The neotropical savanna is the second largest biome in South America, with significant potential for agricultural development. In Colombia, this biome is experiencing rapid land-use change leading to the conversion of seminatural landscapes into intensive agricultural systems. Our Dataset Paper documents the emerging intensive grain production systems. Between 2011 and 2013, we established 336 observatory plots within farmer’s maize, rice, and soybean fields along a 200Km transect from Puerto Lopez (Meta) to Viento (Vichada). From each of these plots, we submit 184 descriptors or variables capturing their location, rotation history, management, and environment. Our specific objective in collecting the data was to identify key factors explaining yield variation, with emphasis on interactions between management and environmental factors potentially informing the development of site-specific management protocols. Beyond this objective, the dataset submitted here is intended to support additional inquiries contributing to the sustainable development of agriculture in the neotropical savannas.

1. Introduction

The neotropical savanna is the second largest biome in South America, occupying about 250 million hectares of land [1–3]. Its soils are notorious for their high acidity and aluminum levels toxic to most crops. Still, Nobel Prize Laureate Norman Borlaug called it “the last agricultural frontier in the world” [4]. Indeed, between 1955 and 2005 the Brazilian savannas experienced an extraordinary frontier expansion leading to the cultivation of over 40 million hectares of land previously considered infertile [5]. Reflecting on this achievement, Borlaug envisioned a similar transformation for the savannas in Colombia, Venezuela, and Southern Africa [5]. Contributing to his vision, our Dataset Paper documents the ongoing land-use conversion of seminatural savannas into intensive grain production systems in Colombia.

Our objective in collecting the data was to help identify key factors explaining yield variation in maize, rice, and soybeans on farmers’ fields. From an operational standpoint, we recognized that these factors were of two types: those that easily lend themselves to agronomic manipulation and those that do not. The first group includes variables like soil pH, which can be adjusted relatively easily by liming. We call this group management factors. The second group includes variables like soil texture, which cannot be changed. We call this group environmental (or zoning) factors. Further, we recognized that the influence of some management factors on yield may depend on one or more environmental factors. For example, the same amount of irrigated water may cause yield improvements in sandy soils and may cause the root to rot in clay soils. Characterizing these types of interactions between management and environmental factors is the foundation of site-specific agriculture. Aware of this potential, our study was designed to provide a stepping stone for the development of site-specific grain agriculture in the Colombian savannas.

Encouraged by multiple requests of our dataset to address additional research questions, we are pleased to formally present it to our community of interest in this Dataset Paper.
2 Dataset Papers in Science

Figure 1: Study site. (a) Tropical savannas in South America. (b) Tropical savannas in Colombia. (c) Locations of EGM plots along the study transect. (d) Representative distribution of EGM plots within a farm.

Paper. The objective of submitting this dataset, therefore, is to encourage and support diverse research inquiries contributing to the sustainable development of agriculture in the neotropical savannas.

2. Methodology

The study was conducted in the Colombian savannas, locally known as “Llanos Orientales,” a region that extends from the Meta Department to the Venezuelan border (Figure 1(c)). Its climate is characterized by a wet season that begins in March and a dry season that begins in December, with an average annual temperature around 26°C [6, 7]. The length of the wet season accommodates two planting seasons for grain crops, one around April and another around September. Soils are mainly Oxisols with low fertility and high acidity and Al saturation [6, 8]. A good ecological characterization of the region is provided by Blydenstein [9].

Our method involved the establishment of observatory plots within farmer’s maize, rice, and soybean fields along a 200 Km transect from Puerto Lopez to Viento from 2011 to 2013. We call these plots “EGM,” after the Spanish acronym for Georeferenced Sampling Station (Figures 2(a) and 2(b)). EGMs were 20 m², a size we chose because it facilitated intensive sampling and matched the experimental field size used by the Colombian Ministry of Agriculture to evaluate and register new cultivars. A series of farm visits, involving unstructured interviews with farm managers and guided field inspections, helped us survey the variability between and within fields with respect to topography, soil texture, rotation history, and yield history. During the inspections, we consultatively established two or more EGMs per field, in such way that captured the greatest perceived variability with respect to the above-mentioned factors. Our sampling within fields was therefore not random but was designed to increase statistical variance with respect to yield and a few of its potentially important environmental determinants.

We relied on three sources of data: farm records, direct measurements, and geographic information systems (GIS) databases. Farm records helped us capture rotation history, crop cultivar, and planting dates. Direct measurements helped us capture soil parameters, plant density, and yield. Immediately before the planting season, we collected three soil subsamples (Figure 2(c)) along a diagonal transect across the EGM, at depths of 0–10 cm and 10–20 cm, and bulked them into a single sample per depth profile. These samples
Figure 2: Sampling methods used in the field. (a) Flags deliniating a 20 m² Georeferenced Sampling Station on a soybean field. (b) Manual harvest of a soybean Georeferenced Sampling Station to estimate yield. (c) One of three soil subsamples (0–10 cm) taken per Georeferenced Sampling Station for chemical analyses. (d) Instruments used to collect 100 cm³ soil core samples for physical analyses.

were submitted for chemical analyses to the soil laboratory at the International Center for Tropical Agriculture (CIAT). In addition, core samples of 100 cm³ volume (Figure 2(d)) were taken from near the center of the EGM, at depths of 0–10 cm and 10–20 cm, and submitted for physical analyses to the Soil Laboratory at the Colombian Corporation of Agricultural Research (Corpoica). Plant density, yield, and grain moisture were measured within two weeks of the field’s intended harvest date. We harvested the EGM manually to measure yield and grain moisture content (Figure 2(b)). We used these two values to adjust yield based on the moisture content desired for storage (i.e., dry yield), which is 14.2% for rice and maize and 12.2% for soybeans. Finally, EGMs were georeferenced using geographic positioning system receivers (GPSMap 76Cx; Garmin, Olathe, Kansas, USA), and the coordinates were used to retrieve 250 m normalized difference vegetation index (NDVI) data from the Moderate Resolution Imaging Spectroradiometer (MODIS, [10]), precipitation data from the Tropical Rainfall Measuring Mission (TRMM; [11]), and interpolated climate data from WorldClim [12].

3. Dataset Description

The dataset associated with this Dataset Paper consists of 2 items which are described as follows.

Dataset Item 1 (Table). Data of the 336 observatory plots (EGM) within farmer’s maize, rice, and soybean fields with 184 descriptors or variables capturing their location, rotation history, management, and environment at Colombian savannas (Llanos Orientales). Each row corresponds to an EGM, and each column corresponds to a descriptor or variable. Broadly, there are five categorical descriptors for location at different scales (storage type: character) and two variables for geographic coordinates (storage type: float), one for plot area (storage type: float), four for rotation history (storage type: character), two for the crop and cultivar sown (storage type: character), five capturing the temporal dimension of the production event (storage types: integer, character, and date), three capturing plant density (storage types: float and integer), one for grain moisture (storage type: float), two for yield (storage type: float), 63 for soil physical and chemical properties at two soil depth profiles (storage type: float), 29 for precipitation data retrieved from TRMM (storage type: float), and 67 for temperature data retrieved from WorldClim (storage type: float). The missing values are represented by blank cells. In the table, the column Grain Yield Standardized presents the grain yield standardized percentage of moisture content desired for storage. Also the column Mean Diurnal Range was calculated as (mean of monthly (max temp − min temp)), the column Isothermality as (BIO2/BIO7) (∗100), the column Temperature Seasonality as (standard deviation ∗ 100), and the column Temperature Annual Range as (BIO5 − BIO6). The column Rainfall Seasonality was measured by coefficient of variation. For more details, see Table 1.

Column 1: Plot Identifier
Column 2: Field Identifier
Column 3: Farm Identifier
...
### Table 1: Variable name.

| Variable name                                      | Unit         |
|----------------------------------------------------|--------------|
| Plot Identifier                                    | na           |
| Field Identifier                                   | na           |
| Farm Identifier                                    | na           |
| Political Subdivision of Department                | na           |
| Political Subdivision of Country                   | na           |
| Latitude of Plot Centroid                          | DD          |
| Longitude of Plot Centroid                         | DD          |
| Area of the Plot                                   | m²           |
| Plant Cover Four Semesters Back                    | na           |
| Plant Cover Three Semesters Back                   | na           |
| Plant Cover Two Semesters Back                     | na           |
| Plant Cover the Preceding Semester                 | na           |
| Crop Common Name                                   | na           |
| Crop Cultivar Name                                 | na           |
| Year of Planting                                   | y            |
| Semester of Planting                               | na           |
| Date When Field Planting Began                     | dd/mm/yy     |
| Date When Field Planting Ended                     | dd/mm/yy     |
| Date When the Plot Was Harvested                   | dd/mm/yy     |
| Spacing between Plant Rows                         | cm           |
| Spacing between Plants within a Row                | cm           |
| Plant Density                                      | Plants ha⁻¹  |
| Grain Moisture Content at Harvest                  | %            |
| Grain Yield Standardized for % Moisture Content    | t ha⁻¹       |
| Desired for Storage                                | t ha⁻¹       |
| Measured Grain Yield at the Harvested Moisture     | t ha⁻¹       |
| Content                                            | t ha⁻¹       |
| Soil Available Water at 0–10 cm Depth              | mm           |
| Soil Bulk Density at 0–10 cm Depth                 | g m⁻³        |
| Soil Particle Density at 0–10 cm Depth             | g m⁻³        |
| Soil Total Porosity at 0–10 cm Depth               | %            |
| Soil Macroporosity at 0–10 cm Depth               | %            |
| Soil Mesoporosity at 0–10 cm Depth                | %            |
| Soil Microporosity at 0–10 cm Depth               | %            |
| Soil Mean Weight Diameter of Aggregates at 0–10 cm Depth | mm    |
| Soil Sand at 0–10 cm Depth                         | %            |
| Soil Silt at 0–10 cm Depth                         | %            |
| Soil Clay at 0–10 cm Depth                         | %            |
| Soil Organic Matter at 0–10 cm Depth              | g kg⁻¹       |
| Soil pH at 0–10 cm Depth                           | na           |
| Soil Base Saturation at 0–10 cm Depth              | %            |
| Soil Cation Exchange Capacity at 0–10 cm Depth     | mol kg⁻¹     |
| Soil Effective Cation Exchange Capacity at 0–10 cm Depth | mol kg⁻¹        |
| Soil Electrical Conductivity at 0–10 cm Depth      | dS m⁻¹       |
| Soil Aluminum at 0–10 cm Depth                     | cmol kg⁻¹    |
| Soil Boron at 0–10 cm Depth                        | mg kg⁻¹      |
| Soil Carbon at 0–10 cm Depth                       | %            |
| Soil Calcium at 0–10 cm Depth                      | cmol kg⁻¹    |
| Soil Calcium to Magnesium Ratio at 0–10 cm Depth   | %            |
| Soil Copper at 0–10 cm Depth                       | mg kg⁻¹      |
| Soil Iron at 0–10 cm Depth                         | mg kg⁻¹      |
| Soil Potassium at 0–10 cm Depth                    | cmol kg⁻¹    |
| Soil Magnesium at 0–10 cm Depth                    | cmol kg⁻¹    |
| Soil Manganese at 0–10 cm Depth                    | mg kg⁻¹      |
| Soil Nitrogen at 0–10 cm Depth                     | %            |
| Soil Sodium at 0–10 cm Depth                       | cmol kg⁻¹    |
| Soil Phosphorus at 0–10 cm Depth                   | mg kg⁻¹      |
| Soil Sulfur at 0–10 cm Depth                       | mg kg⁻¹      |
| Soil Zinc at 10–20 cm Depth                        | mg kg⁻¹      |
| Soil Available Water at 10–20 cm Depth             | mm           |
| Soil Bulk Density at 10–20 cm Depth                | g m⁻³        |
| Soil Particle Density at 10–20 cm Depth            | g m⁻³        |
| Soil Total Porosity at 10–20 cm Depth              | %            |
| Soil Macroporosity at 10–20 cm Depth               | %            |
| Soil Mesoporosity at 10–20 cm Depth                | %            |
| Soil Microporosity at 10–20 cm Depth               | %            |
| Soil Sand at 10–20 cm Depth                        | %            |
| Soil Silt at 10–20 cm Depth                        | %            |
| Soil pH at 10–20 cm Depth                          | na           |
| Soil Base Saturation at 10–20 cm Depth             | %            |
| Soil Cation Exchange Capacity at 10–20 cm Depth    | mol kg⁻¹     |
| Soil Electrical Conductivity at 10–20 cm Depth     | dS m⁻¹       |
| Soil Aluminum at 10–20 cm Depth                    | cmol kg⁻¹    |
| Soil Boron at 10–20 cm Depth                       | mg kg⁻¹      |
| Soil Carbon at 10–20 cm Depth                      | %            |
| Soil Calcium at 10–20 cm Depth                     | cmol kg⁻¹    |
| Soil Calcium to Magnesium Ratio at 10–20 cm Depth  | %            |
| Soil Copper at 10–20 cm Depth                      | mg kg⁻¹      |
| Soil Iron at 10–20 cm Depth                        | mg kg⁻¹      |
| Soil Potassium at 10–20 cm Depth                   | cmol kg⁻¹    |
| Soil Magnesium at 10–20 cm Depth                   | cmol kg⁻¹    |
| Soil Manganese at 10–20 cm Depth                   | mg kg⁻¹      |
| Soil Nitrogen at 10–20 cm Depth                    | %            |
| Soil Sodium at 10–20 cm Depth                      | cmol kg⁻¹    |
| Soil Phosphorus at 10–20 cm Depth                  | mg kg⁻¹      |
| Soil Sulfur at 10–20 cm Depth                      | mg kg⁻¹      |
| Soil Zinc at 10–20 cm Depth                        | mg kg⁻¹      |
| Total Rainfall for the First Half of January       | mm           |
| Total Rainfall for the Second Half of January      | mm           |
| Total Rainfall for the First Half of February      | mm           |
| Variable name | Unit   |
|---------------|--------|
| Total Rainfall for the Second Half of February | mm     |
| Total Rainfall for the First Half of March    | mm     |
| Total Rainfall for the Second Half of March   | mm     |
| Total Rainfall for the First Half of April    | mm     |
| Total Rainfall for the Second Half of April   | mm     |
| Total Rainfall for the First Half of May      | mm     |
| Total Rainfall for the Second Half of May     | mm     |
| Total Rainfall for the First Half of June     | mm     |
| Total Rainfall for the Second Half of June    | mm     |
| Total Rainfall for the First Half of July     | mm     |
| Total Rainfall for the Second Half of July    | mm     |
| Total Rainfall for the First Half of August   | mm     |
| Total Rainfall for the Second Half of August  | mm     |
| Total Rainfall for the First Half of September | mm     |
| Total Rainfall for the Second Half of September | mm     |
| Total Rainfall for the Days 0–15 after Planting | mm |
| Total Rainfall for the Days 16–30 after Planting | mm |
| Total Rainfall for the Days 31–45 after Planting | mm |
| Total Rainfall for the Days 46–60 after Planting | mm |
| Total Rainfall from Planting to Harvest      | mm     |
| Average Monthly Mean Temperature in January  | °C     |
| Average Monthly Mean Temperature in February  | °C     |
| Average Monthly Mean Temperature in March     | °C     |
| Average Monthly Mean Temperature in April     | °C     |
| Average Monthly Mean Temperature in May       | °C     |
| Average Monthly Mean Temperature in June      | °C     |
| Average Monthly Mean Temperature in July      | °C     |
| Average Monthly Mean Temperature in August    | °C     |
| Average Monthly Mean Temperature in September | °C     |
| Average Monthly Mean Temperature in October   | °C     |
| Average Monthly Minimum Temperature in January | °C |
| Average Monthly Minimum Temperature in February | °C |
| Average Monthly Minimum Temperature in March  | °C     |
| Average Monthly Minimum Temperature in April  | °C     |
| Average Monthly Minimum Temperature in May    | °C     |
| Average Monthly Minimum Temperature in June   | °C     |
| Average Monthly Minimum Temperature in July   | °C     |
| Average Monthly Minimum Temperature in August | °C     |
| Average Monthly Minimum Temperature in September | °C |
| Average Monthly Minimum Temperature in October | °C |
| Average Monthly Minimum Temperature in November | °C |
| Average Monthly Minimum Temperature in December | °C |
| Annual Mean Temperature                       | °C     |
| Mean Diurnal Range (Mean of Monthly (Max Temp – Min Temp)) | °C |
| Isothermality (BIO2/BIO7) (× 100)              | na     |
| Temperature Seasonality (Standard Deviation ÷ 100) | °C |
| Max Temperature of Warmest Month              | °C     |
| Min Temperature of Coldest Month              | °C     |
| Temperature Annual Range (BIO5 – BIO6)        | °C     |
| Mean Temperature of Wettest Quarter           | °C     |
Table 1: Continued.

| Variable name                                      | Unit |
|---------------------------------------------------|------|
| Mean Temperature of Driest Quarter                | °C   |
| Mean Temperature of Warmest Quarter               | °C   |
| Mean Temperature of Coldest Quarter               | °C   |
| Annual Rainfall                                    | mm   |
| Rainfall of Wettest Month                         | mm   |
| Rainfall of Driest Month                          | mm   |
| Rainfall Seasonality (Coefficient of Variation)   | na   |
| Rainfall of Wettest Quarter                       | mm   |
| Rainfall of Driest Quarter                        | mm   |
| Rainfall of Warmest Quarter                       | mm   |
| Rainfall of Coldest Quarter                       | mm   |

Column 182: Rainfall of Driest Quarter (mm)
Column 183: Rainfall of Warmest Quarter (mm)
Column 184: Rainfall of Coldest Quarter (mm)

Dataset Item 2 (Table). It consists of time series NDVI data of 202 EGMs (i.e., Plot ID L1) for which this reading could be retrieved.

Column 1: Plot ID L1
Column 2: NDVI Date
Column 3: NDVI

4. Concluding Remarks

This comprehensive Dataset Paper is submitted to support research leading to the sustainable agricultural development of the neotropical savannas. Its specific design, however, responds to our interest in identifying management by environment interactions characterizing the potential for site-specific grain agriculture in the region. Our approach is informed by the rapidly growing literature demonstrating the promise of ecoinformatics approaches to streamline agricultural research [13–18]. Based on these experiences, we believe our Dataset Paper holds significant potential to facilitate a quantum leap in agricultural research for the development of the Colombian savannas.

Dataset Availability

The dataset associated with this Dataset Paper is dedicated to the public domain using the CC0 waiver and is available at http://dx.doi.org/10.1155/2015/625846/dataset.

Conflict of Interests

There is no conflict of interests in the access or publication of this Dataset Paper.

Authors’ Contribution

Soroush Parsa and Jaime Gómez Naranjo contributed equally to the study.

Acknowledgments

The authors’ most sincere gratitude goes to the farmers and farm managers that partnered with them in this study and to Santiago González Venzano from Solapa4 for generously sharing his experience and strategic vision for the project. They also thank Ximena Moreno and Nidia Zuleta for their administrative support; Mariano Tamburrino for kindly retrieving the NDVI data; Mariela Rivera and Harvey Parada for their invaluable recommendations; and Elcio Guimaraes for his strategic leadership. This project was generously funded by the Colombian Ministry of Agriculture and Rural Development.

References

[1] E. P. Guimaraes, J. Sanz, I. Rao, M. C. Amézquita, E. Amézquita, and R. J. Thomas, *Agropastoral Systems for The Tropical Savannas of Latin America*, International Center for Tropical Agriculture (CIAT), Cali, Colombia, 2004.
[2] R. J. Thomas and M. A. Ayarza, *Sustainable Land Management for the Oxisols of the Latin American Savannas: Dynamics of Soil Organic Matter and Indicators of Soil Quality*, International Center for Tropical Agriculture (CIAT), Cali, Colombia, 1999.
[3] L. R. Sanint, C. O. Sere, L. Rivas, and A. Ramírez, *The Savannas of South America: Towards a Research Agenda for CIAT*, International Center for Tropical Agriculture (CIAT), Cali, Colombia, 1992.
[4] N. Bourlag and C. Dowswell, “Feeding a human population that increasingly crowds a fragile planet,” in *Supplement to Transactions of the 15th World Congress of Soil Science*, International Society of Soil Science, Acapulco, Mexico, 1994.
[5] World Food Prize, “Background on Brazil’s Cerrado Region,” 2014, http://www.worldfoodprize.org/en/laureates/20002009_laurate/lobato_mcclung_paolinelli/cerrado/.
[6] P. Lavalle, N. Rodríguez, O. Arguello et al., “Soil ecosystem services and land use in the rapidly changing Orinoco River Basin of Colombia,” *Agriculture, Ecosystems & Environment*, vol. 185, pp. 106–117, 2014.
[7] G. Escobar, G. Rippstein, and F. M. Motta, *Agroecología y Biodiversidad de las Sabanas en los Llanos Orientales de Colombia*, International Center for Tropical Agriculture (CIAT), Cali, Colombia, 2001.
[8] J. Gómez, “Manejo del suelo en la Altillanura de los Llanos Orientales de Colombia,” in *Producción Eco-Eficiente del Arroz en América Latina*, V. Degiovanni, C. Martínez, and F. Motta, Eds., pp. 279–305, International Center for Tropical Agriculture (CIAT), Cali, Colombia, 2010.
[9] J. Blydenstein, “Tropical Savanna vegetation of the Llanos of Colombia,” *Ecology*, vol. 48, no. 1, pp. 1–15, 1967.
[10] C. Justice and J. Townshend, “Special issue on the moderate resolution imaging spectroradiometer (MODIS): a new generation of land surface monitoring,” *Remote Sensing of Environment*, vol. 83, no. 1-2, pp. 1–2, 2002.
[11] C. Kummerow, W. Barnes, T. Kozu, J. Shiue, and J. Simpson, “The tropical rainfall measuring mission (TRMM) sensor package,” *Journal of Atmospheric and Oceanic Technology*, vol. 15, no. 3, pp. 809–817, 1998.

[12] R. J. Hijmans, S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis, “Very high resolution interpolated climate surfaces for global land areas,” *International Journal of Climatology*, vol. 25, no. 15, pp. 1965–1978, 2005.

[13] J. A. Rosenheim, S. Parsa, A. A. Forbes et al., “Ecoinformatics for integrated pest management: expanding the applied insect Ecologist’s tool-kit,” *Journal of Economic Entomology*, vol. 104, no. 2, pp. 331–342, 2011.

[14] J. A. Rosenheim and M. H. Meisner, “Ecoinformatics can reveal yield gaps associated with crop-pest interactions: a proof-of-concept,” *PLoS ONE*, vol. 8, no. 11, Article ID e80518, 2013.

[15] S. Parsa, R. Ccanto, E. Olivera, M. Scurrah, J. Alcázar, and J. A. Rosenheim, “Explaining Andean potato weevils in relation to local and landscape features: a facilitated ecoinformatics approach,” *PLoS ONE*, vol. 7, no. 5, Article ID e36533, 2012.

[16] M. H. Meisner and J. A. Rosenheim, “Ecoinformatics reveals effects of crop rotational histories on cotton yield,” *PLoS ONE*, vol. 9, no. 1, Article ID e85710, 2014.

[17] D. Jiménez, J. Cock, A. Jarvis et al., “Interpretation of commercial production information: a case study of lulo (*Solanum quitoense*), an under-researched Andean fruit,” *Agricultural Systems*, vol. 104, no. 3, pp. 258–270, 2011.

[18] D. Jiménez, J. Cock, H. F. Satizábal et al., “Analysis of Andean blackberry (*Rubus glaucus*) production models obtained by means of artificial neural networks exploiting information collected by small-scale growers in Colombia and publicly available meteorological data,” *Computers and Electronics in Agriculture*, vol. 69, no. 2, pp. 198–208, 2009.
Submit your manuscripts at
http://www.hindawi.com