Empirical regression models using NDVI, rainfall and temperature data for the early prediction of wheat grain yields in Morocco

Balagni, R., Tychon, B., Erens, H. and Jlibene, M.

CGMS-2006 – GEOCLAND meeting
Antoni 23-26 October 2006

JUSTIFICATION

- Cereal production strongly fluctuates from year to year due to an erratic climate, in a situation where 85% of agricultural lands are rainfed.
- Cereals are a strategic food in Morocco, with a consumption of 210 kg per capita, one of the highest in the world (159 kg at world level) which is not entirely covered by local production.
- The coverage ratio of the cereal needs strongly fluctuates from year to year (most of years in deficit).
- There is a need for yield prediction models to manage wheat imports.
- Actually, no specific yield prediction models exist for Morocco.
HYPOTHESIS

1. Wheat yields could be predicted using models based on easily available indices or weather data.

2. The models predict accurately and well in advance yields.

Methodology

\[
\text{NDVI} + \text{Rainfall} + \text{Temperature} \rightarrow \text{Regression model} \rightarrow \text{Physical yield prediction}
\]

\[
\text{NDVI} + \text{Rainfall} \rightarrow \text{Regression model} \rightarrow \text{Agricultural yield prediction}
\]
Methodology

Available Database:

- Wheat production and area by province from 1979 to 2004.
- Dekadal rainfall by province from 1987 to 2004 (23 provinces).
- Monthly temperature by province from 1987 to 2004 (17 provinces).
- Dekadal NDVI/AVHRR from 1990 to 2004 (15 years from which 3 years have been dropped due to bad quality).

The considered provinces in Morocco (Moroccan Sahara not shown) with their average wheat production (1990-2004; Data source: Economic Services of the Ministry of Agriculture) and the location of the meteorological stations.
NDVI/AVHRR for the 2nd dekad of March (average of the years 1990-2004). The non-agricultural zones and irrigated fields have been masked using the Global Land Cover 2000 for Africa (GEC2000, version 5.0).

Regression models:

- Ordinary least squares regressions using:
  - Moving sums of dekad rainfall (2 till 11 dekads, from September to May)
  - Moving means of monthly temperature (from September to May)
  - Sums of dekadal NDVI from February to April (C:NDVI)
  - Stepwise selection of the best predictors, by removing time overlapping predictors if they appear automatically
  - « Leave-one-out » cross-validation in order to verify the replicability of results and checks the prediction performance of the model for "new" years.

\[ R^2_v = R^2 \text{ in validation} \]

\[ \text{Error} = (\text{Kg} \cdot \text{ha}^{-1}) \]

Relative error = (%)

\[ \text{Coefficient of determination between observed and predicted yields} \]
Prediction models:

1. By province, using provincial NDVI, rainfall and temperature (23 provinces = 64% of the national wheat production)
   - National yield prediction using provincial predicted yields weighted by their corresponding agricultural area

2. For the whole country, using national NDVI and rainfall (31 provinces = 98% of the national wheat production)
   - National yield prediction directly

National wheat yield and rainfall variations
Relation between observed wheat yield and NDVI at national level in Morocco.

National wheat yield prediction model using provincial NDVI, rainfall and temperature.
## National wheat yield prediction models using national NDVI and rainfall

<table>
<thead>
<tr>
<th>Model</th>
<th>Date</th>
<th>2NDVI</th>
<th>Rainfall</th>
<th>df</th>
<th>( R^2 )</th>
<th>( R^* )</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2nd Mar</td>
<td>-0.499 + 7.115 2NDVI + 0.022</td>
<td>607</td>
<td>66</td>
<td>0.84</td>
<td>3.4</td>
<td>100 kg/ha</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4)</td>
<td>(0.16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3rd Mar</td>
<td>-0.048 + 0.079</td>
<td>730</td>
<td>66</td>
<td>0.79</td>
<td>7.7</td>
<td>100 kg/ha</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4)</td>
<td>(0.03)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1st Apr</td>
<td>-0.090 + 5.001</td>
<td>835</td>
<td>66</td>
<td>1.00</td>
<td>9.9</td>
<td>100 kg/ha</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4)</td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2nd Apr</td>
<td>-0.109 + 4.447</td>
<td>822</td>
<td>66</td>
<td>1.11</td>
<td>9.3</td>
<td>100 kg/ha</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4)</td>
<td>(0.01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3rd Apr</td>
<td>-0.029 + 0.265</td>
<td>820</td>
<td>66</td>
<td>0.73</td>
<td>6.8</td>
<td>100 kg/ha</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4)</td>
<td>(0.01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1st May</td>
<td>-0.041 + 2.042</td>
<td>851</td>
<td>66</td>
<td>0.82</td>
<td>7.4</td>
<td>100 kg/ha</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4)</td>
<td>(0.01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2nd May</td>
<td>-0.272 - 4.584 2NDVI</td>
<td>851</td>
<td>66</td>
<td>1.05</td>
<td>14.6</td>
<td>100 kg/ha</td>
</tr>
</tbody>
</table>

**\( R^2_p = 0.96^{* * *} \)**

Average error = 73 kg/ha (6.8%) between 8 to 168 kg/ha/year
Conclusion

- Average error = 73 kg/ha (8 to 188 kg/year) using national data ($R^2 = 96.1\%$)
- Average error = 55 kg/ha (1 to 183 kg/year) using provincial data ($R^2 = 97.7\%$)
- Early predictions, starting from March using national model
- $\sum$ NDVI is the most important predictor (85% of yield variability for national model)
- 3 time non-overlapping predictors for national model

Perspectives

- Potential improvements:
  - Use of better quality NDVI images (SPOT/Veget.)
  - Use of higher resolution NDVI images and land cover maps (CORINE)
  - Use of more climatic stations
  - Use of models taking into account phenological phases at province level
› Possibility to extend the models to other crops: durum wheat and barley
› Possibility to extent the approach to countries with similar agroclimatic environments
› Possibility to incorporate the proposed models to MARS regional forecasting system

Thank you