Big Data for Climate Smart Agriculture
Case study: Rice Systems in Latin America

5th Annual Workshop on Understanding Climate Change from Data, Minneapolis, MN
4th - 5th August 2015
What need are we responding to?

¿What, When, Where to grow?

In terms of agricultural research mostly based on a top-down (from controlled conditions – technologies are passed onto farmers)
What we propose?

A complementary bottom-up approach: Information from commercial fields - Taking advantage of modern information technologies

Climate + Soil + Crop management (including varieties) = Crop response (productivity/ha)

% ? + % ? + %? = To Explain (100 %)

Empirical modelling approaches aimed to identify the patterns that lead to either high or low productivities (mostly based on machine learning techniques) as strategy to climate change adaptation!!!
How?

Main site-specific climatic limiting factors

<table>
<thead>
<tr>
<th>Databases used</th>
<th>Initial purpose</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Rice Survey</td>
<td>Keep the crop sector updated, technological changes</td>
<td>1237</td>
</tr>
<tr>
<td>Harvesting records</td>
<td>Monitoring production levels in the regions</td>
<td>6000</td>
</tr>
<tr>
<td>Planting dates experiments</td>
<td>Technical research on the best sowing date</td>
<td>600</td>
</tr>
</tbody>
</table>

Information on: planting and harvesting date, productivity, variety, cropping system

Zones: Caribbean, Andean (Tolima), Plains (Llanos)

Climate
- About 27 weather stations
How?

Main site-specific climatic limiting factors

Relating daily series of main climate variables to every 120 days cropping events yield, we should be able to characterize the climate-crop relationship.

Cock et al., (2011). Agricultural Systems
DAILY CLIMATE DATA

SOWING TO HARVEST

~120 DAYS x N RICE CROPPING EVENTS

How?
Hypothesis: Variation in crop yield is associated with climate

Evolution of the national average rice yield in Colombia

Free trade agreements: sector is threatened by the commercialization of rice produced in more efficient countries
How?

Traditional methods
- Multiple linear regression (OLS)
- Factorial analysis (PCA, MCA, CATPCA...)
- Generalized linear model
- Mixed Models
- Time series

Methods based on machine learning
- Supervised and unsupervised ANNS (MLP, SOM)
- Random Forest
- C-Forest
- Fuzzy Logic

Challenges in data-driven analysis:
Both quantitative and qualitative, noisy, non-linear, incomplete, heterogeneous, often non-parametric, (y) transformation, etc.,
Main site-specific climatic limiting factors


Fedearroz 733 (N=267)

FEDEARROZ 733 - 34 % of productivity variation explained

Cimarron Barinas (N=78)

Cimarron Barinas - 56 % of productivity variation explained

From data to action!!! Varieties perform differently under identical climatic conditions
Main site-specific climatic limiting factors

Climate and analysis based on phenological stages in Saldaña (research station)
Andean zone 2007 – 2012 (N= about 800 cropping events – irrigated rice), ANNs – Relevance metric:

Input perturbation

- The crop sector can suggest to farmers the best planting date (when to grow)
- By assessing the same approach in other stations (environments) – New insights for future breeding
- Adaptation strategy for climate change
Main site-specific climatic limiting factors

Case study: Espinal (Tolima Department)- Data: 2007 – 2013

N= 180 harvesting events- Variety Cimarron Barinas – Irrigated rice

C- Forest– Relevance metric: Partial dependence plots

Average minimum temperature in ripening stage is a critical factor for variety Cimarron Barinas

Average temperature in reproductive stage is a critical factor for variety Cimarron Barinas

Climate accounts for more than 59 % to production variability in irrigated rice – Variety Cimarron Barinas
Main site-specific climatic limiting factors

Case study: llanos  Data: 2007 – 2012
N= 200 harvesting events – Variety F174 – Rainfed rice
ANNs – Relevance metric: Input perturbation

- Average temperature during grain filling is a critical factor for variety 174
Historical profiling

Climatic profiles

Seasonal forecast
## Historical profiling

### Region: Casanare: 2007 – 2014 – Irrigated

Weather stations from FEDEARROZ e IDEAM (N= about 756 cropping events) – 17 clusters

Dynamic Time Wrapping (DTW)

### Yield associated with each historical profile (cluster) presented in the region

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Number of cropping events</th>
<th>Productivity (kg/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18</td>
<td>5,354</td>
</tr>
<tr>
<td>2</td>
<td>238</td>
<td>5,653</td>
</tr>
<tr>
<td>3</td>
<td>51</td>
<td>5,821</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>4,913</td>
</tr>
<tr>
<td>5</td>
<td>148</td>
<td>4,946</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>5,557</td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>6,041</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>5,174</td>
</tr>
<tr>
<td>9</td>
<td>60</td>
<td>6,000</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>5,726</td>
</tr>
<tr>
<td>11</td>
<td>42</td>
<td>5,898</td>
</tr>
<tr>
<td>12</td>
<td>65</td>
<td>4,688</td>
</tr>
<tr>
<td>13</td>
<td>18</td>
<td>5,469</td>
</tr>
<tr>
<td>14</td>
<td>33</td>
<td>5,222</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>5,312</td>
</tr>
<tr>
<td>16</td>
<td>15</td>
<td>5,521</td>
</tr>
<tr>
<td>17</td>
<td>6</td>
<td>5,053</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>765</td>
<td><strong>5,438</strong></td>
</tr>
</tbody>
</table>
Cluster 2 - Scenario 2 NovDecJanFeb

Cluster 2 - Scenario 5 DecJanFebMarch
Historical profiling – What to grow?

Weather stations from FEDEARROZ e IDEAM
(N= about 756 cropping events) – 17 clusters

Performance of each variety within cluster 2

<table>
<thead>
<tr>
<th>Rice variety</th>
<th>Number of cropping events</th>
<th>Productivity (Kg/Ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F2000</td>
<td>39</td>
<td>6,717</td>
</tr>
<tr>
<td>F174</td>
<td>27</td>
<td>5,979</td>
</tr>
<tr>
<td>COPROSEM 304</td>
<td>25</td>
<td>4,973</td>
</tr>
<tr>
<td>F733</td>
<td>23</td>
<td>5,981</td>
</tr>
<tr>
<td>F50</td>
<td>22</td>
<td>5,922</td>
</tr>
<tr>
<td>FORTALEZA</td>
<td>21</td>
<td>4,641</td>
</tr>
<tr>
<td>F369</td>
<td>18</td>
<td>5,244</td>
</tr>
<tr>
<td>INPROARROZ 1550</td>
<td>16</td>
<td>5,180</td>
</tr>
<tr>
<td>F60</td>
<td>12</td>
<td>5,189</td>
</tr>
<tr>
<td>INPROARROZ 216</td>
<td>9</td>
<td>4,104</td>
</tr>
<tr>
<td>LAGUNAS</td>
<td>7</td>
<td>6,598</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>219</strong></td>
<td><strong>5,653</strong></td>
</tr>
</tbody>
</table>
Some of the reasons why this is so exciting!!!
Some of the reasons why this is so exciting!!

Tropics: Climate-smart, site-specific agriculture:

Temperate: Big Data for Climate Smart Agriculture - Enhancing & Sustaining Rice Systems for Latin America and the World:
Some of the reasons why this is so exciting!!!

Big Data Climate Challenge Winners
Last but not least....Interdisciplinary effort

Adaptation across timescales

Ramirez-Villegas and Khoury- Climatic Change (2013)
THANK YOU!!!
BIBLIOGRAPHY


http://dataimpacts.org/project/climate-modeling/