October 2003 FREC SP03-03

PREDICTIVE TIME MODEL OF AN ANGLIA AUTOFLOW MECHANICAL CHICKEN CATCHING SYSTEM

Saravanan Ramasamy,
Operations Research, University of Delaware

Eric R. Benson,
Bioresources Engineering, University of Delaware

John C. Bernard,
Food and Resource Economics, University of Delaware

Garrett L. Van Wicklen,
Bioresources Engineering, University of Delaware

FOOD & RESOURCE ECONOMICS

FREC Staff Paper

Department of Food & Resource Economics • College of Agriculture & Natural Resources • University of Delaware

ABSTRACT

In this project, a predictive time model was developed for an Anglia Autoflow mechanical chicken catching system. At the completion of poultry growout, hand labor is currently used to collect the birds from the house, although some integrators are beginning to incorporate mechanical catching equipment. Several regression models were investigated with the objective of predicting the time taken to catch the chicken. A regression model relating distance to total time (sum of packing time, catching time, movement to catching and movement to packing) provided the best performance. The model was based on data collected from poultry farms on the Delmarva Peninsula during a six-month period. Statistical Analysis System (SAS) and NeuroShell Easy Predictor were used to build the regression and neural network models respectively. Model adequacy was established by both visual inspection and statistical techniques. The models were validated with experimental results not incorporated into the initial model.

INTRODUCTION

Mechanical poultry catching systems have been under investigation for a number of years. The cost and complexity of mechanical catching systems has been a factor in the relatively slow adoption of mechanical catching equipment. These systems have the potential to reduce the labor and the time required for catching.

The objectives of this study were:
1) to develop a predictive time model for an Anglia Autoflow Chicken Catching System.
2) the tools used for developing the model were regression and neural networks.
3) to compare the predictions from regression and neural networks models.

It was assumed that the time required to catch the chicken depends on the distance between the packing unit and the harvester in the chicken house.

LITERATURE REVIEW

Researchers have reported many advantages and problems in automated broiler catching. Automated catching has taken into consideration methods of reducing injury and physiological stress on the bird. However, certain improvements have to be made in the automated process to reduce breakdowns of harvesters and packers.

Farsie et al., (1983) used a long tined scoop mounted on a loader to harvest the birds. The long tined scoop reduced wing and breast bruising relative to manual catching. The scoop also encountered problems during operation, such as difficulty in equipment maneuverability, driving over birds in darkness and scoop prongs digging into litter.

Lacy and Czarick (1998) have found that mechanical catching would improve the bird welfare through reduced injury and stress compared to manual catching. Lacy (1994) tested a mechanical harvester to determine its ability. The harvester successfully caught birds however, certain modifications were needed before it could be used by the poultry industry. This harvester caught too many birds, which caused difficulties for its conveyors. Inconsistent performance of the sensing system, which determined whether the cage doors were open for loading the birds, also caused problems. These problems made it difficult to accurately determine the average speed of catching.

Lacy (1992) evaluated the performance of a mechanical catching system consisting of a self-propelled tractor, a catching mechanism of six rubber-fingered rotors, associated conveyors and a caging mechanism that could fit 5 or 6 live haul cages. The harvester was able to catch and convey over 12,000 broilers per hour. Labor requirements were reduced by more than 50%, but the reliability and maintenance costs had to be determined.

Duncan et al., (1985) investigated the bird physiological stress following mechanical and hand catching. They detected statistically significant differences in the heart trace between the manually-caught and machine-caught birds. The trace in both systems rose to same levels immediately after catching, however the trace of machine caught birds dropped quicker and the average heart rate was significantly lower than in manually harvested birds.

Ramasamy et al., (2003) reported on the efficiencies of a commercial mechanical catching system. Potential improvements to the catching system included truck scheduling and preventive maintenance of machines. Ramasamy et al., also found that using two harvesters instead of one harvester would decrease the total time required to harvest birds. Sources of idle time such as truck availability and equipment repair reflected the inefficiencies of Anglia Autoflow catching system.

Salle et al., (2002) used Artificial Neural Networks (ANN) to estimate the production parameters in broilers. They selected their model based upon high R^2 value and low mean square error (MSE) value. The ANN method supplied tools for the decisions made by the technical staff to be based in objective, scientific criteria. This method also allowed simulations of the consequences related to these decisions, and reports the contribution of each variable to the poultry production parameters under study.

Thus several attempts have been made to improve the mechanical catching process taking into account the stress and injury in birds. In this research a predictive time model is developed where for a particular distance from the packers to the harvesters the total time taken to harvest the birds is predicted.

DATA AND METHODOLOGY

The data pertaining to catching time, packing time, movement to catching and movement to packing time collected from ten poultry farms during a period from January – July 2002 was used in this study. The data was randomly split into two categories: 953 observations (80% of data) for building the model (training data set) and 238 observations (20% of data) for validating the model (validation data set). Regression methods and Neural Networks were used for modeling the total time in Anglia Autoflow Chicken catching system. Model adequacy was established for both techniques. The models were validated for both regression and neural networks and the model performances were compared.

Regression

Multiple regression models were developed from the available data to predict the overall catching time. SAS (SAS Inc., 1990) was used to build the regression models. Movement to packing, movement to catching, packing time and distance were taken as independent variables and catching time was taken as the dependent variable. This model did not predict catching time within an acceptable range (95% confidence interval).

Secondly packing time was taken as dependent variable and the other variables were taken as independent variables. The model was not statistically significant as the p-value was 0.1763 (>0.05) and had a very low F-value (1.4472). The economic model formed was

$$\text{Packing time} = f\,(\text{catching time, movement, distance}) \qquad (1)$$

The regression model formed was

$$\text{Packing time} = \beta_0 + \beta_1 * \text{catching time} + \beta_2 * \text{movement time} + \beta_3 * \text{distance} \qquad (2)$$

where, β_0 is the intercept, $\beta_1, \beta_2, \beta_3$ are the regression coefficients for catching time, movement time and distance respectively. The regression equation for the predicted packing time was given by

$$\text{Packing time} = 59.8299 + 0.1789 * \text{catching time} - 0.0015 * \text{movement time} - 0.1089 * \text{distance} \quad (3)$$

Unfortunately, the F-value for this model was 1.4472 with a p-value of 0.1763, implying that the model was not significant. Finally, a linear regression model was formed taking time as the response variable and distance from catching to packing (travel distance) as the independent variable. The economic model formulated was

$$\text{Time} = f\,(\text{distance}) \quad (4)$$

It was hypothesized that as the length of the travel distance increases, time increases. The level-level model was used to build the relationship between time and distance. As the length of the house increases, more time is spent in travel between catching and packing unit.
The regression model can be expressed as

$$t_i = \beta_0 + \beta_1 * d \quad (5)$$

Where:
t_i = time for the individual run, including catching, packing and movement, seconds.
d = distance from the harvester to the packing unit, feet
β_0 is the intercept
β_1 is the parameter estimate (regression co-efficient) for distance.

Regression diagnostics as discussed here were examined for evaluating model adequacy and regression assumptions. The diagnostic methods used both statistical and visual inspection techniques. The closeness of fit of the model was evaluated by co-efficient of determination (R^2). The F value was used for testing the overall significance of the model. The t-value was used for testing the statistical significance of the regression co-efficient (β_1). Robust White test was used for testing the heteroskedasticity. The residual plots between time and distance are helpful in detecting the behavior of the residuals.

Model adequacy can be established when the t-values of the regression coefficients are all significantly different from zero, the sign of the coefficient is correct and the assumptions dealing with linearity, uniform scatter, independence and normality of errors supported.

REGRESSION RESULTS AND DISCUSSION

The regression results are summarized in Table 1 and Table 2. The regression equation for the predicted time is given by:

$$\text{Time} = 133.8902 + 0.1526 * \text{distance} \quad (6)$$

From Table 2, the value of R^2 was 0.0210. R^2 explains the variation in the time explained by distance. This low value raises question over the validity of the model. The coefficient of distance from (equation 1) was 0.1526, which implies that a unit increase in distance causes 0.1526 unit increase in time.

TABLE 1. REGRESSION OUTPUT FOR THE TRAINED DATA SET

Model	time = f (distance)
Dependent Variable	time
Observations	933
R^2	0.0210
RMSE	119.4330
F value	20.4200
Significance level of F	<0.0001
Chi-square Statistic	1.5700
Significance level of χ^2	0.4559

The F-value for testing the overall significance of the model is 20.42. The p value (minimum) level of significance for rejecting the null hypothesis is < 0.0001which is less than $\alpha = 0.05$, which implies that the model is significant. The chi-square statistic for heteroskedasticity is 1.57 and the p value is 0.4559 (> 0.05). Therefore, we fail to reject the null hypothesis that there is no hetroskedasticity. Table 2 shows that the p value for distance is (< 0.05). This implies that the regression coefficient for distance is statistically significant.

TABLE 2: FITTED REGRESSION MODEL CO-EFFICIENTS.

Variable	distance
Co-efficient	0.15266
Standard Error	0.03378
T value	4.52000
Significance level	<0.0001

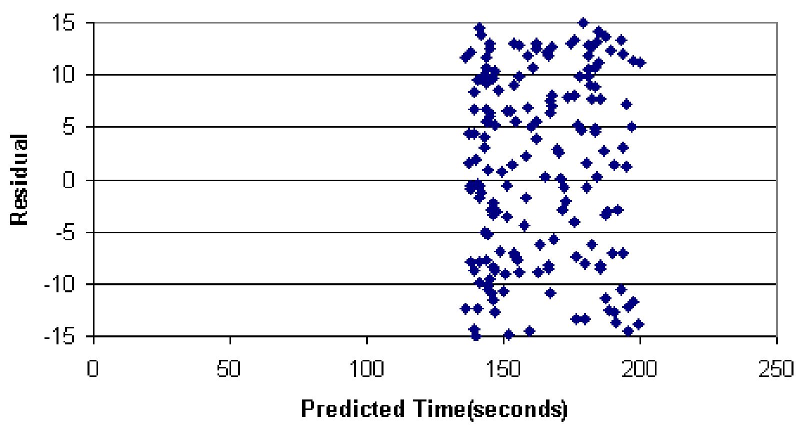

FIGURE 1. RESIDUAL VERSUS PREDICTED TIME IN REGRESSION MODEL

Residual scatter plot shown in Figure 1 was used to check the abnormalities and regression assumptions. From Figure 1, it can be seen that the residuals are randomly scattered around zero and show no discernable pattern. The plot between actual and predicted time is shown in Figure 2.

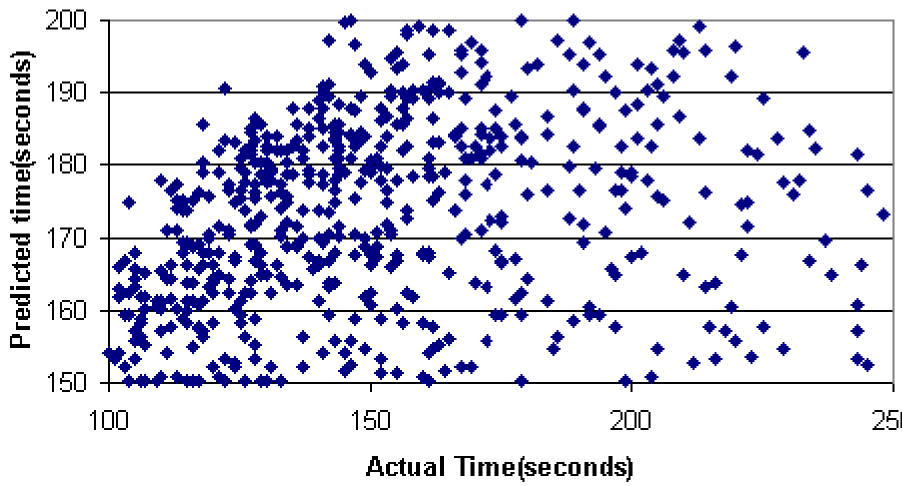

FIGURE 2. PREDICTED TIME VERSUS ACTUAL TIME IN REGRESSION MODEL

A two tailed t-test was done to see if any statistical difference existed between actual time and predicted time for the trained data set. The t-value calculated was 0.0465 and the critical t-value from the table was 1.98 (α = 0.05, df = 951). Since the t-value calculated is less than the tabulated t-value, there was no statistical difference between the actual and predicted values.

Model validation.

For validating the model 238 observations (20% of the data) was used. The mean SSE for the validated data set is less than that of the trained data set, implying that the predictions are better for the validated data set than the trained data set. A two-tailed t -test was used to test whether there was any statistically significant difference between the actual and predicted values of the validated data set. The actual t-value 0.36753 was less than the critical t-value 1.98 (α = 0.05, df = 236). Hence there was no statistically significant difference between the actual and predicted values.

NEURAL NETWORKS

A Neural Network constitutes several interconnected processing elements known as neurons. They cannot be programmed but they imitate the human brain by repeatedly trying to find the relationship between the input and output values. Neural Networks create a model after a sufficient number of learning iterations.

This model could be used to predict for new input values. Neural Networks have been widely used in pattern recognition, classification of noisy data, and market forecasting.

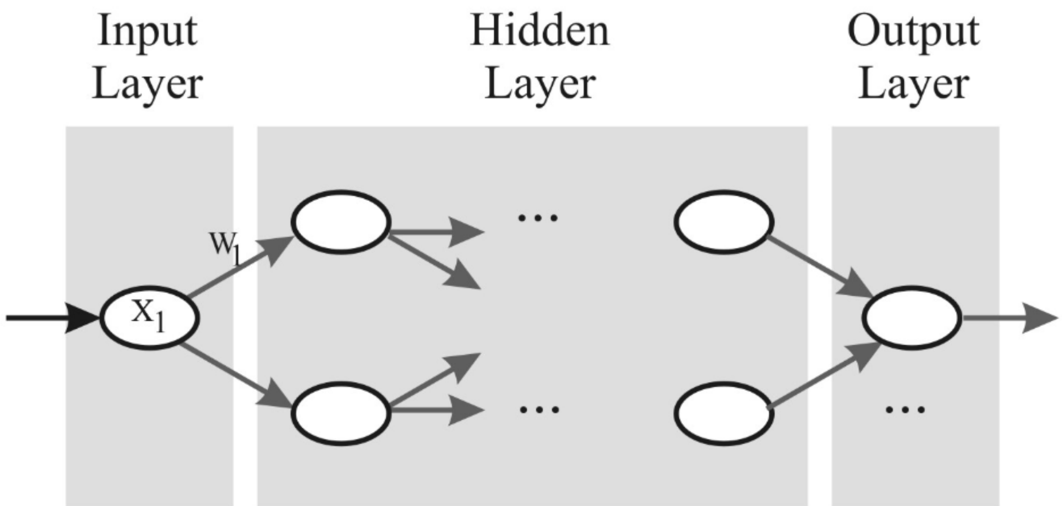

FIGURE 3. SCHEMATIC REPRESENTATION OF NEURAL NETWORK

All neural networks consist of an input layer, one or more hidden layers and an output layer as shown in Figure 3. The network processes a number of inputs from the input layer. These inputs X_i with weights W_i are sent to the hidden layers, which perform the actual processing and weight adjustments by summing the product of inputs and weights. The calculated sum is then transformed by functions, such as threshold or sigmoid. These are sent to the output layer, which give the output values (Chitra, 1992).

NeuroShell Easy Predictor (Ward Systems Group, Inc., 1997) was used to create a model for predicting the total time of the automated chicken catching system. In our study, one hidden layer was used. The input layer refers to distance and the output layer refers to the predicted time.

Neural Networks learn by adjusting the weights. Back-propagation is the most common method used for adjustment. The back-propagation networks typically consist of a sequence of layers fully connected between successive layers. There are always at least three hierarchical layers of neurons: an input layer, one or more hidden layers and an output layer. Every neuron in the input layer sends its output to every neuron in the hidden layer, and every neuron in the hidden layer sends its output to every neuron in the output layer (Chitra, 1992). The number of neurons in the input and output layers typically correspond to the dimension of the input and output vectors respectively. The number of neurons in the hidden layer can be varied based on the complexity of the problem and size of input information. It may not be possible to use an existing pattern if the hidden layer is very large. When the hidden layers are small it drastically extends the number of iterations required to train the network. Conventionally the number of neurons in the hidden layer should roughly equal to the total number of input and output units (Jongh and Wet, 1993).

In back-propagation, the weights adjusted to minimize the squared of the difference between the model output and actual output for an observation in the data set, which is known as the error. The squared error is then propagated backwards through the network and is used to adjust the weights and biases. The adjustment process is repeated until the error converges to a minimum value.

After satisfactory training, the values of the weights represent the neural network's state of knowledge. The trained neural network can be used to predict outputs corresponding to a set of new inputs. A sufficiently trained network is expected to produce outputs that are close to actual outputs. The strength of the network lies in its ability to handle complex nonlinear relationships when the exact nature of the relationship is unknown.

Delurgio (1998) summarized the steps in developing neural networks in four steps:

1) Identification. Designing the neural network by selecting the input and output variables.
2) Estimation. Training the neural network to minimize the error, as to whether the network is simplistic or complex.
3) Diagnosis. Testing the networks with trained data sets and comparing to validated data sets.
4) Forecasting. Using the networks to make predictions.

NeuroShell Easy Predictor is a software program designed to simplify the creation of a neural network application to solve forecasting and pattern recognition problems. The following steps were used to create a predictive model:

1) The data file containing the list of variables is selected.
2) Easy Predictor is informed of the inputs and outputs by specifying the columns, which are inputs (independent variable), and outputs (dependent variable) respectively.
3) A training strategy is chosen (either neural or genetic algorithm).
4) The network is trained.
5) The trained network is applied to the existing data or to the new data to obtain the predictions.

In our model, the same 953 (80%) data points used for building the regression model were used for training the neural network. The remaining 238 (20%) of the data were used for validation. The data was divided into input and output columns. Distance was taken as the input column and the output column was time. The resulting R^2 value was 0.0201. A two-tailed t-test was done to check if there was any statistically significant difference between the actual and predicted values, for the trained data set. The actual t-value 0.5826 was less than the tabulated t-value 1.98 (α = 0.05, df= 951). Therefore, no significant differences were found between the actual and predicted values since we fail to reject the null hypothesis at 95% confidence interval. This implies that the predicted time is in accordance with the actual time. Figure 4 shows the importance of distance in neural network model for predicting the time required for catching the chickens.

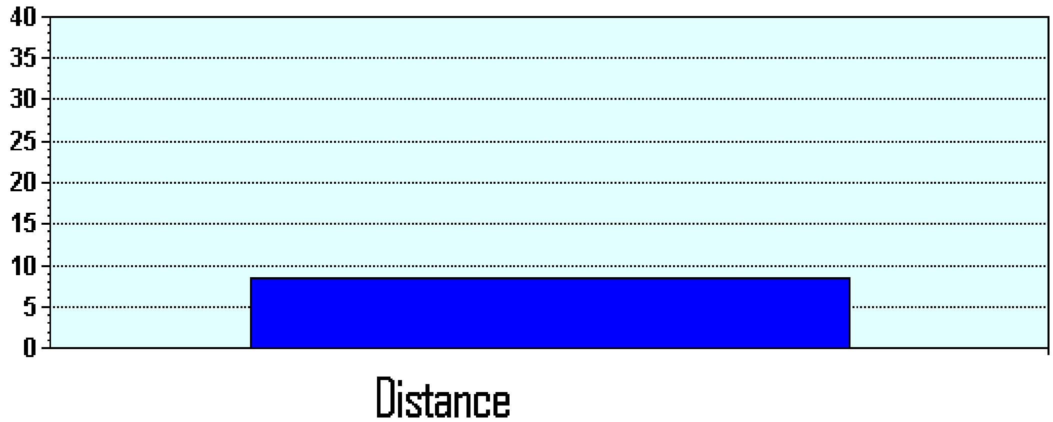

FIGURE 4. IMPORTANCE OF DISTANCE IN NEURAL NETWORK MODEL

Model validation.

The mean SSE for the validated data set is less than that of the trained data set, implying that the predictions are better for the validated data set than the trained data set. A two-tailed t-test was done to see whether there was any statistically significant difference between the actual and predicted values of the validated data set. The actual t-value was 0.01158, which was less than the critical t value 1.98(α = 0.05, df= 236). Hence there was no statistically significant difference between the actual and predicted values.

COMPARISON OF REGRESSION AND NEURAL NETWORK MODEL

The results of the regression model showed that distance was significant, which was also found to be important in neural networks. To compare the accuracy of the two models, Mean Absolute Error (M.A.E), an error statistic was used. Percentage error was taken as the ratio of the M.A.E to the actual value. Percentage accuracy was 100 – percentage error which, was found for the trained and validated models in regression and neural networks respectively. A two-tailed t test was used for finding the differences between the predicted values obtained from the two models for trained and validated data set. The null hypothesis is rejected if the calculated t-value was greater than the tabulated t-value.

The M.A.E calculated from the regression model was slightly smaller than the neural network model, implying that the predictions for the regression model were better than the neural network model. The calculated t-value was less than the tabulated value for both the trained and validated data sets. Therefore the mean difference was not statistically different from zero. Therefore, both models perform equally well. The model statistics pertaining to both the regression and neural network models are shown below in Table 4.

TABLE 4. MODEL STATISTICS FOR REGRESSION AND NEURAL NETWORKS

	Regression Training	Regression Validation	Neural Networks Training	Neural Networks Validation
N (observations)	953	238	953	238
R^2	0.2010	0.1651	0.2005	0.2103
% Accuracy	70.57	69.13	71.54	68.33
MSSE	1548.268	305.616	1937.430	292.755
Mean Actual Time (seconds)	164	205	164	205
Mean Predicted Time (seconds)	152	168	162	164
Mean Absolute Error (seconds)	52	69	53	72

CONCLUSIONS

Two models have been built to predict time, using regression and neural network techniques. The model is adequate for predicting the time because of the significance of F and t statistics. Visual inspection of residual plots confirms adequacy of the model. The inference of t-tests showed that there was no significant difference between the actual and predicted values. The comparisons showed that both models perform well. Although the models seem to predict well, the models explain a very small part of the variation in time. The model predictions could be improved by collecting additional data categories in the field such as number and experience of operators in the house, width of the house, ground conditions, lateral position, obstacles, location, and other factors and adding these variables to the model. As a further extension, comparisons can be made between the two techniques to see if the performance is the same after adding these variables in the model.

References

Chitra, S.P. "Use Neural Networks for problem solving". *Chemical Engineering Progress,* 1993,April 44-52.

Duncan, I.J.H., G. S. Slee, P. Kettlewell, P. Berry and A. J. Carlisle. "Comparison of the stressfulness of catching broiler chickens by machine and hand". *British Poultry Science* 1986, 27: 109-114.

DeLurgio, S. A. "Forecasting Principles and applications". Boston: Irwin McGraw-Hill. pp. 92-132,401-450.

Farsaie, A., E. Carr, and C. J. Wabeck. "Mechanical catching of broilers". *Trans ASAE*: 1983,1650-1653.

Gujarati, D.N. "Basic Econometrics" (3rd Edition) McGraw Hill, New York, NY, USA 1995.

Jongh, P. J. and T. Wet. "An introduction to neural networks". *South African Statistical Journal*, 1993, 27, 103-128.

Lacy, M.P., and M. Czarick. "Evolution of a mechanized broiler catching system". *Poultry Science*, 1992, 71 (Suppl. 1): 100. (Abstr.).

Lacy, M.P., and M. Czarick. "Field-testing of a mechanized broiler catching system". *Poultry Science,* 1994, 73 (Suppl. 1): 41. (Abstr.).

Lacy, M.P., and M. Czarick. "Mechanical catching of broilers". *Poultry Science,* 1998, 77:1794-1797.

Ramasamy, S., E. R. Benson, and G. Van Wicklen. "Efficiency of a Commercial Mechanical Chicken Catching System". *Journal Of Applied Poultry Research* (submitted), 2003.

Salle., C. T. Pippi, Guahyba, Wald, V. Beatriz; Silvia, A. B. da, F. de Oliveira, Nascemento, and V. Pinheiro do. "Use of Artificial Neural Networks to estimate production parameters of broilers breeders in the production phase". *British Poultry Science,* 2002 (Publication information not available) http://www.guahyba.vet.br/trabalhos/trab28.htm

SAS Institute Inc. "*SAS/STAT User's Guide*". (6th ed., Vol.2.). Cary, North Carolina: SAS institute Inc, 1990.

Ward Systems Group Inc. "*Neuro Shell Easy Predictor Instruction*". Maryland: Ward Systems Group, 1997.

The Department of Food and Resource Economics
College of Agriculture and Natural Resources
University of Delaware

The Department of Food and Resource Economics carries on an extensive and coordinated program of teaching, organized research, and public service in a wide variety of the following professional subject matter areas:

Subject Matter Areas

Agricultural Finance
Agricultural Policy and Public Programs
Environmental and Resource Economics
Food and Agribusiness Management
Food and Fiber Marketing
International Agricultural Trade
Natural Resource Management
Operations Research and Decision Analysis
Price and Demand Analysis
Rural and Community Development
Statistical Analysis and Research Methods

The department's research in these areas is part of the organized research program of the Delaware Agricultural Experiment Station, College of Agriculture and Natural Resources. Much of the research is in cooperation with industry partners, other state research stations, the USDA, and other State and Federal agencies. The combination of teaching, research, and service provides an efficient, effective, and productive use of resources invested in higher education and service to the public. Emphasis in research is on solving practical problems important to various segments of the economy.

The department's coordinated teaching, research, and service program provides professional training careers in a wide variety of occupations in the food and agribusiness industry, financial institutions, and government service. Departmental course work is supplemented by courses in other disciplines, particularly in the College of Agriculture and Natural Resources and the College of Business and Economics. Academic programs lead to degrees at two levels: Bachelor of Science and Masters of Science. Course work in all curricula provides knowledge of tools and techniques useful for decision making. Emphasis in the undergraduate program centers on developing the student's managerial ability through three different areas, Food and Agricultural Business Management, Natural Resource Management, and Agricultural Economics. The graduate program builds on the undergraduate background, strengthening basic knowledge and adding more sophisticated analytical skills and business capabilities. The department also cooperates in the offering of an MS and Ph.D. degrees in the inter disciplinary Operations Research Program. In addition, a Ph.D. degree is offered in cooperation with the Department of Economics.

For further information write to:

Dr. Thomas W. Ilvento, Chair
Department of Food and Resource Economics
University of Delaware
Newark, DE 19717-1303

FREC Research Reports are published as a service to Delaware's Food and Agribusiness Community by the Department of Food and Resource Economics, College of Agriculture and Natural Resources of the University of Delaware.