por fatores genéticos, porém, o ambiente (fatores climáticos e aspectos agrônômicos como nutricionais, fitossanitários e condução de lavoura), exerce forte influência na expressão de sua qualidade. O processo de colheita e de beneficiamento também têm papel fundamental para manutenção da qualidade da fibra obtida no campo, interferindo principalmente, no teor de impurezas provenientes da mecanização da colheita (SALGADO et al., 2015).

Para determinar sua qualidade, é amplamente utilizado nos programas de melhoramento, o emprego da análise de fibra feita pelo HVI (*High Volume Instrument*). Porém, certas questões surgem a respeito desta prática, como o impacto do uso destas análises na seleção dos genótipos utilizando mais de um critério intrínseco da fibra e a diminuição da diversidade genética para algumas das características analisadas. Além disso, a eficácia da avaliação morfológica utilizando um critério visual e a variação das notas atribuídas por diferentes avaliadores é questionável. Estudos revelam que a avaliação de características morfológicas por meio de notas visuais é de baixa eficiência (GABRIEL; BLANCO, 2009)

Nesse contexto, a modelagem matemática surge como uma ferramenta auxiliar no melhoramento genético do algodoeiro, auxiliando na classificação dos genótipos. Uma das alternativas, são as redes neurais artificiais (RNA's), que é um conceito da computação que visa trabalhar o processamento de dados de maneira semelhante ao cérebro humano, adquirindo conhecimento através da experiência, possibilitando prever

algo, reconhecer padrões ou estabelecer agrupamentos (HAYKIN, 2008; BRAGA et al., 2011). No melhoramento genético as RNA's tem sido empregadas em estudos de diversidade genética (BARBOSA et al., 2011), predição de valor genético (SILVA et al., 2014; CARNEIRO, 2015) e análises de adaptabilidade e estabilidade (BARROSO et al., 2013; NASCIMENTO et al., 2013). No algodoeiro foram utilizadas em estudos de predição de algumas características relacionadas a fiação na indústria têxtil (JACKOWSKA-STRUMILLO et al., 2004; GHOSH et al., 2005; UREYEN; KADOGLU, 2007; GHAREHAGHAJI; SHANBEH; PALHANG, 2007)

Um dos principais atributos da técnica de RNA's está em sua estrutura não linear aliada à sua capacidade de não requerer informações detalhadas sobre os processos físicos do sistema a ser modelado (SUDHEER et al., 2003). Como método de classificação, as RNA's apresentam certas vantagens, como a de ser não paramétrica e serem tolerantes a perda de dados (KAVZOGLU; MATHER, 2003). Assim sendo, o objetivo deste trabalho foi avaliar o potencial das redes neurais artificiais na seleção de genótipos de algodoeiro com alta qualidade de fibra.

2. MATERIAL E MÉTODOS

Neste trabalho foram utilizados dados de 40 genótipos de algodoeiro avaliados nas safras 2013/14 e 2014/15, que foram cedidos pelo Programa de Melhoramento Genético do Algodoeiro (PROMALG), pertencente à Universidade Federal de Uberlândia, quanto a 8 características intrínsecas da fibra, obtidas com auxílio do aparelho HVI (*High Volume Instrument*), além da qualidade de fibra. Os dados foram obtidos em experimentos conduzidos na Fazenda Capim Branco (18°52'S; 48°20'W e 805m de altitude), pertencente à Universidade Federal de Uberlândia, em Uberlândia, Minas Gerais.

O experimento da safra 2013/14 foi conduzido no delineamento de blocos aumentados com 4 repetições, parcela de quatro linhas de 5 metros (m) e espaçamento de 0,9 m entre si. Da mesma forma se procedeu com o experimento da safra 2014/15, exceto o número de linhas por parcela, onde foi acrescida uma linha. Utilizou-se 8 covas por metro e duas sementes por cova, com posterior desbaste deixando-se uma planta por cova.

As características avaliadas foram: comprimento de fibra, uniformidade do comprimento, resistência da fibra, índice micronaire, alongamento, índice de fibras curtas, índice de maturação (Tabela 1), grau de reflectância e qualidade de fibra.

Tabela 1. Valores referência para as características intrínsecas da fibra avaliadas.

COM	UCOM	RF	MIC	ALO	IFC	MAT
> 31,8	> 85,0	> 34,0	> 5,9	> 7,6	> 17,0	> 88,0
31,8 - 30,0	85,0 - 83,0	34,0 - 31,0	5,9 - 4,9	7,6 - 6,8	17,0 - 14,0	88,0 - 80,0
29,9 - 27,3	82,9 - 80,0	30,9 - 27,0	4,8 - 3,5	6,7 - 5,9	13,9 - 10,0	79,9 - 74,0
27,2 - 23,5	79,9 - 77,0	26,9 - 23,0	3,4 - 3,0	5,8 - 5,0	9,9 - 6,0	73,9 - 65,0
< 23,5	< 77,0	< 23,0	< 3,0	< 5,0	< 6,0	< 65,0

COM – Comprimento de fibra (mm); UCOM – Uniformidade do Comprimento (%); RF – Resistência de fibras (gf/tex); MIC – Índice Micronaire; ALO – Alongamento da fibra (%); IFC – Índice de Fibras Curtas (%); MAT – Índice de Maturação da fibra (%).

Grau de reflectância tem como base o conteúdo de cinza existente em uma amostra de algodão. Quanto mais branca for a amostra de algodão maior será o grau de reflectância (COSTA ET AL., 2006).

Quanto a qualidade de fibra, cada característica possui uma classificação em cinco níveis de qualificação, sejam quais forem. Assim, utilizou-se a escala destas classificações atribuindo-as notas de 1 a 5, de acordo com o menos desejável (nota 1) e mais desejável (nota 5).

Com o objetivo de se identificar as principais características determinantes da qualidade de fibra do algodoeiro foi realizada a análise de regressão múltipla com opção

stepwise para fins de seleção de variáveis para ajuste do modelo que incluía, originalmente, 12 características mensuradas nos genótipos nas duas avaliações, safras de 2013/14 e de 2014/15. A análise de regressão múltipla, com estratégia de seleção *stepwise*, foi realizada com o auxílio do software GENES (CRUZ, 2016).

Os dados das variáveis de interesse foram submetidos à análise de variância, conforme o modelo de blocos aumentados. Posteriormente, foi realizada a análise de variância conjunta para as safras. Para todas as análises, todos os efeitos foram considerados como fixos, exceto o erro. Para as análises com RNA's foram utilizados os dados dos genótipos em cada repetição para a obtenção de maior tamanho amostral. Os genótipos foram alocados em dois grupos estabelecidos qualidade de suas fibras. O primeiro grupo foi composto por genótipos com notas até 2,5 e o segundo por genótipos com notas superiores a 2,5. Os genótipos que foram alocadas em grupos diferentes nas repetições e, ou, safras foram desconsideradas nas análises.

Assim, 20 dos 40 genótipos avaliados em 2013/14 e 2014/15 foram utilizados nas análises com RNA's (Tabela 2), perfazendo um total de 60 observações por ano de avaliação, uma vez que foram usados os dados de cada repetição, para treinamento e validação das RNA's.

Tabela 2. Genótipos selecionados para análises com RNA's e seus respectivos grupos de classificação.

Genótipo	Grupo	Genótipo	Grupo
PA UFU – S	1	PA UFU - T	2
PA UFU – M	1	PA UFU - N	2
PA UFU – C	1	PA UFU - E	2
FM 966	1	PA UFU - R	2
PA UFU – Z	1	DP 555	2
PA UFU – D	1	PA UFU - F	2
PA UFU – H	1	PA UFU - OB	2
PA UFU – L	1	PA UFU - A	2
PA UFU – P	1	PA UFU - 18	2
PA UFU – G	1	PA UFU - 7	2

A análise de Redes Neurais Artificiais foi utilizada para prever a qualidade de fibra dos genótipos na safra de 2014/15 com RNA's baseadas nos dados da safra 2013/14 de comprimento de fibra individualmente ou em conjunto com índice de fibras curtas, maturação da fibra e índice micronaire. No treinamento das RNA's, os dados relativos às repetições do experimento de 2013/14 foram submetidos ao processo de ampliação, conforme citado a seguir, obtendo informações de 300 genótipos por grupo com as

mesmas propriedades (vetor de médias e matriz de variâncias e covariâncias) dos genótipos originais. A validação ocorreu com os dados das repetições (60 observações) utilizadas no processo de ampliação e a predição com os dados individuais de repetições (60 observações) e de média de repetições (20 observações) da safra de 2014/15, conforme se segue:

1 – Safra 2013/14 – Treinamento e validação

Safra 2014/15 – Predição (60 observações – dados de repetições)

2 – Safra 2013/14 – Treinamento e validação

Safra 2014/15 – Predição (19 observações – dados de média de repetições)

Para este cenário, foram estimadas as taxas de erro aparente global (TEA) para o treinamento, a validação e a predição das RNA's. A TEA foi obtida pelo percentual de classificação incorreta, considerando os grupos de alocação dos genótipos. Também foram estimadas as taxas de erro aparente por grupo para validação e para as predições.

Em relação à ampliação dos dados, foram simuladas 300 novas informações por grupo a partir dos dados referentes ao cenário proposto. Esses novos conjuntos de dados apresentaram as mesmas propriedades (média, variância e covariância) dos conjuntos de dados originais. O processo de ampliação foi realizado com o auxílio do software GENES (CRUZ, 2016).

Os dados dos experimentos das safras de 2013/14 e 2014/15 foram submetidos à análise de RNA's, realizadas com o auxílio do software MATLAB (BEALE et al., 2015). Para o treinamento das RNA's foram utilizados os 600 dados simulados ampliados (300 de cada um dos grupos) considerando a arquitetura de *perceptron* multicamadas com as seguintes descrições para as topologias:

- a. Número de camadas ocultas. Foram consideradas 3 camadas ocultas.
- b. Número de neurônios. Foram consideradas as combinações de 3 a 12 neurônios para cada camada oculta.
- c. Função de ativação. Foi adotada a função linear de ativação para a camadas de saída. Para as camadas ocultas foram investigadas a adequação de todas as combinações possíveis das funções linear, logística e tangente hiperbólica.

d. Número de ciclos de treinamento: Foi fixado em 5000 épocas. Teve-se o cuidado de limitar o número de interações, para que esse não se tornasse excessivo, o que poderia levar à perda do poder de generalização.

e. Função de treinamento: *trainbr* – *Backpropagation* que é uma função de treinamento da rede que atualiza os valores de peso e de viés de acordo com a otimização de Levenberg-Marquardt. Isso minimiza uma combinação dos quadrados dos erros e pesos, e, em seguida, determina a combinação correta de modo a produzir uma rede com boa capacidade de generalização, cujo processo é denominado de regularização bayesiana.

3. RESULTADOS

Com base na análise de regressão múltipla, com opção *stepwise* de seleção de variáveis, observou-se que o comprimento da fibra (COM), o índice de fibras curtas (IFC), maturação da fibra (MAT) e o índice micronaire (MIC) foram os caracteres determinantes da qualidade de fibra (QF) do algodoeiro. Nas análises com RNA's foi utilizado comprimento de fibra individualmente ou em conjunto com o índice de fibras curtas, maturação de fibra e o índice micronaire, visando a predição da qualidade de fibra dos genótipos de algodoeiro.

Os resumos das análises de variância individuais referentes aos caracteres QF, COM, IFC, MAT e MIC em 20 genótipos de algodoeiro nas safras de 2013/14 e 2014/15 estão apresentados na Tabela 2. Os coeficientes de variação experimental (CVe's) dos experimentos de 2013/14 e de 2014/15 situaram-se abaixo de 10% para os caracteres avaliados, indicando boa precisão experimental. Os valores de CVe's obtidos estão de acordo com os relatados para experimentos desta natureza com a cultura do algodoeiro (BONIFÁCIO; MUNDIM; SOUSA, 2015).

Tabela 3. Resumos das análises de variância dos experimentos realizados com 20 genótipos de algodoeiro.

QUADRADOS MÉDIOS											
FV	GL	SAFRA 2013/14					SAFRA 2014/15				
		QF	MIC	MAT	COM	IFC	QF	MIC	MAT	COM	IFC
G	19	9,33*	0,16*	0,00*	22,94*	33,55*	10,27*	0,22*	0,00*	24,35*	33,04*
Cve		8,64	6,47	3,76	2,51	8,32	9,78	7,93	4,52	3,40	7,64
h ²		57,88	60,46	89,65	97,89	97,35	59,63	61,42	89,92	98,33	96,77
χ		2,83	3,86	0,83	26,74	10,97	3,22	4,06	0,85	28,48	9,24

** Significativo a 1%, pelo teste F (P<0,01); CVe – Coeficiente de variação experimental; h² - Coeficiente de determinação genotípico; χ – Média; FV – Fonte de variação; GL – Graus de liberdade; QF – Índice de Qualidade de fibra; COM – Comprimento de fibra; IFC – Índice de Fibras Curtas; MAT – Índice de Maturação da fibra; MIC – Índice Micronaire.

Foram observados efeitos significativos (P<0,01) para o efeito de genótipos nos dois experimentos (Tabela 3), indicando a existência de variabilidade genética entre os genótipos para os 5 caracteres avaliados nas duas safras. Os coeficientes de determinação genotípicos (h²) dos caracteres COM, IFC e MAT foram, para ambos os experimentos, de alta magnitude.

Os resumos das análises conjuntas para os caracteres QF, COM, IFC, MAT e MIC, avaliados tanto em 2013/14 quanto em 2014/15, são apresentados na Tabela 4.

Observou-se efeitos significativos ($P < 0,01$) de genótipos sobre os caracteres QF, COM, IFC, MAT e MIC. Já para a fonte de variação ambientes, safras no caso, observou-se efeitos significativos ($P < 0,01$) para os caracteres COM, IFC, MAT e MIC. Houve efeito significativo ($P < 0,05$) de interação genótipos x ambientes para os caracteres QF, COM e IFC.

Tabela 4. Resumo das análises de variância conjuntas dos caracteres avaliados em 20 genótipos de algodoeiro nas safras 2013/14 e 2014/15.

FV	GL	QUADRADOS MÉDIOS				
		QF	MIC	MAT	COM	IFC
Genótipos	19	19,25**	0,37**	0,01**	45,87**	67,02**
Ambientes	1	0,48 ^{ns}	2,46**	0,12**	142,57**	234,83**
GxA	19	2,39*	0,14 ^{ns}	0,23 ^{ns}	23,33*	45,17*
CVe(%)		9,21	7,28	4,14	2,87	7,68
h ² (%)		64,81	69,33	92,25	99,27	98,66
Média		3,03	3,96	0,84	27,61	10,11

**, * Significativo a 1 e 5%, pelo teste F, respectivamente; CVe – Coeficiente de variação experimental; h² - Coeficiente de determinação genotípico; FV – Fonte de variação; GL – Graus de liberdade; QF – Índice de Qualidade de fibra; COM – Comprimento de fibra; IFC – Índice de Fibras Curtas; MAT – Índice de Maturação da fibra; MIC – Índice Micronaire.

Na Tabela 5 são apresentados os resultados obtidos pelas RNA's, utilizando o comprimento da fibra individualmente ou em conjunto com índice de fibras curtas, maturação da fibra e índice micronaire. Observou-se maior TEA para treinamento, validação e predição baseando-se apenas em COM. Também se observou menor TEA na predição dos dados de média (16,78% e 9,27%) do que dos dados de repetições (26,89% e 14,55%) tanto para as RNA's utilizando COM individualmente quanto em conjunto com as demais características.

Tabela 5. Taxas de erro aparente calculadas nas análises com redes neurais.

Procedimentos	Taxa de Erro Aparente - TEA (%)			
	Treinamento	Validação	Predição 1	Predição 2
RNA (COM)	10,58	16,81	26,89	16,78
RNA (COM+IFC)	8,47	13,33	14,55	9,27
RNA (COM+IFC+MAT)	5,43	10,29	10,22	4,86
RNA (COM+IFC+MAT+MIC)	2,11	6,55	8,74	1,92

COM – Comprimento de fibra; IFC – Índice de Fibras Curtas; MAT – Índice de Maturação da fibra; MIC – Índice Micronaire.

Considerando a classificação dos genótipos quanto aos grupos (Tabela 6), observou-se que na validação das RNA's, utilizando os caracteres COM, IFC, MAT e

MIC, houve maior percentual de acerto (91,33%) na alocação dos genótipos do grupo 1 do que utilizando somente o COM (73,78%). Resultado similar foi observado para a classificação dos genótipos do grupo 2, com 88,96% de acerto, considerando COM, IFC, MAT e MIC, contra 74,55% com base apenas em COM. Ao submeter os dados de repetição dos genótipos avaliados na safra 2014/15 à predição, observou-se que as RNA's utilizando COM individualmente ou em conjunto com as demais características alocaram corretamente todos os genótipos do grupo 1, enquanto para o grupo 2 a RNA baseada em COM, IFC, MAT e MIC se demonstrou superior, com 89,43% de acerto contra 73,28% considerando apenas COM.

Tabela 6. Resumo do percentual de acerto da classificação utilizando redes neurais artificiais.

Grupos	Classificação (%)		
	Validação	Predição 1	Predição 2
1	73,78 ^a	100,00 ^a	100,00 ^a
1	81,48 ^b	100,00 ^b	100,00 ^b
1	87,12 ^c	100,00 ^c	100,00 ^c
1	91,33 ^d	100,00 ^d	100,00 ^d
2	74,55 ^a	73,28 ^a	81,37 ^a
2	79,98 ^b	77,42 ^b	84,24 ^b
2	84,61 ^c	82,67 ^c	88,50 ^c
2	88,96 ^d	89,43 ^d	92,45 ^d

* a – RNA (COM); b – RNA (COM+IFC); c – RNA (COM+IFC+MAT); d – RNA (COM+IFC+MAT+MIC).

Quando se submeteu os dados de médias dos genótipos avaliados na safra 2014/15 à predição, as RNA's utilizando COM individualmente ou em conjunto com as outras características também alocaram corretamente os genótipos do grupo 1, enquanto que para o grupo 2 a RNA baseada em COM, IFC, MAT e MIC foi superior com 92,45% de acerto contra 81,37% considerando apenas COM. Quando comparadas as predições, observou-se que ao submeter os dados de repetições ou de médias dos genótipos avaliados na safra 2014/15 à predição, as RNA's baseadas em COM individualmente alocaram corretamente todas os genótipos do grupo 1. Porém, a predição com base nos dados de médias foi superior com 81,37% de acerto contra 73,28% utilizando os dados de repetições. Baseando-se em COM em conjunto com IFC, MAT e MIC, ao submeter os dados de repetições ou de médias à predição, as RNA's também alocaram corretamente os genótipos do grupo 1. Novamente, a predição com base nos dados de médias foi superior com 92,45% de acerto contra 89,43% utilizando os dados de repetições.

Na Tabela 6 são apresentadas as topologias, considerando a arquitetura de *perceptron* multicamadas, quanto ao número de neurônios e função de ativação nas camadas ocultas, das RNA's que apresentaram menor TEA na validação utilizando o comprimento da fibra individualmente ou em conjunto com índice de fibras curtas, maturação da fibra e índice micronaire.

Tabela 7. Topologia das RNA's, quanto ao número de neurônios e função de ativação nas camadas ocultas (O1, O2 e O3), em relação aos caracteres avaliados em 20 genótipos de algodoeiro.

COM						COM + IFC					
Neurônios			FUNÇÃO DE ATIVAÇÃO			Neurônios			FUNÇÃO DE ATIVAÇÃO		
O1	O2	O3	O1	O2	O3	O1	O2	O3	O1	O2	O3
3	3	3	Tansig	Logsig	Tansig	3	3	3	Logsig	Purelin	Purelin
COM + IFC + MAT						COM + IFC + MAT + MIC					
Neurônios			FUNÇÃO DE ATIVAÇÃO			Neurônios			FUNÇÃO DE ATIVAÇÃO		
O1	O2	O3	O1	O2	O3	O1	O2	O3	O1	O2	O3
3	3	3	Purelin	Purelin	Purelin	3	3	3	Purelin	Purelin	Purelin

COM – Comprimento de fibra; IFC – Índice de Fibras Curtas; MAT – Índice de Maturação da fibra; MIC – Índice Micronaire.

O maior número de neurônios por camada foi observado para a RNA baseada em COM individualmente comparadas às RNA's baseadas em COM, IFC, MAT e MIC. As topologias das RNA's baseadas apenas em COM apresentaram função de ativação nas camadas ocultas mais complexas comparadas às RNA's baseadas em COM e outras características da fibra, por apresentarem predominância de funções como *logsig* e *tansig*, enquanto a associações com o comprimento de fibra predominavam funções de ativação lineares como *purelin*.

4. DISCUSSÃO

Para se obter maior acurácia na seleção de genótipos de algodoeiro com qualidade de fibra superior exige informações de outros caracteres tecnológicos da fibra, envolvidos na sua expressão fenotípica. Neste trabalho observou-se que os caracteres determinantes da qualidade de fibra foram o comprimento de fibra, índice de fibras curtas, maturação da fibra e índice micronaire.

Por ser a qualidade de fibra do algodoeiro um caráter governado por muitos genes e que sofre grande influência ambiental, a expressão dos genes em diferentes fases de desenvolvimento da fibra de algodão indica que há uma grande quantidade de alelos envolvidos no desenvolvimento da fibra e na sua determinação de qualidade. Neste caso, a seleção indireta para qualidade de fibra com base em caracteres auxiliares é uma possibilidade real para os melhoristas do algodão. A qualidade da fibra já foi relacionada ao comprimento e maturidade da fibra (ZABOT, 2007). O uso de tais características associadas à expressão fenotípica da qualidade de fibra do algodoeiro em análises discriminatórias será efetiva se baseada em caracteres de elevada acurácia para o processo de seleção. Somada à qualidade, é importante alcançar elevados níveis de produtividade, fator esse que está intimamente ligado à alta tecnologia (ROSOLEM, 2001).

As RNA's, baseadas em COM, IFC, MAT e MIC, foram superiores às RNA's baseadas somente em COM; COM e IFC; COM, IFC e MAT, uma vez que apresentaram menor TEA para as etapas de treinamento, validação e predição. Além disso, as RNA's baseadas em COM, IFC, MAT e MIC apresentaram TEA's, em todas as etapas, inferiores a 9%, que neste trabalho representou a classificação errônea de apenas dois dos 20 genótipos avaliados, o que evidencia o elevado potencial de generalização das RNA's (BRAGA et al., 2011; CARNEIRO, 2015).

Na predição, quando se utilizou os dados de médias dos genótipos, as RNA's baseadas em COM, IFC, MAT e MIC também foram superiores às RNA's baseadas somente em COM e demais características, uma vez que as TEA's foram muito inferiores.

No melhoramento do algodoeiro visando qualidade de fibra, se selecionaria, de modo geral, as plantas com notas superiores a 2,5, que corresponderiam aos genótipos alocados no grupo 1. Assim, considerando as predições, observou-se que as RNA's baseadas nos caracteres COM, IFC, MAT e MIC foram superiores às análises baseadas somente no COM e demais associações, uma vez que apresentaram para o grupo 1 o

mesmo percentual de classificação correta e maior percentual de classificação correta dos genótipos do grupo 2.

A maior acurácia utilizando-se os dados de médias dos caracteres COM, IFC, MAT e MIC para a predição se deve ao fato dos efeitos ambientais tenderem a se cancelar com o uso das médias. Na avaliação dos genótipos utilizando a escala de notas constatou-se que 20 dos 40 genótipos apresentavam contradição quanto às notas de qualidade de fibra dentro dos mesmos experimentos e, ou, em experimentos diferentes. Considerando estas contradições como erros de avaliação, constata-se, nestes experimentos, uma taxa de erro de 50% associada à avaliação. Esta taxa de erro foi muito superior à taxa de erro de predição das RNA's baseadas em COM, IFC, MAT e MIC, o que evidencia o potencial de uso das RNA's no melhoramento do algodoeiro visando qualidade de fibra. As RNA's têm se mostrado muito eficazes na solução de problemas de predição, reconhecimento de padrões e agrupamentos (HAYKIN, 2008), que também são dificuldades encontradas nas diferentes etapas de um programa de melhoramento.

5. CONCLUSÕES

As RNA's se mostraram muito eficazes na solução de problemas de predição, reconhecimento de padrões e agrupamentos dos genótipos de algodoeiro.

Utilizar dados de médias na predição por RNA's geram resultados mais confiáveis quanto a seleção para qualidade de fibra dos genótipos de algodoeiro.

Menor número de variáveis explicativas para treinamento e validação requerem RNA's com arquiteturas de maior complexidade.

REFERÊNCIAS

- COTTON INCORPORATED. **Did you know?**. Disponível em: <<http://www.cottoninc.com/>> Acesso em 10 mar. 2017.
- BARBOSA, C. D.; VIANA, A. P.; SILVA, S.; QUINTAL, R.; PEREIRA, M. G. Artificial neural network analysis of genetic diversity in *Carica papaya* L. **Crop Breeding and Applied Biotechnology**, 11(3), p. 224-231, 2011.
- BARROSO, L. M. A.; NASCIMENTO, M.; NASCIMENTO, A. C. C.; SILVA, F. F.; FERREIRA, R. P. Uso do método de eberhart e russell como informação a priori para aplicação de redes neurais artificiais e análise discriminante visando a classificação de genótipos de alfafa quanto à adaptabilidade e estabilidade. **Revista Brasileira de Biometria**, 31(2), p. 176-188, 2013.
- BEALE, M. H.; HAGAN, M. T.; DEMUTH, H. B. **Neural Network Toolbox™ User's Guide How to Contact MathWorks**. Natick: The MathWorks, Inc, 2015.
- BONIFÁCIO, D. O. C.; MUNDIM, F. M.; SOUSA, L. B. Variabilidade genética e coeficiente de determinação em genótipos de algodoeiro quanto a qualidade da fibra. **Revista Verde de Agroecologia e Desenvolvimento Sustentável**, v. 10, n. 3, p. 66-71, 2015.
- BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDEMIR, T. B. **Redes Neurais Artificiais - Teoria e Aplicações** (2nd ed.). Rio de Janeiro: LTC, 2011.
- CARNEIRO, V. Q. **Rede neural e lógica fuzzy aplicadas no melhoramento do feijoeiro**. 2015. 91f. Dissertação (Mestrado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa. 2015.
- COSTA, J. N.; SANTANA, J. C. F.; WANDERLEY, M. J. R.; ANDRADE, J. E. O.; SOBRINHO, R. E. **Padrões Universais para Classificação do Algodão**. Campina Grande, Embrapa Algodão, p. 22, 2006.
- CRUZ, C. D. Genes Software-extended and integrated with the R, Matlab and Selegen. **Acta Scientiarum. Agronomy**, v. 38, n. 4, p. 547-552, 2016.
- DE ARAÚJO, L. F.; BERTINI, C.H.C.M.; BLEICHER, E.; VIDAL NETO, F.C.; ALMEIDA, W.S. Características fenológicas, agronômicas e tecnológicas da fibra em diferentes cultivares de algodoeiro herbáceo. **Revista Brasileira de Ciências Agrárias**, v. 8, n. 3, p. 448-453, 2013.
- GABRIEL, D.; BLANCO, F. M. G. Efeito de genótipos com características morfológicas mutantes sobre o bicus e a produção do algodoeiro. **Arquivos do Instituto Biológico**, p. 211-215, 2009.
- GHAREHAGHAJI, A. A; SHANBEH, M.; PALHANG, M. Analysis of two modeling methodologies for predicting the tensile properties of cotton-covered nylon core yarns. **Textile Research Journal**, 77:565–571, 2007.

GHOSH, A.; ISHTIAQUE, S.; RENGASAMY, S.; MAL, P.; PATNAIK, A. Predictive models for strength of spun yarns: An overview. **AUTEX Research Journal**, 5(1):20–29, 2005.

HAYKIN, S. **Neural Networks and Learning Machines** (3rd ed.). Hamilton: Pearson - Prentice Hall, 2008.

JACKOWSKA-STRUMILLO, L.; JACKOSWKI, T.; CYNIAK, D.; CZEKALSKI, J. Neural model of the spinning process for predicting selected properties of flax/cotton yarn blends. **FIBRES & TEXTILES in Eastern Europe**, 12(4):17–21, 2004.

KAVZOGLU, T.; MATHER, P. The use of backpropagation artificial neural networks in land cover classification. **International Journal of Remote Sensing**, 24(23), 4907–4938, 2003.

MORELLO, C. L.; FREIRE, E. C. **Estratégias para o melhoramento genético do algodoeiro no Brasil**. In: Congresso Brasileiro de Algodão. 2005.

NASCIMENTO, M.; PETERNELLI, L. A.; CRUZ, C. D.; CAMPANA, A. C. N.; FERREIRA, R. P.; BHERING, L. L.; SALGADO, C. C. Artificial neural networks for adaptability and stability evaluation in alfalfa genotypes. **Crop Breeding Applied Biotechnology**, 13, p. 152-156, 2013.

ROSOLEM, C.A. **Informações Agronômicas no 95**: Ecofisiologia e manejo da cultura do algodoeiro. Botucatu, SP: Faculdade de Ciências Agronômicas, UNESP, 2001. p.1.

SALGADO, C.C.; CASTRO, L.H.S.; LEMES, E.M.; SILVA JÚNIOR, E.G. Melhoramento genético do algodoeiro visando à qualidade da fibra. In: SILVA, J.C.; SILVA, A.A.S.; ASSIS, R.T.; FRAVET, P.R.F. (Org.) **Agronegócio sustentável**. Uberlândia: Composer, 2015. p.147-172.

SILVA, G. N.; TOMAZ, R. S.; CASTRO, I.; ANNA, S.; NASCIMENTO, M.; BHERING, L. L. Neural networks for predicting breeding values and genetic gains. **Scientia Agricola**, 71(6), p. 494-498, 2014.

SUDHEER, K. P.; GOSAIN, A. K.; RAMASASTR, K. S. Estimating Actual Evapotranspiration from Limited Climatic Data using Neural Computing Technique. **Journal of Irrigation and Drainage Engineering**, 129(3), 214–218, 2003.

UREYEN, M. E.; KADOGLU, H. The prediction of cotton ring yarn properties from afis fibre properties by using linear regression models. **FIBRES & TEXTILES in Eastern Europe**, 15(4):63–67, 2007.

ZABOT, L. **A cultura do algodão: (*Gossypium hirsutum* L.)**. Santa Maria, RS: Centro de Ciências Rurais, p.15-37, 2007.