An Application of Neural Networks: Predicting Corn Yields

by

J. William Uhrig, Bernard A. Engle, and W. Lance Baker

AN APPLICATION OF NEURAL NETWORKS: PREDICTING CORN YIELDS

J. William Uhrig, Bernard A. Engel and W. Lance Baker[1]

Abstract

Neural networks consist of highly parallel, interconnected, simple processing units. These systems differ radically from conventional computing systems; no programming is required, and knowledge is stored in the topology of the net and in the connection matrix, rather than explicitly coded in defined data structures. They offer an alternative to rule-based expert systems for developing intelligent applications. Computer algorithms allow these systems to learn from examples and generalize this learned knowledge for each unique situation. They provide an extremely powerful method for storing and recovering relational information in symbolic and numeric domains. Neural network software was used on weather data, soil moisture data and a trend yield variable to predict corn yields. Modeling corn yields allows an alternative to yield projections made by other methods, and can provide an early forecast of corn yields.

INTRODUCTION

Corn ranks second in terms of overall production of the world's food crops. In 1990 the United States produced approximately 43 percent of the world's total, with about 9 percent of the U.S. total being produced in Indiana. About 60 percent of the corn produced in the U.S. is fed to livestock. Up to 25 percent of the U.S. corn crop is exported. In 1990, the U.S. exported 2.37 billion bushels of corn, almost 71 percent of world corn exports. That was nearly 2.5 times the amount of corn exported by all foreign producers.

During the past 10 years, U.S. average corn yields have varied from a low of 81.1 bushels per acre during the 1983 drought to a record high of 119.8 bushels per acre in 1987. By definition, corn production is the product of the U.S. average corn yield times the corn acreage harvested for grain. Variation in corn yields has been a major factor in the large fluctuation in annual production. Variation in U.S. planted acreage of corn has been heavily influenced by government programs designed to control production by limiting acreage.

The price of corn is determined by worldwide supply/demand conditions. Variations in U.S. production, especially changes in corn yield, can greatly influence world feed grain prices.

The National Agricultural Statistical Service (NASS) of the U.S. Department of Agriculture (USDA) has a long history of gathering data important to agriculture. NASS has the responsibility of making crop production estimates. The first official production estimate for corn, based on a survey of producers, is released in August. In addition to a producer survey, NASS also conducts an objective yield survey, and looks at data gathered by satellite and garnered from computersimulated crop/weather models.

The objective of this paper is to demonstrate the application of neural network software in predicting corn yields. A crop reporting district in Indiana (9 counties in size) is the basic unit of this study. Most states have 9 or 10 crop reporting districts. Yields for each crop reporting district were weighted by acreage to predict state corn production. The same procedure could be extended to other major corn-producing states to arrive at an estimate of U.S. production.

FACTORS IN CORN PRODUCTION

Because corn is the number one cash crop and typically the most profitable crop grown in the Cornbelt, agronomists have devoted many resources to determining the critical factors in corn production. Cowan and Milthorpe (1968) offer the following as a list of the most important environmental factors affecting corn growth: 1) temperature, 2) light, 3) water, and 4) soil mineral nutrient supply.

Air temperature measures the amount of heat energy in the air. Kiniry and Keener (1982) attributed 95 percent of the variability in corn development to temperature indices alone.

Water is a necessary part of nearly every biochemical process in plants. The availability of water to corn (and the interaction with other climatic factors) is often the most limiting factor in corn production. Lack of moisture may retard seed germination after planting, slow or temporarily stop growth during the vegetative stage, and delay tasseling and silking 4 to 5 days.

Hewitt and Smith (1974) found many plant nutrients were needed for normal corn plant growth and development. They found that nitrogen generally effected corn growth to a much greater extent than other nutrients. It is an essential element in the synthesis of amino acids and proteins, and is generally regarded as the most deficient micronutrient. Nutrient supply is the only primary growth factor that is controllable (in non-irrigated production areas).

Other factors may limit corn production. Shaw and Newman (1984) list freeze prior to physiological maturity as the single greatest risk to a corn crop during grain fill. Potential losses range up to 50 percent. Ullstrup (1977) estimates that diseases account for 2-7 percent of the U.S. corn crop. Corn insects may cut yields as well as reduce the acreage harvested.

PREVIOUS MODELING EFFORTS

In 1914, Smith studied the effects of summed weekly and monthly temperature and precipitation on corn yield in Ohio. determined precipitation from the middle of July through the

The middle of August to be the most dominant weather factor controlling final yield. In 1958, Runge and Odell found similar results in central Illinois.

Corn yields in the U.S. have increased steadily since the 1930's and more rapidly since the 1960's. Thompson (1969) used technology-trend variables for two time periods: 1) a linear yield increase from 1930-1960, and 2) a quadratic increase from 1960-1967. Using quadratic regressions on yield on monthly precipitation and mean temperatures for five midwestern states, he found that below-average mean temperatures in July and August and above-average July rainfall were associated with highest yields.

Nelson and Dale (1978) used applied nitrogen rates to account for technological trends. Soil moisture has also been used as the moisture variable rather than precipitation. Dale and Shaw (1965) modeled the daily soil moisture in the corn root zone and correlated the number of non-moisture stress days with experimental plot corn yields. Leeper et al. (1974) found that under high levels of fertility and management, the potential of a soil to produce high corn yields was due largely to its waterholding characteristic.

NEURAL NETWORKS

Neural networks offer an alternative to rule-based systems for research application. These systems have been available for several decades (Hebb, 1949, Rosenblatt, 1961). Our understanding of neural networks has grown during the past few years and fostered many diverse applications. The July 1991 issue of AI *Expert* lists 40 neural-net simulation packages, neural-net shells, and chip-based networks. Several basic network architectures have emerged--all are based on the concept of connecting a large number of simple processing elements in a massively-parallel, highly-distributed processing environment.

Two basic neural network architectures have emerged. These include feedback and feedforward networks (Bolte, 1989). Each have particular characteristics which make them applicable to particular types of problems. (For additional information on the differences, see Bolte, 1989.)

Uhrig and Botkin describe neural networks as follows: Neural network applications can be classified into two broad categories of recognition and generalization (Dutta, 1989). For both classes, the neural network is first trained on a set of input/output pairs: (I,O), (I,O),...,(I,O). In recognition problems, the trained network is tested with the input I(1<=j<=n) corrupted by random noise. The trained network is expected to reproduce the output O corresponding to the input I, in spite of the presence of noise. Examples of recognition problems are shape recognition (Hinton, 1987) and speech generation (Senjowski, 1986).

Conversely, in generalization problems, the trained neural network is tested with input I, which is distinct from the inputs 1, I,..., I used for training the network. The network is expected to correctly predict the output O for the input I from the paradigm it has learned in the input/output training session. Classification and prediction are common examples of generalization problems. Within the generalization category there are various applications, some of which are well defined while others are less specific. Electrical circuit analyses on the one hand, and the diagnosis of diseases from symptoms on the other, are examples of these extremes.

Neural networks are composed of many simple processing elements joined together through numerous interconnections. Processing elements, the neural network equivalent of biological neurons, are generally uncomplicated devices that receive a number of inputs (x) which are weighted with weights (w). This sum is processed in an activation function to determine the activation level of the receiving neuron. The activation function may take many forms. Typically a threshold function (sending 1 if the input sum exceeds some threshold, or 0 otherwise) or a sigmoid function (ranging continuously between 0 and 1) is used to compute the activation level of a particular neuron (Bolte, 1989). The output value of the transfer function is generally passed directly to the output path of the processing element. This output signal (y) is then sent to many other processing elements as input signals via the interconnections between processing elements. Due to the interconnected nature of the network, these calculations must proceed in parallel to accurately determine the state of a network.

Processing elements are usually organized into groups called layers. A typical network consists of a sequence of layers with full or random connections between successive layers. Input layers are connected to output layers through numerous junctions with a hidden layer. "Learning" is the process of adapting or modifying the connection weights in response to stimuli being presented at the input layer and (optionally) the desired outputs of those inputs (Klimasauskas, 1989). A "trained" network is referred to as hetero-associative if the desired output is different from the input.

An essential characteristic of any network is its learning rule which specifies how weights adapt in response to a learning example. The parameters governing a learning rule may change over time as the network progresses in its learning. In supervised learning, for each input stimulus, a desired output stimulus is presented to the system and the network gradually configures itself to achieve that desired input/output relationship.

Most neural networks employ an algorithm termed "backpropagation" or "generalized delta rule" algorithm. In short, when the network is presented with a training set, the difference between the predicted output (calculated from the current network state) and the training output (from actual data) is computed. The error, transformed by the derivative of the transfer function, is then propagated backwards from the output layer through the hidden layer to the input layer, with connection strengths adjusted so that output error is minimized. All data in the training set is presented until the network is able to successfully duplicate the training set with some predefined error tolerance. At this point, the network is considered trained and new input data is presented for testing purposes. Back-propagation is a very powerful technique for constructing non-linear transfer functions between several continuously valued inputs and one or more continuously valued output.

Back-propagation networks configure themselves in such a way as to establish a relationship between input and output variables that minimizes the error between the network-generated output and the actual, or target, output (in the sense of least squares fitting). After minimizing the error or training the network, the network has learned the input/output relationship. The technique used to establish the coefficients in a network is described as a set of differential equations which modify the coefficients in such a way as to reduce the mean-square-error of the output for the training set. These are derived by defining the error for a single pass through a training set (which consists of several presentations of input and expected, or desired, output) as the sum of the squares of the difference between the actual and expected outputs. This is called the square error.

The derivative of each weight, or coefficient, in the network is computed as a function of the error. For most cases, these differential equations do not have a closed-form solution. As such, the solution is derived by using numerical methods. The numerical techniques compute the direction to change the weight coefficients (gradient) and then changes them slightly in that direction. Today's computational technology allows rapid, multiple iterations of differential equations which result in the appropriate error-minimizing solution.

Conventional statistical techniques can be applied to well defined domains. Regression analysis has often been applied in modeling agricultural phenomena. However, conventional modeling efforts are quickly complicated when the underlying functional form of the data is not obvious. Because a neural network does not require a priori specification of a particular functional form, its use may present improvement over currently used statistical methods. Furthermore, insight into correct data relationships may be gained.

NEURAL NETWORK MODEL

Indiana corn yields were analyzed using a trained backpropagated neural network. This network was assembled in three layers consisting of 25 input processing elements (weekly data from May 10 through September 15), one hidden layer with 50 nodes, and one output element for corn yields. Data was collected for the variables deemed important in determining corn yields Variables used in this paper include:

1. Weekly high and low temperature by crop reporting districts.
2. Two soil moisture variables, from layers 0-1 and 1-2[2] meters in depth.
3. Cumulative growing degree days (GDD) calculated.
4. Trend yield calculated by
5. The output data was bushels yield per acre.

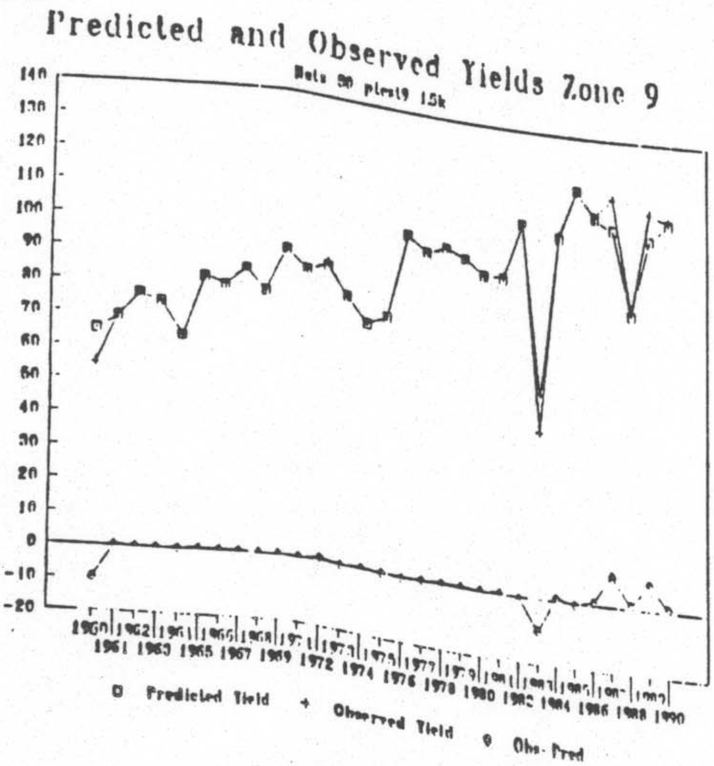
Both the input and output variables were normalized between .1 and .9. Corn plant population and nitrogen fertilizer application were included in the early phase of this research, but eliminated after several simulations.

The analysis was conducted using NETS, a back-propagation neural network software package developed by the National Aeronautics and Space Administration (NASA).

The training set consisted of 25 years of data, during the period 1960 through 1989. Data from 1970 was discarded as an outlier. Yields in 1970 were severely impacted by the corn blight.

Network training was accomplished by presenting examples to the network corresponding to actual input/output data randomized and presented to the network two-thousand times to provide sufficient "learning". No formal relationships were specified. To maintain an independent data set for validation, the last four years were chosen independently for testing. The results are shown in Figure 1.

Figure 1.



After the model has been validated, the network was trained over the data from 1960-1989. After the training was accomplished, the input data for 1990 was subjected to the network to predict 1990 corn yield (the output).

The neural network model was then trained with weekly input data for June and July in an attempt to provide an early forecast of corn yield. The results are very encouraging.

Table 1 shows the yield forecast by crop reporting districts of Indiana for 1990. UPDATENET is the neural network model trained through 1989. Columns 3 and 4 are the August and final USDA crop estimates respectively.

The percent error columns in Table 1 are the difference between forecasted yields and USDA final estimates in proportion to USDA final estimates. The table indicates that the network models have a lower error than those of USDA final estimates for zones 3, and 4, while forecasting yields higher than USDA final estimates for the other zones. The network correctly forecast yield for zone 6 to the nearest bushel. Zone 9 yields were correctly forecast to within one bushel by the network. USDA forecasts were under USDA final estimates for zones 3 and 7; over for zones 1, 4, and 8; and correct to the nearest bushel for zones 2, 5, 6, and 9.

Table 1. Yield Forecasts for 1990. (bu. per acre)

Zone	UPDATENET	USDA	USDA FINAL ESTIMATE	%ERROR UPDATENET	%ERROR USDA
1	136	134	133	2.26	0.75
2	135	132	132	2.27	0
3	122	121	126	-1.59	-3.97
4	132	137	136	-2.9	0.74
5	138	135	135	2.22	0
6	124	124	124	0	0
7	127	113	118	7.63	-4.24
8	116	110	107	8.41	2.8
9	119	118	118	0.85	0

Table 2. August, 1990 USDA and Nets Production Forecasts for Indiana.

	Production (1,000 bu.) Forecasts			Final Estimate	Residuals 1,000 bu.			Residuals % of USDA Final Estimate		
Zone	NETSA	NETSR	USDA	USDA	NETSA	NETSR	USDA	NETSA	NETSR	USDA
1	117,471	127,623	115,230	121,518	-4,047	6,105	-6,288	-3.33	5.02	-5.17
2	99,074	91,337	97,050	95,034	4,035	-3,702	2,011	4.25	-3.90	2.12
3	58,101	62,650	59,992	61,194	-3,093	1,456	-1,202	-5.05	2.38	-1.96
4	86,848	85,603	94,166	94,685	-7,837	-9,082	-519	-8.28	-9.59	-0.55
5	159,061	156,650	159,186	157,454	1,607	-804	1,732	1.02	-0.51	1.10
6	49,744	49,157	49,233	51,275	-1,531	-2,118	-2,042	-2.99	-4.13	-3.98
7	91,551	89,839	76,993	75,663	15,888	14,176	1,330	21.00	18.74	1.76
8	22,937	24,618	21,618	21,279	1,658	3,339	339	15.69	5.72	1.59
9	24,952	24,881	24,132	24,943	9	-62	-811	0.04	-0.25	-3.25
all	709,739	720,731	697,600	703,050	6,689	7,186	-5,450	0.95	1.02%	-0.78

NETSA = 3 Year Average NETSR = Regression

To calculate production, an estimate of harvested acreage was needed. NETSR represents a regression estimate based on only 6 years of data. Because of limited historical data, a second estimate (NETSA) was calculated using a 3 year average, column 4 is the USDA final acreage estimate.

The errors for NETSR ranged from a high of 18.74% for zone 7 to a low of 0.25% for zone 9. USDA errors ranged from a high of -5.17% for zone 1 to a low of 0.55% for zone 4. The state errors were 0.78% for USDA, 0.95% for NETSA and 1.02% for NETSR.

What would we do differently if starting over?

The following changes are very subjective and may or may not improve the results. Suggested changes include:

1. Choosing a software package that is user friendly. A package including graphic capabilities would be desirable.
2. Use state data rather than crop reporting districts.
3. Include 10-12 variables or changes in variables and let the neural network model choose which ones are relevant.
4. Do not use cumulative data.
5. Use the change in data instead of the real value.
6. Normalize the data to allow the maximum range of each variable.

SUMMARY

This research project has applied neural network software to an agricultural application: predicting corn yields in a crop reporting district in Indiana. The same procedure will be applied to the other eight crop reporting districts to predict the state average yield. The next step is to apply the procedure to eight other Midwest states which are important in U.S. corn production (and for which weather data is readily available) to predict a national corn yield.

Using the neural network model with weekly data for June and July can provide a relatively inexpensive method of predicting the final crop yield by early August.

At this stage, choosing the right neural network software package, the design of the architecture of the network model in terms of number of layers, the number of nodes in the hidden layer, the format of the data, etc. is more of an art than a science. This project is a demonstration of what is possible using neural network software. The future research appears to be limited only by our imagination and ingenuity.

REFERENCES

Anderson, J.A., 1987. "Corn Yield Prediction with a Daily Energy-Crop Growth Variable For Counties in Indiana." PhD Thesis, Purdue University, West Lafayette, Indiana.

Baker, W. Lance, "An Application of Neural Networks To Forecasting Corn Yields," unpublished MS thesis, Purdue University, West Lafayette, Indiana, December 1991.

Bolte, John P. 1989. "Applications of Neural Networks in Agriculture," selected paper, American Society of Agricultural Engineers, annual meeting, New Orleans, Louisiana, December 12-15.

Cowan, I.R. and F.L. Milthorpe, 1968. "Physiological Responses in Relation to the Environment Within the Plant Cover." Pages 107-130 in F.E. Eckhardt (ed.) *Functioning of Terrestrial Ecosystems of the Primary Production Level*. Proc. Kobenhaven symposium. UNESC), page 516.

Dutta, S. and S. Shekhar, 1989. "Bond Rating: A Non-Conservative Application of Neural Networks." Working paper. Computer Science Division, University of California, Berkeley, California.

Hinton, G.E., 1987. "Learning to Recognize Shapes in Parallel Networks," in M. Imbert (ed.) *Proceedings of the Fyssen Conference*, Oxford: Oxford University Press.

Kiniry, J. R. and M.E. Keener, 1982. "An Enzyme Kinematic Equation to Estimate Maize Development Rates." *Agronomy Journal*, 74: Page - 114-119.

Klimasauskas, C. et al. "Neural Works Professional II: Users Guide." Neural Ware, Inc., Pittsburgh, Pennsylvania.

Kunkel, K.E., 1990, "Operational Soil Moisture Estimation for the Midwestern United States," Notes and Correspondence, Journal of *Applied Meteorology,* Volume 29, No. 11: Page 1158-1166.

Leeper, R.A., E.C.A. Runge, and W.M. Walker, 1974. "Effect of Plant-Available Shared Moisture on Corn Yields." *Agronomy Journal*, 66: Page - 723-733.

Nelson, W.L. and R.R. Dale, 1978. "Effect of Trend or Technology Variables and Record Period on Prediction of Corn Yields with Weather Variables." *Journal of Applied Meteorology*, 17: Page - 926-933.

Senjowski, T.J., et.al., 1986. "Net Talk: A Parallel Network That Learns to Read Aloud." Technical Report 86-01, EECS Department, John Hopkins University, Baltimore, Maryland.

Shaw, R.H. and J.E. Newman, 1984. "Weather Stress in the Corn Crop." From *National Corn Handbook,* NCH-18, Cooperative Extension Service, Purdue University, West Lafayette, Indiana.

Smith, J.W., 1914. "The Effect of Weather Upon the Yield of Corn," *Monday Weather Review,* 42: Page - 78-87.

Thompson, L.M., 1969. "Weather and Technology in the Production of

Corn in the U.S. Corn Belt." *Agronomy Journal,* 61: Page 453-456.

Uhrig, J.W. and C.J. Botkin, 1990. "An Application of Neural Networks in Predicting Planted Acreage of Corn and Wheat in the United States," paper presented at the American Agricultural Economics Association annual meeting, Vancouver, British Columbia, August 4-8, 1990.

1. The authors are Professor of Agricultural Economics, Assistant Professor of Agricultural Engineering, and Graduate Research Assistant at Purdue University, West Lafayette, Indiana, respectively.

2. "An operational soil moisture monitoring capability for the Midwestern United States is developed using a multilayer soil water balance model which incorporates daily weather data to calculate precipitation, soil evaporation, plant transpiration, runoff, and drainage through the soil profile. The effects of vegetation on soil evaporation and plant transpiration are incorporated through the use of a model for the growth and development of corn. Data requirements include daily observations of maximum temperature, minimum temperature, precipitation and hourly observations of cloud cover, humidity, and wind speed; this data are collected in real time and aggregated on a climate division scale. The average characteristics of the dominant soils in each climate division are used as representative of that climate division. Using these weather and soil data, the model makes estimates of the current soil moisture status on a climate division basis updated daily. Historical soil moisture estimates using this same model were generated for the period 1949-89 to provide an historical perspective on current soil moisture estimates. This information is accessible to the public through a dial-up computer information system," Kunkel, 1990.