Published July 11, 2016

Wheat Yield Gap in the Pampas: Modeling the Impact of Environmental Factors

Josefina L. De Paepe* and Roberto Álvarez

ABSTRACT

As global grain demand is expected to keep on rising, productive but underachieving regions like the Argentine Pampas play a key role. Reduction of yield gaps in these regions would allow an increase in global food production. The objectives were to model the spatial patterns of the wheat (*Triticum aestivum* L.) yield gap in the Pampas and relate it to environmental factors. The study comprised an area of approximately 45 Mha during a 40-yr interval. Attainable yield was estimated by a stochastic frontier production function adjusted on statistical data generated at county scale. Yield gap was calculated for each combination of climate and soil variables as the difference between attainable yield and the average yield estimated using an artificial neural network (ANN) model. Yield gap was then modeled by another ANN using as inputs climate and soil factors. Average yield gap was 865 kg ha^{-1} (25% of average attainable yield), ranging from 740 kg ha^{-1} (26%) in humid environments to 1140 kg ha^{-1} (42%) under semiarid ones during the last 5 yr. Yield gap could be adequately modeled with an ANN (R^2 = 0.745, RMSE = 144 kg ha^{-1}). The model showed that soil factors deeply impacted yield gap and minimum values were obtained in soils with medium to high organic C contents and available water storage capacity. Yield gap and a soil productivity index developed locally were negatively correlated. The methodology developed for yield gap analysis can be used for other crops and regions.

Core Ideas

- Yield gap calculation combined two modeled yield levels, applying a frontier production function and an artificial neural network.
- Climate partially defined yield gaps; in semiarid environments these were largest.
- Soil properties explained 50% of yield gap variability at regional scale.
- Soil organic carbon and available water holding capacity interacted positively defining a minimum yield gap.
- Yield gap reducing efforts should be focused in low productivity soils.

Published in Agron. J. 108:1367–1378 (2016)
doi:10.2134/agronj2015.0482
Received 29 Sept. 2015
Accepted 2 Mar. 2016

Copyright © 2016 by the American Society of Agronomy
5585 Guilford Road, Madison, WI 53711 USA
All rights reserved

Yield gaps are defined as the difference between production levels over a specified scale of interest (Evans and Fischer, 1999; Van Ittersum and Rabbinge, 1997). In rainfed cropping agroecosystems there is a common, often large, yield gap between attainable and actual yields (Hochman et al., 2013). Attainable yield represents yield under water and nutrient-limiting conditions but without impacts of reducing factors like pest, diseases, etc; it can be used as a reference value in agroecosystems where crops go through water limiting periods during their growing cycle as is the case in many rainfed agroecosystems, and it relates to the best current management practices (Hochman et al., 2012). Attainable yield is soil type dependent as it relates to soil properties like water holding capacity (Van Ittersum et al., 2013). Actual yield is the production level obtained under production conditions where not only water and nutrients reduce yield, but also many other factors like pests, diseases, erroneous management practices, etc. (Lobell et al., 2009). Actual yield is the lowest production level used for yield gap calculation. This latter production level is regulated by environmental constraints under common management practices and is usually lower than attainable yield (Cassman et al., 2003).

Yield gap identification for particular crops and production regions provides a framework to prioritize research and policy efforts to narrow it (Tittonell et al., 2008), and in consequence to meet the continued rise in global food demand (Lobell et al., 2009). Yield gap analysis was focused on its quantification (Lobell, 2013) and spatial pattern description (Hochman et al., 2013), but a better understanding of the environmental factors that define it is required to develop improved management practices to narrow it.

Some causal relationships between yield gaps and climate or management factors were detected previously, but only a few also included soil properties. Wheat yield gap was positively associated to rainfall in western Australia (Anderson, 2010) and central United States (Patrignani et al., 2014). Also the rice (*Oryza sativa* L.) yield gap was reported to be greater in humid environments of Sub-Saharan Africa (Sileshi et al., 2010). Regarding the management factors, rice yield gap decreased when fields were fully irrigated in Cote d'Ivoire (Becker and Johnson, 1999) and also with N fertilization according to relief position in Thailand (Boling et al., 2011). Finally, a minority of works also looked for causal relationships between soil properties and yield gap, for instance wheat yield gap increased with larger soil evaporation

J.L. De Paepe and R. Álvarez, Faculty of Agronomy-University of Buenos Aires, Av. San Martín 4453, Buenos Aires, AR 1417, Argentina, CONICET. *Corresponding autor (depaepe@agro.uba.ar).

Abbreviations: ANN, artificial neural network, PET, potential evapotranspiration; SAWSC, soil available water storage capacity.

(Sadras and Angus, 2006) and with increasing soil N at sowing (Peake et al., 2014) in semiarid Australia.

National statistics information is usually published at county scale and can represent the actual yield production level (Egli and Hatfield, 2014). For yield gap analysis, estimation of attainable yield production level is required. Attainable yield can be estimated by process-based models (Grassini et al., 2011) but in many areas validation can be difficult due to the lack of quality data. It can be estimated as well by the stochastic frontier production function (Aigner et al., 1977; Meeusen and Van den Broeck, 1977). Attainable yield estimated with the frontier function represents the maximum value of the dependent variable for a determined set of independent variables (Pesaran and Schmidt, 1999) and thus results from the combination of these variables and obtained under current limiting conditions (Neumann et al., 2010; Patrignani et al., 2014). When actual is lower than attainable yield, this is caused by production inefficiency (Coelli et al., 2005). The error term of the frontier function term is split into inefficiency and a stochastic component (Coelli et al., 2005). The latter may be caused by statistical noise of data errors and uncertainties inherent to reporting and sampling of information. An advantage of the frontier function is the consistent use of one data set of national statistics yield information for attainable yield determination (Neumann et al., 2010).

Even if there have been attempts of finding causal relations between yield gaps and independent variables, to our knowledge no yield gap empirical model has been developed yet. Artificial neural networks (ANNs) have become a popular statistical technique because they can be used to construct an explicative yield gap model. This technique simulates the neural working of the human brain and does not require an a priori data structure (Özesmi et al., 2006). Predictions have significantly smaller errors than the more traditional linear regression models for yield prediction (Alvarez, 2009). The relations that link input data (independent variables) to output data (yield gaps) are obtained and implemented in an iterative calibration procedure (Schaap et al., 1998). Complex input–output relations are modeled using ANN as these are capable of detecting patterns and learning relationships comprised of curvilinear effects and interaction between variables (Dai et al., 2011).

The vast fertile Argentine plain called Pampas (Hall et al., 1992) plays a key role in the international food security situation as the agricultural area for export increased its production during the last four decades (FAOstat, 2015; Imhoff et al., 2004). Rainfed wheat production in this region started a century ago and has the largest spatial coverage, both under humid and semiarid conditions (Hall et al., 1992), and concentrates 95% of national production (MinAgri, 2015). Because of the wide spatial distribution of this crop in the Pampas across areas with contrasting climate and soil characteristics it offers a good opportunity for studying environmental factor impacts on the yield gap. Underperforming regions like the Pampas with favorable rainfed conditions represent good opportunities to improve yields and bridge the gap between attainable and actual yields (Cassman, 1999; Foley et al., 2011).

Although the wheat yield gap has not been determined in the Pampas in local studies, some global-scale research has addressed this region and some coarse comparisons can be made. A wheat yield gap of 15% was calculated for Argentina as the difference between modeled attainable yield and actual yield based on statistical data without the identification of possible factors influencing the magnitude (Liu et al., 2007). Other studies estimated attainable yield using the highest values from historical statistical records and calculated a wheat yield gap rounding 40% for the whole Pampas (Licker et al., 2010; Mueller et al., 2012). Regarding the spatial distribution of this yield gap, some contradictory results were found. One of these estimations reported that the average maize (*Zea mays* L.)–wheat–rice gap decreased from 30 to 50% in the humid portion of the region to approximately 10% in some semiarid areas; implying that in the semiarid environments of the Pampas actual yield levels are close to the attainable (Mueller et al., 2012). Using a stochastic frontier production function approach and statistical data, a yield gap of approximately 50% was also estimated for Argentina and with larger yield gaps in the more humid environments of the Pampas (Neumann et al., 2010). Using spatial datasets around the year 2000 larger yield gaps in the semiarid portion of the Pampas were also detected (Licker et al., 2010). A common result of all previous global scale yield gap estimations for wheat is that yield gap in the Pampas was much larger than in some other important production regions of the World (Licker et al., 2010; Mueller et al., 2012; Neumann et al., 2010).

Due to the large scale of all this previous research, the resolution of the database used was low or incomplete and soil property information was barely accounted for. This study thus aimed to model the spatial patterns of the wheat yield gap in the Pampas and relate it to environmental factors. The focus was to analyze possible effects of soil properties on the yield gap by separating them from climate influences.

MATERIALS AND METHODS

Study Region

The Argentine Pampas (28° S and 40° S and 57° W and 68°W) comprises a approximately 60 Mha temperate vast plain, with a flat or slightly rolling relief, where the natural vegetation consists of grasslands with dominant graminaceous plant species (Hall et al., 1992). Mean annual temperature ranges from 23°C in the North down to 14°C in the South. Annual rainfall varies from 1200 to 200 mm following a Northeast–Southwest gradient. The majority of the soils are fertile and well-drained, mainly Mollisols, formed on loess-like materials (Alvarez and Lavado, 1998). Rainfed crops are widespread in the Pampas, mostly in the humid and semiarid portions where annual rainfall exceeds 500 mm. The humid portion corresponds to the eastern side of the region and includes the northeastern areas were annual rainfall can exceed 1000 mm while the semiarid portion is located at the western side were annual rainfall is low and variable. Approximately 65% of the area, on well drained soils, is under agricultural use, and the most important grain crops are soybean [*Glycine max* (L.) Merr.], corn, and wheat (MinAgri, 2015). In this area forage crops are rotated with grain crops. The area under wheat represents approximately 10 to 15% of the whole agricultural surface, depending on the year. Only after approximately 1995 fertilization became a common practice but application rates compared to other countries are low (FAO, 2004). The natural fertility of soils combined with the short cropping history, where annual crop production was alternated with livestock production on leguminous pastures, sustained this low input scenario (Alvarez et al., 2015). Usually a fallow period of approximately 3 mo, from April to June, preceded wheat cultivation. Wheat growing cycle starts in June or July, depending on the sowing date, and ends in December with crop harvest.

Database Description

Wheat yield data of 150 counties were calculated using national statistics on annual grain production and harvested area per county covering a time interval of 40 yr, specifically 1967 to 2006 (MinAgri, 2015). These counties were distributed along the provinces of the Pampas of Buenos Aires, Córdoba, Entre Ríos, La Pampa, and Santa Fe, where wheat is a common crop rotation component. The covered area was of approximately 45 Mha which represents about 75% of the Pampas and where 95% of the national wheat production is generated. Forty of these counties were discarded because hydromorphic soils dominated and <30% of the total county area was under wheat production during the comprised time interval. Because of enormous differences in county areas, the largest county was up to 30 times larger than the smallest one; data were aggregated for processing and statistical analysis. This spatial aggregation resulted in the generation of 41 geographic units of 1 ± 0.5 Mha as was previously described (De Paepe and Álvarez, 2013). The aggregation was performed accounting for relief, type of landscape, and dominant soil classes of the subregions of the Pampas described in previous works (Alvarez and Lavado, 1998; Hall et al., 1992). Weighted averages of county yield, climate, soil, and management information were calculated for all variables to aggregate information up to the geographic unit scale.

Climate records of monthly maximum and minimum temperature and rainfall covering the 1967 to 2006 time interval from approximately 80 weather stations well distributed over the study area and its surroundings, were obtained from the Servicio Meteorológico Nacional (SMN, 2015) and the Instituto Nacional de Tecnología Agropecuaria (INTA, 2015). County-scale averages of these three climate variables were spatially interpolated with good results using the inverse distance weighting method using the Spatial Analyst extension of ArcGIS 9.1 (ESRI) as was reported previously (De Paepe and Álvarez, 2013). This spatial interpolation method consists of assigning values to unknown points calculated as weighted averages of the values available at the known points, specifically the meteo-stations in our analysis (Zimmerman et al., 1999). In total 960 interpolation maps were generated and monthly temperature was calculated as the average of maximum and minimum monthly values. Estimated climate variables were validated with measured information at 30 meteo-stations spread across the Pampas. The best estimation performance was obtained for maximum and minimum temperature and a difference of 5% was detected when comparing measured with estimated rainfall values, probably due to the large variability of rainfall (De Paepe and Álvarez, 2013). The crop growing cycle was split into a vegetative growing phase, from July to September, and a flowering-to-maturity phase, from October to November. Both crop phases were preceded by a fallow period before crop seeding took place, during April to June.

Potential evapotranspiration (PET) was calculated with a modified version of the Penman formula (Linacre, 1977) using the estimated monthly temperature and rainfall. Locally adjusted k_c coefficients (a crop coefficient that accounts for differences in crop type, cultivar, and development stage that should be considered when assessing evapotranspiration) were applied to estimate wheat PET (Doorenbos and Pruitt, 1977; Totis and Perez, 1994). A k_c coefficient of 0.5 was assumed for the fallow period as no coefficients for this period were available. Ratios of rainfall to PET were calculated monthly and also integrated per crop phases like in a previous study (Alvarez, 2009).

The photothermal quotient was estimated for the critical crop period of 1 mo before wheat crop anthesis as the ratio of incoming radiation and mean temperature above 4.5°C as described in Fischer (1985). Anthesis dates along the 150 counties varied with latitude from 30 September in the North to 10 November in the South, this was accounted for taking published information from experiments (Magrin et al., 1993). A locally developed modification of the Hunt et al. (1998) model was used to estimate atmospheric transmittance (Alonso et al., 2002) and solar radiation at the top of the atmosphere was calculated with RadEst 3.00 algorithms (Donatelli et al., 2003). These methods were described in detail previously (Alvarez, 2009).

Descriptions and corresponding influence areas from more than 1000 soil profiles published in the soil surveys of the provinces of Buenos Aires, Córdoba, Entre Ríos, La Pampa, and Santa Fe were used to spatially aggregate information up to the geographic unit scale as was described previously (De Paepe and Álvarez, 2013). The following soil variables were included in the analysis: organic C, textural composition, pH, carbonate C, and depth to petrocalcic horizon when present within the upper 1 m of the soil profile. Soil variables were described per genetic horizons and were modeled in layers of 25 cm up to 1 m depth by fitting functions with Table Curve WD (Systat Software) with good results (R^2 > 0.90) and the models used were generally of the exponential or the potential type. Using the method of Álvarez and Lavado (1998) soil information at the profile level was aggregated to the cartographic scale, and afterward to the county and geographical unit scale accounting for corresponding influence areas. As the soil surveys used for data acquisition were performed mainly during the 1960 to 1980 period; soil organic C content data obtained from those surveys was used for yield gap modeling during the first years of our analysis (1967–1976). Organic C data generated during a recent soil sampling during 2007–2008 (Berhongaray et al., 2013) was used as the representative county soil C content at the end of the analyzed interval (1997–2006). For intermediate time periods linear interpolation was used for organic C estimation.

Soil available water storage capacity (SAWSC) was calculated as the difference between estimated soil field capacity, at a matric potential of −0.033 MPa, and wilting point, at a matric potential of −1.50 MPa using the Rawls et al. (1982) equations. The effects of the linear regression equations described by Rawls et al. (1982) included sand, clay, and organic matter percentages for field capacity estimation and clay and organic matter percentages for wilting point estimation. These equations were used to calculate SAWSC per layer of 25 cm to the upper 1 m of the soil profile or up to the upper limit of the petrocalcic horizon when it appeared in the first meter of the soil profile. Gravimetric water was transformed to volumetric water using bulk density values estimated by the Rawls (1983) method that used organic matter percentage and mineral bulk density derived from soil textural composition. In a previous work it was detected that this bulk density estimation method overestimated soil density of soils of the Pampas by 4% (Berhongaray et al., 2013); therefore a uniform correction factor of 0.96 was applied to all estimated values.

The management factors included in the analysis were fertilizer rate, tillage system, and relative genetic improvement. An official survey performed on more than 3000 farms with fertilizer rate data corresponding to the 2002 to 2006 interval was available (RIAN, 2015). Mean N and P county rates were estimated as the mean applied fertilizer rate by all the farmers per county included in the survey. Fertilizer rate information was not available for about 20% of the counties therefore it was estimated using the information from nearby counties with the inverse distance weighting factor mentioned earlier. As no past data were available, the national fertilizer consumption trend was used to estimate past fertilizer rates (Alvarez et al., 2015).

Predominant tillage system per decade and area of the Pampas was obtained from expert opinion (M. Taboada, G. Studdert, A. Bono and C. Quintero, personal communications, 2014). Tillage systems accounted for were: conventional tillage (moldboard or disk plow) and conservation tillage (harrow disk, chisel tillage, and no-till). The fraction of county under each tillage system was calculated based on this information and aggregated to the unit scale. Genetic improvement was estimated accounting for the spatial distribution of potential yield that represents the biophysical yield ceiling at a given location achieved under nonlimiting water and nutrient conditions (Evans and Fischer, 1999) and serves as a proxy of the mentioned improvement as no other factor can be reducing yield. Potential yield, estimated with CERES-Wheat, was obtained from a map elaborated by Magrin (2004) at regional scale and all values were calculated relative to the maximum. Using the locally estimated average annual genetic improvement of 0.74% yr^{-1} (Calderini et al., 1995), past potential yield was estimated per geographic unit by linear interpolation. By this procedure the evolution of relative potential yield across time and space was calculated in the Pampas. This new variable was used for subrogating the factor genetic improvement in our analysis.

Attainable Yield

Frontier functions were originally developed for econometric analysis to calculate firm efficiencies (Aigner et al., 1977; Meeusen and Van den Broeck, 1977) and were applied to estimate attainable yields (Battese and Coelli, 1995). Inefficiency of agricultural production results in yield values that differ from attainable yields as these are located under the limit defined by the frontier function. Statistical noise, caused by data errors and uncertainties, is estimated along with the inefficiency (Coelli et al., 2005). In frontier analysis, the Cobb–Douglas function is the most widely used model (Battese and Coelli, 1995):

$$q_i = \beta x_i + \nu_i - \mu_i \quad [1]$$

where, for our analysis, q_i is the attainable yield of a unit (i = 1, 2, ..., 41); β is the vector of the unknown parameters, x_i is a vector that comprises the independent variables; ν_i is a random (i.e., stochastic) error to account for statistical noise that can be positive or negative with zero mean; and μ_i is a non-negative variable associated with the technical inefficiency and is independent of ν_i. The symmetric random error term ν_i comprises the statistical noise which arises from unnoticed omission of relevant variables in vector x_i and measurement errors. Frontier function outputs are limited from above by the random or stochastic variable exp $(x_i' \beta + \nu_i)$. The term μ_i represents a decrease in yield that results from management factors that are insufficient or not applied properly or environmental constraining conditions. Estimated attainable yields can lie above or below the deterministic frontier function depending on the ν_i error term. The model is fitted using maximum log likelihood methods. The attainable yield fitted using the frontier function to our yield data set is equivalent to the attainable yield concept and represents the best attained yield for a combination of production factors. Soil factors used as independent variables for estimating attainable yield were soil organic C, SAWSC, clay, silt, and sand content, pH, and soil inorganic C at different layers. Climate factors tested in the frontier function were monthly temperature, monthly rainfall, monthly PET, the ratio rainfall/PET, photothermal quotient and radiation during crop critical period, and the weighed combination of climate variables during fallow, crop vegetative, and flowering periods. Time was also used as an independent variable subrogating the effect of technology improvement because all management factors tested in the analysis were autocorrelated (see below). A positive coefficient in the final model indicated that the tested variable had a positive influence on the dependent variable; a negative coefficient indicated the opposite. The model was estimated using FRONTIER 4.1 software (Coelli, 1994).

Yield Gap Quantification and Modeling

Yield gap at the geographic unit scale was calculated as the difference of: (i) attainable yield–statistical observed yield and (ii) attainable yield–mean ANN yield derived from a previously developed regional yield estimation model (De Paepe and Álvarez, 2013). Mean ANN yield comprised the average wheat yield estimated with a local fitted ANN model. It included the same combinations of environmental factors for attainable yield determined with the frontier function. A data set of 1640 yield gap values was calculated. These yield gaps were used later as dependent variable for the development of a yield gap estimation model using again the ANN method for evaluating regional soil and climate factor impacts.

When testing ANNs as modeling methods for regional yield gap, the gap was the output and the same variables tested for defining attainable yield were used as inputs. Tested network architecture comprised three neuron layers: input, hidden, and output. Linear transfer functions were used from the input layer to the hidden layer and also from the output layer to the network output, and sigmoid functions connected the hidden to the output layer (Lee et al., 2003). A feed-forward ANN was used as this type of network was proven to be suitable for yield analysis at various scales (Kaul et al., 2005). Supervised learning procedure for weight fitting was performed with the back propagation algorithm (Rogers and Dowla, 1994). Network architecture simplification, scaling methods, learning rate, and epoch size were similar to those described by Alvarez (2009) and De Paepe and Álvarez (2013). Selected independent variables had a sensitivity ratio greater than one according to a sensitivity analysis (Miao et al., 2006). The data set was partitioned into 50% for training, 25% for validation, and 25% for testing to avoid over learning (Özesmi et al., 2006). Models were adjusted with the training set and early stopping of weight fitting was achieved when the R^2 of the validation set was lower than the R^2 of the training set (Kleinbaum and Kupper, 1979). The test set was used for an independent model validation. Artificial neural networks were fitted using Statistica Neural Networks (version 2011, Stat Soft.).

Yield gap was related to regional productivity by plotting it against a soil productivity index developed in the Pampas for wheat (De Paepe and Álvarez, 2013). This normalized index represented the historical yield production per geographic unit of the Pampas modeled at regional scale by a previously developed ANN model that included defining soil and climate factors. Regression and correlation analysis were performed for searching associations between variables testing significance by the F test ($P < 0.05$). Model performances were compared using R^2 and root mean squared error (RMSE) (Kobayashi and Salam, 2000). Possible differences between the R^2 values were tested by a specific test using Fisher's Z transformation (Kleinbaum and Kupper, 1979). Both the attainable yield and yield gap were graphically presented by maps using Quantum GIS Development Team (2015). Sensitivity analysis of the effects of environmental factors on the yield gap was performed as indicated in Alvarez and Grigera (2005) for estimating their effects on the gap.

RESULTS

The rainfall/PET ratio during the fallow and crop cycle showed a marked longitudinal spatial pattern with values as high as 1.5 in the East side to 0.6 in the West border of the region (Fig. 1A). As the photothermal quotient was related to temperature it followed

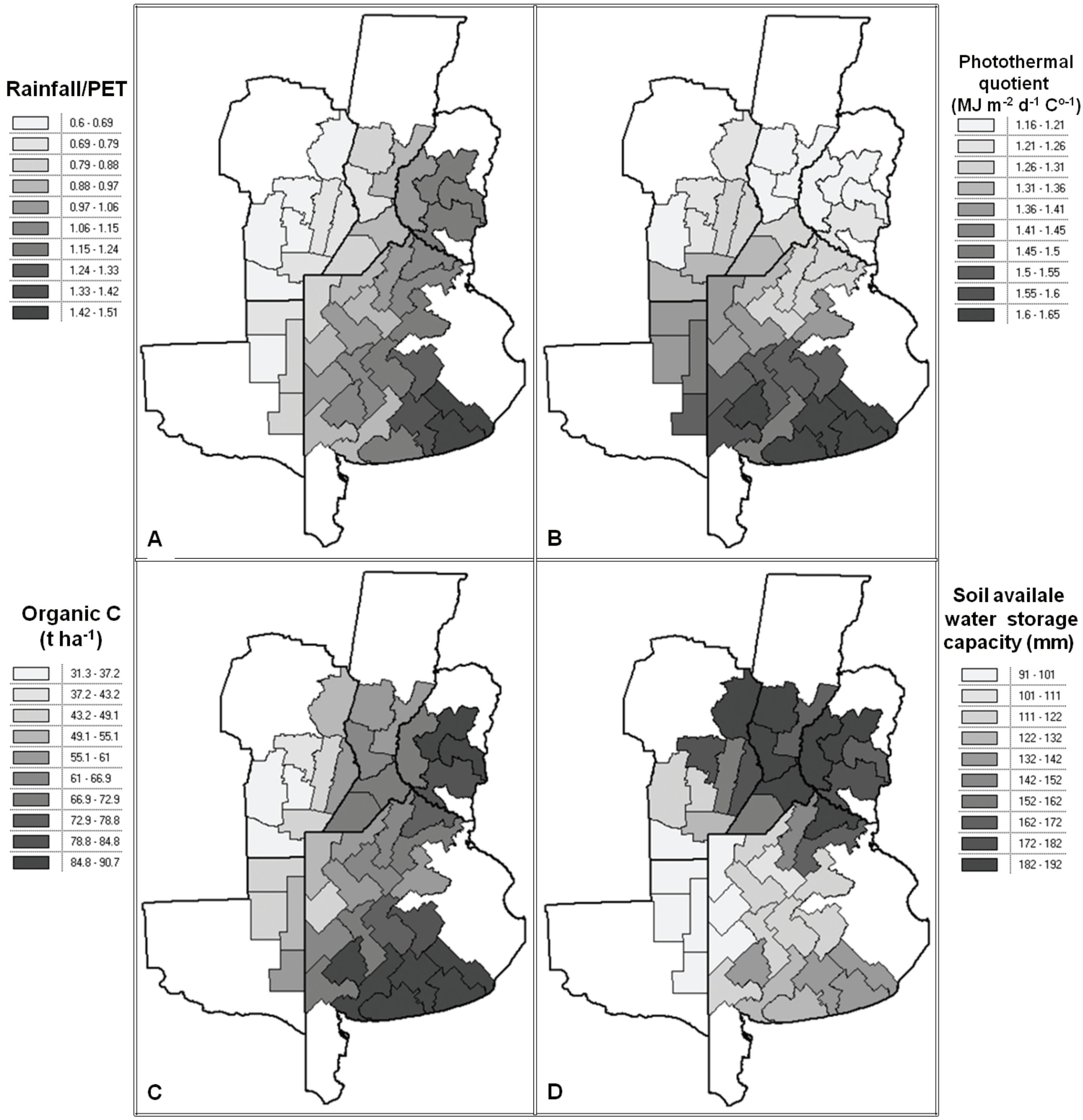

Fig. 1. Geographical distribution of: (A) the average rainfall/potential evapotranspiration (PET) corresponding to fallow and crop growing cycle periods, (B) photothermal quotient during the crop critical period, (C) average soil organic C content in the upper 50 cm of the soil profile for the total time series accounted for, and (D) soil available water storage capacity up to 1-m depth. The blank sections correspond to areas within the Pampas where wheat was not an important crop and therefore were not accounted for in this analysis. The presented information corresponds to the 1967 to 2006 period.

Table 1. Pearson correlation coefficients between time, management-related variables, and crop yield. All significant at $P < 0.05$.†

Variable	Time	Yield	N fert.	Conserv. tillage
	yr	—— kg ha^{-1} ——		%
Yield, kg ha^{-1}	0.612			
N fert. kg ha^{-1}	0.834	0.647		
Conserv. tillage, %	0.766	0.534	0.724	
Genetic improve., %	0.681	0.679	0.683	0.482

† Fert., fertilizer; Conserv, conservation; Improve., improvement.

Table 2. Coefficients for the parameters of the yield frontier function at the geographic unit scale (n = 1640).

Parameter and variable	Coefficient	SE	$P > z$
β_1 Year	37.3	1.173	0.00
β_2 SAWSC†	2.32	0.35	0.00
β_3 Rainfall/PET fallow‡	163.8	21.2	0.00
β_4 Rainfall/PET vegetative growth§	210	25.2	0.00
β_5 Photothermal quotient¶	419	47.7	0.00

† SAWSC, soil available water storage capacity 0 to 100 cm (mm).
‡ PET, potential crop evapotranspiration from April to June.
§ Potential crop evapotranspiration from July to September.
¶ During crop critical period (MJ m^{-2} d^{-1} °C^{-1}).

a North–South gradient but its variation was much lower than the rainfall/PET ratio (Fig. 1B). No temporal trends along the 40 yr study period were detected in temperature and radiation during the crop growing cycle. For example, average temperature during the growing cycle in the 1967 to 1972 period was 15.8°C and in the 2001 to 2006 period 15.7°C, meanwhile average radiation was 16.0 MJ m^{-2} d^{-1} and 15.7 MJ m^{-2} d^{-1}, respectively.

The soil organic C content had a spatial variation associated to the rainfall/PET ratio ($R^2 = 0.30$, $P < 0.05$) (Fig. 1C). Soils of humid environments had larger C contents than under semiarid conditions. Conversely, SAWSC presented a different spatial variation, being greater in the North of the Pampas, in areas with higher temperature, but the association between both variables was weak ($R^2 = 0.08$, $P < 0.05$) (Fig. 1D).

Fertilizer application started in the 1990's in the Pampas with average rates rounding 5 to 10 kg N ha^{-1} reaching 80 kg N ha^{-1} in some geographical units during the last years of this analysis. As N and P rates were highly correlated ($R^2 = 0.83$, $P < 0.05$) we used only N data to represent regional fertilizer use. Nitrogen rates showed a significant positive trend with time (Table 1). Conventional tillage dominated in the past but in recent years in more than 50% of total cropped area crops were produced under conservation tillage practices. Conservation tillage use and genetic improvement were also correlated with time (Table 1). Consequently, as all three management factors considered in this analysis were correlated among them and with time we could not isolate effects from each other and used only time as a surrogate of all these variables, considering it as a variable descriptive of management improvement.

A frontier function could be fitted for defining attainable yield which included significant β parameters for the variables time and some climate and a soil variable (Table 2). Coefficients were positive in all cases indicating that these independent variables affect attainable yield positively. Sensitivity analysis showed that time had the largest impact on attainable yield, followed by the climate variables and SAWSC.

Attainable yield estimated with the frontier model ranged from 1820 to 3900 kg ha^{-1}. The relation between attainable and observed yield had an intercept different from zero and a slope equal to one ($R^2 = 0.453$; $P < 0.05$). As expected, attainable yields were generally larger than observed yields, with the exception of approximately 5% of the cases (Fig. 2A). These cases belong mainly to a humid subregion of the Pampas that corresponded to environments with very high productivity. When comparing attainable against mean ANN yields a similar trend was observed ($R^2 = 0.679$, $P < 0.05$) but in only two cases mean ANN yield was greater than attainable yield (Fig. 2B). Dispersion, calculated as the difference between maximum and minimum values, decreased in the following order: attainable < mean ANN modeled < observed yields (Fig. 3A, 3B, 3C). Slopes of these relations were similar implying that yield increases per year showed the same pattern in observed, mean ANN modeled, and attainable yield. Because of these results, further comparisons were performed only between attainable and ANN yield. Attainable yields were greater and yield gaps smaller in high productivity soils (Fig. 4A, 4B). Variability of attainable yields did not show a marked pattern while yield gap variability decreased where the soil productivity was larger. Spatial analysis showed that the largest estimated attainable yields corresponded to the southeastern portion of the Pampas with average values of

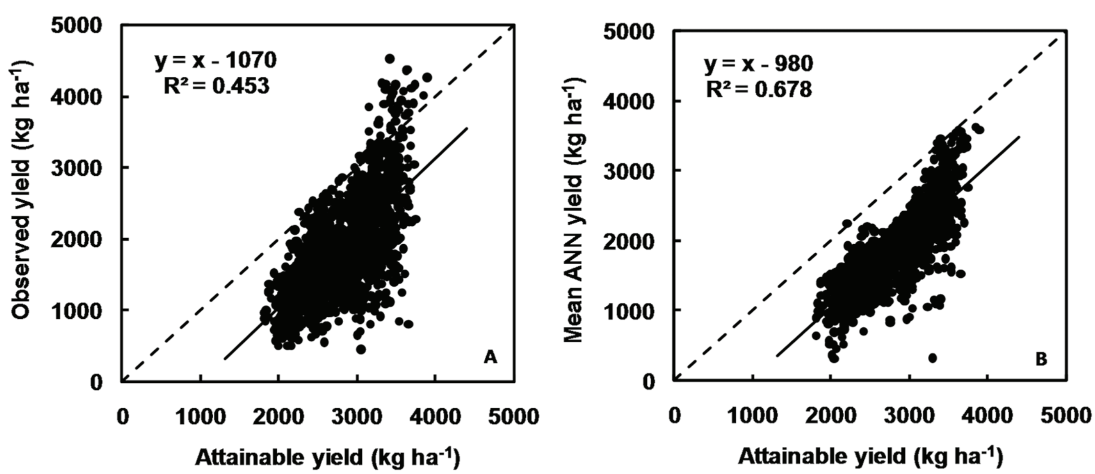

Fig. 2. Attainable yields vs. (A) observed yields and vs. (B) mean artifical neural network (ANN) yields.

3100 kg ha^{-1} while the lowest values corresponded to the semiarid portion with an average of 2600 kg ha^{-1} (Fig. 5A). Yield gap presented an opposite spatial trend (Fig. 5B).

The yield gap could be modeled by an ANN model (R^2 = 0.745; RMSE = 144 kg ha^{-1}, $P < 0.05$) (Fig. 6A, 6B). The fitted network had six neurons in the hidden layer and a good generalization capacity as no significant differences were detected in R^2 between training + validation and test sets. Slopes of observed vs. estimated values were not different from 1 and ordinates were equal to 0 ($P < 0.05$). Selected input variables were: time, soil organic C (0–50 cm), SAWSC (0–100 cm), average rainfall/PET during fallow and crop growing periods, and photothermal quotient. Contributions of other variables on the yield gap prediction were not significant over this model. All the inputs showed curvilinear effects and strong interactions. This regional model allowed estimating input impacts on the yield gap. When rainfall/PET ratio was one yield gap had the lowest value and with smaller and larger climate ratios it increased. Soil variables interacted and a minimum yield gap was detected at medium to high values of organic C and SAWSC (Fig. 7). Under an average climate scenario, represented by average rainfall/PET and photothermal quotient values per geographic unit corresponding to the 1967 to 2006 period, this optimal yield gap rounded 300 kg ha^{-1} in units where soil organic C content was 60 to 70 t ha^{-1} in combination with a SAWSC ranging 120 to 150 mm. The yield gap increased to more than 1000 kg ha^{-1} in soils with a high organic C content and low SAWSC or vice versa.

DISCUSSION

We disentangled some of the factors that determine wheat yield gap in the Pampas by combining frontier analysis and an ANN approach. The analysis relies on the integration of biophysical factors at regional scale for modeling attainable yield, average yield, and the yield gap between these production levels. As available data on crop production were generated at county scale from field surveys, and as in these datasets uncertainties were previously detected (Paruelo et al., 2004; Sadras et al., 2014), the information was aggregated up to groups of counties with similar area, to eliminate outliers and decrease variability (Bakker et al., 2005; Grassini et al., 2015). Conversely, climate and soil data were estimated using information from 960 climate interpolation maps and 1000 modeled soil profiles of seven soil variables accounting for their influence area. The uncertainty of these estimations was small, rounding for example approximately 5% for the variable soil organic C (Berhongaray et al., 2013). Because of the wide range of climate and soil properties included in the analysis, environmental factor effects could be weighted under very contrasting combinations.

A limitation arose when addressing management effects on the yield gap. Detailed management information is lacking in the Pampas as was previously detected in other regional assessments (Neumann et al., 2010). Consequently, we were forced to use a combination of the available present information at regional scale, national time trends, and expert knowledge to estimate present and past management factors. Fertilizer rate, conservation tillage

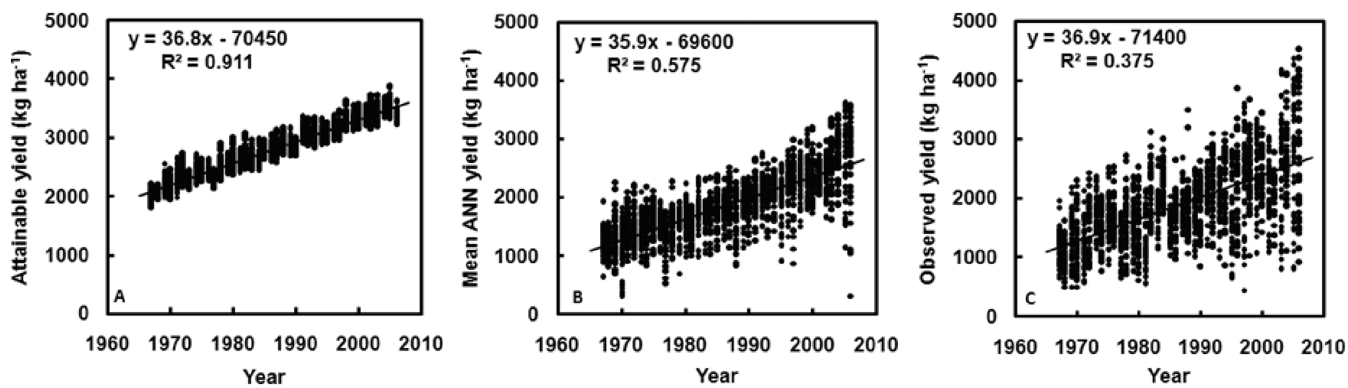

Fig. 3. Time trends of (A) attainable, (B) mean artifical neural network (ANN), and (C) observed yields for the study period of 40 yr (1967–2006).

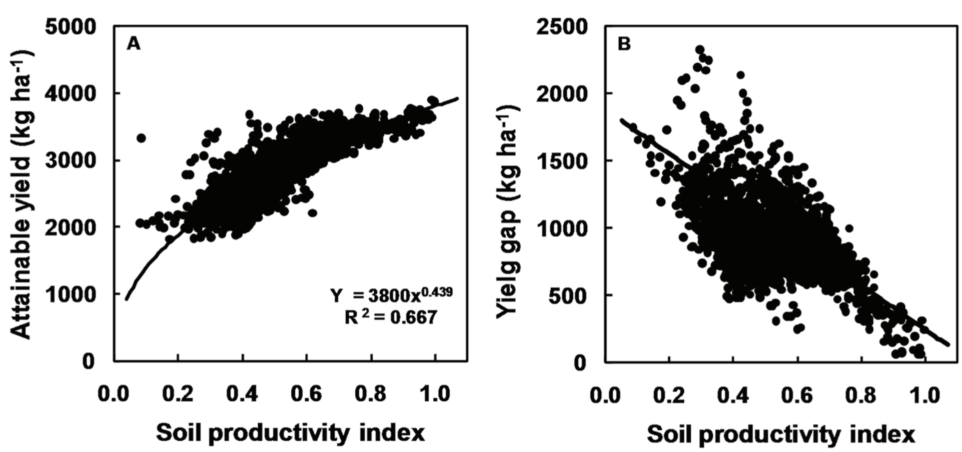

Fig. 4. Association between a soil productivity index and (A) average attainable yield, and (B) the average yield gap. The soil productivity index was developed locally and defined wheat yield related to environmental factors.

use, and genetic improvement were the management factors that we attempted to split into independent variables but due to the high correlation between them and with time we could not determine their individual effects. The only alternative was using time as a variable that partially surrogated management improvements as was done previously by others (Lobell and Burke, 2010).

The frontier model allowed a simple estimation of attainable yield based on available national statistics. The estimated values resulted from the highest attained yields for each combination of climate and soil conditions during a determined year. The constant term was not significant and therefore no assumptions had to be made (Singh et al., 2003). The method assumed that yield reducing factors like frosts, diseases, etc. were absent. Considering that under production conditions these reducing factors are usually present at different intensities, it seemed reasonable to assume that their impact was minimal when the attainable yield production level was reached (Egli and Hatfield, 2014). Attainable yields increased in deep soils with large SAWSC when climate conditions were favorable for the crop. This resulted in greater attainable yields in soils with larger productivity. Attainable yields were also time dependent showing a strong influence of technological improvements on wheat yield in the Pampas.

We used a new yield gap estimation approach by subtracting a mean ANN modeled yield from the attainable yield production level. This modeled mean ANN yield represented the expected average yield for combinations of environmental variables. Its use reduced the variability in the yield gap estimation and eliminated some inconsistencies detected when comparing observed and

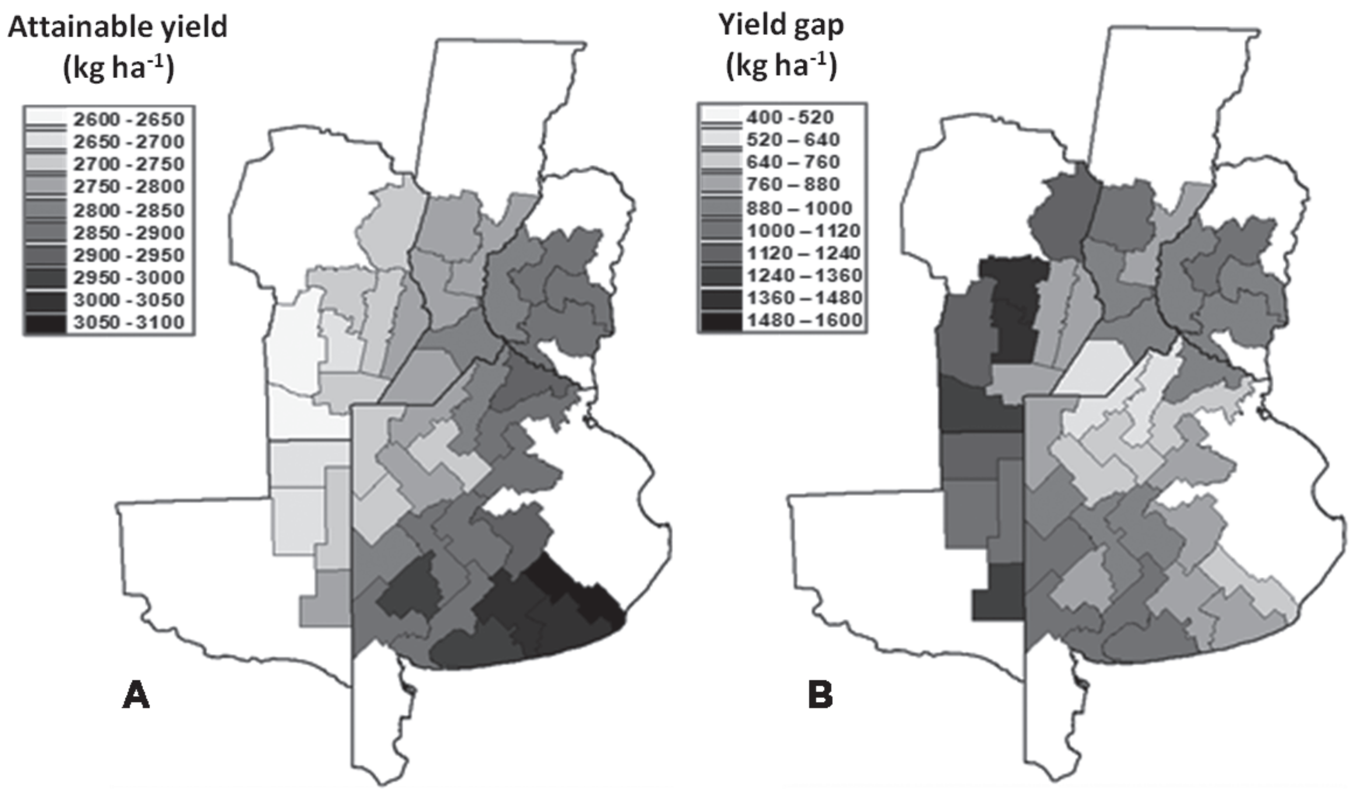

Fig. 5. (A) Average estimated attainable wheat yield and (B) average yield gap. The presented information corresponds to the 1967 to 2006 period.

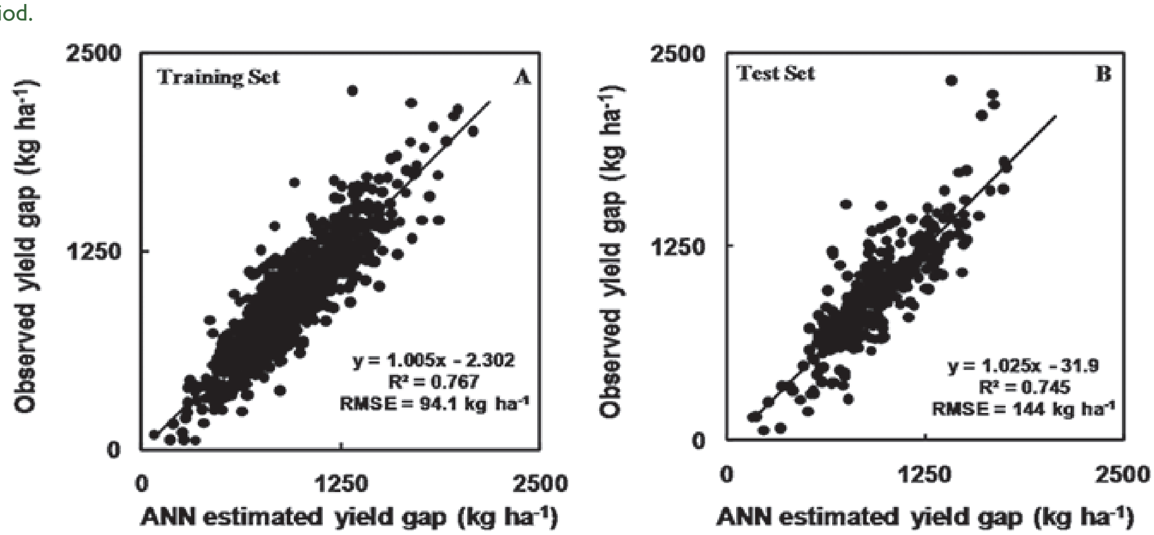

Fig. 6. Relation between observed and estimated yield gap for the training + validation (A) the test data set (B).

attainable yield, given that larger observed yields were detected in some areas when compared to attainable yields. These inconsistencies have been observed in other studies with process based model estimations compared to observed yields and were attributed to system misspecifications (e.g., soil, climate), incorrect reported yields, or model constraints (Carberry et al., 2013; Van Ittersum et al., 2013). As these inconsistencies disappeared in our data set when using mean ANN yield for yield gap calculation it seemed more probable that in our case they were produced by a bias in the available observed yield data set. Yield gap size and its trends in different subregions of the Pampas were generally similar, both using the observed or the mean ANN yield (data not shown), but the yield gap based on the modeled mean ANN yield was less affected by unknown extreme reducing factors and possible errors that impacted the reported observed yield values.

Attainable yield and yield gap were negatively correlated ($R^2 = 0.18$, $P < 0.01$). In the geographic units of the Pampas were optimal climate conditions in combination with large soil productivity resulted in large estimated attainable yield levels during most of the years resulting in lower yield gaps. Conversely, where attainable yield levels were reached only during some years because of variable and low-rainfall conditions combined with low soil productivity yield gaps were larger. Similar results, that is yield gaps were larger where actual yields were lower, have been reported in other wheat production regions of Europe (Abeledo et al., 2008) and Asia (Lu and Fan, 2013) but, conversely, another study found no relation between wheat yield gap and yield levels in Australia (Peake et al., 2014).

For the 2002 to 2006 time interval mean yield gap of wheat in the Pampas was 865 kg ha^{-1} (25% of the attainable yield production level) and varied between 200 and 1800 kg ha^{-1} depending on the geographic unit. During the mentioned time interval it ranged from 26% in humid environments to 42% in semiarid ones. The semiarid western portion of the Pampas and the very humid northeastern border had the lowest attainable yields and the largest

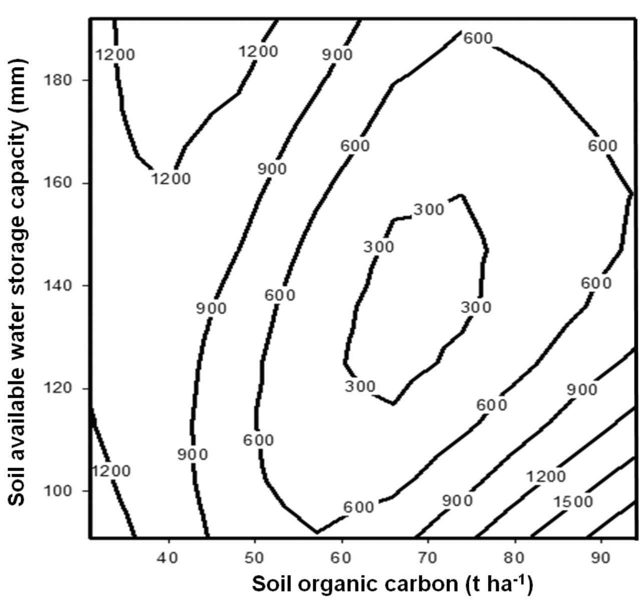

Fig. 7. Interaction between the two soil variables, soil available water storage capacity and organic C, which partially determine wheat yield gap in the Pampas. Isolines indicate identical yield gaps for the combination of both variables. Numbers on the lines show yield gap values (kg ha^{-1}).

yield gaps. This spatial pattern may be partially attributed to the high rainfall variability of the semiarid Pampas (Hall et al., 1992) that can explain why average yields largely differ from attainable yields. In the northeastern geographic units excessive rains can induce wheat diseases (Annone, 2001) and lodging problems (De San Celedonio et al., 2014), meanwhile under semiarid environments favorable years allow yields to get close to attainable yield but rainfall variability leads to high yield gaps many years.

By applying a similar approach than ours, that is by calculating differences between attainable yield estimated by frontier analysis fitted on statistical data and observed yields, Neumann et al. (2010) reported a yield gap that ranged from 500 to 2000 kg ha^{-1} in the major part of the Pampas. This reported yield gap showed a relatively similar spatial pattern than that of our study. The results of the estimated yield gap of the Pampas and that of Neumann et al. (2010) tended to be greater in low productivity areas decreasing in favorable environments, and contrast with results from Mueller et al. (2012) that estimate greater yield gap in the semiarid-low productivity semiarid Pampa. The discrepancy may be the consequence of the different level of detail in the baseline information used for the analysis.

To our knowledge, this was the first time that an ANN model was used as a tool for identifying some of the regional yield gap environmental controlling factors. Beyond the very strong effect of the rainfall/PET ratio, which accounted for 46% of the yield gap variability, as revealed by a sensitivity analysis, soil properties also regulated it, accounting for 52% of its variability. Minimum yield gaps were modeled for combinations of soil organic C contents and SAWSC for which maximum wheat productivity was detected in the Pampas (De Paepe and Álvarez, 2013). Yield gap increases not only in low productivity areas but especially in low productivity soils located in the semiarid and in the upper humid East areas. There are only a few reports of measurable effects of soil properties on yield gap. For example, in the semiarid Pampas when the water retained in soils increased the yield gap decreased in sunflower (*Helianthus annuus* L.) (Grassini et al., 2009) and yield gaps of maize grown with inorganic fertilizer in Sub-Saharan Africa decreased when clay contents were below 20% or above 40% (Sileshi et al., 2010). The results of our study suggested that reducing yield gap efforts must be focused in the Pampas with low productivity soils. Soil available water storage capacity is a property that suffers only minimum changes by management but soil C can be deeply impacted by agricultural practices, for instance through crop rotations. Local studies have determined that increasing the proportion of perennial forage crops in rotations (Studdert et al., 1997) or replacing soybean by other crops like corn, that leave more residues in the soil (Álvarez et al., 2011), can increase the soil organic C content. In our study the wheat yield gap was significantly affected by soil organic C but it is not probable that wheat yield has a significant impact on soil C because this crop accounts for no more than 10 to 15% of rotation time (Caride et al., 2012), thus the main contributors to soil C are the other rotation components as stated above.

The present study represents the first attempt to disentangle the effects of soil and climate variables on the wheat yield gap in the Pampas using a combination of frontier analyses and ANN approaches. This methodology can be applied to other crops and regions. The modeling methods, although statistically demanding, can be performed when only statistical yield data are available and are simpler than process based models. They do not require a parameterization and validation process and the information needed can be obtained from common soil surveys, climatic records, and national statistics.

CONCLUSIONS

A novel approach was applied in this analysis for yield gap calculation combining two modeled yield values, a frontier yield, and an average ANN yield, to avoid large variability observed in national statistic yield information. The results showed that yield gap and its geographical pattern was not only associated with climate factors but also largely with soil factors and this could be very well modeled by an ANN method. More than 50% of the gap variability was explained by two soil variables: organic C and SAWSC. Yield gap modeling allowed the identification of a positive interaction between these soil variables that defined a minimum yield gap at average values. The combination of very low or high values of C and SAWSC leads to gap increases. Lowest yield gaps matched soils with a large productivity while low productivity soils were associated to the largest yield gaps.

ACKNOWLEDMENTS

We thank Prof. R. Ceulemans and his research group in Plant Ecology at the University of Antwerp for previous revisions of this manuscript. This research was funded by the University of Buenos Aires (G004, G033, and 20020100617), CONICET (PIP 02050 and PIP 02608) and FONCyT (PID-BID 37164-49).

REFERENCES

Abeledo, L.G., R. Savin, and G.A. Slafer. 2008. Wheat productivity in the Mediterranean Ebro Valley: Analyzing the gap between attainable and potential yield with a simulation model. Eur. J. Agron. 28:541–550. doi:10.1016/j.eja.2007.12.001

Aigner, D., C.A.K. Lovell, and P. Schmidt. 1977. Formulation and estimation of stochastic frontier production function models. J. Econom. 6:21–37. doi:10.1016/0304-4076(77)90052-5

Alonso, M., R. Rodriguez, S. Gomez, and R. Giagnoni. 2002. Un método para estimar la radiación global con la amplitud térmica y la precipitación diarias. Revista de la Facultad de Agronomia de la Universidad de Buenos Aires 22:51–56.

Alvarez, R. 2009. Predicting average regional yield and production of wheat in the Argentine Pampas by an artificial neural network approach. Eur. J. Agron. 30:70–77. doi:10.1016/j.eja.2008.07.005

Alvarez, R., and S. Grigera. 2005. Analysis of soil fertility and management effects on yields of wheat and corn in the Rolling Pampa of Argentina. J. Agron. Crop Sci. 191:321–329. doi:10.1111/j.1439-037X.2005.00143.x

Álvarez, R., and R.S. Lavado. 1998. Climate, organic matter and clay content relationships in the Pampa and Chaco soils, Argentina. Geoderma 83:127–141. doi:10.1016/S0016-7061(97)00141-9

Alvarez, R., H.S. Steinbach, C. Alvarez, and J.L. De Paepe. 2015. Fertilizer use in pampean agroecosystems: Impact on productivity and nutrient balance. In: S. Sinha et al., editors, Chemical Engineering Series. Studdium Press, New Delhi, India. p. 352–368.

Álvarez, R., H.S. Steinbach, and A. Bono. 2011. An artificial neural network approach for predicting soil carbon budget in agroecosystems. Soil Sci. Soc. Am. J. 75:965–975. doi:10.2136/sssaj2009.0427

Anderson, W.K. 2010. Closing the gap between actual and potential yield of rainfed wheat. The impacts of environment, management and cultivar. Field Crops Res. 116:14–22. doi:10.1016/j.fcr.2009.11.016

Annone, J.G. 2001. Criterios empleados para la toma de decisiones en el uso de fungicidas en trigo. Revista Tecnológica Agropecuaria. 6:16–20.

Bakker, M.M., G. Govers, F. Ewert, M. Rounsevell, and R. Jones. 2005. Variability in regional wheat yields as a function of climate, soil and economic variables: Assessing the risk of confounding. Agric. Ecosyst. Environ. 110:195–209. doi:10.1016/j.agee.2005.04.016

Battese, G.E., and T.J. Coelli. 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. Empir. Econ. 20:325–332. doi:10.1007/BF01205442

Becker, M., and D.E. Johnson. 1999. Rice yield and productivity gaps in irrigated systems of the forest zone of Côte d'Ivoire. Field Crops Res. 60:201–208. doi:10.1016/S0378-4290(98)00123-3

Berhongaray, G., R. Alvarez, J.L. De Paepe, C. Caride, and R. Cantet. 2013. Land use effects on soil carbon in the Argentine Pampas. Geoderma 192:97–110. doi:10.1016/j.geoderma.2012.07.016

Boling, A.A., B.A.M. Bouman, T.P. Tuong, Y. Konboon, and D. Harnpichitvitaya. 2011. Yield gap analysis and the effect of nitrogen and water on photoperiod-sensitive Jasmine rice in north-east Thailand. NJAS– Wagen. J. Life Sci. 58:11–19.

Calderini, D.F., M.F. Dreccer, and G.A. Slafer. 1995. Genetic improvement in wheat yield and associated traits. A re-examination of previous results and the latest trends. Plant Breed. 114:108–112. doi:10.1111/j.1439-0523.1995.tb00772.x

Carberry, P.S., W. Liang, S. Twomlow, D.P. Holzworth, J.P. Dimes, T. McClelland et al. 2013. Scope for improved eco-efficiency varies among diverse cropping systems. Proc. Natl. Acad. Sci. USA 110:8381–8386. doi:10.1073/pnas.1208050110

Caride, C., G. Pineiro, and J.M. Paruelo. 2012. How does agricultural management modify ecosystem services in the argentine Pampas? The effects on soil C dynamics. Agric. Ecosyst. Environ. 154:23–33. doi:10.1016/j.agee.2011.05.031

Cassman, K.G. 1999. Ecological intensification of cereal production systems: Yield potencial, soil quality, and precision agriculture. Proc. Natl. Acad. Sci. USA 96:5952–5959. doi:10.1073/pnas.96.11.5952

Cassman, K.G., A. Dobermann, D.T. Walters, and H. Yang. 2003. Meeting cereal demand while protecting natural resources and improving environmental quality. Annu. Rev. Environ. Resour. 28:315–358. doi:10.1146/annurev.energy.28.040202.122858

Coelli, T. 1994. A computer program for stochastic frontier production and cost function estimation. A Guide to Frontier 4.1. CEPA Working Paper 96/07. Centre for Efficiency and Productivity Analysis, University of New England, Armidale, Australia.

Coelli, T., P. Rao, C.J. O'Donnell, and G.E. Battese. 2005. An introduction to efficiency and productivity analysis. Kluwer Academic Publishers, Norwell, Massachusetts.

Dai, X., Z. Huo, and H. Wang. 2011. Simulation for response of crop yield to soil moisture and salinity with artificial neural network. Field Crops Res. 121:441–449. doi:10.1016/j.fcr.2011.01.016

De Paepe, J.L., and R. Álvarez. 2013. Development of a regional soil productivity index using an artificial neural network approach. Agron. J. 105:1803–1813. doi:10.2134/agronj2013.0070

De San Celedonio, R.C., J.R. Micheloud, L.G. Abeledo, D. Miralles, and G.A. Slafer. 2014. Riesgo de anegamiento en trigo (*Triticum aestivum* L.) para distintas localidades de la región triguera argentina. Cienc. del Suelo 32:233–246.

Donatelli, M., G. Bellocchi, and F. Fontana. 2003. RadEst3.00: Software to estimate daily radiation data from commonly available meteorological variables. Eur. J. Agron. 18:363–367. doi:10.1016/S1161-0301(02)00130-2

Doorenbos, J., and W. Pruitt. 1977. Crop water requirements. Irrigation and drainage paper. FAO, Rome, Italy.

Egli, D., and J. Hatfield. 2014. Yield gaps and yield relationships in central US soybean production systems. Agron. J. 106(2):560–566. doi:10.2134/agronj2013.0364

Evans, L.T., and R.A. Fischer. 1999. Yield Potential: Its Definition, Measurement, and Significance. Crop Sci. 39(6):1544–1551. doi:10.2135/cropsci1999.3961544x

FAO. 2004. Fertilizer use by crop in Argentina. Food and Agriculture Organization on the United Nations, Rome, Italy.

FAOstat. 2015. Food and Agriculture Organization of the United States-Statistics Division. www.faostat.com (accessed 14 Apr. 2016).

Fischer, R.A. 1985. Number of kernels in wheat crops and the influence of solar radiation and temperature. J. Agric. Sci. 105:447–461. doi:10.1017/S0021859600056495

Foley J.A., N. Ramankutty, K.A. Brauman, E.S. Cassidy, J.S. Gerber, M. Johnston et al. 2011. Solutions for a cultivated planet. Nature (London) 478:337–342. doi:10.1038/nature10452

Grassini, P., A.J. Hall, and J.L. Mercau. 2009. Benchmarking sunflower water productivity in semiarid environments. Field Crops Res. 110:251–262. doi:10.1016/j.fcr.2008.09.006

Grassini, P., J. Thorburn, C. Burr, and K.G. Cassman. 2011. High-yield irrigated maize in the Western U.S. Corn Belt: I. On-farm yield, yield potential, and impact of agronomic practices. Field Crops Res. 120:142–150. doi:10.1016/j.fcr.2010.09.012

Grassini, P., L.G.J. van Bussel, J. Van Wart, J. Wolf, L. Claessens, H. Yang et al. 2015. How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. Field Crops Res. 177:49–63. doi:10.1016/j.fcr.2015.03.004

Hall, A.J., C. Rebella, C. Guersa, and J. Culot. 1992. Field-crop system of the Pampas. In: C.J. Pearson, editor, Field crop ecosystems. Elsevier, Amsterdam, the Netherlands.

Hochman, Z., D. Gobbett, D. Holzworth, T. McClelland, H. van Rees, O. Marinoni et al. 2012. Quantifying yield gaps in rainfed cropping systems: A case study of wheat in Australia. Field Crops Res. 136:85–96. doi:10.1016/j.fcr.2012.07.008

Hochman, Z., D. Gobbett, D. Holzworth, T. McClelland, H. van Rees, O. Marinoni et al. 2013. Reprint of Quantifying yield gaps in rainfed cropping systems: A case study of wheat in Australia. Field Crops Res. 143:65–75. doi:10.1016/j.fcr.2013.02.001

Hunt, L.A., L. Kuchar, and C.J. Swanton. 1998. Estimation of solar radiation for use in crop modelling. Agric. For. Meteorol. 91:293–300. doi:10.1016/S0168-1923(98)00055-0

Imhoff, M.L., L. Bounoua, T. Ricketts, C. Loucks, R. Harriss, and W.T. Lawrence. 2004. Global patterns in human consumption of net primary production. Nature 429:870–873. doi:10.1038/nature02619

INTA. 2015. Instituto Nacional de Tecnologia Agropecuaria. Ministerio de Agroindustria de la Nación. www.inta.gob.ar/instdeclimayagua (accessed 6 July 2015).

Kaul, M., R.L. Hill, and C. Walthall. 2005. Artificial neural networks for corn and soybean yield prediction. Agric. Syst. 85:1–18. doi:10.1016/j.agsy.2004.07.009

Kleinbaum, D.G., and L.L. Kupper. 1979. Applied regression analysis and other multivariable methods Duxbury Press, North Scituate, MA.

Kobayashi, K., and M.U. Salam. 2000. Comparing simulated and measured values using mean squared deviation and its components. Agron. J. 92:345–352. doi:10.2134/agronj2000.922345x

Lee, J.H.W., Y. Huang, M. Dickman, and A.W. Jayawardena. 2003. Neural network modelling of coastal algal blooms. Ecol. Modell. 159:179–201. doi:10.1016/S0304-3800(02)00281-8

Licker, R., M. Johnston, J.A. Foley, C. Barford, C.J. Kucharik, C. Monfreda, and N. Ramankutty. 2010. Mind the gap: How do climate and agricultural management explain the ‘yield gap’ of croplands around the world? Glob. Ecol. Biogeogr. 19:769–782. doi:10.1111/j.1466-8238.2010.00563.x

Linacre, E. 1977. A simple formula for estimating evapotranspiration rates in various climates, using temperature data alones. Agric. Meteorol. 18:409–424. doi:10.1016/0002-1571(77)90007-3

Liu J., J. R. Williams, A. J. B. Zehnder, and H. Yang. 2007. GEPIC-Modelling wheat yield and crop water productivity with high resolution on a global scale. Agr. Syst. 94:478–493.

Lobell, D.B. 2013. The use of satellite data for crop yield gap analysis. Field Crops Res. 143:56–64. doi:10.1016/j.fcr.2012.08.008

Lobell, D.B., and M.B. Burke. 2010. On the use of statistical models to predict crop yield responses to climate change. Agric. For. Meteorol. 150:1443–1452. doi:10.1016/j.agrformet.2010.07.008

Lobell, D.B., K.G. Cassman, and C.B. Field. 2009. Crop Yield Gaps: Their Importance, Magnitudes, and Causes. Annu. Rev. Environ. Resour. 34:179–204. doi:10.1146/annurev.environ.041008.093740

Lu, C., and L. Fan. 2013. Winter wheat yield potentials and yield gaps in the North China Plain. Field Crops Res. 143:98–105. doi:10.1016/j.fcr.2012.09.015

Magrin G. O. 2004. Rendimientos potenciales de trigo en las distintas zonas trigueras. Sistema de Información Clima y Agua- INTA: 1-7.

Magrin, G., A.J. Hall, G. Baldi, and M. Grondona. 1993. Spatial and interannual variations in the phototermal quotient: Implications for the potential kernel number of wheat crops in Argentina. Agric. For. Meteorol. 67:29–41. doi:10.1016/0168-1923(93)90048-M

Meeusen, W., and J. Van den Broeck. 1977. Efficiency estimation from Cobb-Douglas production functions with composed error. Econ. Rev. 18:435–444. doi:10.2307/2525757

Miao, Y., D. Mulla, and P. Robert. 2006. Identifying important factors influencing corn yield and grain quality variability using artificial neural networks. Precis. Agric. 7:117–135. doi:10.1007/s11119-006-9004-y

MinAgri. 2015. Sistema integrado de información agropecuaria. http://www.siia.gov.ar/_apps/siia/estimaciones/estima2.php (accessed 14 Apr. 2016).

Mueller, N.D., J.S. Gerber, M. Johnston, D.K. Ray, N. Ramankutty, and J.A. Foley. 2012. Closing yield gaps through nutrient and water management. Nature 490:254–257. doi:10.1038/nature11420

Neumann, K., P.H. Verburg, E. Stehfest, and C. Müller. 2010. The yield gap of global grain production: A spatial analysis. Agric. Syst. 103:316–326. doi:10.1016/j.agsy.2010.02.004

Özesmi, S.L., C.O. Tan, and U. Özesmi. 2006. Methodological issues in building, training, and testing artificial neural networks in ecological applications. Ecol. Modell. 195:83–93. doi:10.1016/j.ecolmodel.2005.11.012

Paruelo, J.M., J.P. Guerschman, G. Baldi, and C. Di Bella. 2004. La estimacion de la superficie agricola. Antecedentes y una propuesta metodologica. Interciencia 29:421–427.

Patrignani, A., R.P. Lollato, T.E. Ochsner, C.B. Godsey, and J. Edwards. 2014. Yield gap and production gap of rainfed winter wheat in the Southern Great Plains. Agron. J. 106:1329–1339. doi:10.2134/agronj14.0011

Peake, A., N. Huth, P. Carberry, S. Raine, and R. Smith. 2014. Quantifying potential yield and lodging-related yield gaps for irrigated spring wheat in sub-tropical Australia. Field Crops Res. 158:1–14. doi:10.1016/j.fcr.2013.12.001

Pesaran, M.H., and P. Schmidt. 1999. Handbook of applied econometrics. Blackwell Publ., Cambridge.

Quantum GIS Development Team. 2015. Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org (12 Apr. 2016).

Rawls, W. 1983. Estimating soil bulk density from particle size analysis and organic matter content. Soil Sci. 135:123–125. doi:10.1097/00010694-198302000-00007

Rawls, W., D. Brakensiek, and K. Saxton. 1982. Estimation of soil water properties. Trans. ASAE 25:1316–1332. doi:10.13031/2013.33720

RIAN. 2015. Red de Informacion Agropecuaria Nacional. Inst. Nacional de Tecnologia Agropecuaria. www.rian.gov.ar (accessed 10 Jan. 2015).

Rogers, L.L., and F.U. Dowla. 1994. Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling. Water Resour. Res. 30:457–481. doi:10.1029/93WR01494

Sadras, V., and J. Angus. 2006. Benchmarking water-use efficiency of rainfed wheat in dry environments. Aust. J. Agric. Res. 57:847–856. doi:10.1071/AR05359

Sadras, V., P. Grassini, R. Costa, L. Cohan, and A.J. Hall. 2014. How reliable are crop production data? Case studies in USA and Argentina. Food Security 6:447–459. doi:10.1007/s12571-014-0361-5

Schaap, M.G., F.J. Leij, and M.T. van Genuchten. 1998. Neural Network Analysis for Hierarchical Prediction of Soil Hydraulic Properties. Soil Sci. Soc. Am. J. 62:847–855. doi:10.2136/sssaj1998.03615995006200040001x

Sileshi, G., F.K. Akinnifesi, L.K. Debusho, T. Beedy, O.C. Ajayi, and S. Mong'omba. 2010. Variation in maize yield gaps with plant nutrient inputs, soil type and climate across sub-Saharan Africa. Field Crops Res. 116:1–13. doi:10.1016/j.fcr.2009.11.014

Singh, H., D. Mishra, N.M. Nahar, and M. Ranjan. 2003. Energy use pattern in production agriculture of a typical village in arid zone India: Part II. Energy Convers. Manage. 44(7):1053–1067. doi:10.1016/S0196-8904(02)00115-2

SMN. 2015. Servicio Meteorológical Nacional. Secretaria de Ciencia, Tecnologia y Producción. http://datosclimaticos.mincyt.gob.ar/contacto.php (accessed 15 July 2015).

Studdert, G.A., H.E. Echeverría, and E.M. Casanovas. 1997. Crop-pasture rotation for sustaining the quality and productivity of a typic Argiudoll. Soil Sci. Soc. Am. J. 61:1466–1472. doi:10.2136/sssaj1997.03615995006100050026x

Tittonell, P., B. Vanlauwe, M. Corbeels, and K.E. Giller. 2008. Yield gaps, nutrient use efficiencies and response to fertilisers by maize across heterogeneous smallholder farms of western Kenya. Plant Soil 313:19–37. doi:10.1007/s11104-008-9676-3

Totis, L., and O. Perez. 1994. Relaciones entre el consumo de agua máximo de la secuencia de cultivo trigo/soja y la evapotranspiración potencial para el cálculo de la dosis de riego. In: INTA-Pergamino, editor, Carpeta de Producción Vegetal. Inst. Nacional de Tecnologia Agropecuaria (INTA), Pergamino City, Argentina. p. 1–4.

Van Ittersum, M.K., K.G. Cassman, P. Grassini, J. Wolf, P. Tittonell, and Z. Hochman. 2013. Yield gap analysis with local to global relevance- A review. Field Crops Res. 143:4–17. doi:10.1016/j.fcr.2012.09.009

Van Ittersum, M.K., and R. Rabbinge. 1997. Concepts in production ecology for analysis and quantification of agricultural input-output combinations. Field Crops Res. 52:197–208. doi:10.1016/S0378-4290(97)00037-3

Zimmerman, D., C. Pavlik, A. Ruggles, and M. Armstrong. 1999. An experimental comparison of ordinary and universal Kriging and inverse distance weighting. Math. Geol. 31:375–390. doi:10.1023/A:1007586507433