Forecasting with Associative Memories: An Application to Optimal Control

Jing-Jung Chen, Jeffrey H. Dorfman, and Michael E. Wetzstein

Ⅰ. Forecasting with Associative Memories: An Application to Optimal Control

Many of the dynamic systems of interest to economists are fundamentally nonlinear. In most control problem formulations, describing the state of the system with a nonlinear model greatly complicates the solution procedure. Further,the specification of the precise form of the nonlinearities in a system can be difficult. However, if data generated from the system being studied is obtainable, a new type of neural network can be used to produce a linear-in-the- parameters approximate model of the system. This approximate model can then be employed within an optiml control framework to forecast the future state of the system subject to the values of the control variables. This allows control problems with nonlinear systems to be solved for approximately optimal controllers with much simpler solution procedures than are necessary to obtain fully optimal policies. It further insulates the proposed control policies from influences due to model misspecification, since no exact model needs be selected.

In this paper, we will demonstrate that a nonlinear extension of a type of neural network called a multicriteria associative memory model, which was developed by Kalaba and Tesfatsion (1991), can successfully forecast the future state of a highly-nonlinear, dynamic, computer-generated system for peanut plant growth. The approximate model produced by employing

作者分別為東海大學食品科學研究所副教授，美國喬治亞大學農業暨應用經濟學系副教授及教授。

the associative memory techniques is then applied to an irrigation control problem designed to select the net revenue maximizing sequence of irrigation applications.

II. Multicriteria Associative Memory Models

A memory is associative if, having stored the words $Y^{(1)}$, ...,$Y^{(N)}$ at the respective addresses $X^{(1)}$, ..., $X^{(N)}$, it retrieves a word Y close to $Y^{(i)}$ when it is presented with an address X close to $X^{(i)}$. There is considerable flexibility in what it means for one word to be close to another and for one address to be close to another, but generally speaking, associative memories can be programmed to map similar inputs to similar outputs (kohonon, 1989). An underlying assumption of associative memories is the fundamental properties of real numbers (Buck, 1978); without this assumption the theory of associative memories cannot be applied.

Specifically, an associative memory is a system that recalls a particular data vector from a corresponding key vector. This suggests a number of applications, including pattern recognition, vector quantization, error correction, and associative search (chou, 1989).

Linear Multicriteria Associative Memory Models

Linear associative memories are built by finding a matrix M such that

$$y_i = Mx_i, \qquad (1)$$

for a set of vector pairs $\{(x_i, y_i): i = 1, ..., q\}$. M is called an associative memory matrix, which associates the stimulus (input) vectors x_1, ..., x_q with the corresponding response (output) vectors y_1, ..., y_q, in the sense that $\hat{M}x_i$ is "close" to y_i. The x_i vector is an (m x 1)key vector and the y_i vector is an (n x 1) data vector. Horizontal concatenation of x_i's and y_i's form the X and Y matrices,

$$X = [x_1, x_2, ..., x_q]_{mxq}, \quad (2)$$

and

$$Y = [y_1, y_2, ..., y_q]_{nxq}, \quad (3)$$

where X and Y are called the key (or stimulus) matrix and data (or response) matrix, respectively. Together they form the training set for the matrix M. Kohonen and Ruohonen (1972) proposed that the matrix M which satisfies the following matrix equation be used as the matrix operator M of equation (1)

$$Y = MX. \quad (4)$$

solving (4) for M yields

$$\hat{M} = YX^+, \quad (5)$$

where X^+ is the Moore-Penrose generalized inverse of X (Albert, 1972). Onec the memory matrix M is constructed using a finite set of training cases, it can be used to generate an estimate for the actual vector of response vectors on the basis of actual stimulus vector observations. Specifically, given any observed stimulus matrix X^*, an estimate Y for the response matrix Y is determined by

$$\hat{Y} = \hat{M}X^*. \quad (6)$$

As long as the training cases used to construct the associative memory matrix M remain relevant, new response estimates $\hat{Y}$ can be generated for new (m x q) stimulus matrices X^* by repeated application of the simple matrix operation(6).

Unfortunately, a serious practical difficulty is the possibility of the matrix $\hat{M}$ containing large components, which may result in unstable estimates (Kalaba and Tesfatsion, 1991). In particular, the elements of the memory matrix $\hat{M}$ can have large orders of magnitude relative to the components of the training vectors. This problem can occur if the positive semidefinite matrix XX' has one or more small nonzero eigenvalues. When the problem does occur, the resulting estimates $\hat{Y}$ are highly sensitive to observation noise. Furthermore, round-off errors in the computation of the large elements of $\hat{M}$ can induce large estimation errors. As a solution, Kalaba and Tesfatsion (1991) suggest a reformulated $\hat{M}$

$$\hat{M}(\alpha) = \alpha YX'[\alpha XX' + (1 - \alpha)I]^{-1}, \tag{7}$$

Where α is a suitably chosen scalar multiplier between 0 and 1. The matrix $\hat{M}(\alpha)$ converges to $\hat{M} = YX^{+}$ as α converges to 1 (Kalaba and Tesfatsion, 1991).

Let X^{*} be any given (m x q) observation matrix. For each α in [o,1], the parameter vector estimate corresponding to the matrix $\hat{M}(\alpha)$ is

$$\hat{Y}(\alpha) = \hat{M}(\alpha)X^{*}. \tag{8}$$

Kalaba and Tesfatsion (1991) refer to the estimate in (8) as a Multicriteria Associative Memory (MAM) estimate because it is optimal if one minimizes a two- criteria objective function with components representing the magnitude of the matrix's elements (size cost) and the accuracy of the model's forecasts (associative cost). The weight factor α is a tuning device which can be adjusted up or down relative to the importance of the two criteria. Kalaba and Tesfatsion (1991) use α to control for noise in the system output vectors. The objective is to determine, through the training process, a range of values for α which result in accurate forecasts of the system's response to a stimulus.

Nonlinear Multicriteria Associative Momory (NMAM) Models

The MAM approach may not work well in all cases, and therefore, we suggest introducing a nonlinear associative scheme proposed by Poggio (1975). The Poggio method involves the transformation of each stimulus vector into a processed vector that includes up through k^{th}-order distinct products of the components of the stimulus vector. For example, if the stimulus vector x consists of the three scalar components a, b, and c, then the second-order processed vector for x takes the form

$$x^2 = (a, b, c, a^2, ab, ac, b^2, bc, c^2). \tag{9}$$

Concatenate x_i^2's to get an X^{**} stimulus matrix, apply the X^{**} and Y matrices to equation (7), and a second-order polynomial multicriteria associative memory matrix $\hat{N}(\alpha)$ is determined. Kalaba et al. (1992) use a second-order polynomial associative memory matrix in an image processing application. However, they do not combine the polynomial transform with the Kalaba and Tesfatsion (1991) multicriteria objective function.

The covariance matrix for $\hat{N}(\alpha)$ can be derived as follows. The transpose of $\hat{N}(\alpha)$ from (7) is

$$\begin{aligned}\hat{N}'(\alpha) &= [\alpha X^{**}X^{**'} + (1-\alpha)I]^{-1}X^{**}Y'\alpha \\ &= \{(\alpha/\alpha)X^{**}X^{**'} + [(1-\alpha)/\alpha]I\}^{-1}X^{**}Y' \\ &= \{X^{**}X^{**'} + [(1-\alpha)/\alpha]I\}^{-1}X^{**}Y'. \end{aligned} \tag{10}$$

The right hand side of (10) is the same as a ridge regression estimator with k = $(1-\alpha)/\alpha$.The covariance of N' (α) is therefore (Judge et al., 1985)

$$\mathrm{cov}[\hat{N}'(\alpha)] = \sigma^2\{X^{**}X^{**\prime} + [(1-\alpha)/\alpha]I\}^{-1}X^{**}X^{**\prime}\{X^{**}X^{**\prime} + [(1-\alpha)/\alpha]I\}^{-1}$$

$$= \alpha^2\sigma^2[\alpha X^{**}X^{**\prime} + (1-\alpha)I]^{-1}X^{**}X^{**\prime}[\alpha X^{**}X^{**\prime} + (1-\alpha)I]^{-1}, \quad (11)$$

Where σ^2 is the variance of system noise $\sigma^2 = \mathrm{var}(y \mid x^{**})$.

III. Applying Nonlinear Multicriteria Associative Memories to Peanut Growth Forecasting

The yield of a crop at a specific location is the consequence of environmental and management factors. Environmental factors include: soil type and moisture content, pests, blights, and the pattern of daily weather during the growing season. Management factors include: the chosen plant variety, plant density, planting date, fertilization, irrigation, and chemical treatment. Plant scientists have studied the problem of predicting crop yield for many years. A common approach is to employ regression techniques to relate weather variation during the growing season to yield, given crop variety and soil type, and other factors with data collected from carefully controlled experimental field plots. An objective is to derive a model whose predicted yield accurately estimates observed yield, given the weather which occurred.

Research from these studies has resulted in crop growth simulation models which integrate the environmental and management effects in order to predict crop growth. Initially, interest centered on the development of a leaf rather than the growth of an individual plant, a particular crop, or an entire production system. In recent years, crop growth models have been developed which can be used to analyze the effects of a current management decision against various probable future events, thus aiding in the determination of the optimal course of action.

Many agricultural economists have used simple linear models to investigate optimal levels of input use. In recent years, crop response analysis has reconsidered a functional form originally suggested by von Liebig around 1840 (Paris and Knapp, 1989). Paris and Knapp (1989) argue that the correct function involves a linear response followed by a plateau (LRP) or von Liebig model, inspired by his famous "law of the minimum". In the LRP model, a plant will respond linearly to the addition of a limiting input until a different input becomes limiting. While the smooth, concave functions allow for substitution between inputs to plant growth, the von Liebig model reflects a complete lack of substitution possibilities. Berck and Helfand (1990) investigate why other Models seem to work. They find smooth models may work due to heterogeneous application of inputs. However, these models may not have the complexity necessary to model crop growth over the course of the growing season.

In contrast, the equations used in crop growth models are often highly nonlinear both in the states and parameters. In some cases crop growth models are based on experimental data with only general relations among variables provided in constructing a specific set of system equations. The estimation of parmeters in the equations of crop growth models can be formulated as nonlinear optimization problems.

Generation of the Training Sets

To generate data for the training of the associative memory matrices, we used a computer simulation model of peanut growth. PNUTGRO V1.02 is a process- oriented peanut crop growth model which predicts crop development, dry matter growth, leaf area index, and a final peanut yield.

The primary inputs are soil parameters and daily weather data. Other inputs required are cultivar choice, planting date, row and plant spacing, and irrigation management options (Boote et al., 1989). A validation analysis of PNUTGRO V1.02 was performed by comparing its output to the observed results from field experiments (Lacey, 1989).

The geographical region assumed for generating training sets is the southeast Georgia coastal plain. This region has loose friable sandy soils conducive to efficient peanut production and receives an average 42 to 45 inches of rainfall per year. Surveys indicate that 45 to 50 percent of the peanut acreage is irrigated. The peanut crop was assumed to be planted under the following conditions: a planting date of May 1, peanuts of the Florunner variety, row spacing of 0.914m, plant spacing of 0.085m, and a soil type of lake land fine sand.

The daily weather input file was developed from data collected at the U.S. Weather Station in Tifton, Georgia, for the period 1975 to 1989. These data include daily values for precipitation, maximum and minimum temperature, and solar radiation. The simulator was used to simulate current yields based on historical weather patterns. The irrigation decision is assumed to be made every four days (Johnson et al., 1987). This irrigation decision period divides the peanut growing season into 39 discrete time periods. Peanuts are generally not irrigated for the first 50 days, but if irrigation is necessary, 19 mm (.75 inches) of water is assumed to be applied. After 50 days, if irrigation is necessary, 38 mm (1.5 inches) of water is assumed for each irrigation.

The PNUTGRO simulator will calculate irrigation demand according to soil profile data, weather data and crop condition. A threshold of 2.54 mm (0.1 inches) was set for soil moisture inside the simulator. If the irrigation demand is less than the threshold, no irrigation will be applied.

To develop training sets for the associative memory model, fifteen years of historical weather data were given to the simulator and it calculated suggested irrigation schedules. Each irrigation schedule was shifted forward one period (4 days) resulting in 15 new sequences of irrigation schedules and shifted backward one period for another 15 sequences of irrigation schedules. Including the original 15 sequences, a total of 45 (15×3) sequences of irrigation schedules were constrcuted in this manner. Each irrigation schedule was then paired with all 15 years of weather data and all these combinations were fed to the simulator. In this manner, a total of 675 (45×15) sets of output data were obtained from the PNUTGRO simulation model. These data are assumed to possess enough variation to allow successful training of the associative memory model. The data sets were then separated by a random number generator an in-sample group containing 386 observations and an out-of-sample group containing 289 observations. The in-sample observations comprise the training set for the nonlinear multicriteria associative memory matrices.

$\hat{N}(\alpha)$ matrices were then estimated over the in-sample data set for each four day period and tested for validity over the out-of-sample data set. The use of different matrices, $\hat{N}_t(\alpha)$, for each period allows the plant's responses to inputs to vary throughout the growing season.

Lagged biomass, rainfall, irrigation, average temperature, and solar radiation were selected from the set of state variables in the computer simulator. These state variables were assumed to represent the essential influences of the peanut biomass. To avoid excessive parameterization and multicollinearity in the training cases, a partial second order expansion of the five state variables was employed. Thus, the expansion is accomplished by adding temperature squared, the cross product of temperature and solar radiation, and solar radiation squared to the five original state variables. The NMAM model now can be defined as follows

$$Y_t = [Y_t^1, Y_t^2, \ldots, Y_t^{386}]'_{386\times1},$$

$$X_t^{**j} = [Y_{t-1}^j, R_t^j, i_t^j, T_t^j, S_t^j, T_t^j \cdot S_t^j, (T_t^j)^2, (S_t^j)^2]_{1\times8},$$

$$X_t^{**} = [X_t^{**1}, X_t^{**2}, \ldots, X_t^{**386}]'_{386\times8}$$

where $Y_t \equiv$ Biomass (kg/ha) period t,

$R_t \equiv$ Rainfall (mm) period t,

$i_t \equiv$ Irrigation (mm) period t,

$T_t \equiv$ Average temperature (℃) period t,

$S_t \equiv$ Solar radiation (MJ/m2) period t,

j denotes pairing of irrigation and weather, $j = 1, \ldots, 675$,

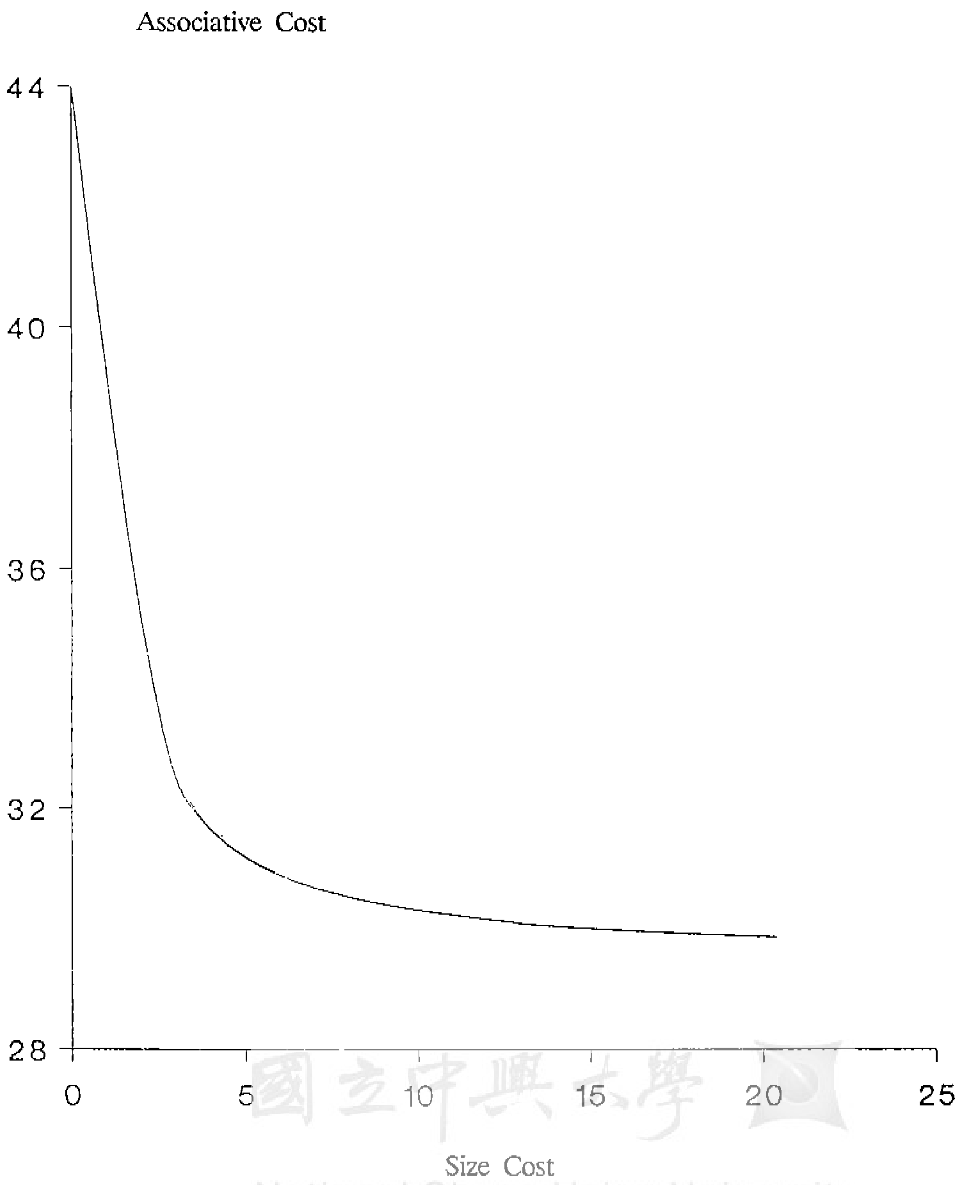
$t \equiv 11, \ldots, 39$.

In the above NMAM model, Y_t is a vector in a selected training set (response matrix) and X_t^{**} is a matrix of system inputs (stimulus matrix). Estimation begins in period 11 because biomass does not change sufficiently in the first ten periods.

Empirical Results

Using the 386 in-sample training cases, NMAM matrices were estimated for a variety of α values. As α varies between 0 and 1, the resulting MSE and norm trace out a cost efficiency frontier (Kalaba and Tesfatsion, 1991). This convex surface depicts the tradeoffs between in-sample forecasting accuracy and the size of the elements of the associative memory matrix. Due to multicollinearity- induced ill conditioning of many data matrices in economics, some reduction

Figure 1. Cost Frontier for NMAM Matrix



Values shown are for period 36.

Table 1. R^2-values for the NMAM with $\alpha = 0.05$

Time (t)	in sample	out-of sample
11	.6618	.6656
12	.7935	.7909
13	.8804	.8852
14	.8993	.8993
15	.9217	.9077
16	.9224	.9164
17	.9197	.9191
18	.9440	.9384
19	.9540	.9493
20	.9503	.9483
21	.9308	.9194
22	.9368	.9269
23	.9546	.9511
24	.9702	.9666
25	.9777	.9771
26	.9734	.9702
27	.9699	.9640
28	.9855	.9860
29	.9829	.9795
30	.9795	.9797
31	.9846	.9841
32	.9835	.9853
33	.9884	.9872
34	.9879	.9860
35	.9840	.9820
36	.9865	.9851
37	.9793	.9803
38	.9844	.9840
39	.9968	.9957

in the size of estimated coefficients is often desirable (Hill and Judge, 1987). The cost efficiency frontier for the NMAM matrix for time period 36 is shown in figure 1. After examining the tradeoff between size and MSE and calculating the standard errors of the coefficients for different valuse of α , a value of α =0.05 was chosen as "optimal" for the peanut forecasting application.

The R^2-values for the NMAM model with α =0.05 are shown in table 1 for all 29 time periods. The R^2-values increase through the growing season, reaching more than 0.99 by the last period. The NMAM model shows essentially no drop off in forecasting accuracy in the out-of-sample data. This performance validates the forecasting ability of the model and suggests that the nonlinear associative memory matrices can successfully mimic the highly complex PNUTGRO simulator in the forecasting of peanut biomass.

Table 2 displays the coefficients of the NMAM model and their t-values. Eighty-six percent of the coefficients are significant at the 0.10 level, including all of the coefficients on lagged biomass and 25 (out of 29) of the coefficients on irrigation. The coefficients for the privious period biomass ($\hat{N}_t[1,1]$) are generally positive in sign and close to 1. The model is very consistent in this result. In general, the coefficients for rainfall and irrigation ($\hat{N}_t[1,2]$ and $N_t[1,3]$) are positive with occasional negative signs in time periods where increased water could harm the plant. Temperature and solar radiation accelerate the evaporation rate, such that stress on these two factors will cause the plant weight (biomass) to decrease in some periods (coefficients have negative sign). Therefore, there are a number of negative coefficints for these two parameters and their cross product.

The empirical results for forecasting peanut growth with a partial second- order mulicriteria associative memory model are acceptable for use in an irrigation scheduling model.

Table 2. Nonlinear Multicriteria Associate Memory Matrix Estimates (α =0.5)

Time (t)	$\hat{N}[1,1]$	$\hat{N}[1,2]$	$\hat{N}[1,3]$	$\hat{N}[1,4]$	$\hat{N}[1,5]$	$\hat{N}[1,6]$	$\hat{N}[1,7]$	$\hat{N}[1,8]$
11	1.2419	2.0867	3.3576	-29.045	36.650	1.7315	-0.1725	-1.8064
	(22.177)	(8.6946)	(3.9971)	(-5.3002)	(6.3190)	(4.8097)	(-0.7188)	(-7.5267)
12	1.4129	1.0579	4.5843	-51.456	45.511	-0.8474	1.4268	-0.4423
	(27.171)	(1.8891)	(4.5843)	(-6.7352)	(5.4180)	(-4.2370)	(5.9450)	(-2.7644)
13	1.3610	5.3939	4.2102	10.5075	2.6667	-0.1720	-0.5374	0.0353
	(48.607)	(5.6186)	(5.5397)	(1.2217)	(0.2554)	(-0.4300)	(-1.6794)	(0.2206)
14	1.1623	6.2591	3.1047	-54.077	80.995	-3.4170	2.12125	0.1980
	(48.29)	(10.432)	(5.1745)	(-5.4733)	(6.2113)	(-6.1018)	(5.3030)	(0.9900)
15	1.2404	1.8833	2.8482	5.58715	6.1278	-2.6709	0.5570	1.7475
	(62.020)	(4.7083)	(6.4732)	(0.8569)	(0.8606)	(-4.4515)	(1.5472)	(4.3688)
16	1.6659	4.8657	3.6408	-3.6541	46.778	0.1921	-0.2592	-1.7201
	(58.325)	(5.7925)	(7.0015)	(-0.3035)	(3.2850)	(0.2825)	(-0.5400)	(-4.7781)
17	1.0653	2.3770	2.4765	-2.1079	4.0829	0.7936	0.2236	-0.6373
	(44.387)	(2.7011)	(4.1275)	(-0.1267)	(0.1944)	(0.7348)	(0.2942)	(-1.3277)
18	1.1232	0.6130	1.6246	20.6945	-20.013	-1.5323	0.5857	1.3291
	(70.200)	(0.8066)	(2.7077)	(1.2202)	(-0.9908)	(-1.4188)	(0.7321)	(3.0207)
19	1.0445	6.1686	1.0003	57.8575	54.012	2.2427	-2.9966	-2.6155
	(65.281)	(4.2838)	(1.5630)	(3.7667)	(3.0276)	(2.4377)	(-4.4068)	(-4.3592)
20	1.0381	6.8199	7.0953	-37.847	90.019	-12.741	5.3347	6.0962
	(86.508)	(6.8199)	(10.434)	(-2.4965)	(4.2948)	(-6.9243)	(6.0622)	(5.0802)
21	1.0286	5.6534	7.2213	83.2075	-24.467	1.7935	-2.8124	-0.9793
	(64.281)	(6.1450)	(8.5968)	(3.6882)	(-08689)	(0.8968)	(-2.6041)	(-0.7898)
22	1.0096	1.4682	5.7378	152.275	3.4723	0.4512	-5.049	-0.3948
	(63.100)	(0.5647)	(6.8307)	(7.0364)	(0.1262)	(0.4029)	(-7.0132)	(-0.7050)
23	1.0210	6.6656	6.8172	38.8705	69.807	2.0607	-2.4216	-3.5618
	(85.083)	(3.3328)	(8.5215)	(2.3248)	(3.6974)	(1.7173)	(-4.3243)	(-4.6866)
24	1.0603	0.4931	3.4483	53.4525	-0.4119	-0.1941	-1.8864	0.4322
	(88.358)	(0.6488)	(5.0710)	(3.8733)	(-0.0218)	(-0.1565)	(-3.3686)	(0.5687)
25	1.0146	1.0271	0.9600	-1.3695	40.632	0.7178	0.1279	-1.9770
	(126.82)	(1.6048)	(1.6000)	(-0.1214)	(3.2247)	(0.9445)	(0.2907)	(-4.4932)
26	0.9896	2.6416	1.8416	13.7285	53.604	-3.6968	0.7064	1.4457
	(123.70)	(1.7849)	(2.5578)	(0.8891)	(2.9980)	(-3.8508)	(1.1773)	(2.7802)
27	0.9998	1.8317	4.7636	70.1935	11.734	-1.4696	-1.865	0.90613
	(83.316)	(3.0528)	(5.9545)	(4.3761)	(0.6984)	(-0.8961)	(-3.1092)	(0.8583)
28	0.9952	-0.4258	0.9342	66.0365	-12.808	2.3170	-2.4978	-1.0694
	(124.40)	(-0.8871)	(1.6682)	(5.7724)	(-0.9853)	(2.6330)	(-4.8035)	(-1.6709)
29	0.9923	-1.2121	1.8956	37.677	64.370	-5.9224	0.3507	2.8697
	(124.03)	(-1.5151)	(3.1593)	3.1503)	(4.0741)	(-4.3547)	(0.5480)	(3.2610)
30	1.0035	2.3535	3.1316	70.8225	-67.861	0.5944	-1.3173	2.5125
	(125.43)	(1.5484)	(4.6053)	(3.9699)	(-3.1015)	(-0.7430)	(-2.0583)	(5.2344)
31	1.0058	-0.5189	3.7183	37.4455	37.495	-2.3671	-0.7057	0.7879
	(125.72)	(-0.6828)	(6.6398)	(2.7452)	(1.9986)	(-2.6899)	(-1.3571)	(1.7907)
32	0.9815	0.2825	3.2163	127.875	-103.55	2.6360	-3.8804	1.3970
	(122.68)	(0.5885)	(5.3605)	(8.4352)	(-4.9405)	(2.8652)	(-7.4623)	(2.3283)
33	0.9682	1.4512	1.5237	79.1205	46.991	-0.7266	-2.3944	-0.6793
	(242.05)	(2.5914)	(2.9302)	(5.8694)	(2.9969)	(-0.5860)	(-3.1505)	(-0.8087)
34	0.9558	-0.4058	0.5752	95.938	49.086	1.1766	-3.2713	-2.6293
	(119.47)	(-0.6763)	(1.0271)	(7.5423)	(2.8739)	(0.9192)	(-5.1114)	(-3.1301)
35	0.9532	0.6779	0.9236	90.2675	11.537	0.4415	-2.5148	-0.7544
	(119.15)	(1.1298)	(1.5393)	(6.6766)	(0.5557)	(0.4626)	(-3.9294)	(-0.9926)
36	0.9419	1.7662	1.4504	130.125	-45.697	3.8537	-4.3916	-1.3786
	(235.47)	(2.3239)	(2.1329)	(11.374)	(-2.7662)	(4.8171)	(-9.9809)	(-2.8721)
37	0.9605	-0.5900	1.9029	-38.502	125.82	-3.2109	2.3698	-2.2618
	(80.041)	(-0.3207)	(2.2654)	(-3.2409)	(10.245)	(-5.3515)	(5.3859)	(-3.5341)
38	0.9914	-16.252	1.8748	108.205	84.496	6.6223	-5.6236	-11.225
	(82.616)	(-1.1381)	(2.1305)	(7.6414)	(6.7274)	(1.8813)	(-4.2603)	(-4.1268)
39	0.9586	2.0074	-2.0671	0.81225	-0.0257	-3.8788	2.6178	-1.98423
	(119.82)	(3.1366)	(-3.0399)	(0.3329)	(-0.6425)	(-3.0303)	(4.3630)	(-3.1005)

T-values in Parenthesis.

IV. The Irrigation Scheduling Model

Optimization of irrigation management strategies is actively pursued by many producers. The decision to begin irrigation is at the discretion of a manager who must base the decision on an analysis of available information and a subjective consideration of future enents. Irrigation scheduling is a particularly important problem in Georgia. While annual rainfall is adequate for most agricultural crops in the state, the distribution of precipitation is often highly unpredictable. Large portions of the growing season are characterized by sporadic rainfall, excessive rainfall, or long periods of drought. The primary effect of irrigation is to offset the impact of rainfall variability on crop yields.

Earlier Work on Optimal Irrigation Schedules

Considerable previous research exists on the allocation of irrigation water and its effect on yield. The management of farm irrigation systems involves the choice of methods, time, and quantity of water applied. A wide range of modeling procedures has been used for irrigation scheduling. These methods include the use of inventory models (Fogel, Duckstein, and Kisiel, 1976), mathematical programming, dynamic programming, simulation, stochastic dynamic programming, optimal control theory and stochastic dominance. Boggess et al. (1983) reviewed approximately 50 articles to determine the specific irrigation management objectives and how the issue of variability (risk or uncertainty) was addressed.

Although limitations due to uncertainties imposed by the crop growth models have been expressed, dynamic programming has been used for more than 20 year in crops irrigation problems. Burt (1966) was among the first to apply dynamic programming to groundwater management problems, and provides an excellent discussion of the potential usefulness of the technique in firm analysis. One of the earlier irrigation scheduling models, presented by Burt and Stauber (1971), is designed to analyze the feasibility of irrigation investment in a subhumid climate. A major objective of their study is to develop and illustrate the methodology of deriving optimal irrigation polices using stochastic dynamic programming. However, dynamic programming studies are limited because to maintain computational tractability, the number of state variables describing the system must be kept to a minimum. Reducing the number of state variables in normal econometric or computer simulation models may seriously limit the accuracy in forecasting the state of the system. The use of neural networks such as the multicriteria associative memory model presented here may allow greater use of dynamic programming (as well as the optimal control techniques used below) by allowing accurate state forecasts based on a smaller number of state variables.

The Peanut Market

Marketing quotas and acreage allotments have been in effect for peanuts since 1949 and import quotas since 1953. Producers are guaranteed a price considerably above the world price for their quota peanuts. Producers can produce in excess of their quota, within their acreage allotments, but the excess quantity, referred to as additionals, receives a lower, market-determined price. These additionl peanuts are usually destined to be crushed for oil or exported.

Although quota and additional peanuts enter different marketing channels, they are usually contracted simultaneously. The contracting of quota deliveries is not required, but additionals must be contracted before August 1 or each crop year. Farmers are generally reluctant to contract their quota unless premiums are offered. Shellers prefer that farmers contract quota and additionals together as a means to reduce sheller's risk. For the simulations presented below a 1988 quota support price of \$618.83 per ton and a price for additionals of \$216.67 per ton were assumed. For simplicity in this study, a grower will contract quota and aditional peanuts simultaneously. Also, production is assumed to always exceed the quota amount.

The Optimization Model

Farmers are assumed to maximize net revenues per acre. Since the NMAM model above forecasts peanut biomass, prediction of peanut yield is based on plant biomass. Biomass is the weight of the whole peanut plant, including plant leaf, stem, root, and flower. In the irrigation control model, the biomass is converted to peanut yield at the final time period. The biomass to yield conversion ratio is assumed to be 0.467, which is the 15-year average of the biomass to yield conversion ratio from the simulation output using historical weather data.

Because the computer simulation model used to generate the training sets for the NMAM crop growth model assumes all nonwater inputs are present in optimal amounts, we cannot use this model to select other variable inputs in conjunction with irrigation. Therefore, the optimization problem has irrigation, i_t, as the single control variable and can be written as

$$\max\ 0.467[p_s y_s + p_m(Y_T - y_s)](1+r)^{-T} - \sum_{t=11}^{T} W_t i_t (1+r)^{-t} - VC \tag{12}$$

$$\text{s.t.}\ Y_t = \hat{N}_t Z_t + \hat{n}_t i_t + \varepsilon_t \tag{13}$$

$$W_t = \gamma + \delta i_t \tag{14}$$

where p_s is the quota support price (\$/kg), y_s is the biomass (kg/ha) necessary to fill the farmer's quota (after converting to yield), p_m is the price (\$/kg) for additionals, Y_T is actual biomass (kg/ha) at harvest, i_t is the irrigation (mm)applied in period t, z_t contains all the state variables in the NMAM model except for irrigation, N_t and n_t contain the respective coefficients from the estimated NMAM model for that period, w_t represents the cost of irrigation (\$/ha- mm), and r is the discount rate. The first term of equation (12) is gross revenue, the second term is irrigation variable cost, and the third term VC represents other production costs such as tractors, land, and seeds. Because the farmer is assumed to always fill his quota and the other variable costs can be ignored. The problem can be simplified further by substituting the irrigation cost equation in (14) into the objective function in place of w_t. This gives

$$\max\ 0.467 p_m Y_T (1+r)^{-1} - \sum_{t=11}^{T} (r i_t + \delta i_t^2)(1+r)^{-t} \tag{15}$$

$$\text{s.t.}\ Y_t = z_t \hat{N}_t + \hat{n}_t i_t + \varepsilon_t. \tag{16}$$

Given the estimates of N_t and n_t from the nonlinear multicriteria associative memory model, the above problem is solved as a linear-quadratic, closed-loop optimal control problem (cf. Chow, 1975). The result is a linear feedback rule which expresses the optimal irrigation amount as a function of Y_{t-1}, z_{t-1}, and i_{t-1}.

Since biomass does not change in the first ten 4 day periods, the control problem started with period 11 (t=11, 12, ..., 39). Because irrigation can only be applied in discrete amounts, simple assumptions were used to discretize the continuous values of the optimal irrigation amount provided by solving the control problem. For periods 11 and 12, if irrigation demand (i_t) is positive and the rainfall in that period is less than 19 millimeters, 19 millimeters of irrigation is applied. For periods 13 through 39 (or harvest), if the demand for irrigation in a period is positive and the rainfall in that period is less than 38 millimeters, 38 millimeters of water is applied. The smaller amount of water applied in the first two periods is in keeping with actual farming practices and the set amounts of irrigation are due to the difficulties in adjusting the settings on the standard type of center-pivot irrigation typically used in South Georgia peanut farming.

Simulation Results

A quota of 1.503 tons per acre (3369 kg/ha) belonging to a specific peanut grower in southeastern Georgia is used for calculating net revenue (to include the extra price received on the quota peanuts in the objective function). Converting the quota price and the market price for additionals to $/kg gives p_s=$0.6821/kg and p_m=$0.2388/kg. The total operating costs of an open hole well irrigation system is assumed to be $0.37 per ha-mm with a fixed cost of $190.90 per hectare (Westberry, 1989). Therefore, δ is set equal to 0.00003 and γ is set at 0.369. An annual discount rate of eight percent is assumend; in the model this is adjusted to reflect the four day period.

Results for optimal irrigation schedules, timing of harvest, estimated yields, amount of net revenues, and total amount of rainfall in the growing season are presented in table 3 for the 15 years of data used in estimating the crop growth model (1975-1989). As indicated from this table, the most active optimal irrigation schedules are concentrated in periods 11 to 16 and periods 20 to 23. Experience and research have shown three critical periods for crop water use by peanuts. The first occurs with planting and germination. Producers can wait to plant until the soil moisture is adequate for germination, or they can irrigate before planting with 1/2 to 3/4 inches of water. This period is not included in the optimal control model ,and thus it is assumed that a producer will wait or irrigate to obtain adequate moisture for planting. As a practical fact, farmers usually do not irrigate during the first 50 days after planting. A second critical water requirement occurs during the pod setting period which is usually from 50 to 70 days after planting, and a third critical period is from 80 days until harvest as pods mature (Johnson et al., 1987). As the results of the optimal control model indicate, there is a gap between periods 17 and 20. This period corresponds to the peanut flowering and pegging season. During this period, too much water can delay formation of flowers or, depending on the stress severity, completely inhibit flowering. Specifically, the optimal irrigation schedule seems to be consistent with actual developing and growing stages of peanuts. Generally, 8 to 10 irrigations were suggested by the control problem.

Peanut maturity greatly influences yield, quality, sound mature kernel content, and subsequent dollar value per ton. "Shell-out," the most commonly used method of determining peanut maturity, involves the removal and shelling of all pods from several plants randomly selected from a field. The variety of peanuts assumed in our simulation generally requires between 130 and 145 days (periods 33 to 37) from planting to maturity (Johnson et al., 1987), which is consistent with the results. The criterion of harvest timing is based on the predicted yield from the peanut growth simulation model; peanuts are harvested in the highest yield period. Most of the crop years are harvested in periods 36, 37, and 38.

To estimate the gains to farmers from using the proposed feedback rule, three altrenative irrigation rules were postulated: 1) always irrigate, 2) never irrigate, and 3) irrigate if it did not rain in the current 4 day period. The net revenues under each of these alternatives were simulated for the same fifteen years. The irrigation amounts were identical to those of the optimal control model (19 or 38 mm). For all 15 years, the net revenues of these three irrigation decision rules are lower than that derived from the optimal feedback rule. The values are displayed for all four rules in table 4. When the sample variances of the net revenue simulations are calculated for the four rules, the optimal feedback rule is found to have a lower variance than two of the alternative rules, but the lowest variance is obtained with the always irrigate rule. Thus, a risk averse farmer faces a

Table 3. Optimal Irrigation Schedules

Time (t)	Year '75	'76	'77	'78	'79	'80	'81	'82	'83	'84	'85	'86	'87	'88	'89
11	X	X	X			X	X	X		X	X	X	X	X	X
12		X	X	X	X	X	X	X	X	X	X		X	X	X
13	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
14	I	I	I	I	I	I	I	I	I	I	I		I	I	I
15	I	I		I	I	I	I		I	I	I	I	I	I	I
16	I	I	I	I	I	I	I	I	I	I	I	I	I		I
17															
18															
19															
20	I	I	I	I	I	I	I		I	I	I	I	I	I	I
21	I	I	I		I	I	I	I	I	I	I	I	I		I
22	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
23	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
24															
25															
26															
27															
28															
29															
30															
31															
32															
33															
34															
35										H					
36				H								H	H		
37	H		H					H	H		H				H
38							H							H	
39		H			H	H									
Yield[a]	5629	6304	5524	5577	6294	5695	5825	5820	5263	5570	5985	5320	5209	5765	5915
Rainfall[b]	500	696	591	466	470	445	448	514	440	462	487	380	493	520	445
NR[c]	2438	2582	2421	2442	2587	2442	2474	2503	2354	2422	2514	2383	2437	2488	2497

[a]:kg/ha.

[b]:Total rainfall from time period 1 to 39 millimeters.

[c]:Net return $/ha.

X and I: Irrigation applied with 19 mm and 38 mm, respectively.

H: Harvest.

Table 4. Comparison of Net Revenues under Alternative Irrigation Rules

		Irrigation Rule			
	Year	Opt. Feedback	Never	Always	If no Rain
	1975	2438	2349	2367	2381
	1976	2582	2503	2460	2505
	1977	2421	2321	2338	2352
	1978	2442	2362	2397	2359
	1979	2587	2524	2502	2518
	1980	2442	2363	2356	2374
	1981	2474	2387	2376	2367
	1982	2503	2411	2431	2412
	1983	2354	2273	2293	2321
	1984	2422	2313	2364	2313
	1985	2514	2414	2432	2415
	1986	2383	2296	2331	2267
	1987	2437	2233	2263	2259
	1988	2488	2385	2398	2419
	1989	2497	2398	2416	2436
mean		2466	2369	2382	2380
variance		4246	6154	3921	5599

Figure 2. Irrigation & Rainfall Patterns

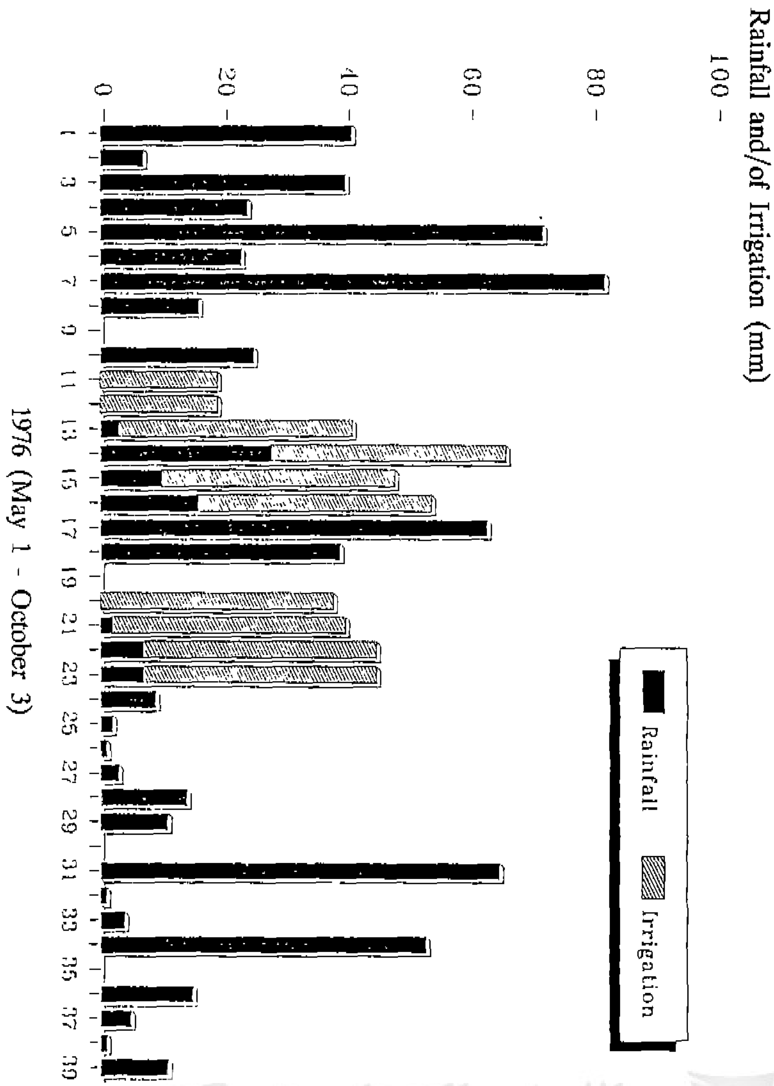
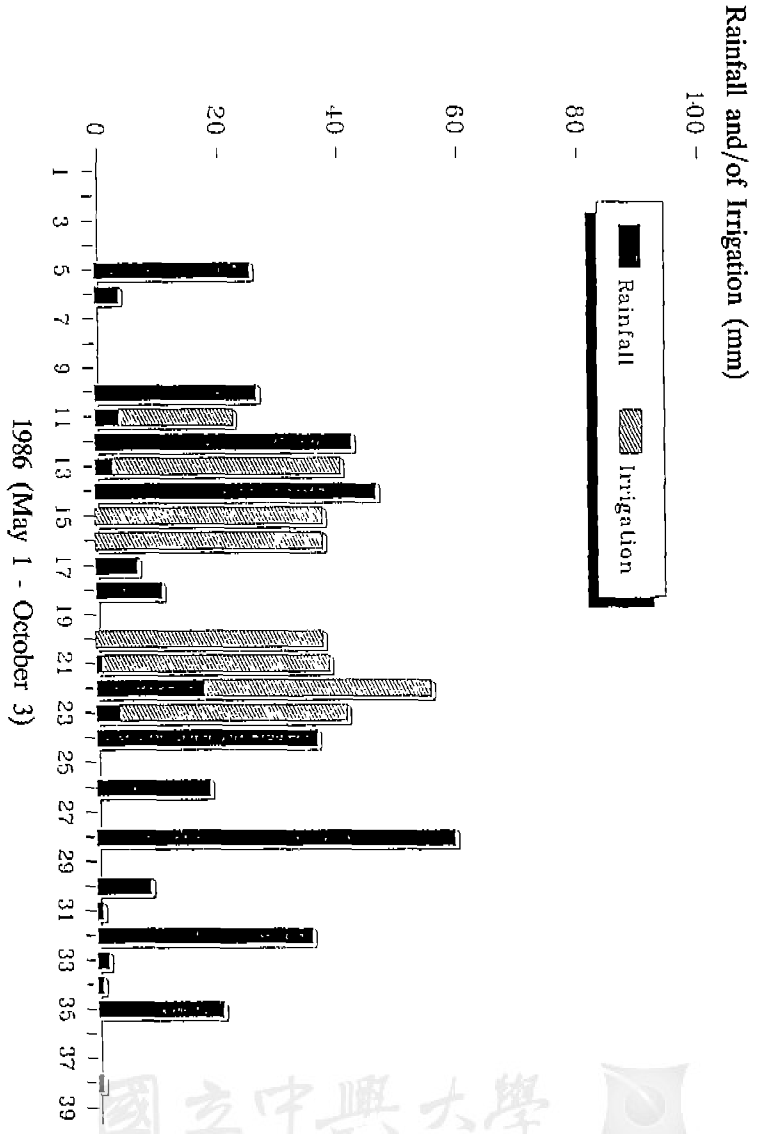


Figure 3. Irrigation & Rainfall Patterns



tradeoff between the higher expected net revenue of the optimal feedback rule and the lower variance of net revenue (risk) of the always irrigate rule. The other two irrigation rules are dominated in a mean-variance sense by the optimal feedback and the always irrigate rules.

In order to examine the relationship between rainfall and the optimal irrigation schedules, two specific years are selected as examples and their results are presented in figures 2 and 3. These two years are 1976 (a wet year) and 1986 (a dry year) with precipitation in the peanut growth time period 1 to 39 (May 1 to October 3) of 696 mm and 380 mm, respectively. Note that due to the distribution of rainfall, the model actually suggests irrigation ten times in the "wet" year compared to only 8 times in the "dry" year. This is because the rain in the wet year is concentrated at the start and end of the growing season, while the critical period when water stress can hinder plant growth occurs in the middle of the growing season.

V. Conclusions

This paper performs the first application of nonlinear multicriteria associative memories, combining techniques developed by Kalaba and Tesfatsion (1991) and Poggio (1975). This new technique is used to forecast peanut crop growth over a sequence of four day horizons. The model is trained with data generated by a complex computer simulation model with 26 state variables and over 4000 equations, many of which are nonlinear. The NMAM model forecasts peanut biomass with a high degree of accuracy using only 8 state variables and is validated by out-of-sample forecasting using observations not in the training set with no appreciable drop-off in accuracy. Thus, this procedure provides a much simpler model than the computer simulator with little resulting loss of precision.

Because the NMAM model is linear in biomass and irrigation, the crop growth model allows optimal irrigation schedules to be determined using a standard linear-quadratic closed-loop control formulation. If the original simulator had been employed, complex iterative routines would have been needed to solve for the optimal controllers. The simplicity gained is important because linear-quadratic control provdes a linear feedback rule for the optimal controller. Such a rule could be implemented in a spreadsheet environment, such as LOTUS, making the methodology developed here available to farm managers without the need for them to learn how to run a complex computer simulation or optimization model.

The 15 years of simulations performed with the irrigation control model produced irrigation decisions which appeared reasonable. The optimal feedback rule outperformed three alternative irrigation decision rules in terms of net revenue in all fifteen years. The optimal rule also had a favorable income- stabilizing effect, having a lower variance of net revenue than all the alternatives except an always irrigate rule.

References

1. Albert, A. *Regression and the Moore-Penrose Pesudoinverse* (New York: Academic Press, 1972).

2. Berck, P. and G. Helfand. "Reconciling the von Libing and Differentiable Crop Production functions." *Amer. J. Agr. Econ.* 72(1990):985-996.

3. Boggess, W. G., G. G. Lynne, J. W. Jones, and D. P. Swaney. "Risk-Return Assessment of Irrigation Decisions in Humid Regions." *Southern J. Agr. Econ.* 15(1983):135-143.

4. Boote, K. J., J. W. Jones, G. Hoogenboom, G. G. Wilkerson, and S. S. Jagtap. "Peanut Crop Growth Simulation Model." Florida Agricultural Experiment Station Journal No. 842(1989). Agricultural Engineering Dept., University of Florida, Gainsville, FL 32611.

5. Burt, O. R. "Economic Control of Groundwater reserves." *Amer. J. Agr.*

Econ. 8(1966):632-647.

6. Burt, O. R. "Dynamic Programming: Has Its Day Arrived?" *West. J. Agr. Econ.* 7(1982):381-393.

7. Burt, O. R. and M. S. Stauber. "Economic Analysis of Irrigation in Subhumid Climates." *Amer. J. Agr. Econ.* 53(1971):33-46.

8. Chou, P. A. "The Capacity of the Kanerva Associative Memory." *IEEE TRans. Inform. Theory* 35(1989):281-298.

9. Chow, G. c. *Analysis and Control of Dynamic Economic Systems* (New York: John Wiley, 1975).

10. Hill, R. C., and G. Judge. "Improved Prediction in the Presence of Multicollinearity." *J. econometrics* 35(1987):83-100.

11. Johnson, W. C. et al. *Georgia Peanut Production Guide.* Cooperative Extension Service, The University of Georgia, Athens 1987.

12. Judeg, G. G., W. E. Griffiths, R. C. Hill, H. Lutkepohl, and T.-C. Lee. *The Theory and Practice of Econometrics* 2nd Edition (New York: John Wiley Sons, 1985).

13. Kalaba, R., Z. Lichtenstein, T. Simchony, and L. Tesfatsion. "Linear and Nonlinear Associative Memories for Parameter Estimation." *Inform. Sci,* (1992)

14. Kalaba, R. and L. Tesfatsion. "Obtaining Initial Parameter Estimates for Nonlinear Systems Using Multicriteria Associative Memories." *Comp. Sci. in Econ. Manag,* 3(1991)

15. Kohonen, T. *Self-Organization and Associative Memory.* 3rd Edition (Berlin: Springer-Verlag, 1989).

16. Kohonen, T. and M. Ruohonen. "Representation of Associated Data by Matrix Operator." *IEEE Trans. Comput.* 21(1972):353-359.

17. Lacey, D. K. *Peanut Insect Pest Management Using a Crop Growth Simulation Model Based Expert System,* unpublished Masters thesis, U niversity of Georgia, 1989.

18. Paris, Q. and K. Knapp. "Estimation of von Liedig Response Functions." *Amer. J. Agr.* Econ. 71(1989):178-186.

19. Poggio. T. "on Optimal Nonlinear Associative Recall." *Boil. Cybernetics* 19 (1975):201-209.

20. Westberry, G. O. *Economic Analysis of Sprindler Irrigation Systems* (Athens: Cooperative Extension Service/Univ. of Georgia, July 1989),

相聯記憶體巨陣預測法：最適控制之應用實例

Jing-Jung Chen, Jeffrey H. Dorfman,

and Michael E. Wetzstein

摘 要

應用動態模型來找出最適解通常需要找出整個系統中控制狀變數的最佳策略，尤其是非線性模型更需要用到數值趨近的方法，其假設條件也比較複雜。利用相聯記憶體巨陣，可以利用線性模型及線性二次模型來預測非線性系統之未來狀態。使用相關記憶體巨陣方法之優點在於不必精確的描述整個系統爲何型態的非線性式，大大的簡化了預測之假設條件。本篇報告系利用相職憶巨陣方法來預測美國喬治亞洲花生的整個生長過程，並以最適控制法求出花生生長過程中之最佳灌溉條件。