Predicting Regional Crop Production: Applications
James W. Jones
University of Florida
Questions

- Can models successfully predict variations in crop yield due to climate variations over time in a region in which soils, weather, and management vary?
- If so, how can they best be used?
- What are the requirements for this approach?
- Are there examples?
Objective

Evaluate approaches for predicting crop yield response to climate variations, regional scale, using crop models

Hypotheses:

1. Use of local data is necessary for accurately predicting yield

2. Yield predictions are improved when local yield data are used to adjust estimates of soil properties relative to using general soil inputs

Irmak, Jagtap and Jones. 2002. (In Preparation)
Material and Methods

- **Region:** Tift County, GA
- **Data:**
  - 24 years of county yield data.
    - 17 years for calibration
    - 7 years for independent validation (randomly selected)
  - Yields were de-trended
  - 0.5 degree soils & weather data from VEMAP database
- **Management:** Rainfed practices based on NASS data, surveys
- **Simulations:** Average of 9 planting dates, 3 varieties to account for within county variations
- Compare root mean square errors of prediction (independent data)

Irmak, Jagtap and Jones. 2002. (In Preparation)
Case 1: Base Case, no parameter estimation

\[
RMSEP = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (Y_{ij} - y_{ij})^2}
\]

- RMSEP = root mean square error for county j
- Where \( m \) = # of validation years
- \( Y_{ij} \) = Observed yield (kg/ha) for county j, year i
- \( y_{ij} \) = Simulated yield (kg/ha) for county j, year i
- \( j \) = county
- \( i \) = year

Irmak, Jagtap and Jones. 2002. (In Preparation)
Case 2: Simulated deviations used to predict yield for year I, county j

\[ P_{ij} = d_{ij} \cdot \bar{Y}_{Fj} + \bar{Y}_{Fj} \]

Predictions for each independent year

\[ d_{ij} = \frac{y_{ij} - \bar{Y}_{Fj}}{\bar{Y}_{Fj}} \]

Compute deviations for years for which observed data are available, using simulated values only

\[ \bar{Y}_{Fj} = \frac{1}{n} \sum_{i=1}^{n} Y_{ij} \]

Mean observed yield for years for which observed data are available

Irmak, Jagtap and Jones. 2002. (In Preparation)
Case 3: Regression model adjustment of simulated values to account for bias (systematic and non-systematic)

\[
\hat{y}_{ij} = a_j (y_{ij}) + \sum_k c_{jk} \cdot V_{jk} + b_j
\]

Fitted yield using F dataset

\[
Pv_{ij} = a_j (y_{ij}) + \sum_k c_{jk} \cdot V_{jk} + b_j
\]

Simulated yield for validation dataset

- Where \(a_j, b_j, c_{jk}\) = regression coefficients
- \(V_{jk}\) = \(k\) variables such as rainfall at planting, and rainfall at harvest

Irmak, Jagtap and Jones. 2002. (In Preparation)
Case 4: Yield correction factor to account for average bias

\[ Y_{cf} = \frac{Y_{ij}}{y_{ij}} \]  
Yield Correction Factor = Observed/Simulated

\[ P_{v_{ij}} = Y_{cf} (y_{ij}) \]  
Simulated yield for validation dataset

- Where \( y_{ij} \) = Simulated yield for grid j, year i
- \( Y_{ij} \) = Observed yield for grid j, year i

Irmak, Jagtap and Jones. 2002. (In Preparation)
Case 5: Combined model-regression approach; simultaneous estimation of soil parameters and regression coefficients

\[ \hat{y}_{ij} = a_j \cdot y_{ij}(CN_i, DUL_i) + b_i \cdot R_{jp} + c_i \cdot R_{jH} + d_j \]

• Where \( a_j, b_j \) and \( c_j \) = Regression coefficient for grid j,
• \( y_{ij} \) = Simulated yield with estimated soil parameters for grid j, year i
• \( R_{jp} \) and \( R_{jH} \) = Cumulative (3 weeks) rainfall at planting and at harvest for grid j, respectively.
• Simulated annealing is used for parameter estimation

Irmak, Jagtap and Jones. 2002. (In Preparation)
Case 1, base case

Irmak, Jagtap and Jones. 2002. (In Preparation)
Case 1, base case

\[ y = 1.9953x + 329.73 \]

\[ r^2 = 0.5176 \]

RMSEP = 1588 kg/ha

Irmak, Jagtap and Jones. 2002. (In Preparation)
Case 2, Simulate deviations

**Calibration**

- **RMSE**\(_{\text{fitting}} = 164 \text{ kg/ha}\)

**Validation**

- **RMSE**\(_{\text{validation}} = 348 \text{ kg/ha}\)

Irmak, Jagtap and Jones. 2002. (In Preparation)
Case 2:

Calibration:
\[ y = 0.9432x + 64.251 \]
\[ r^2 = 0.7062 \]

Validation:
\[ y = 0.6611x + 514.4 \]
\[ r^2 = 0.2062 \]

Irmak, Jagtap and Jones. 2002. (In Preparation)
Case 5, Estimate crop model and regression parameters

Irmak, Jagtap and Jones. 2002. (In Preparation)
Case 5:

Calibration
\[ y = 1.0645x - 127.96 \]
\[ r^2 = 0.7681 \]

Validation
\[ y = 1.1034x + 34.366 \]
\[ r^2 = 0.6186 \]

Irmak, Jagtap and Jones. 2002. (In Preparation)
Preliminary Conclusions

• Variations in yield associated with climate variations over time can be estimated with accuracies on the order of 5-10%
• Use of simulated deviations and historical yields to predict yield variations is a good first approximation (errors on the order of 10-20%)
• Availability of historical yields for an area can improve predictions, reducing errors by half or more

Irmak, Jagtap and Jones. 2002. (In Preparation)
What is Required?

- Soil profile inputs, each grid or polygon
- Daily weather data, each grid or polygon
- Management inputs, each grid or polygon
- Crop model, evaluated using experiment station data for some location(s) in region

And, if available:

- Historical yield data for a number of years
Example for Georgia, USA

- Soybean study by Jagtap et al.
- 18 years of historical yield data (NASS database)
- Daily weather data provided at 50 km grid (VEMAP database developed by NCAR)
- Dominant soil(s) for each 50x50 km grid cell (VEMAP)
- Management practices (NASS database & state Extension Specialists)
Experience in applying current crop growth models to predict regional productions and its variability is limited.

The interest is now increasing to use models more and more for policy and decision making at regional scales that are relevant to farmers, grain traders and other decision makers.

A number of challenges arise when applying dynamic crop models to regional scales.

Jagtap and Jones, 2002, Agr. Ecosystems & Env.
Known are…

☑ Historical county yields
☑ 50-km digital soil and weather data base
☑ <7% of area cultivated

Unknown are..

Spatial variability in
Inputs
outputs
Georgia Soybean Yield Predictions

@ Plot level
- Soil
- Plant Population
- Variety
- Inputs
- Weather

@ Field/Regional
- Input Variations
- Crop Regions
- Planting Calendar
- Weather

Jagtap and Jones, 2002, Agr. Ecosystems & Env.
Georgia Soybean Yield Predictions

@ Plot level
- Soil
- Plant Population
- Variety
- Inputs
- Weather

Requires perfect aggregation of
Perfect models
Perfect inputs
across range of variability

Aggregation errors distort spatial
Mean values of prediction
Year-to-year variability

@ Field/Regional
- Input Variations
- Crop Regions
- Planting Calendar
- Weather

To reduce errors
Sampling input variability
Calibration of inputs
Calibration of outputs

Deal with Data Limitations

Jagtap and Jones, 2002, Agr. Ecosystems & Env.
County Yields vs. Experimental Field Yields, Tifton, Georgia

Jagtap and Jones, 2002, Agr. Ecosystems & Env.
Crop Models Simulate Yield at Field Scale: Bias Exists

Jagtap and Jones, 2002, Agr. Ecosystems & Env.
Why Does the Bias Exist?

Yield Gap between..

1. Station and on-farm yields
2. Varies from year-to-year
3. Simulated yields more variable due to imperfect model and aggregation errors (due to lack of specific information, accuracy of inputs,...)

To minimize the errors

• Simulated results and or inputs must be calibrated
2. Model must be improved
3. Better sources of data

Jagtap and Jones, 2002, Agr. Ecosystems & Env.
RMSE ranged from 0.24 to 0.65 t/ha or 16 to 47% of grid yields. RMSE during validation was 38-60% of the 1990-95 mean yields.

![Graph showing actual and un-calibrated yields over years (1970-1995).](image-url)
Before Removing Bias, High RMSE for Validation Sites

Jagtap and Jones, 2002, Agr. Ecosystems & Env.
After Adjusting for Bias, – Yield Predictions at a Validation Site are Improved

Jagtap and Jones, 2002, Agr. Ecosystems & Env.
After Adjusting for Bias, – Yield Predictions at a 2\textsuperscript{nd} Validation Site are Improved

Jagtap and Jones, 2002, Agr. Ecosystems & Env.
Yield Bias Variations

Jagtap and Jones, 2002, Agr. Ecosystems & Env.
After Adjusting for Bias, RMSE for Independent Years Improved Greatly

RMSE

- <5 %
- 6 - 10 %
- 11 - 16 %

Jagtap and Jones, 2002, Agr. Ecosystems & Env.
After Adjusting for Bias, Predictions for Calibration Years and Validation Years
Aggregation Reduced Errors in This Study, Averaged over Fitting and Validation Years
Conclusions

• Yield bias must be removed when using crop models to predict regional yields
• RMSE averaged 14% when predicting yields for independent years, after correcting for bias
• Bias varied considerably across the state of Georgia, implying variations in management and other factors not accounted for in the model
• Other approaches may further reduce bias (see Irmak et al.)
Yield Forecast Tool

- Conditioned on:
  - Climatology
  - Climate Forecast

Time

- History – Use recorded weather
- Future – Use climatology or forecast

Start of Season  Today
Famine Early Warning in Burkina Faso -- A Case Study of Millet Production

• Scope for improvement in yield forecasting methods for millet and sorghum in the Sahel

• Potentially high value of information because of the policy implications (alleviation measures)

• A proof-of-concept case study, involving the assembly of soils, satellite rainfall data, crop

How it was Done

SOILS
Profile Data

METEOSAT IMAGE
Dekadal Rainfall Estimate (RFE)

MILLET MODEL
Unique Combinations of Soil and Weather

PROVINCIAL YIELDS
GEOGRAPHIC INFORMATION SYSTEM

PROVINCIAL STATISTICS
Production and Yield Averages

PROVINCIAL PRODUCTION
Dekadal Estimate of Current Season Production Anomaly

SATELLITE-DERIVED CROP USE INTENSITY
Provincial Area Planted to Millet

WEATHER
Long-term Historical Data

Rainfall Estimate: Dekad 25, 1990

Conclusions

• Satellite-derived rainfall, coupled with generated temperature and solar radiation values, have potential for early warning yield forecasts
• Yields forecast at mid season were within 15% of final yields
• Although more work is needed, this approach shows considerable promise

Regional Scale Crop Yield Predictions:

Discussion