Early Maize Yield Forecasting From Remotely Sensed Temperature/Vegetation Index Measurements

Mauro E. Holzman and Raúl E. Rivas

***Abstract*—High and low soil moisture availability is one of the main limiting factors-affecting crops productivity. Thus, determination of the relationship between them is crucial for food security and support importing– exporting strategies. The aim of this work was to analyze the aptitude of temperature vegetation dryness index (TVDI) to forecast maize yield. MODIS/AQUA enhanced vegetation index and land surface temperature (LST) at 1 km were used to calculate TVDI and maize yield over a large agricultural area of Argentine Pampas. The comparison between TVDI and official yield statistics was carried out to derive regression models in two agro-climatic zones, obtaining linear and quadratic adjustments. The models account for between 73% and 83% of yield variability, with the best prediction in the humid zone. The RMSE values ranged from 14% to 19% of average yield. The bias showed a slightly higher difference between predicted and observed yield data in semi-arid zone. The models showed aptitude to estimate yield with reasonable accuracy 8–12 weeks before harvest. In addition, the TVDI-maize yield relationship and the impact of submonthly water stress were evaluated at field scale using yield measurements to ensure the analysis on maize. The highest R^2 (0.61) was obtained using monthly values suggesting that the entire critical stage should be taken into account for yield forecasting. Although these results would not be directly extrapolated to other agricultural regions in the world, the proposed model is promising for forecasting spatial yield in other regions with poor data coverage several weeks before harvest.**

***Index Terms*—Optical–thermal, soil moisture, stress index, temperature vegetation dryness index (TVDI).**

I. Introduction

EXTREMELY high and low water availability affects the dynamics of agro-ecosystems. In addition, the ability of society to adapt to climate change may be put to the test through extreme events, which might increase even under modest changes in climate [1]. To solve economic and social problems related to the impact of climate variability on crops and food security, there is an urgency to improve monitoring of agricultural land [2]. Soil water availability influences the transpiration rate of vegetation and it is a frequent limiting factor for rainfed crops. Thus, there is an increasing interest to estimate soil moisture and the impact on crop productivity at regional scale [3] for crop yield prediction, which is crucial for decision makers such as agricultural, insurance, and governmental agencies.

Soil moisture status can be appropriately measured at field scale by traditional methods (e.g., lysimeters and soil moisture probes) [4]. However, given that soil moisture is highly dynamic in space and time due to diverse factors like types of soil and water table depth, the assessment of this variable at regional scale is complex. Satellite-derived information has great potential for the estimation of water availability in the soil-vegetation system and to evaluate the relationship with crop yield.

During the last three decades different methods based on low resolution (about 1 km), remotely sensed (RS) data have been developed to estimate crop yield at regional scale. In general, these methods consider RS indicators of vegetation performance anomalies that are related to soil moisture [5], [6]. Early studies mainly included crop monitoring approaches based on vegetation indices (VI) from the NOAA AVHRR [7], [8] and Landsat series [9], [10]. The last decade has been characterized by studies using high temporal resolution sensors like MODIS TERRA/AQUA that allow a proper evaluation of the temporal fluctuations of water availability and the influence on crop condition [11], [12]. A group of methods, called qualitative or semi-quantitative, is based on the comparison between the actual crop condition and average situation [13]. These approaches were traditionally applied in arid and semi-arid regions using normalized difference VI (NDVI). Fewer works have used land surface temperature (LST) as indicator of water availability for early warnings of crop damage due to water scarcity [14], [15]. Typically, these methods can be useful in large areas with poor or no availability of agricultural data.

Another group of crop yield estimation methods is represented by regression models that quantify the expected yield. These models need to be calibrated with consistent series of crop yield data as reference information and the main advantage is their simplicity. Many studies have used cumulated NDVI of different growing seasons and historical yields [11], [16]. Most of them have concluded that the best results are obtained during the heading and grain filling stage of crops. Some studies have improved these methods incorporating agro-climatic variables like rainfall and temperature [17], [18]. However, it should be noted that several agro-climatic variables are not independent of VI [13].

A sophisticated group of methods includes modeling of crop physiology, which describe the physiological mechanisms of crop growth and the interaction with environmental variables (e.g., soil moisture and temperature) using equations [13]. A simplified method considers the strong relationship between

Manuscript received March 18, 2015; revised November 12, 2015; accepted November 18, 2015. Date of publication January 11, 2016; date of current version January 28, 2016.

Mauro E. Holzman is with the Consejo Nacional de Investigaciones Científicas y Técnicas, Instituto de Hidrología de Llanuras "Dr. Eduardo J. Usunoff", Azul B7300, Argentina (e-mail: mauroh@faa.unicen.edu.ar).

Raúl E. Rivas is with the Comisión de Investigaciones Científicas de la provincia de Buenos Aires, Instituto de Hidrología de Llanuras "Dr. Eduardo J. Usunoff", Tandil B7000, Argentina.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2015.2504262

1939-1404 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

net primary production of crops and the amount of fraction of absorbed photosynthetically active radiation (fAPAR) [13], which is evaluated seasonally with NDVI. Thereafter, these models are extended to analyze light-use efficiency and the correlation with temperature, photosynthesis, and respiration [13]. A common limitation that restricts the applicability at field scale is the low temporal resolution of satellite missions with appropriate spatial resolution images (e.g., Landsat and ASTER). Finally, deterministic crop growth modeling, agro-meteorological or soil vegetation atmosphere (SVAT) modeling are the most complex approaches. The physiological mechanisms of crop growth are modeled taking into account the relationship between crop condition and environmental variables [13]. Given that RS data can incorporate the spatial dimension in the model (e.g., LAI, soil moisture, and crop type), such data are especially useful with distributed numerical models [19]. The main advantages of these methods are the fact that they are dynamic (updating the state variables with input data) and their capability to capture the soil-environment-plant interactions [13]. However, the main disadvantage is the high data requirements, physiological and environmental parameters that are not easily available.

As mentioned, in extensive and rainfed agricultural areas, soil moisture is a frequent limiting factor for crops. Thus, simple RS techniques of soil moisture assessment can contribute to predict crop yield. Although the regression methods need to be calibrated for application in new regions, they became valuable in areas where agro-meteorological data are poor at regional scale. In case crop yield data are not available, thresholds of water scarcity could be defined using RS techniques, beyond which the crops are seriously damaged. One of the most widely used methods to estimate soil water utilizes the electromagnetic spectrum in microwave bands. These methods have a coarse resolution (20–40 km), being useful to monitor near-surface soil moisture (0–10 cm) in large areas [20]. Near-surface soil moisture is not decisive for crop yield, because plants can extract water from deeper soil layer in advanced stages of growing season, including groundwater [3], [21]. In addition, in soils with high water permeability (e.g., sandy soils), the decoupling between shallow and deep horizons is common [21]. Moreover, soil moisture data retrieved from microwaves are frequently more biased under fully vegetated areas [5]. Therefore, a reliable method to monitor the effect of soil water availability on vegetation condition should consider decoupling between near-surface and root zone soil moisture.

Methods based on optical–thermal infrared data from medium-resolution sensors (e.g., MODIS, 250 m–1 km) are suitable for studies at landscape and regional scales. Also, they can give a proper assessment of vegetation water status through the strong correlation between evapotranspiration and soil moisture availability in the root zone [21], [22]. In general, optical data are useful to obtain spectral VI that indicate the amount of bare soil and vegetation and the photosynthetic capacity of vegetation. Hence, several works have analyzed crops condition and the impact on yield through VI [7], [11], [23], [24]. However, it should be noted that these indices are conservative indicators, because they decrease in advanced stage of the stress process when photosynthetic systems are structurally damaged [25]. On the other hand, thermal data give information about energy status since evapotranspiration largely controls the LST according to the surface energy balance

$$Rn = LE + H + G \tag{1}$$

where Rn is the net radiation, LE is the latent heat flux (evapotranspiration), H is the sensible heat flux, and G is the soil heat flux [20]. The lower LE, the higher H is according to the higher energy available for sensible heating of the surface (LST). Given that plants can exert physiological control over the stomatal resistance to transpiration according to soil water content in the root zone, with a given available energy incident at the surface (Rn-G), the stomatal resistance to transpiration determines the partitioning of radiation into H and LE [20]. In this manner, the ability of vegetation to transfer heat away from the surface into the atmosphere strongly affects the LST. Therefore, the combination of optical (visible and near infrared) and thermal data in LST/VI methods can be used to evaluate short and long-term variations in the evapotranspiration process in relation to root zone soil moisture, hence the effects on crop yield.

LST/VI methods have been frequently used to estimate surface soil moisture [20], [26]–[29]. In [30], we showed the strong correlation between the TVDI, which is based on the relationship between LST and VI, and surface soil water content (10–20 cm depth) in Northern hills of Argentine Pampas. Ref. [26] also reported a strong relationship between TVDI, rainfall data and surface soil moisture (10–20 cm) in China. Nevertheless, few studies have analyzed subsurface soil water with optical–thermal methods. In Canada, Ref. [31] estimated water content at different soil depths (up to 100 cm) with the temperature vegetation wetness index from MODIS 8-day composites, reporting low errors ($\pm 20\%$) between RS-derived soil water content and ground-based measurements. Ref. [32] estimated the evaporative fraction (ratio between latent flux and available energy at the land surface) in semi-arid ecosystems of Africa with optical and thermal data from MODIS products. They found good determination coefficient ($R^2 = 0.54$) between field evapotranspiration measurements and estimated evaporative fraction, which is essentially controlled by water availability in the root zone. In [21], we validated the TVDI/subsurface soil moisture relationship in vegetated areas of two agro-climatic zones of Argentine Pampas. Analyzing rainfed maize crop and native grassland over sandy loam and silty loam-silty clay soils, we obtained high determination coefficient ($R^2 > 0.69$) between daily TVDI from MODIS and soil water content up to 120 cm depth.

Despite their aptitude for deep soil water and vegetation water condition assessment, LST/VI methods have been scarcely used for crop yield estimation. In southern Africa, Ref. [15] found a statistically significant correlation between maize yield, weekly vegetation condition index (VCI) and temperature condition index (TCI), which are based on normalized NDVI and LST computed from AVHRR [14], respectively. On the other hand, Ref. [33] found high correlation coefficients between rice productivity in Bangladesh and weekly TCI, VCI, and VHI (vegetation health index, a linear combination of VCI and TCI) using AVHRR data. In addition,

these authors found higher correlation using VHI than TCI and VCI used separately. In [30], we reported a strong relationship ($R^2 > 0.68$) between soybean and wheat yield and monthly TVDI data during the critical growth stage (flowering, grain filling, and setting pod) in four different agro-climatic zones spread over the Argentine Pampas. In addition, a preliminary model of crop yield and soil water availability relationship at regional scale was proposed. Nevertheless, the performance of that model needs to be evaluated at finer spatial–temporal scales in different crops and environmental conditions. Moreover, the influence of submonthly water availability on crop yield and the limitations/significance of the approach should be analyzed. The overall aim of the present study was to analyze the aptitude of the TVDI to forecast spatial maize yield at regional and field scales. In addition, the impact of 8-day and 16-day soil moisture availability on maize yield is analyzed in the cornbelt region of Argentina to determine the most suitable temporal scale for yield forecasting. Finally, the limitations of the model are analyzed.

II. Study Area, Material and Methods

A. Study Area

Argentine Pampas are characterized by a humid and subhumid temperate climate, flat terrain, fertile soils, and extensive rainfed crops. Agriculture is based on few crops, among which maize and soybean are the most representative during the summer. Two different agro-climatic zones were selected to evaluate the relationship between LST/VI data and maize yield at regional scale: Northern hills and semi-arid Pampas [Fig. 1(a)]. In Northern hills, the mean annual rainfall (P) is about 900 mm, the reference evapotranspiration (ET_0) is 1000 mm, with usual water excess during autumn and winter [Fig. 1(c)]. The mean annual temperature is about 16 °C. The dominant soil type is Typic Argiudoll with silty loam and silty clay loam soil texture, showing better soil conditions for cropping than semi-arid Pampas because of the high water retention capacity and high fertility. The principal summer crops are soybean, sunflower, and maize covering 50% of total area. The average field size is about 100 has and crops production takes place mostly without irrigation, which is applied only at local level without an effect at regional scale. In semi-arid Pampas, P is about 600 mm, highly variable, with peaks during spring and summer. The ET_0 is around 1200 mm and water deficits are common during the summer (December–February) [Fig. 1(d)]. The mean annual temperature is 20 °C. The predominant soil type is Haplustoll with sandy loam texture, low water retention capacity, and low organic matter. These agro-climatic characteristics are common limiting conditions for several crops, because they are frequently exposed to water stress and high temperatures, causing high yield variability from year to year. The dominant summer crops are maize, sunflower, sorghum, and soybean (30% of total area) and the mean field average size is about 100 has. Agricultural lands are mainly located in the central-northern area, where rainfall (≈ 750 mm/year) and soil quality increase. In both agro-climatic areas, the critical growth stage (period of highest crop yield sensitivity to water fluctuations) can include November, December, and January while harvest usually occurs in late-February in semi-arid Pampas and late-March in Northern hills [34].

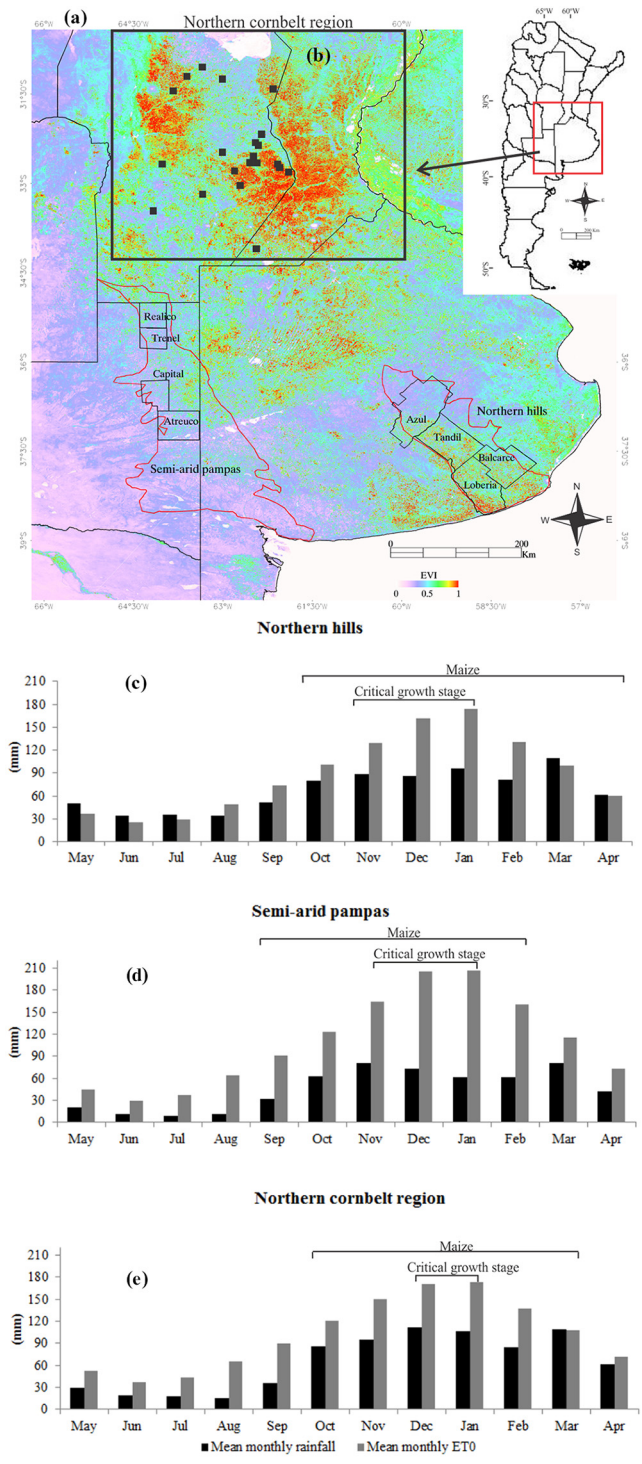

Fig. 1. (a) Study area in central-eastern Argentina, location of the two agro-climatic zones and counties considered for the regional analysis (EVI image, January 2011). (b) Location of test sites used for the analysis at field scale. Typical maize growth period and representative meteorological characteristics (monthly rainfall and reference evapotranspiration) in (c) Northern hills, (d) semi-arid Pampas, and (e) Northern Cornbelt region (data from International Irrigation Management Institute, 1997).

On the other hand, the TVDI-maize yield relationship was analyzed at field scale in Northern cornbelt region [Fig. 1(b)]. The climate is characterized by a mean annual precipitation

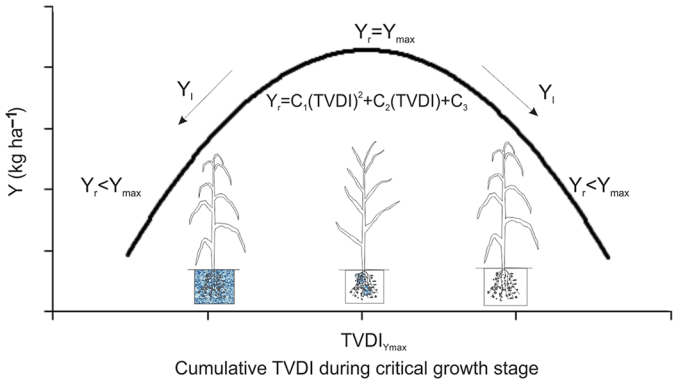

Fig. 2. Generalized model of crop yield and water stress index (TVDI) relationship (adapted from [30]).

of 800 mm with maximum values during spring–summer, ET_0 around 1000 mm and sporadic water deficits [Fig. 1(e)]. The mean annual temperature is 19 °C . The primary soil types are Haplustoll and Argiudoll with high water availability during spring and summer, producing high productivity of dominant summer crops (maize and soybean). Typically, the critical growth stage of maize is between December and January, depending on planting date, and the harvest date is between mid-March and mid-April.

At regional scale, it should be noted that an intensification of farming has been observed during the last century, resulting in important changes in land use and tradeoffs among productivity, stability, and sustainability (e.g., floods in the northern sector of semi-arid Pampas) [35], [36]. However, the processes that occur at such large scale are still not entirely understood, thus the development of methods to evaluate the spatial dynamic of these systems is required.

B. Calibration Methodology

1) Generalized Crop Yield Prediction Model: Considering that soil moisture availability in rainfed cropland is the main factor causing crop yield variability at regional scale, in [30], we proposed and validated a preliminary model of the relationship between an RS water stress index (TVDI) and crop yield (Fig. 2). If soil water availability satisfies the demand of crop during the critical growth stage, yield would reach a maximum value (Y_{max}) for intermediate TVDI. Under these conditions there are no water limitations and the real evapotranspiration is close to the potential evapotranspiration ($\lambda E \approx Rn$) [37]. Frequently, moisture supply is less than crop demand (high TVDI), and real yield (Y_r) is lower than Y_{max}. At this point, low VI and high LST indicate vegetation damage and water stress. This condition would be expected not only under limited rainfall, but also in soils with low water retention capacity. Also, yield should decrease under water excess (low TVDI) in poorly drained areas and during periods with excessive rainfall [30]. This behavior is typical of the subhumid Pampas of Argentina and has relevance for agricultural productivity at regional scale. In this manner, the generalized model is described by a quadratic function

$$Y_r = C_1(TVDI)^2 + C_2 TVDI + C_3 \quad (2)$$

where TVDI is the cumulative TVDI during critical growth stage of the crop, $C_1, C_2,$ and C_3 are coefficients derived from the regression between crop yield and TVDI data being characteristic of each agro-climatic region according to soil type, crop type, and rainfall. C_1 is negative (concave shape of the relationship) indicating that yield can decrease with water deficit and excess. High values of C_1 indicate a narrow parabola showing that yield is highly sensitive to soil moisture variability (e.g., subhumid areas where yield is strongly affected by both saturated soils and water deficit). $C_1 = 0$ in case of predominant water stress conditions (e.g., semi-arid regions and crops with high water demand) where an increase in soil moisture only produces an increase of yield. C_2 is positive (parabola located on the right of y-axis), increasing with the productive capacity of the area. C_3 is negative in case of a quadratic relationship and positive in case of $C_1 = 0$, taking theoretical values characteristic of the regression. Given that the relationship between root zone soil moisture and TVDI over vegetated surfaces depends on soil physical limitations, depth soil exploration by roots and the sensitivity of vegetation to water stress, the $C_1, C_2,$ and C_3 coefficients would be different for each soil type, climatic region, and crop types.

According to this approach, the relative loss of yield (Y_l) can be estimated as function (f_1) of Y_r and Y_{max} [38]

$$Y_l = f_1\left(\frac{Y_r}{Y_{max}}\right) = 1 - \frac{Y_r}{Y_{max}} \quad (3)$$

Integrating (2) and (3), we can obtain

$$Y_l = 1 - \frac{Y_r}{Y_{max}} = 1 - \frac{C_1(TVDI)^2 + C_2 TVDI + C_3}{Y_{max}} \quad (4)$$

To determine Y_{max}, the first-order derivative of (2) can be computed to find the point where the slope of the function is equal to 0. Then, thresholds of Y_l could be defined operationally for each productive region and thus, early warnings can be triggered from TVDI images previous to harvest. If extreme values of yield were considered in the calibration process for a study region, the model is valid for every year unless new technology, agricultural management strategies (e.g., agrochemical application, tillage system), or long-term climatic changes produce an important effect on regional productivity. It should be noted that the model estimates expected crop yield and other events like plant diseases or temperature induced heat stress after the forecast date are not taken into account.

2) Satellite Data and Estimation of TVDI: Monthly TVDI and estimation of maize yield were achieved using LST and EVI data from MODIS/AQUA. MODIS/AQUA 8-day composite LST, level 3, version 5, at 1 km spatial resolution (MYD11A25) and 16-day composite vegetation index level-3, version 5, 1 km spatial resolution (MYD13A25) were obtained from the National Aeronautics and Space Administration's Earth Observing System Data and Information System (http://reverb.echo.nasa.gov/reverb). Given that MODIS products provide atmospherically corrected images, they were used to reduce data processing to obtain surface reflectance and LST, thus facilitating future uses of the method by nontechnical decision makers. In addition, Ref. [39] reported errors of

± 1 K in MODIS LST data in the study area, which is appropriate for studies of large areas with high LST variability. On the other hand, TVDI was calculated using EVI instead of NDVI, because the former takes into account the canopy background and atmospheric influences with improved sensitivity in high biomass surfaces [40]. AQUA data were used to consider periods of maximum atmospheric evaporative demand during the day (approximate satellite overpass: 2:30 P.M.) to ensure that fluctuations in canopy LST are mainly explained by water availability variations. Two adjacent images were mosaicked to cover the study area. Cloudy observations were excluded from the analysis based on cloud masks accompanying the MYD11A25 product. Then, four LST and two EVI composite images were averaged to obtain monthly images [30]. Ref. [41] has shown the positive filtering effect on composite images to eliminate atmospheric noise. Given that current image products, which are globally available have no automated process for data smoothing to reduce noise [42], monthly images allowed the smoothing of data series. The study period included drought (2007–2008), normal (2009–2010, 2010–2011), and high (2002–2003) water availability conditions, allowing the analysis over a wide range of stress level.

On the other hand, the analysis at field scale in Northern cornbelt region was carried out computing an 8-day LST image with a 16-day EVI image to obtain 8 day-TVDI values. We assumed that 16-day EVI values are representative of the 8-day period, given the conservative nature of VI [25]. 16-day TVDI was obtained averaging 2 LST images per EVI image.

Several authors [29], [30], [32], [43] demonstrated the relationship between LST and VI in the form of a scatterplot. If pixels of a heterogeneous area with wide range of soil wetness and fractional vegetation cover (wet to dry and from bare soil to fully vegetated zones) are considered, the scatterplot of LST and VI frequently shows a triangular shape [26]. Following this concept, the TVDI having values of 1 (indicating limited water availability) and 0 (maximum soil wetness and potential evapotranspiration) can be defined [29]

$$TVDI = \frac{LST - LSTmin}{LSTmax - LSTmin} \tag{5}$$

where LST is the observed surface temperature at a given pixel, $\mathrm{LST}_{\min}$ is the minimum temperature (maximum LE). $\mathrm{LST}_{\max} = \mathrm{aVI} + \mathrm{b}$ is the maximum temperature for a given VI, modeled as a linear fit to VI. The "a" and "b" parameters are the intercept and slope of the linear regression of $\mathrm{LST}_{\max}$. Both $\mathrm{LST}_{\min}$ and $\mathrm{LST}_{\max}$ refer to minimum and maximum LST, respectively, during the study period of a study area with regular atmospheric forcing.

A number of possible error sources in estimation of TVDI should be noted.

1) The parameters of TVDI have to be estimated on the basis of pixels from a region with uniform atmospheric forcing. Otherwise, the sensitivity of TVDI to reflect changes in soil moisture is limited.
2) The definition of TVDI parameters involves a large degree of uncertainty, because the triangular shape is visible if different conditions of fractional vegetation and soil moisture are considered [44]. Otherwise, if wet/dry conditions are predominant the $\mathrm{LST}_{\max}$ and $\mathrm{LST}_{\min}$ can be underestimated and overestimated, respectively. This limitation of the method is often overcome with the use of medium-resolution sensors (>250 m spatial resolution), whose large swath width easily offers the required heterogeneity of water availability and vegetation cover. With images of higher resolution, other models should be applied to estimate water status (e.g., trapezoidal relationship between LST and VI) because the concept that LST decreases with an increase of VI is usually not linear.
3) Cloud cover and shadow restrict the calculation of TVDI because of the lack of data or variations in net radiation. Therefore, the use of high quality images is required.

To compute TVDI by using (5) in both agro-climatic zones, monthly $\mathrm{LST}_{\max}$ and $\mathrm{LST}_{\min}$ were determined from LST/VI scatter plots. On the other hand, Refs. [45] and [46] explained that the true $\mathrm{LST}_{\max}$ represents zero water availability and evapotranspiration, with LST reaching a physical maximum when complete stomatal closure occurs. Although this theoretical parameter is rarely observed, monthly $\mathrm{LST}_{\max}$ was obtained based on the LST/EVI scatter plot of semi-arid zone. In such zone, high LST is easily noticeable and the remotely sensed $\mathrm{LST}_{\max}$ frequently reflects the true one [45]. Then, $\mathrm{LST}_{\max}$ was obtained using the method of least squares, with a significance level of 5%. Monthly $\mathrm{LST}_{\min}$ was obtained based on the LST/EVI scatter plots of humid areas of Argentine Pampas by averaging a group of points with minimum LST for different EVI intervals. In such humid areas, numerous pixels show potential evapotranspiration allowing a suitable definition of $\mathrm{LST}_{\min}$. Although Ref. [26] reported that the $\mathrm{LST}_{\min}$ modeled as a horizontal line parallel to the VI axis may be inappropriate for diverse land cover and climatic characteristics, in [21] and [30], we showed that averaging minimum LST is suitable to define the $\mathrm{LST}_{\min}$ in Argentine Pampas. Finally, given that TVDI values have seasonal influence [29] and $\mathrm{LST}_{\min}$, "a" and "b" parameters of TVDI are characteristics of each image, monthly parameters were compared to define extreme $\mathrm{LST}_{\max}$ and $\mathrm{LST}_{\min}$. Thus, comparable TVDI values throughout the study period were obtained using the extreme (maximum slope and intercept) and $\mathrm{LST}_{\min}$ (minimum $\mathrm{LST}_{\min}$) (Fig. 3) [30]. Previously, it has been proved that these extreme parameters are useful for the normalization of data in multitemporal studies [21].

3) Crop Yield Data: Data of maize yield ($\mathrm{kg\,ha^{-1}}$) in Argentina at regional scale come from Statistics Argentina [47]. The official statistics, which are only available at county level (administrative districts in which the provinces are organized), were extracted for each season of the study period. Previous works [48], [49] showed that developing regression models for each soil and climatic type increased the correlation between RS data and crop yield. Given that soil types in Argentine Pampas tend to follow the agro-climatic zones, a regression analysis was done for maize crop in both agro-climatic zones. The correlation between yield data (dependent variable) and TVDI (independent variable) was evaluated in the following counties: Azul, Tandil, and Lobería (Northern hills), Trenel, Capital, and Atreucó (semi-arid Pampas) (Table I). The pixels were averaged to produce a county-level estimate of TVDI.

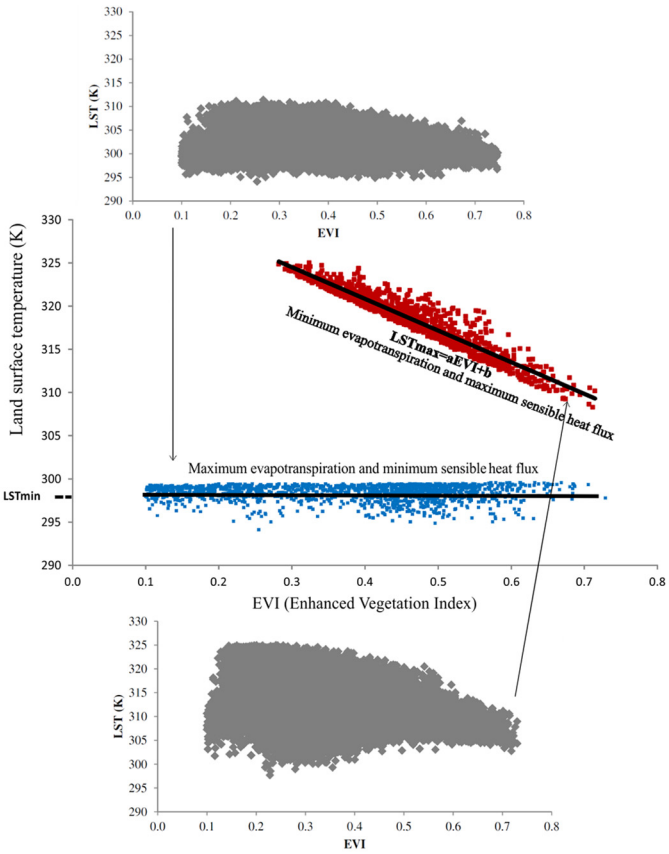

Fig. 3. LST/EVI scatter plot and monthly extreme $\mathrm{LST_{max}}$ (December 2010) and $\mathrm{LST_{min}}$ (December 2009) used to compute TVDI (5). The value of $\mathrm{LST_{min}}$ and the linear equation representing the $\mathrm{LST_{max}}$ was obtained based on the monthly LST/EVI scatter plots of humid and semi-arid zones of Argentine Pampas, respectively.

Based on EVI images of each analyzed season, only cultivated areas were taken into account for regression analysis to eliminate the influence of nonagricultural land in the TVDI signal. These areas were defined removing water bodies and natural grassland (areas where plots were not identifiable). These areas were verified with MODIS/AQUA 16-day composite EVI at 250 m (MYD13Q1) images for each analyzed season. Crop type masks are not currently available in Argentina and it was not possible to isolate the areas cultivated with maize from other summer crops at the analyzed scale. Therefore, a study was carried out at field scale using only yield measurements of maize fields in the Northern cornbelt region to ensure the analysis on maize [Fig. 1(b)]. Maize yield data of 26 plots (plot size $\approx$ 100 has) during 2009–2010 season were compared with the corresponding TVDI on 8-day, 16-day, and monthly basis to determine the most suitable temporal scale to forecast yield. Thus, the performance of the model for maize yield estimation at regional scale was also complemented with the analysis at field scale.

It should be noted that the effect of water availability on crop yield is different in each phase of the growth process. There is a specific stage during which soil moisture fluctuations impact adversely resulting in low yields. This critical growth stage is the period of highest crop yield sensitivity to soil moisture and it includes the grain filling phase [50]. Fortunately, this stage is consistent with high fractional vegetation, thus the evaporation of soil being reduced; therefore the TVDI signal comes principally from vegetation transpiration, which strongly determines the final productivity [51]. Hence, the background effect is reduced and the TVDI is more suitable for crop yield monitoring. According to [34], for maize, this stage can involve November, December, or January in Northern hills and semi-arid Pampas. We investigated the month of the highest crop sensitivity to TVDI for maize yield estimation. It should be noted that a season can be delayed or advanced according to soil moisture availability or weather forecast. Thus, the regional planting date for each year should be useful to define the critical month and to apply the method in an area of interest. The data of the fields analyzed in Northern cornbelt region allows us to define the critical growth stage between mid-December to mid-January in such zone, being the harvest date between mid-March and mid-April.

C. Validation of Maize Crop Yield Estimates

To assess the robustness of the proposed regression models, grain yields were estimated for counties or years previously not considered in the calibration process and they were compared with yield data in: Azul and Balcarce (Northern hills), Trenel, Capital, and Realicó (semi-arid Pampas) (Table I). Thus, different data were used to build the models and evaluate their aptitude for prediction. As in the calibration process, cultivated areas were taken into account to eliminate the influence of nonagricultural land in the TVDI signal during the validation analysis. The results were evaluated using root-mean-square error (RMSE), the bias and index of agreement (d) [53]

$$d = 1 - \frac{\sum_{i=1}^{N} (O_i - E_i)^2}{\sum_{i=1}^{N} \left(|E_i - \bar{O}| + |O_i - \bar{O}|\right)^2} \tag{6}$$

where N is the number of observations, $\mathrm{O_i}$ and $\mathrm{E_i}$ are observed and estimated yield values, respectively. The bias indicates overestimation (underestimation) in case of negative (positive) values. It should be noted that while $\mathrm{R^2}$ measures the proportion of the total data variation explained by the regression model, the d index is used especially for validating prediction models [54]. If all modeled values fit the observed values, d equals 1. In case of no relationship between estimated and observed data, d shows values close to 0.

III. Results and Discussion

A. Estimation of TVDI

1) Parameters of $\mathrm{LST_{min}}$ *and* $\mathrm{LST_{max}}$ *:* The temporal evolution of monthly slope and intercept of $\mathrm{LST_{max}}$ follows the typical behavior of temperature, rainfall, and incident energy at the surface in Argentina Pampas [Fig. 4(a) and (b)] [55]. In other words, the slope shows increasingly negative values [Fig. 4(a)] and the intercept increasingly positive values from spring (October and November) to summer, reaching peaks in December and January [Fig. 4(b)]. This behavior indicates a

TABLE I
CHARACTERISTICS OF ANALYZED COUNTIES FOR TVDI/MAIZE YIELD RELATIONSHIP CALIBRATION (A) AND VALIDATION (B)

Agro-climatic zones	County	Central coordinates		Total area (km^2)	Cultivated area (km^2)	Average yield ($kg\ ha^{-1}$)*
		Longitude	Latitude			
Northern hills	Azul (a, b)	59°53′ W	37°02′ S	1947	1369(±68.2)	5800(±1250.5)
	Tandil (a)	59°14′ W	37°18′ S	3504	2464(±123.4)	6000(±1300.3)
	Balcarce (b)	58°25′ W	37°43′ S	2847	2002(±100.7)	5800(±1200.7)
	Lobería (a)	58°41′ W	38°06′ S	4881	4881(±244.3)	5700(±1200.6)
Semi-arid Pampas	Trenel (a, b)	64°10′ W	35°34′ S	1989	1399(±96.3)	3700(±1200.2)
	Capital (a, b)	64°08′ W	36°30′ S	2500	1758(±120.8)	3800(±450.8)
	Realicó (b)	64°11′ W	35°12′ S	2643	1859(±130.6)	4100(±1100.9)
	Atreucó (a)	63°42′ W	37°02′ S	3386	2284(±160.5)	3700(±650.7)

*Average yield for 2000–2011.

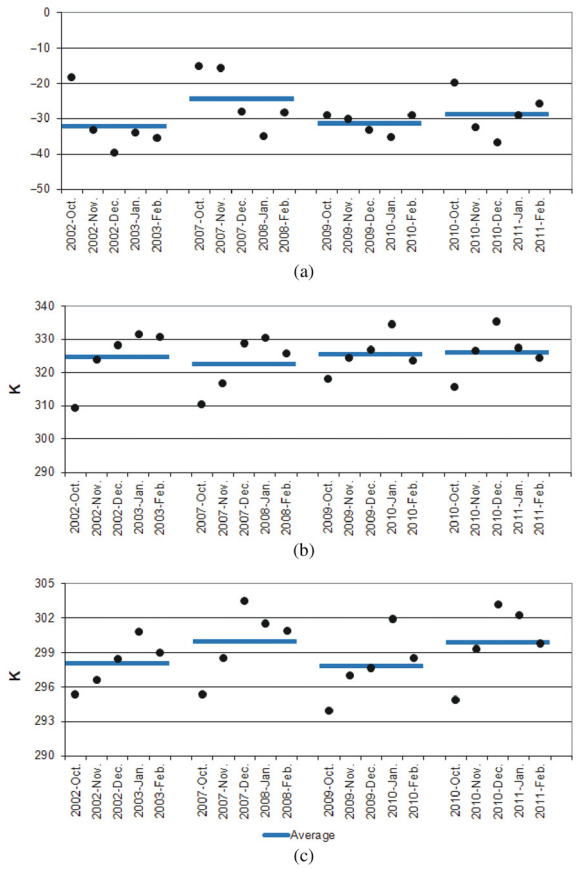

Fig. 4. Evolution (a) slope and (b) intercept parameters of calculated monthly LST_{max} and (c) monthly LST_{min} for 2002–2003, 2007–2008, 2009–2010 and 2010–2011 periods. The values are expressed in kelvin (K). Comparable TVDI values were computed using the extreme LST_{max} (December 2010) and LST_{min} (December 2009).

general drying process during the summer. An opposite trend was observed since February according to decreasing incident energy and atmospheric evaporative demand during the beginning of autumn. These results are consistent with [26], [30], and [32], which reported flat LST_{max} during rainy seasons and periods of low-incoming solar radiation in the Mekong Delta, Argentine Pampas, and Northwest Africa, respectively. This intraannual seasonality of TVDI parameters should be taken into account in multitemporal analysis to evaluate the changes in land surface processes. Otherwise, fluctuations in TVDI should be caused by atmospheric forcing and not by soil moisture availability. Moreover, high variation of the parameters of TVDI, especially in the slopes, is evident in accordance with [29], being the slopes steeper than expected in January 2003 and November 2007. This variability may be due to different factors, including highly variable atmospheric conditions (sun radiation, wind, and rainfall) and surface conditions like fractional vegetation cover or antecedent wetness of soil [43]. Ref. [56] also showed high variability in slope of in mixed grassland of the central United States. Ref. [32] reported high variation in TVDI parameters in semi-arid areas of Africa between dry and wet seasons, showing the seasonal influence in TVDI. On the other hand, LST_{min} (horizontal line of LST/EVI relationship) also shows a strong seasonal behavior, with low values in spring and high values during the summer [Fig. 4(c)]. This parameter also shows high interannual variability, with the highest maximum values in dry period (December 2007) and lowest maximum during the wet period (January 2003). On the other hand, the 298 ± 3 Kaverage of for each period would be related to the El Niño Southern Oscillation. Although the teleconnections of this phenomenon are not strong in the study area, general wet (2002–2003, 2009–2010) and dry conditions (2007–2008, 2010–2011) would be associated to El Niño and La Niña occurred during such years, respectively [57], [58].

Given that TVDI parameters are specific to each image, comparable values of this index for critical growth stage of maize were obtained using monthly extreme LST_{max} and LST_{min}. The former was defined taking into account the maximum slope and intercept, which indicate the lowest water availability [30], [56]. The extreme LST_{max} was $LST_{max} = -37EVI\ (\pm 2) + 336(\pm 3)$ (December 2010). The value of the extreme or minimum LST_{min} that indicates the highest water availability was (December 2009).

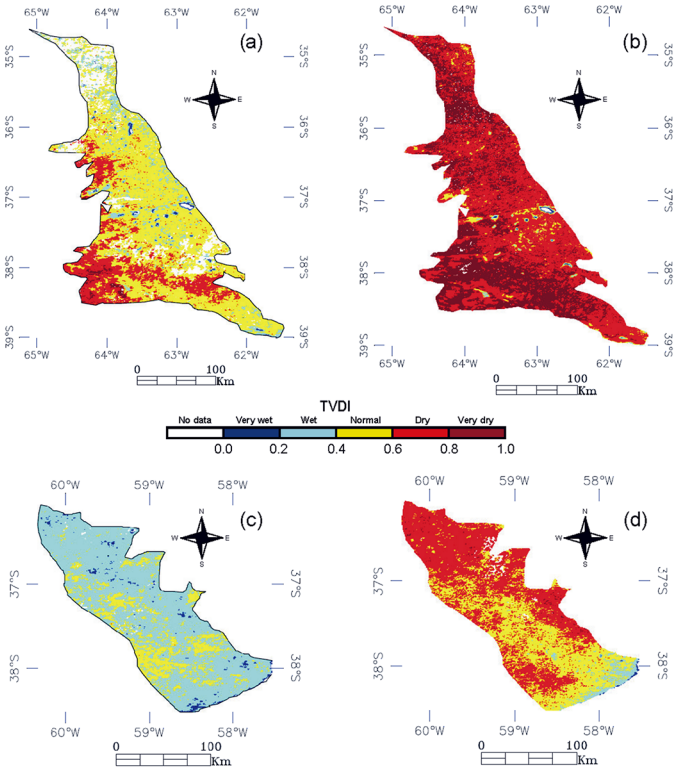

Fig. 5. Monthly TVDI maps obtained for critical growth stage of maize (December) in Semi-arid Pampas during (a) 2002, (b) 2007 and Northern hills during (c) 2002, (d) 2007.

2) Spatio-Temporal Evolution of TVDI: Monthly TVDI values were obtained for the study period using the calculated extreme parameters of (5). Then, a density slice was applied to TVDI images to show classes of soil moisture conditions during critical growth stage of maize (December). Considerable differences between spatial distribution of TVDI during wet (2002) and dry (2007) periods were observed (Fig. 5). In both agro-climatic zones, December 2002 was wet [Fig. 5(a) and (c)], especially in Northern hills where low TVDI values (0–0.4) covered 80% of the area, showing high and uniform soil regional moisture availability. Water scarcity was evident during December 2007 in both zones, particularly in semi-arid Pampas where dry and very dry conditions (TVDI > 0.6) covered 94% of the area [Fig. 5(b)]. In the semi-arid zone high TVDI levels are mainly explained by low water retention capacity of sandy soils which favors moisture-stressed conditions. Also, the rapid soil drying produces a large vertical gradient of soil water content, opposed to a slowly varying fluctuation at greater depths [59]. Thus, poor agreement between near-surface and deeper soil moisture should not be unusual in sandy soils because of the decoupling between them. Given that in arid and semi-arid areas the proportion of dry bare soil viewed by the sensor is usually high, the TVDI method could increase the error in subsurface soil moisture estimation. In case of cultivated surfaces, the canopy temperature reflects the root zone soil moisture content [21].

The high spatial and temporal variability of TVDI is due not only to the variability of rainfall events as shown by [26], [29] or atmospheric evaporative demand but also to subsurface conditions like shallow groundwater table. In general, in low lands (e.g., around the lakes) of semi-arid Pampas, shallow groundwater can contribute to transpiration of vegetation during dry periods, resulting in high crop yield. In particular, in certain areas, the arid climate has produced saline soils not suitable for crop production. In Northern hills, low TVDI values were mainly observed during December 2007 in moist areas close to coastline, related to shallow water table [Fig. 5(d)]. It should be noted that in Northern hills normal soil moisture contents (TVDI 0.4–0.6) are highly variable in space and time and are associated with the best water conditions for wheat yield (the main winter crop in such zone) [30]. Given this high variability of water status, the spatial monitoring of soil moisture availability is essential for crop yield estimation to support regional and national importing–exporting strategies. Finally, it should be noted that cloud cover was more extended in 2002 (semi-arid Pampas), which may restrict the use of TVDI during wet periods for land processes studies.

TABLE II
COEFFICIENT OF DETERMINATION OF THE RELATIONSHIP BETWEEN MAIZE YIELD AND TVDI FOR DIFFERENT MONTHS OF THE GROWING SEASON

Months	R^2 northern hills	R^2 semi-arid Pampas
Sep	0.10	0.26
Oct	0.18	0.37
Nov	0.01	0.57
Dec	0.83	0.73
Jan	0.06	0.19
Feb	0.19	0.03

B. TVDI/Maize Yield Relationship Estimation and Validation of Yield Estimates

To define the month of highest crop sensitivity to water availability, the comparison between TVDI and maize yield was done for different months of the growing season (Table II). For both agro-climatic zones, December was the month of the highest correlation between TVDI and maize yield, suggesting that the TVDI method should be applied on maize taking into account the month of critical stage. Otherwise, considering the whole growing period accuracy of the method should be low. Although the number of data is limited, the dataset shows a clear trend in the TVDI/yield relationship (Fig. 6). Consistent with [30] and [38], the high values of the coefficient of determination ($R^2 > 0.73$) corroborate that maize is highly sensitive to water availability and yield was directly correlated with the RS stress index. These results strengthen the idea that TVDI could be a useful indicator not only for drought early warning but also for crop yield forecasting at regional scale. It should be noted that during critical growth stage (late vegetative and grain filling) the crop reaches the maximum canopy cover, rooting depth, and water requirement [6], [37] and uses mainly the root zone soil moisture for transpiration. During this stage, evapotranspiration is dominated by the transpiration process and is affected by the evaporative demand of the atmosphere and the stomatal resistance of vegetation [60]. Generally, during this stage, micro wave-derived soil moisture has low accuracy by canopy effect [27] while TVDI fluctuations respond to

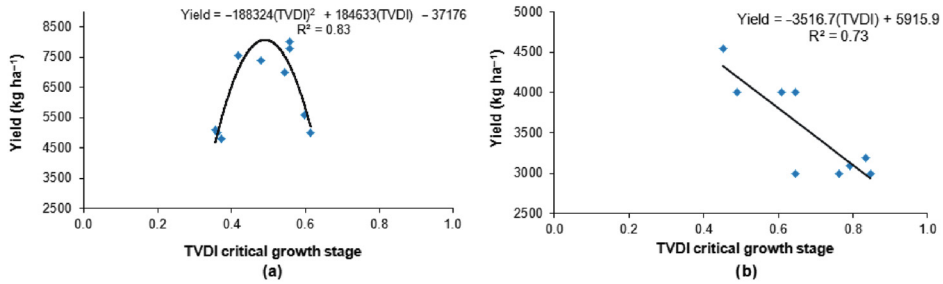

Fig. 6. Regression models of maize yield (kg ha^{-1}) as a function of TVDI in (a) Northern hills and (b) Semi-arid Pampas. In humid zones (a) a quadratic function shows that yield decreases not only due to water stress but also water excess. Linear relationship in semi-arid zone (b) shows a strong effect of water stress on yield.

subsurface soil moisture making it more suitable for assessment of water condition in cropland.

The quadratic adjustment depicting the TVDI/maize yield relationship in Northern hills agreed with the generalized model shown in Section II-B1 [Fig. 6(a)]. Such adjustment shows that, in addition to the effect of water stress, the impact of water excess is evident in the humid climate of Northern hills even over summer crops. Principally, yield decreased under water excess because of the more frequent diseases (e.g., fungi) and decrease of photosynthetic active radiation by high cloud cover [30]. In addition, the shallow water table and dense soil horizon (Bt) in such zone would reduce the vertical movement of water in the soil and would increase poor aeration [21], [61]. The quadratic adjustment found in such zone indicates that TVDI values around 0.5 (normal conditions) would be desirable to obtain the maximum yield and that a minimum and maximum TVDI threshold around 0.40 and 0.55 would be identified beyond which maize yield loss increases. These results, together with the maps of Fig. 5, show that the central area of this agro-climatic zone is the most stable showing normal soil moisture condition and probably efforts should be made so that production becomes more efficient and sustainable in terms of natural cycle of nutrients and soil fertility. In this sense, Ref. [62] has reported an increasing weakness of organic compartments in nitrogen and phosphorus cycles in Argentine Pampas during the last decades due to intensification of farming.

On the other hand, the linear model of TVDI/maize yield relationship in semi-arid Pampas shows a strong effect of water stress on maize yield [Fig. 6(b)]. In Argentine Pampas, water retention of soils decreases from humid to semi-arid areas being a limiting factor for efficient use of rainfall by crops. Although the coefficients of regression adjustments can differ depending on crop, soil types and environmental conditions, these results strengthen the assumption that in zones with predominant water stress, a linear trend would be expected (see Section II-B1). This is consistent with [51], who reported that maize on sandy soils experienced more frequent water stress than in clay soils. Using a soil water balancing, [37] also found a linear relationship between a cumulative soil-moisture index and pearl millet yield in the arid zone of India. On the other hand, in the central-eastern area of semi-arid Pampas, the intensification of farming has resulted in an important conversion of native forests into cultivated grasslands and croplands. Ref. [63] has reported tradeoffs between ecosystem services and productivity in these farm systems in response to increased human intervention. Such situation, together with the frequent high levels of water stress, shows that agro-ecological alternatives should be explored to increase sustainability in these sensitive systems.

Regarding previous works, analyzing maize yield in Zimbabwe with temperature-vegetation condition indices from AVHRR, Ref. [15] reported similar results (linear regression, R^2 between 0.64 and 0.93), obtaining expected yield 6 weeks before harvest. Given that harvest usually occurs in late-February in semi-arid Pampas and late-March in Northern hills [34], our results indicate that the proposed model could predict maize yield 8–12 weeks before harvest, depending on the agro-climatic zone. Moreover, it should be noted that this period varies because harvest can be delayed up to 2 months in the study area depending on diverse environmental conditions after yield estimation (e.g., spatial heterogeneity of seed maturation, diseases, weeds, and water stress). Nevertheless, the key point of the method is to define the critical growth stage of maize. The R^2 values were found to be comparable with recent results obtained using more complex models based on satellite data at regional/landscape scale [5], [12], [64], [65]. It should be noted that the lower coefficient of determination in semi-arid Pampas could be explained by the spatial heterogeneity and the inclusion of other crop/vegetation types, which may degrade the model performance. The mixed nature of low resolutions pixels may be a constraint of the method for such zone [13] and regions with sparse cultivated areas or where different crop types coexist. [50], [66]–[68] have indicated the effect of the subpixel heterogeneity in scattered cultivated areas on low-resolution sensors. In this sense, the next generation of sensors with high spatial resolution (e.g., SENTINEL series, resolution $= 300$ m) will improve the aptitude for yield forecasting. However, it should be noted that the LST/VI triangle method requires a wide range of surface conditions (dry to saturated bare soil, water stressed to well watered vegetation) to define the model parameters [20], [29], which is easily achieved using sensors with a large swath width.

A validation of the proposed models was carried out to quantify their accuracy for maize yield estimation comparing satellite-derived and official statistics of maize yield (kg ha^{-1}) in counties or dates previously not considered in the models. Low RMSE values indicate a good performance of our proposed models for maize yield estimation (Table III). Obtained RMSE values are similar to the results found in [30] for different crops in the study area (12%–13% for soybean and

TABLE III
VALIDATION RESULTS BETWEEN MAIZE CROP YIELD ESTIMATED FROM OUR APPROACH AND OFFICIAL STATISTICS REPORTED BY SISTEMA INTEGRADO DE INFORMACIÓN AGROPECUARIA AT COUNTY LEVEL

Agro-climatic zone	RMSE (kg ha^{-1})	Bias (kg ha^{-1})	d
Northern hills	900 (14)	110	0.85
Semi-arid Pampas	700 (19)	–180	0.98

For comparison, numbers in brackets are percentages of average yield.

14%–22% for wheat). As we mentioned previously, the spatial heterogeneity of semi-arid zone would explain the slightly higher RMSE in relation to the more uniform humid zone. In this sense, Ref. [69] estimating yield from MODIS products data reported the effect of different radiation use efficiency of crops and mixed pixels over soybean and maize. In addition, bias values indicate that yield was underestimated in Northern hills and overestimated in semi-arid Pampas, respectively. The d index values showed a good correspondence between estimated and observed yield. Although data were limited by cloud cover during the study period (n $= 8$ in both zones), the overall parameters of validation indicated the aptitude of the models for maize yield forecasting in both agro-climatic zones. These results show an acceptable level of accuracy of the proposed approach to estimate spatial maize yield at regional scale where detailed data are not available. The largest errors were observed in semi-arid Pampas, where the sub-pixel heterogeneity within an MODIS $1\ \text{km} \times 1\ \text{km}$ pixel would be more in sparse agricultural areas [67]. In addition, the presence of shallow petrocalcic horizons in small cropland areas can produce yield variations at landscape scale that cannot be captured at 1 km resolution. Also, given the high temporal variation of meteorological conditions in semi-arid Pampas, the errors may have been caused by the fact that the sowing dates vary from year to year depending on weather conditions and these periods were considered static [70]. Under such conditions, an updated crop type mask that allows to isolate the areas cultivated with maize would improve yield estimates.

Although the TVDI method cannot differentiate crop type, which can affect the performance of the model, the good correlation obtained in both agro-climatic zones would indicate that all vegetation in a cultivated region integrates the accumulated effect of water availability in some manner. These results agree with observations by [71] for maize in the United States using historical correlations between NDVI from AVHRR and yield. These authors showed higher correlation between NDVI and maize yield in areas dominated by maize and soybean with high spatial homogeneity. In addition, the correlation would increase by the fact that critical stage of maize (December) is not coincident with soybean (January–February) and sunflower (January). If available, a crop-specific mask that allows to consider only TVDI information pertaining to the crop of interest would improve the crop yield forecasting ability at regional scale as shown by [72], [73]. Crop type classification using medium/coarse resolution images remains a challenge in sparsely cultivated fields where different crops coexist. Also, in multiple years, analysis a major problem relates to the widespread practice of crop rotation [42]. The tradeoff between spatial and temporal resolution of current satellite missions that allow information of complete growing season is a key factor in this sense. Although it was not the aim of the work, feasible approaches as resampling (e.g., [39]) or image masking using historical correlation (e.g., [71]) may be tested in future works.

C. TVDI/Maize Yield Relationship at Field Scale

Although the obtained results at regional scale were promising, a test was carried out at field scale in Northern cornbelt region using maize yield measurements to ensure the analysis on maize (Fig. 7). The size of analyzed fields ($\approx$100 ha, $1\ \text{km} \times 1\ \text{km}$) and the high spatial homogeneity of crop type in such region allowed us to use TVDI values at 1 km for analysis and there was no need to employ any spatial resampling technique. In addition, submonthly TVDI values within the critical stage were compared with yield data to evaluate the effect of short variations of soil moisture on yield. The planting date of analyzed fields was between late-September and mid-October, the critical growth stage was between mid-December and mid-January and the harvest date was between mid-March and mid-April. 8-day TVDI values showed low and mid correlation ($R^2 = 0.10 - 0.45$) with final maize yield, showing linear and quadratic relationships, thus suggesting that weekly water availability is not decisive for maize yield [Fig. 7(a)]. Higher correlations were obtained during the central period of critical stage (December 27–31 and January 1–8). Future works may evaluate possible weekly periods during midcritical stage, which could determine maize yield. 16-day TVDI values showed higher correlation with yield than 8-day values ($R^2 = 0.48$ and 0.53) and quadratic adjustments in both cases, also indicating the effect of water excess on yield [Fig. 7(b)]. Finally, monthly TVDI values were analyzed for the month of critical stage and the previous one [Fig. 7(c)]. No correlation was obtained considering the previous month to the critical stage (19 November–18 December) suggesting that water availability during precritical stage could determine plants density but it is not decisive for grain number and grain weight, which determine yield. The monthly period including the critical stage (19 December–16 January) showed the best correlation ($R^2 = 0.61$) with a quadratic adjustment that indicates the negative effect of water excess in a zone with high water retention capacity of soils and rainfall periods concentrated during the critical stage of maize [see also Fig. 1(e)]. These results would indicate that the analysis of the entire month of critical stage should be more appropriate. In [21], we showed the daily variability of vegetation response to water stress according to vegetation and soil types. However, these short-time fluctuations would not have a considerable impact on yield and the cumulative effect of water availability during the entire critical stage should be taken into account.

On the other hand, Ref. [42] has reported that despite the effort to make the VI products globally available in near real-time (e.g., MODIS), data smoothing to reduce noise is an important step that is not yet automated. Thus, a depression in VI time series could not be necessarily caused by water stress [74]. In this sense, averaging TVDI to obtain bi-weekly

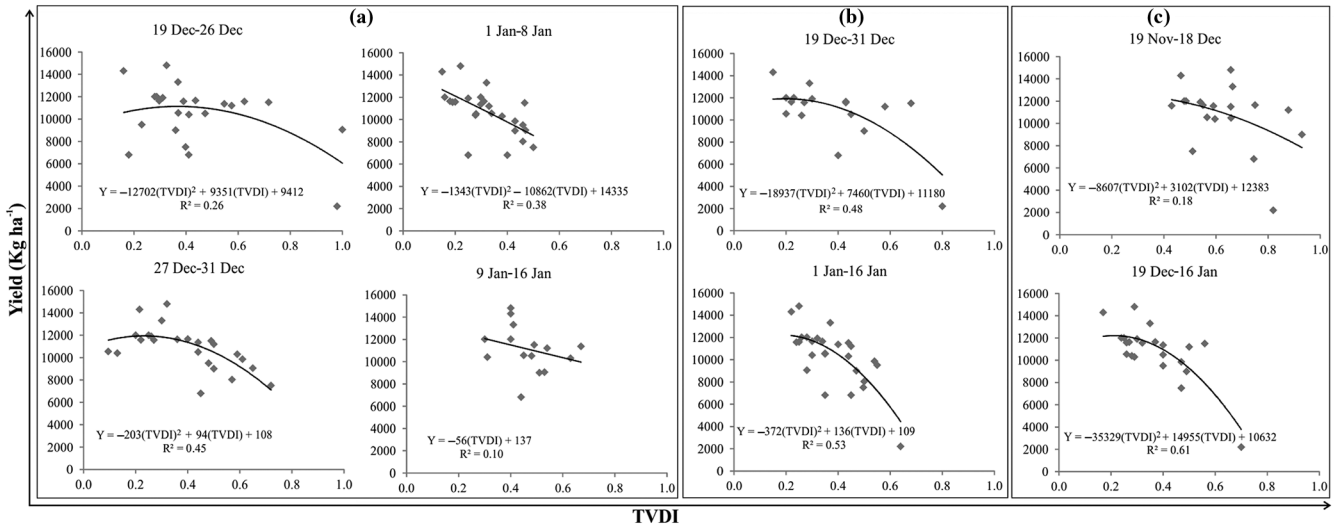

Fig. 7. Relationship between maize yield and (a) 8-day TVDI, (b) 16-day TVDI, and (c) monthly TVDI in maize fields of Northern cornblet region of Argentine Pampas.

or monthly data may result in loss of information about short-time variations of water stress but the correlation with maize yield data would increase. Finally, the slightly lower coefficient of determination obtained at field scale in comparison to the results at regional scale would show that, although the analysis on maize was ensured, other factors such management techniques, topography, maize varieties, and the different response to water availability would affect the performance of the model with finer spatial resolution sensors. However, the obtained results at field scale confirm the aptitude of TVDI for maize yield forecasting and an acceptable level of correlation was achieved 8–10 weeks prior to harvest.

IV. Conclusion

Regression models for forecasting maize yield at regional scale using TVDI from MODIS/Aqua EVI and LST products were evaluated against official statistics. The proposed approach is simple, not data intensive and, once calibrated against yield statistics, requires only remotely sensed data to provide maize yield forecasts. Also, the spatial loss and surplus of yield can be derived prior to harvest in areas with low-ground data coverage.

Results showed that adjustments are consistent with a preliminary quadratic/linear model that we previously reported in Argentine Pampas for different crops (soybean and wheat). Although more studies should be conducted using a larger dataset, the proposed approach is promising to be used operationally for maize yield forecasting with reasonable accuracy in semi-arid and humid large cultivated areas with rainfed crops. Although the results found here would not be directly extrapolated to other agricultural regions in the world, maize crops would respond similarly to TVDI and the method may be useful to forecast maize yield several weeks before harvest. However, in places where yield data are not available for calibration process, the TVDI can be valuable to identify spatial vulnerability to water stress/excess and to estimate critical thresholds as a part of a drought monitoring early warning system.

Statistically significant relationships between the TVDI of critical stage and yield were found in the analyzed counties, although the strength of the correlation varied with the agroclimatic zone. The values were 0.73 and 0.83 for semi-arid and humid area, respectively. The model predictions of maize yield similarly matched the official statistics. The RMSE ranged from 14% to 19% and bias values were lower than 200 kg ha^{-1}. These results are comparable with validation data reported for crop yield estimation using remotely sensed data, which show the potential of using the stress index as an early indicator of crop yield. In addition, the models showed ability to predict maize yield 8 to 12 weeks before harvest in the study area. The analysis at field scale on maize fields suggested that monthly TVDI is more suitable than submonthly values for maize yield forecasting, showing that the cumulative effect of water availability during the entire critical stage should be taken into account. Water availability before and after critical stage seemed to have no effect on yield. Therefore, given that the planting dates can vary from year to year according to weather conditions, a source of error may be caused by considering noncritical stage. Phenological indicators as the start of growing season can be properly estimated from remotely sensed data, which will be useful in other study areas where such information is not available.

Regarding sources of error, short historical data series of yield could represent a limitation to apply this method. If extreme yield values are not well represented in the dataset used for the calibration process, then these extremes are inadequately forecast when the model is applied. On the other hand, a possible inclusion of other crop types within the agricultural areas may degrade the model performance. This would explain the slightly lower accuracy of the model in the heterogeneous semi-arid zone (RMSE $= 19\%$ of average yield) than in the humid zone (RMSE $= 14\%$ of average yield). However, the good correlation obtained at regional scale would indicate that all vegetation in a cultivated region integrates the accumulated effect of water availability in some manner. Using a dynamic crop type mask that allows to isolate

the area cultivated with maize every year would increase the model accuracy, especially in zones with several crop types. In regions where the predominant production system is monoculture, errors should be low. Nowadays, crop type masks are not available in Argentina. Future work should test the proposed approach at landscape scale (e.g., 250 m) and incorporate more detailed characterization of areas covered by maize and growing period.

Finally, the obtained results are promising and show the preharvest maize yield forecasting is operationally feasible given the current and future satellite missions that provide optical and thermal data. These progresses will have significant impacts not only on food security but also on export strategies.

ACKNOWLEDGMENT

The authors would like to thank Consejo Nacional de Investigaciones Científicas y Técnicas, Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, and Instituto de Hidrología de Llanuras "Dr. Eduardo Usunoff" for supporting this research.

REFERENCES

[1] N. S. Diffenbaugh and M. Ashfaq, "Intensification of hot extremes in the United States," *Geophys. Res. Lett.*, vol. 37, no. 15, pp. 1–5, Aug. 2010.

[2] Z. Zhuo, C. Gao, and Y. Liu, "Regional grain yield response to climate change in China: A statistic modeling approach," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 11, pp. 4472–4479, Nov. 2014.

[3] S. Chakrabarti, S. Member, T. Bongiovanni, J. Judge, and S. Member, "Assimilation of SMOS soil moisture for quantifying drought impacts on crop yield in agricultural regions," *IEEE J. Sel. Topics Appl. Earth Observ Remote Sens.*, vol. 7, no. 9, pp. 3867–3879, Sep. 2014.

[4] E. L. Clawson, S. A. Hribal, G. Piccinni, R. L. Hutchinson, R. V. Rohli, and D. L. Thomas, "Weighing lysimeters for evapotranspiration research on clay soil," *Agron. J.*, vol. 101, no. 4, pp. 836–840, 2009.

[5] A. V. M. Ines, N. N. Das, J. W. Hansen, and E. G. Njoku, "Assimilation of remotely sensed soil moisture and vegetation with a crop simulation model for maize yield prediction," *Remote Sens. Environ.*, vol. 138, pp. 149–164, Nov. 2013.

[6] A. Araya, L. Stroosnijder, G. Girmay, and S. D. Keesstra, "Crop coefficient, yield response to water stress and water productivity of teff (Eragrostis tef (Zucc.)," *Agric. Water Manage.*, vol. 98, no. 5, pp. 775–783, Mar. 2011.

[7] M. J. Hayes and W. L. Decker, "Using NOAA AVHRR data to estimate maize production in the United States Corn Belt," *Int. J. Remote Sens.*, vol. 17, no. 16, pp. 3189–3200, 1996.

[8] S. W. Running and R. R. Nemani, "Relating seasonal patterns of the AVHRR vegetation index to simulated photosynthesis and transpiration of forests in different climates," *Remote Sens. Environ.*, vol. 24, pp. 347–367, 1988.

[9] T. L. Barnett and D. R. Thompson, "Large-area relation of Landsat MSS and NOAA-6 AVHRR spectral data to wheat yields," *Remote Sens. Environ.*, vol. 13, pp. 277–290, 1983.

[10] P. J. Pinter, R. D. Jackson, S. B. Idso, and R. J. Reginato, "Multidate spectral reflectance as predictors of yield in water stressed wheat and barley," *Int. J. Remote Sens.*, vol. 2, no. 1, pp. 43–48, 1981.

[11] M. S. Mkhabela, P. Bullock, S. Raj, S. Wang, and Y. Yang, "Crop yield forecasting on the Canadian Prairies using MODIS NDVI data," *Agric. For. Meteorol.*, vol. 151, no. 3, pp. 385–393, Mar. 2011.

[12] T. Sakamoto, A. A. Gitelson, and T. J. Arkebauer, "Near real-time prediction of U.S. corn yields based on time-series MODIS data," *Remote Sens. Environ.*, vol. 147, pp. 219–231, May 2014.

[13] F. Rembold, C. Atzberger, I. Savin, and O. Rojas, "Using low resolution satellite imagery for yield prediction and yield anomaly detection," *Remote Sens.*, vol. 5, no. 4, pp. 1704–1733, Apr. 2013.

[14] F. N. Kogan, "Application of vegetation index and brightness temperature for drought detection," *Adv. Space Res.*, vol. 15, no. 11, pp. 91–100, 1995.

[15] L. S. Unganai and F. N. Kogan, "Drought monitoring and corn yield estimation in southern Africa from AVHRR data," *Remote Sens. Environ.*, vol. 63, pp. 219–232, 1998.

[16] J. Huang, H. Wang, Q. Dai, and D. Han, "Analysis of NDVI data for crop identification and yield estimation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 11, pp. 4374–4384, 2014.

[17] O. Rojas, "Operational maize yield model development and validation based on remote sensing and agro-meteorological data in Kenya," *Int. J. Remote Sens.*, vol. 28, no. 17, pp. 3773–3793, 2007.

[18] A. K. Prasad, L. Chai, R. P. Singh, and M. Kafatos, "Crop yield estimation model for Iowa using remote sensing and surface parameters," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 8, no. 1, pp. 26–33, Jan. 2006.

[19] Z. Jiang *et al.*, "Application of crop model data assimilation with a particle filter for estimating regional winter wheat yields," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 11, pp. 4422–4431, Nov. 2014.

[20] K. Mallick, B. K. Bhattacharya, and N. K. Patel, "Estimating volumetric surface moisture content for cropped soils using a soil wetness index based on surface temperature and NDVI," *Agric. For. Meteorol.*, vol. 149, no. 8, pp. 1327–1342, Aug. 2009.

[21] M. E. Holzman, R. Rivas, and M. Bayala, "Subsurface soil moisture estimation by VI–LST method," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 11, pp. 1951–1955, Nov. 2014.

[22] R. Fensholt, S. Huber, S. R. Proud, and C. Mbow, "Infrared reflectance data from polar orbiting and geostationary platforms," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 3, no. 3, pp. 271–285, Sep. 2010.

[23] A. Dobermann and J. L. Ping, "Geostatistical integration of yield monitor data and remote sensing improves yield maps," *Agron. J.*, vol. 96, no. 1, pp. 285–297, 2004.

[24] S. S. Panda, D. P. Ames, and S. Panigrahi, "Application of vegetation indices for agricultural crop yield prediction using neural network techniques," *Remote Sens.*, vol. 2, no. 3, pp. 673–696, Mar. 2010.

[25] G. D. Farquhar and T. D. Sharkey, "Stomatal conductance and photosynthesis," *Annu. Rev. Plant Physiol.*, vol. 33, no. 1, pp. 317–345, 1982.

[26] C.-F. Chen, N.-T. Son, L.-Y. Chang, and C.-C. Chen, "Monitoring of soil moisture variability in relation to rice cropping systems in the Vietnamese Mekong Delta using MODIS data," *Appl. Geogr.*, vol. 31, no. 2, pp. 463–475, Apr. 2011.

[27] J. Cho, Y.-W. Lee, and H.-S. Lee, "Assessment of the relationship between thermal-infrared-based temperature–vegetation dryness index and microwave satellite-derived soil moisture," *Remote Sens. Lett.*, vol. 5, no. 7, pp. 627–636, Aug. 2014.

[28] Y. Han, Y. Wang, and Y. Zhao, "Estimating soil moisture conditions of the greater Changbai Mountains by land surface temperature and NDVI," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 6, pp. 2509–2515, Jun. 2010.

[29] I. Sandholt, K. Rasmussen, and J. Andersen, "A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status," *Remote Sens. Environ.*, vol. 79, no. 2–3, pp. 213–224, Feb. 2002.

[30] M. E. Holzman, R. Rivas, and M. C. Piccolo, "Estimating soil moisture and the relationship with crop yield using surface temperature and vegetation index," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 28, pp. 181–192, May 2014.

[31] M. S. Akther and Q. K. Hassan, "Remote sensing based estimates of surface wetness conditions and growing degree days over northern Alberta, Canada," *Boreal Environ. Res.*, vol. 16, no. 5, pp. 407–416, 2011.

[32] F. Nutini, M. Boschetti, G. Candiani, S. Bocchi, and P. Brivio, "Evaporative fraction as an indicator of moisture condition and water stress status in semi-arid rangeland ecosystems," *Remote Sens.*, vol. 6, no. 7, pp. 6300–6323, Jul. 2014.

[33] A. Rahman, L. Roytman, N. Y. Krakauer, M. Nizamuddin, and M. Goldberg, "Use of vegetation health data for estimation of aus rice yield in bangladesh," *Sensors*, vol. 9, no. 4, pp. 2968–2975, Jan. 2009.

[34] Oficina de Riesgo Agropecuario. (2013, Oct. 10). *MAGyP-Argentina* [Online]. Available: http://www.ora.gov.ar/

[35] E. F. Viglizzo, F. Lértora, A. J. Pordomingo, J. N. Bernardos, Z. E. Roberto, and H. Del Valle, "Ecological lessons and applications from one century of low external-input farming in the pampas of Argentina," *Agric. Ecosyst. Environ.*, vol. 83, no. 1–2, pp. 65–81, Jan. 2001.

[36] E. F. Viglizzo and F. C. Frank, "Ecological interactions, feedbacks, thresholds and collapses in the Argentine Pampas in response to climate and farming during the last century," *Quat. Int.*, vol. 158, no. 1, pp. 122–126, Dec. 2006.

[37] V. K. Boken, "Improving a drought early warning model for an arid region using a soil-moisture index," *Appl. Geogr.*, vol. 29, no. 3, pp. 402–408, 2009.
[38] J. C. J. Chuan, Q. Q. Q. Qiming, Z. L. Z. Lin, N. P. N. Peng, and A. Ghulam, "TVDI based crop yield prediction model for stressed surfaces," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2007, pp. 4656–4658.
[39] M. I. Bayala and R. E. Rivas, "Enhanced sharpening procedures on edge difference and water stress index basis over heterogeneous landscape of sub-humid region," *Egypt. J. Remote Sens. Space Sci.*, vol. 17, no. 1, pp. 17–27, Jun. 2014.
[40] H. Q. Liu and A. Huete, "Feedback based modification of the NDVI to minimize canopy background and atmospheric noise," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 2, pp. 457–465, Mar. 1995.
[41] C. Atzberger and P. H. C. Eilers, "Evaluating the effectiveness of smoothing algorithms in the absence of ground reference measurements," *Int. J. Remote Sens.*, vol. 32, no. 13, pp. 3689–3709, 2011.
[42] C. Atzberger, "Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs," *Remote Sens.*, vol. 5, no. 2, pp. 949–981, Feb. 2013.
[43] E. Valor and V. Caselles, Towards the use of temperature in desertification monitoring: Results of DeMon-I Project, in *Remote Sensing '96: Integrated Applications for Risk Assessment and Disaster Prevention for the Mediterranean*, Spiteri, Ed. Balkema, Rotterdam, 1997, pp. 305–310, ISBN 905410855X.
[44] R. Tang, Z.-L. Li, and B. Tang, "An application of the Ts–VI triangle method with enhanced edges determination for evapotranspiration estimation from MODIS data in arid and semi-arid regions: Implementation and validation," *Remote Sens. Environ.*, vol. 114, no. 3, pp. 540–551, Mar. 2010.
[45] S. Stisen, I. Sandholt, A. Nørgaard, R. Fensholt, and K. H. Jensen, "Combining the triangle method with thermal inertia to estimate regional evapotranspiration—Applied to MSG-SEVIRI data in the Senegal River basin," *Remote Sens. Environ.*, vol. 112, no. 3, pp. 1242–1255, Mar. 2008.
[46] M. S. Moran, T. R. Clarke, Y. Inoue, and A. Vidal, "Estimating crop water deficit using the relation between surface-air temperature and spectral vegetation index," *Remote Sens. Environ.*, vol. 49, no. 3, pp. 246–263, Sep. 1994.
[47] Sistema Integrado de Información Agropecuaria, 2013 [Online]. Available: http://www.siia.gov.ar/_apps/siia/estimaciones/estima2.php, accessed on Oct. 10, 2013.
[48] C. B. Holzapfel *et al.*, "Estimating canola (Brassica napus L.) yield potential using an active optical sensor," *Can. J. Plant Sci.*, vol. 89, no. 6, pp. 1149–1160, 2009.
[49] B. L. Ma, L. M. Dwyer, C. Costa, E. R. Cober, and M. J. Morrison, "Early prediction of soybean yield from canopy reflectance measurements," *Agron. J.*, vol. 93, pp. 1227–1234, 2001.
[50] T. Sakamoto, B. D. Wardlow, and A. A. Gitelson, "Detecting spatiotemporal changes of corn developmental stages in the U.S. corn belt using MODIS WDRVI data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 1926–1936, Jun. 2011.
[51] J. Barron, J. Rockström, F. Gichuki, and N. Hatibu, "Dry spell analysis and maize yields for two semi-arid locations in east Africa," *Agric. For. Meteorol.*, vol. 117, no. 1–2, pp. 23–37, 2003.
[52] L. Borrás and M. E. Westgate, "Predicting maize kernel sink capacity early in development," *Field Crops Res.*, vol. 95, no. 2–3, pp. 223–233, 2006.
[53] C. J. Willmot, "On the validation of models," *Phys. Geogr.*, vol. 2, no. 2, pp. 184–194, 1981.
[54] S. Jin, *Remote Sensing of Natural Resources*. New York, NY, USA: Taylor & Francis, 2013.
[55] F. Carmona, R. Rivas, and V. Caselles, "Estimation of daytime downward longwave radiation under clear and cloudy skies conditions over a sub-humid region," *Theor. Appl. Climatol.*, vol. 115, no. 1–2, pp. 281–295, Apr. 2013.
[56] S. J. Goetz, "Multi-sensor analysis of NDVI, surface temperature and biophysical," *Int. J. Remote Sens.*, vol. 18, no. 1, pp. 71–94, 1997.
[57] M. E. Holzman and R. E. Rivas, "Utilización de imágenes de temperatura radiativa e índice de vegetación mejorado para el estudio de las condiciones hídricas en la región pampeana," *Rev. Geol. Apl. a la Ing. y al Ambient.*, vol. 25, pp. 25–33, 2012.
[58] M. E. Holzman and R. E. Rivas, "Enso effects on hydric conditions of Pampa region: A preliminary evaluation using LST and EVI," in *Proc. Anais XV Simósio Brasileiro de Sensoriamento Remoto*, 2011, p. 2242.
[59] T. N. Carlson, R. R. Gillies, and T. J. Schmugge, "An interpretation of methodologies for indirect measurement of soil water content," *Agric. For. Meteorol.*, vol. 1923, no. 95, pp. 191–205, 1995.
[60] R. G. Allen, L. S. Pereira, D. Raes, and M. Smith, *Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements*. Food and Agricultural Organization of the United Nations (FAO), Rome, p. 300, 1998.
[61] N. C. Brady and R. R. Weil, "Soil architecture and physical properties," in *The Nature and Properties of Soils*, 14th ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2008, pp. 148–157.
[62] E. F. Viglizzo, A. J. Pordomingo, M. G. Castro, and F. A. Lertora, "Environmental assessment of agriculture at a regional scale in the Pampas of Argentina," *Environ. Monit. Assess.*, vol. 87, no. 2, pp. 169–195, 2003.
[63] E. F. Viglizzo and F. C. Frank, "Land-use options for Del Plata Basin in South America: Tradeoffs analysis based on ecosystem service provision," *Ecol. Econ.*, vol. 57, no. 1, pp. 140–151, Apr. 2006.
[64] D. M. Johnson, "An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States," *Remote Sens. Environ.*, vol. 141, pp. 116–128, Feb. 2014.
[65] Q. Xin *et al.*, "A production efficiency model-based method for satellite estimates of corn and soybean yields in the midwestern US," *Remote Sens.*, vol. 5, no. 11, pp. 5926–5943, Nov. 2013.
[66] L. Busetto, M. Meroni, and R. Colombo, "Combining medium and coarse spatial resolution satellite data to improve the estimation of sub-pixel NDVI time series," *Remote Sens. Environ.*, vol. 112, no. 1, pp. 118–131, 2008.
[67] C. Atzberger and F. Rembold, "Mapping the spatial distribution of winter crops at sub-pixel level using AVHRR NDVI time series and neural nets," *Remote Sens.*, vol. 5, no. 3, pp. 1335–1354, 2013.
[68] G. Pan, G.-J. Sun, and F.-M. Li, "Using QuickBird imagery and a production efficiency model to improve crop yield estimation in the semi-arid hilly Loess Plateau, China," *Environ. Model. Softw.*, vol. 24, no. 4, pp. 510–516, Apr. 2009.
[69] Q. Xin *et al.*, "A production efficiency model-based method for satellite estimates of corn and soybean yields in the midwestern US," *Remote Sens.*, vol. 5, no. 11, pp. 5926–5943, Nov. 2013.
[70] W. Kowalik, K. Dabrowska-Zielinska, M. Meroni, T. U. Raczka, and A. de Wit, "Yield estimation using SPOT-VEGETATION products: A case study of wheat in European countries," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 32, pp. 228–239, Oct. 2014.
[71] J. H. Kastens, T. L. Kastens, D. L. A. Kastens, K. P. Price, E. A. Martinko, and R. Y. Lee, "Image masking for crop yield forecasting using AVHRR NDVI time series imagery," *Remote Sens. Environ.*, vol. 99, no. 3, pp. 341–356, 2005.
[72] P. C. Doraiswamy and P. W. Cook, "Spring wheat yield assessment using NOAA AVHRR data," *Can. J. Remote Sens.*, vol. 21, no. 1, pp. 43–51.
[73] F. Maselli, S. Romanelli, L. Bottai, and G. Maracchi, "Processing of GAC NDVI data for yield forecasting in the Sahelian region," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3509–3523, 2000.
[74] M. S. Mkhabela, M. S. Mkhabela, and N. N. Mashinini, "Early maize yield forecasting in the four agro-ecological regions of Swaziland using NDVI data derived from NOAA's-AVHRR," *Agric. For. Meteorol.*, vol. 129, no. 1–2, pp. 1–9, Mar. 2005.
[75] International Irrigation Management Institute, "The world water and climate atlas for agriculture," *CGIAR News*, vol. 4, no. 2, pp. 16–17, 1997.

Mauro E. Holzman received the Ph.D. degree in geography from the Universidad Nacional del Sur, Bahía Blanca, Argentina, in 2013.

His research interests include earth observation hydrology, soil moisture and water vegetation condition, drought early warning systems, crop yield, and evapotranspiration, and also include validation process based on ground measurements from test sites located in Argentine Pampas of central Argentina.

Raúl Rivas received the M.S. degree in environmental physics and the Ph.D. degree in physics from the University of Valencia, Valencia, Spain, in 2003 and 2004, respectively.

He is currently an Independent Researcher and Member of Directorate of Comisión de Investigaciones Científicas, Buenos Aires, Argentina. His research interests include remote sensing and hydrology (more than 20 years experience), thermal infrared, surface energy balance, evapotranspiration, and monitoring networks for hydrology applied research.