Ensemble machine learning prediction of drivers affecting rice and wheat yield, greenhouse gas emissions, and yield-scaled emissions in Bangladesh

Project Title: P1595 - Big data analytics to identify and overcome scaling limitations to climate-smart agricultural practices in South Asia (BigData2CSA)

Description of the innovation: We made use of a comprehensive secondary dataset collected in 2015 of agronomic management information (n=13,684 farmers) and environmental data (50,000 soils datum) for rice and wheat. Data were analyzed by developing an ensemble machine learning model combining deep neural networks, random forest, extreme gradient boosting, and cubist methods to predict yield, GHGss, and yield-scaled emissions and the drivers affecting them. Adaptability/resilience indicators are now being analyzed. Presentation links: (1) https://www.dropbox.com/s/1x6hg17t5a23o7a/200304%20ML_CSA_CIMMYT.pptx?dl=0, and also: (2) https://www.dropbox.com/s/a9pfbc2dbwaxue4/191019%20CCAFS%20TEMPLATE%20P1595-BALI%20-2019%20-%20Timothy%20Krupnik%20-%20Copy.pptx?dl=0

New Innovation: No

Stage of innovation: Stage 1: discovery/proof of concept (PC - end of research phase)

Innovation type: Biophysical Research

Geographic Scope: <Not Defined>

Number of individual improved lines/varieties: <Not Applicable>

Description of Stage reached: Although very preliminary, to our knowledge, this is the first attempt at ensemble machine learning applied to agricultural research and the analysis of 'big data' from farms. In addition, methods have been developed in R to graphically explore the relationships between drivers and predicted outcomes using partial dependency plots.

Name of lead organization/entity to take innovation to this stage: CIMMYT - Centro Internacional de Mejoramiento de Maíz y Trigo

Names of top five contributing organizations/entities to this stage:

- CSSRI - Central Soil Salinity Research Institute
- DAE - Department of Agriculture Extension (Bangladesh)
- NARC - Nepal Agricultural Research Council
- PAD - Precision Agriculture for Development
- CIMMYT - Centro Internacional de Mejoramiento de Maíz y Trigo

Milestones: No milestones associated

Sub-IDOs:

- 45 - Increased capacity for innovations in partner research organizations
- 26 - Agricultural systems diversified and intensified in ways that protect soils and water
- 10 - Closed yield gaps through improved agronomic and animal husbandry practices

Contributing Centers/PPA partners:

- CIMMYT - Centro Internacional de Mejoramiento de Maíz y Trigo

Evidence link: https://tinyurl.com/y7keuwsq

Deliverables associated:

- D9575 - Initial database integrating crop cut, management practice, remotely sensed and farmer survey information to evaluate and inform CSA technologies and practices (https://tinyurl.com/y3a8t4mc)
- D9574 - Interactive web-based dashboards presenting post-season research results and providing CSA management recommendations. (https://tinyurl.com/yxqvlq79)

Contributing CRPs/Platforms:

- BigData - Platform for Big Data in Agriculture
- CCAFS - Climate Change, Agriculture and Food Security