

PREVISÃO AGRÍCOLA DA CULTURA DA MAÇÃ: USO DE REDES NEURAIS E PROGRAMAÇÃO LINEAR.

Hugo José Braga⁶ e Edgar Augusto Lanzer⁷

INTRODUÇÃO

Nos últimos anos, tem surgido uma nova técnica denominada redes neurais que juntamente com sistemas especialistas, ambos fazendo parte da inteligência artificial, têm sido propalados como o maior avanço tecnológico desde o descobrimento do transistor, conforme **LAWRENCE (1992)**. Redes neurais (ou neuronais) artificiais são formadas de centenas ou milhares de neurônios que são conectados de forma similar aos neurônios do cérebro, as quais são capazes de aprender de maneira semelhante às pessoas. Um simulador de rede neural é um programa computacional que cria um modelo de neurônios e de conexões entre eles e então treina tal modelo. As redes neurais têm sido aplicadas em várias áreas do conhecimento nos últimos anos. Entre elas podemos enumerar: previsões financeiras, análise e processamento de sinais, controle de processos, robótica, reconhecimento de padrões, filtros de ruídos, reconhecimento da fala, diagnósticos médicos e de decisões de negócios, entre outras. Neste contexto, questões ficam evidentes, tendo em vista a flexibilidade que as redes neurais atuais possuem para estruturar e esquematizar modelos: não poderiam ajustar melhor os *inputs* previsores de rendimentos ou de safras? Qual o tamanho mínimo da série de dados necessária para treinar adequadamente a rede? Ou qual o número de neurônios da camada oculta (*hidden*) seria adequado para o ajustamento da rede previsora? Assim, a utilização de redes neurais pode ser um importante passo na busca de melhores ajustamentos nas previsões de fases fenológicas relevantes de culturas diversas, além de estimativa de rendimentos baseada em dados de potencial de produção (inerente à cultura) e das condições climáticas reinantes que antecedem a colheita. Portanto, este trabalho objetiva dar um novo enfoque à previsão agrícola propondo e comparando modelos polinomiais contra redes neurais artificiais (Back-propagation) de estádios fenológicos importantes de cultura perene (maçã), como plena floração e início de colheita, bem como de rendimentos e safras. Os modelos têm fundamentação agrometeorológica com combinação de estimadores que contemplam informações inerentes ao potencial produtivo da espécie e cultivares (frutos/planta, metragem de copa). A combinação de tais modelos também é enfocada, pelo uso da programação linear, buscando a minimização dos desvios ou erros relativos de previsão. As cultivares de maçã estudadas foram: Gala, Golden Delicious e Fuji.

MATERIAL E MÉTODOS

A área de estudo escolhida para este trabalho engloba as duas regiões mais produtiva da cultura da maçã de Santa Catarina: Fraiburgo e São Joaquim, contribuindo com mais de 50% da produção nacional. Na região de Fraiburgo/SC,

⁶ Eng^o Agr^o, Dr., EPAGRI/SDA - Fax - 048.234.30.48, Cx.P. 5160 - CEP 88034-901 Florianópolis/SC.

⁷ Eng^o Agr^o, PhD., Professor da Engenharia de Produção - CTC/USFC.

séries históricas de produção de pomares comerciais e da pesquisa foram utilizados, como os das empresas: Pomifrai-Fruticultura S/A, RENAR S/A e Estação Exp. de Caçador/EPAGRI. Da região de São Joaquim/SC, utilizaram-se dados da Estação Experimental da EPAGRI/Secretaria da Agricultura. As séries de dados de produção e meteorológicos utilizados abrangeram safras de 1.973/74 até 1.993/94.

MODELOS DE PREVISÃO: FENOFASES E RENDIMENTOS

Os modelos numéricos desenvolvidos neste trabalho baseiam-se, fundamentalmente, em conceitos da bioclimatologia, agrometeorologia e na inteligência artificial (redes neurais), ou seja, modelam relações entre o desenvolvimento vegetativo da maçã, por intermédio de estádios fenológicos relevantes (plena floração e início de colheita), com variáveis meteorológicas básicas (temperatura, insolação,...) ou derivadas (balanço hídrico, somas térmicas, horas de frio,...). No caso da área de estudo deste trabalho, leva-se em consideração nos modelos apenas as variáveis relacionadas com a planta (macieiras) e o tempo/clima. Os fatores ou variáveis edáficas, fitossanitárias entre outras, individualmente, não são consideradas pelo fato de serem otimizadas pelos produtores envolvidos, pois são plantios tecnificados de empresas agrícolas que exploram esta cultura.

Assim, inicialmente, dentro deste contexto são propostos modelos polinomiais (quadráticos, raiz quadrada e exponenciais) para a caracterização das relações entre estádios fenológicos e quais variáveis climáticas e condições possuem significância e após, emprega-se inteligência artificial, via redes neurais, para treinamento e ajustamento de modelos. A floração foi considerada como função de quatro fatores bioclimáticos, ou seja, $F = f(HF, GD, I, TM)$. Para a determinação das horas de frio - HF, dois métodos são utilizados e varios períodos de acumulos de valores entre os meses de outono inverno: método preconizado por ANGELOCCI et al (1979) e SHALTOUT e UNRATH (1983) modificado por EBERT et al (1986). O método de cálculo de graus-dia GD, é o proposto por OMETTO (1981) com a definição as temperaturas basais (T_b e T_B) pelos métodos utilizados por BRUNINI (1971) e CASTONGUAY (1984). Para estimativa de deficiência hídrica é emprega a relação ETR/ETP (evapotranspiração real/evapotranspiração potencial ou de referência). Tal relação é definida pelo emprego dos modelos de PENMAN (1948) e o balanço hídrico de THORNTHWAITE E MATHER (1953), ambos adaptados para aplicações regionais por BRAGA (1982) e SILVA E BRAGA (1987). A capacidade de água disponível - CAD do solo utilizada foi de 125mm, para uma profundidade média do sistema radicular de 1,25m, conforme POLA e BIASI (1993).

Para os modelos de rendimento, as variáveis estimadoras são de duas fontes: a primeira está relacionada à densidade de frutos caracterizada pelo número médio de frutos deixados para produzir por planta (**carga de frutos**) após o processo de raleio manual ou químico, efetuado nos meses de outubro ou novembro de cada ano. Para a capacidade de produção (CP), condicionada pelo vigor do conjunto de plantas de um pomar foi considerada a idade e variáveis *dummies*, em função da não existência de estimador mais adequado (m^2 da silhueta da copa e por pomar - ha); a segunda fonte está relacionada ao crescimento do fruto (CF), função das condições ambientais (clima) e da espécie considerada.

Para o desenvolvimento dos modelos neurais, utilizou-se a rede neural de propagação para frente, *Back Propagation*, denominada **NeuralWorks Explorer, versão 4.0**, da **NeuralWare Inc.** As entradas (inputs) utilizadas na rede, bem como a saída (output), são as mesmas propostas para os modelos polinomiais. O número de neurônios usados na camada oculta (hidden 1), corresponde em geral, de 1/2 a 1,5 vezes as entradas e saída (uma só) da rede. A regra de aprendizagem é a **Delta**, com a função de transferência **Sigmóide**.

Objetivando melhorar a performance dos modelos previsivos polinomiais e neurais, utiliza-se de **programação linear** combinando aqueles mais promissores.. Desta forma, por meio da programação linear minimiza-se os desvios ou erros relativos dos modelos polinomiais e neurais com a utilização do sistema **GAMS (General Algebraic Modeling System) versão 2.25** e metodologia adaptada de **NARULA et al (1993)**. Os modelos mistos, neste trabalho, foram direcionados somente para as previsões de rendimentos.

RESULTADOS E DISCUSSÃO

MODELOS NUMÉRICOS DE PREVISÃO DE FENOFASES

Os modelos polinomiais mais significativos obtidos para a previsão da data da plena floração e início de colheita tiveram seus coeficientes de determinação (r^2) variando entre 0,82 e 0,94, erro médio entre 2 e 4,0 dias e maior desvio absoluto de 9,3 dias para a cultivar Gala; de 0,81 a 0,90, erro médio entre 1,5 a 3,0 dias e maior desvio absoluto de 7,7 dias para a cultivar Golden Delicious; de 0,86 a 0,95, erro médio entre 1,0 a 2,4 dias e maior desvio absoluto de 7,7 dias para a cultivar Fuji. Pelo ajustamento de redes neurais o r^2 ficou entre 0,75 e 0,93, com erro médio entre 1,6 a 3,1 dias e maior desvio absoluto de 9,7 dias para a cultivar Gala; de 0,67 a 0,89, com erro médio entre 1,8 a 4,7 dias e maior desvio absoluto de 16,8 dias para a cultivar Golden Delicious; de 0,72 a 0,91, com erro médio de 1,8 a 3,2 dias e maior desvio absoluto de 8,0 dias para a cultivar Fuji. Dependendo do modelo e da cultivar, a previsão pode ser prevista de 14 a 105 dias de antecedência com bons resultados.

As variáveis independentes ou estimadores mais significativos da plena floração, conforme observações práticas, confirmaram-se nos modelos, tais como: somatório de horas de frio - $\sum HF$, somatório de graus-dia - $\sum GD$, insolação (I) e temperatura média mensal (TM).

MODELOS NUMÉRICOS DE PREVISÃO DE RENDIMENTOS

Os modelos polinomiais mais significativos obtidos para a previsão de rendimentos tiveram seus coeficientes de determinação (r^2) variando entre 0,94 e 0,96, erro relativo médio entre 6,7 e 17,4% para a cultivar Gala; de 0,97 a 0,98, erro relativo médio entre 5,2 a 10,1% para a cultivar Golden Delicious; de 0,91 a 0,93, erro relativo médio entre 9,4 a 11,7% para a cultivar Fuji. Pelo ajustamento de redes neurais o r^2 ficou entre 0,94 e 0,98, com erro relativo médio entre 6,7 a 16,3% para a cultivar Gala; de 0,94 a 0,98, com erro relativo médio entre 6,2 a 11,5% para a cultivar Golden Delicious; de 0,89 a 0,94, com erro relativo médio de 9,5 a 12,4% para a cultivar Fuji. Quanto aos modelos mistos (neurais+polinomiais), todos apresentam um erro médio (%) menor do que qualquer modelo individual participante, seja polinomial ou neural. Os modelos mistos foram aplicados apenas para a previsão de rendimentos. O número de anos/safras utilizadas para o ajustamento de todos os modelos (91) variou de 21 a 48.

A variável ou estimador de maior importância na composição do rendimento (kg/ha) e mesmo do coeficiente de determinação (r^2) dos modelos polinomiais (+75%) é o número médio de frutos/planta, deixados após o raleio dos pomares (outubro-novembro). Os modelos propostos para as três cultivares variam, dependendo da combinação dos estimadores e período de previsão envolvido, de lineares, raiz quadrada, quadráticos e exponencias.

Alguns modelos apresentam variáveis *dummies* indicadoras de possíveis diferenças entre pomares, notoriamente a capacidade de produção, uma vez que são compostos por dados de pomares com idades diferentes, locais diferentes, porta-enxertos diferentes, densidade de plantas/área diferentes. Os modelos neurais tanto para a previsão de plena floração, início de colheita e rendimento foram treinados com mesmo conjunto de estimadores usados para os modelos polinomiais, com a ajuda de 2 a 10 neurônios (camada oculta 1), correspondentes a metade até 1,5 vezes o número de estimadores de entrada (inputs) mais a saída (output). Apresentam desempenho semelhante aos polinomiais, com pequena vantagem para aqueles em alguns casos, com relação ao erro médio (abs. ou %) e maior desvio absoluto (dias ou kg/ha). Neste aspecto, ressalta-se que foi possível treinar a rede com número relativamente pequeno de dados (repetições), <50 anos/safras, ficando claro que o ajuste dos modelos depende mais da qualidade do que da quantidade de informações. Dependendo do modelo e da cultivar, a previsão pode ser prognosticada com 15 a 120 dias de antecedência com bons resultados, obviamente melhorando a sua performance a medida em que se aproxima da data do evento.

VALIDAÇÃO DOS MODELOS DE PREVISÃO

Para testar a performance dos modelos propostos, quanto a sua habilidade previsória, separou-se os dados da safra 93/94 de Fraiburgo/SC, quanto aos pomares da área de estudo (Pomifrai e Renar S.A.). Evidentemente que para a validação conclusiva dos modelos propostos outras safras e locais seriam necessários. Os modelos compostos (dados de Fraiburgo/SC + São Joaquim/SC), tanto neurais como polinomiais, de maneira geral, apresentaram um desempenho similar, com uma diferença máxima de 5 dias para a data da floração e início de colheita, levando-se em conta os cinco pomares testados (1 Renar+4 Pomifrai) e as três cultivares estudadas. Também quanto a previsão de rendimentos, os modelos polinomiais e neurais tiveram desempenho semelhante. Três modelos compostos (**R8**, **NWR8** e **MM1**), tiveram desempenho superior aos demais com possibilidade de previsão antecipada da cultivar Gala de ± 56 dias para safra 93/94, com erro relativo de previsão menor do que 8,9%. O modelo polinomial (**R8**) superestimou a safra 93/94 da empresa Pomifrai em 8,24% (3.436.578kg). Três modelos compostos destacaram-se para a cultivar Golden Delicious na previsão de rendimento, sendo 2 neurais e 1 polinomial (linear com *dummies*), com erro relativo menor do que 30%, para a cultivar Golden Delicious. O modelo neural, mais ajustado, superestimou em 17,5% em relação a safra obtida pela Pomifrai em 93/94 (388.400kg), com ± 30 dias de antecedência. Modelos simples (dados apenas de Fraiburgo/SC) polinomiais, neurais e misto, tiveram ajustamento melhor da safra 93/94 da cultivar Fuji, com uma diferença relativa menor do que 23,7%. Considerando-se a safra 93/94 da Pomifrai e o modelo de melhor desempenho (polinomial), houve uma diferença a maior de apenas 5,9% da produção total colhida de 3.267.346kg, com ± 25 dias de antecedência.

CONCLUSÕES E RECOMENDAÇÕES

Os modelos propostos deixam claro que as variáveis climáticas efetivamente decisivas ao processo ou desencadeamento da floração são energéticas, representadas simplificada e pela somatória de horas de frio, somatória de graus-dia, insolação e temperaturas médias mensais. A somatória de graus-dia mostrou-se como o melhor variável estimadora da data de início de colheita.

A combinação de estimadores que caracterizam o potencial de produção (frutos/m² de área cultivada, m² de silhueta de copa => *dummies*) com variáveis do tempo/clima relacionadas com o crescimento dos frutos, proporcionaram modelos de previsão mais eficientes do que modelos isolados destes mesmos estimadores. O estimador de maior peso na composição do rendimento (kg/ha) caracterizou-se pelo número médio de frutos/m² de área cultivada, deixados após o raleio (out.- nov.), explicando mais de **75%** dos modelos polinomiais (r^2). Outro estimador importante ligado a capacidade de produção média de cada pomar (vigor), foi ajustado em alguns modelos pelo uso da variáveis *dummies*, com bom desempenho, tendo em vista que a idade dos pomares não caracterizou-se como estimador eficiente;

Os modelos neurais foram treinados com a ajuda de **2 a 10** neurônios (hidden 1), correspondendo a **1/2 a 1,5** vezes o número de estimadores de entrada (inputs) mais a saída (output). Apresentaram desempenho, de forma genérica, semelhante aos polinomiais (alguns casos até superior), nas previsões efetuadas (testes de validação e safra 93/94). Neste aspecto, concluiu-se que é possível efetuar-se previsões agrícolas com número relativamente pequeno de repetições, **<50** anos/safras, ficando claro que o ajuste dos modelos depende mais da qualidade e representatividade dos dados do que a sua quantidade;

A utilização da programação linear proporcionou a geração de modelos mistos (polinomiais + neurais) com erro relativo médio das previsões menor do que os modelos individuais. Tais modelos tiram proveito das vantagens individuais dos modelos simples, melhorando seu desempenho conjunto dos estimadores de rendimento, tornando os modelos previsivos mais seguros. Dependendo da cultivar e do modelo, boa previsão pode ser feita com até **120** dias de antecedência, cujo acerto é maior à medida que se aproxima da data da colheita.

BIBLIOGRAFIA CONSULTADA

- ANGELOCCI, L. R. et al. Estimativa do total de horas abaixo de determinada temperatura-base através das medidas diárias da temperatura do ar. Bragantia, São Paulo, 38(4): 27-36, 1979.
- BRAGA, Hugo J. Caracterização da seca agrônômica através de novo modelo de balanço hídrico, na região de Laguna, litoral sul do Estado de Santa Catarina. Piracicaba, ESALQ, 1982: 157 p. (Tese de mestrado).
- BRUNINI, O. Determinação de índices biometeorológicos e sua aplicação para o Estado de São Paulo. Jaboticabal, Fac. Med. Vet. e Agron., 1971.
- CASTONGUAY, Y., BOISVERT, J., DUBÉ, P.A. Comparaison de techniques statistiques utilisées dans l'élaboration de modèles prévisionnels phénoclimatiques. Agricultural and Forest Meteorology, 31(1984): 273-288.

- EBERT, Andreas et al. First experiences with chill-unit models in southern Brasil. Modelling in Fruit Research. Acta Horticulturae, 184(1986): 74-86.
- LAWRENCE, Jeannette. Introduction to neural networks and expert systems. Sylvia Luedking and Janell Fredrickson editors, Nevada City/EUA, 4th edition, 4th edition, 1992: 264 p.
- NARULA, S., SPOSITO, V. A., WELLINGTON, J. F. Intervals which leave the minimum sum of absolute errors regression unchanged. Royal Statistical Society. Appl. statist. 42 (2), 1993: 369- 378.
- OMETTO, José C. Bioclimatologia Vegetal. São Paulo, Editora Agronômica Ceres Ltda, 1981: 425 p.
- PENMAN, H. L. Natural evaporation from open water, bare soil, and grass. Proceedings of the Royal Society, Series A, 193 (1948): 120-145.
- POLA, Augusto Carlos, BIASI José. Primeiros resultados experimentais de irrigação do alho em Caçador - SC. Agropecuária Catarinense, Florianópolis, SC, v.6, n.2, 1993: 18-20.
- SHALTOUT, A. D., UNRATH, C. R. Rest completion prediction model for 'Starkrimson Delicious' apples. J. Amer. Soc. Hort. Sci., 108(6), 1983: 957-961.
- SILVA, L. M. da, BRAGA, H. J. Sistema agrometeorológico para microcomputador. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA, 4, Belém-Pará. Coletânea de trabalhos apresentados. 1987. 405-406.