Predicting Cotton Yield for Southeastern USA using Climate Indices

T. B. Pathak [a], J. W. Jones [a], C. W. Fraisse [a]

[a] Agricultural & Biological Engineering Department. University of Florida. Gainesville, FL 32611-0570, USA University of Florida

INTRODUCTION

Variability in the agricultural yield over the region dominantly depends on climatic. Diverse literature exists on the exploiting the use of climate information to forecast crop yield. Large-scale seasonal climate indices has shown good teleconnection with crop such as wheat, sugarcane etc. (Hsieh et al., 1999; Everingham et al., 2003). In this study pre-season climate indices were explored for their use to forecast cotton yield for the southeastern USA.

OBJECTIVES

The main objective of this study was to forecast cotton yield for the southeastern USA using pre-season climate indices.

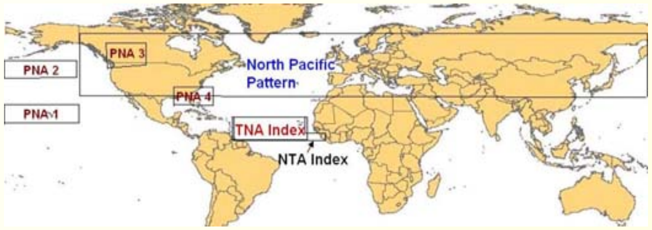
MATERIALS AND METHODS

Statistical Description of Historic Cotton Yield Data

Historic county level yield data for cotton were obtained from National Agricultural Statistical Services (NASS, 1997) for the period 1966 to 2001.Technological trend was removed from the data with linear transformation. Using the estimated linear trend the percentage deviation of yield from the trend line (% residual) was computed for each year.

Climate Indices

Significantly correlated pre-season monthly (January-April) climate indices were used as a predictor for this analysis. Five climate indices shown below were used for this study.



TNA Index: Tropical Northern Atlantic Index is an anomaly of the average of the monthly SST from 5.5 N to 23.5 N and 15W to 57.5 W. Index is created with NOAA OI 1x1 datasets

North Pacific Pattern is the area-weighted sea level pressure over 30N To 65N, and 160E to 140W.

NTA Index: North Tropical Atlantic SST Index is the time series of SST anomalies averaged over 60W to 20 W, 6N to 18N and 20 W to 10W, 6N to 10N. NTA index is created from COADS and NCEP dataset

PNA Index: Pacific North American Index is:

PNA = Z*(PNA 1)– Z*(PNA 2)+ Z*(PNA 3)– Z*(PNA 4)

Z* denotes monthly mean 500mb height anomaly

NOI Index: Northern Oscillation Index is an index of climate variability based on the difference in SLP anomalies at the north pacific high and near Darwin, Australia.

Principal Component Regression Analysis

The main purpose of using principal component analysis was to transform original highly correlated variables in to a new set of independent variables that could be readily used in multiple regression models to forecast cotton yield. Principal components of all climate indices were obtained and were used as predictors for the regression model. Issue of multicolinearity was resolved by using backward stepwise regression.

Leave-One-Out (LOO) Cross Validation

Observed dataset are iteratively and exhaustively used for both model calibration and model testing.

Estimate of model predictive performance is more reliable than estimates from the two-group partition method and less biased than estimates derived from calibration-dependent dataset (Jones and Carberry, 1994). n-1 dataset was used to estimate the parameters for the regression model and one left data point was used for model evaluation. By iterative process all the data points were used for validation. Model evaluation was carried out by calculating the root mean square error of prediction between observed and predicted data. RMSEP was computed by:

$$\mathrm{RMSEP} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_{ij} - y_{ij})^2}$$

RESULTS AND DISCUSSION

Correlation between Climate Indices and Cotton Yield

County cotton yield for AL and GA were significantly correlated with all indices. NOI and NTA indices showed spatial trend in correlations with cotton yield (fig1). Trend could be due to different management practices for different counties.

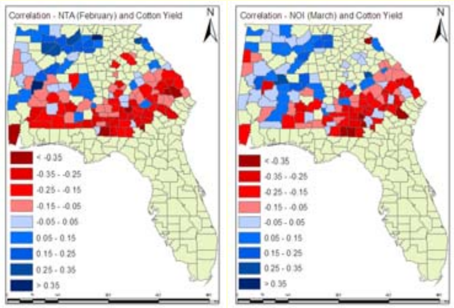


Fig.1 Pearson product correlation of cotton yield with NOI –March (right) and NTA- February (left) respectively

Correlation between Climate Indices and Monthly Rainfall

In order to see if the trend in correlation of yield with indices was due to rainfall, correlation of above indices with rainfall was also explored (fig.2).

Although correlation of indices with rainfall was significant, no clear spatial trend in correlation was observed. All indices were significantly correlated with rainfall and temperature for all counties but varied with months.

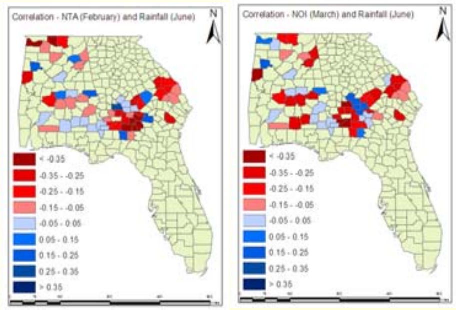


Fig.2 Pearson product correlation of July rainfall with NOI –March (right) and NTA- February (left) respectively

Correlation between best fit model and actual cotton yield

All processed counties showed significant correlation at 99% confidence interval. Maximum correlation of 0.89 was obtained for Houston (AL) county and lowest correlation of 0.50 was obtained for Sumter (GA) county. Each county has distinct principal component regression model and all best fit model showed very high correlations.

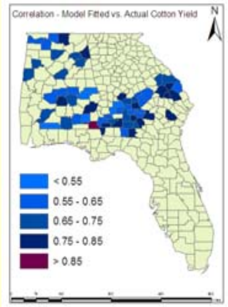


Fig.3 Pearson product correlation of best fit model yield and actual cotton yield

Yield prediction and model evaluation

Highest and the lowest correlated counties for best fit model were chosen for cross validation. Correlation for Sumter for best fit model was 0.50 and 0.36 with calibration. Whereas for Houston the correlation was 0.89 for best fit model 0.72 with calibration.

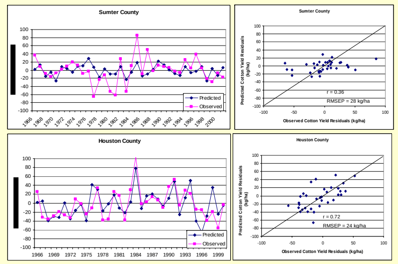


Fig.4 1:1 Line and time series of observed vs. predicted cotton yield for Sumter and Houston Counties

CONCLUSIONS

Pre-season climate indices showed promising predictions of cotton yield for the southeastern USA. Total of sixty six counties were analyzed and all of them showed significant correlation with actual cotton yield for the best fit model. Forecasting yield five-six months in advance with significant level of confidence could be very helpful for the growers to make policy related decisions and make strategic planning. Results show the significance of using climate indices to forecast cotton yield for the southeast USA.

REFERENCES

Hsieh JD, Tang W, Garnett ER. 1999. Teleconnections between Pacific sea surface temperatures and Canadian prairie wheat yield. Agricultural *and Forest Meteorology* **96**: 209–217.

Everingham, YL, Munchow RC, Stone, RC. 2003. Using southern oscillation index phases to forecast sugarcane yields: A case study for northeastern Australia. Intn. Journal of Climatology 23: 1211-1218.

National Agricultural Statistics Service. http://www.nass.usda.gov/

Jones, P. N. and P. S. Carberry. 1994. A technique to develop and validate simulation models. *Agricultural Systems*. 46(4):427-442.