European Journal of Agronomy 99 (2018) 186–194

Contents lists available at ScienceDirect

European Journal of Agronomy

journal homepage: www.elsevier.com/locate/eja

# Environmental and management variables explain soybean yield gap variability in Central Argentina

Guido Di Mauro[a,*], Pablo A. Cipriotti[b], Santiago Gallo[c], José L. Rotundo[a]

[a] Instituto de Investigaciones en Ciencias Agrarias de Rosario – CONICET, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, S2125AA, Zavalla, Santa Fe, Argentina

[b] Universidad de Buenos Aires – CONICET, Facultad de Agronomía – IFEVA, Departamento de Métodos Cuantitativos y Sistemas de Información, C1417DSE, Ciudad Autónoma de Buenos Aires, Argentina

[c] Asociación Argentina de Consorcios Regionales de Experimentación Agrícola, C1041AAZ, Ciudad Autónoma de Buenos Aires, Argentina

ARTICLE INFO

*Keywords:*
*Glycine max* (L.) Merr.
Actual farmers' yield
Water-limited yield potential
Regression tree
Spatial analysis

ABSTRACT

Assessing yield gap (Yg) is required to identify opportunities for future yield increases. Central Argentina is one of the most productive soybean regions in the world. In this region, soybean is planted after a winter fallow period (from now on soybean as single crop) or after the harvest of a winter crop (from now on soybean as second crop). Information regarding options for obtaining even higher yields is limited. The objectives of this paper are: i) to estimate Yg of soybean as single or second crop, ii) to identify management and environmental variables associated with soybean Yg variability, and iii) to assess the spatial distribution of soybean Yg. A farmers' survey with ~22,500 field observations from 2003 to 2015 was compiled. Water-limited yield potential (Ywlim) was estimated as the 95th percentile of actual farmers' yield (Ya) across years. Yield gap was the difference between Ywlim and Ya, expressed as a percentage of Ywlim. Factors associated with Yg were evaluated using regression trees. Ordinary kriging was used to explore spatial patterns of Yg. Average Ywlim were 5095 and 4337 kg $ha^{-1}$ for single and second crop, respectively. Average Yg were 28.7 and 33.5% for single and second crop, respectively. Yield gap showed a wide range of variation. Management accounted for 66 and 91% of explained variation in Yg for single and second crop, respectively. Gap closing for single crop was associated with earlier planting and maize as previous crop. Gap closing for second crop was associated with foliar fungicide utilization, P fertilization, and earlier planting. Single crop Yg was spatially auto-correlated, whereas no auto-correlation was observed for second crop. The spatial structure of single crop was represented by an exponential model, with 81% of total variation explained by the spatial structure and a maximum range of auto-correlation of approximately 120 km. This result is consistent with the observed spatial auto-correlation of variables explaining Yg in single crop. Our approximation allowed the characterization of the magnitude, possible explaining factors, and spatial dependence of soybean Yg in one of the most productive regions in the world. Although average gaps are relatively small compared to those in other regions, there are still opportunities for future yield improvements.

## 1. Introduction

The increase in global crop production will play a crucial role to satisfy food demand in coming years (Godfray et al., 2010). Attaining this goal requires increasing yield per unit land area given that new farming land is currently lacking (Foley et al., 2011). One alternative for increasing yield is closing yield gaps (Yg) at the farm level. Estimating Yg at farm level requires comparing actual farmers' yield (Ya) to some measure of potential yield (or water-limited yield potential in rainfed cropping systems, Ywlim) (Van Ittersum et al., 2013). Potential yield can be estimated by crop models, maximum-yield field experiments, or maximum farmers' yields. These three measures of potential yield, when compared to Ya, allow the calculation of model-based Yg, experiment-based Yg, and farmer-based Yg, respectively (Lobell et al., 2009). Even though model-based Yg analysis is the standard approximation (Van Ittersum et al., 2013; Van Wart et al., 2013a), farmer-

*Abbreviations:* Ya, actual farmers' yield; Ywlim, water-limited yield potential; Yg, yield gap

* Corresponding author at: Instituto de Investigaciones en Ciencias Agrarias de Rosario, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, Campo Experimental Villarino s/n, Zavalla S2125AA, Argentina.

*E-mail address:* dimauro@iicar-conicet.gob.ar (G. Di Mauro).

https://doi.org/10.1016/j.eja.2018.04.012
Received 4 September 2017; Received in revised form 17 April 2018; Accepted 24 April 2018

based analysis has been also widely used for global or regional analysis of different field crops (i.e. Licker et al., 2010; Egli and Hatfield, 2014a,b; Tanaka et al., 2015; Ernst et al., 2016). Our general objective was to assess farmer-based Yg for soybean production. According to Lobell et al. (2009) farmer-based Yg analysis is only appropriate in intensively managed cropping systems and when analyzing many fields, in order to increase the chances of attaining at least one field with yield close to potential. We focused our analysis on ~22,500 field observations from a subgroup of farmers with high level of technology adoption from the Central region of Argentina. The estimation of farmer-based Yg provides an important measure of opportunities to improve crop production under current cropping systems technology.

Yield gap analysis can help to identify regions or production systems where highest priority should be given to successfully increase crop productivity (van Oort et al., 2017). The possibilities of increasing yield are highest in situations where Yg is large enough (> 20%) (Lobell et al., 2009). Bhatia et al. (2008) showed model-based soybean Yg of 54% for farmers in India. Zhang et al. (2016) found model-based Yg of 16% in soybean across years with different levels of water supply in China. Egli and Hatfield (2014a) found that average soybean farmer-based Yg ranged from 9 to 24% across three states in the U.S. Midwest across a 40-yr period. Grassini et al. (2015a) and Rattalino Edreira et al. (2017) showed that model-based combined with farmer-based Yg estimations of soybean producers in U.S. were 32 and 22% under rainfed conditions, respectively. Sentelhas et al. (2015) found that average soybean model-based Yg was 13% in rainfed systems of Brazil. Aramburu Merlos et al. (2015), using Global Yield Gap Atlas approach (www.yieldgap.org), represented the first attempt to evaluate soybean Yg in Central Argentina and found an average of 25% model-based Yg under rainfed conditions. However, information regarding potential causes (management or environmental factors) of Yg variation and spatial distribution within Central Argentina are scarce. Yield gap analysis has recently been expanded to double cropping systems to identify possibilities of yield improvement or design new farming systems (Guilpart et al., 2017). However, information regarding environmental causes of Yg of soybean as second crop after the harvest of a winter crop in Argentina is currently limited (Andrade and Satorre, 2015).

Identifying management and environmental variables associated with Yg is critical for decision-making regarding Yg closure. Different techniques can be utilized to this end. Regression trees have been used to explore explanatory variables of wheat Yg (Ernst et al., 2016). However, regression tree approach has been successfully utilized to identify variables associated with yield of different crops. For instance, they were utilized to explore factors associated to yield variability in wheat (Lobell et al., 2005), maize (Tittonell et al., 2008), rice (Tanaka et al., 2015), sugarcane (Ferraro et al., 2009) and soybean (Mourtzinis et al., 2018; Zheng et al., 2009). This approach has several advantages for being used to analyze field surveys at regional scale (De'Ath and Fabricius, 2000). Briefly, regression trees are easy to interpret, variable selection is unbiased, non-linear relationships between variables can be unraveled, and there are no distributional assumptions of the response variable. Additionally, the trees handle both categorical and continuous variables and allow missing data. Therefore, a regression tree approach will be utilized to explore management and environmental explaining factors of soybean Yg across ~22,500 field observations in Central Argentina.

Yield gap analysis could be conducted under different spatial scales (Sadras et al., 2015). Previous studies focused on methodologies to scale up location-specific Yg estimations to larger spatial areas (van Bussel et al., 2015; Van Wart et al., 2013a,b). This protocol, based on determining homogeneous areas with respect to environmental conditions, was used in Yg analysis in Argentina (Aramburu Merlos et al., 2015). An interesting alternative is to incorporate a geostatistical approach at more detailed spatial scales (e.g. farms or paddocks) to improve the spatial resolution and accuracy of regional Yg analysis (Lobell and Ortiz-Monasterio, 2006; Steinbuch et al., 2016). Therefore, we propose this alternative method to identify spatial patterns in Yg magnitude through geostatistical techniques. Mapping this variability can help the development of spatially specific agronomic strategies aimed at closing Yg for specific areas as was shown for maize in Bangladesh (Schulthess et al., 2013) and Africa (Van Dijk et al., 2012).

There is a clear need to synthesize crop yield, climate, soil and management data from different areas to identify crop production limitations (Lobell and Asner, 2003). Yield gap analysis using farmers' survey constitutes an opportunity to achieve this objective (Beza et al., 2017). Local studies are needed to understand and dissect the role of agricultural system characteristics and biophysical conditions in closing Yg (Rattalino Edreira et al., 2017). In this context, we used field observations across the main soybean production area of Central Argentina to accomplish the following objectives: i) estimate soybean Yg, ii) identify environmental and management variables associated with soybean Yg; and iii) explore spatial distribution of soybean Yg across Central Argentina.

## 2. Materials and methods

### *2.1. Study area and farmers' survey description*

The study area is part of Central Argentina, and soybean is the main crop in this area. Soybean can be planted after a fall/winter fallow period (April–September) with the previous crop being another summer crop (e.g. soybean or maize) that grew during the previous warm growing season (September–April). Therefore, soybean is planted after a fall/winter fallow period which begins after previous summer crop harvest. Usual planting dates range from October to mid-December. This soybean crop is referred as "single crop". Alternatively, soybean can be planted after a winter grain crop is harvested or after a cold-season grass crop is dried with herbicide or used directly as animal feed. The most common winter grain crop in this region is wheat (*Triticum aestivum* L.), while the cold-season grass crop can also be wheat (terminated before full maturity) or rye-grass (*Lolium multiflorum*). Winter crops for grain are usually harvested in late November to December, while grass crops are dried/fed earlier. Soybeans are therefore planted from early December to mid-January. This soybean crop is referred as "second crop".

The study area has a monsoonal climate with rainfall concentrated in the summer season (December-February) (Hall et al., 1992). There is substantial interannual rainfall variability associated with the El Niño Southern Oscillation phenomenon (Podestá et al., 1999). Soils are predominantly Mollisolls (USDA, 1975) having no major physical or chemical limitations.

Farmers' surveys under analysis were provided by farmers' members of Southern Santa Fe Region of Argentine Association of Regional Consortiums for Agricultural Experimentation (AACREA). Soybean yield and management data were compiled from 2003 to 2015. Post-2003 data were entirely based on no-till conditions and herbicide-resistant GMO soybean production. The widespread and rapid adoption of a no-till management strategy and Roundup Ready® soybean germplasm (Monsanto Company, St. Louis, MO) represent major changes in production technology (Satorre, 2011). The timeframe analyzed ensured that approximately the same overall technology was used to avoid abrupt productive leaps across years, while also allowed the construction of a sufficient large climatic data set. Each observational unit corresponded to a specific field in a particular year. All fields were managed with available farmer technology under no-till and rainfed conditions. Separate farmers' surveys were maintained for single (n = 15,522) and second soybean crops (n = 7,112).

Management variables extracted from farmers' surveys were: previous crop, sowing date, row spacing, plant population, maturity group, nutrient rate applied by fertilization, and fungicide and insecticide use (Table 1). Each observation was georeferenced using the closest