Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Forecasting maize yield at field scale based on high-resolution satellite imagery

Rai A. Schwalbert [a,b,c,*], Telmo J.C. Amado [b,c], Luciana Nieto [c], Sebastian Varela [c], Geomar M. Corassa [a,c], Tiago A.N. Horbe [b], Charles W. Rice [c], Nahuel R. Peralta [d], Ignacio A. Ciampitti [c,**]

[a] Agricultural Engineering Department, Federal University of Santa Maria, Rural Science Centre, Santa Maria, RS, Brazil

[b] Soil Department, Federal University of Santa Maria, Rural Science Centre, Santa Maria, RS, Brazil

[c] Department of Agronomy, Kansas State University, 2004 Throckmorton Plant Science Center, Manhattan, KS 66506, USA

[d] Monsanto Argentina, Department of Technology and Development of Corn and Sorghum, Pergamino, Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 25 October 2017
Received in revised form
7 March 2018
Accepted 25 April 2018

Keywords:
Yield forecasting models
Maize
Satellite imagery
Yield maps
Model validation
Sentinel-2

Estimating maize (*Zea mays* L.) yields at the field level is of great interest to farmers, service dealers, and policy-makers. The main objectives of this study were to: i) provide guidelines on data selection for building yield forecasting models using Sentinel-2 imagery; ii) compare different statistical techniques and vegetation indices (VIs) during model building; and iii) perform spatial and temporal validation to see if empirical models could be applied to other regions or when models' coefficients should be updated. Data analysis was divided into four steps: i) data acquisition and preparation; ii) selection of training data; iii) building of forecasting models; and iv) spatial and temporal validation. Analysis was performed using yield data collected from 19 maize fields located in Brazil (2016 and 2017) and in the United States (2016), and normalised vegetation indices (NDVI, green NDVI and red edge NDVI) derived from Sentinel-2. Main outcomes from this study were: i) data selection impacted yield forecast model and fields with narrow yield variability and/or with skewed data distribution should be avoided; ii) models considering spatial correlation of residuals outperformed Ordinary least squares (OLS) regression; iii) red edge NDVI was most frequently retained into the model compared with the other VIs; and iv) model prediction power was more sensitive to yield data frequency distribution than to the geographical distance or years. Thus, this study provided guidelines to build more accurate maize yield forecasting models, but also established limitations for up-scaling, from farm-level to county, district, and state-scales.

* *Corresponding author*. Agricultural Engineering Department, Federal University of Santa Maria, Rural Science Centre, Santa Maria, RS, Brazil.

** *Corresponding author*. Department of Agronomy, Kansas State University, 2004 Throckmorton Plant Science Center, Manhattan, KS 66506, USA

E-mail addresses: rai.schwalbert@gmail.com (R.A. Schwalbert), ciampitti@ksu.edu (I.A. Ciampitti).

https://doi.org/10.1016/j.biosystemseng.2018.04.020
1537-5110/

1. Introduction

Precise and reliable yield forecast tools could play a fundamental role in supporting policy formulation, and decision-making process in agriculture (e.g. storage and transport) (Córdoba, Bruno, Costa, Peralta, & Balzarini, 2016; Kantanantha, Serban, & Griffin, 2010; Stone & Meinke, 2005). Historically, most models developed for yield forecasting purposes are focused to large domains (between-field variability) (DiRienzo, Fackler, & Goodwin, 2000; Doraiswamy, Moulin, Cook, & Stern, 2003; Hamar, Ferencz, Lichtenberger, Tarcsai, & Ferencz-Arkos, 1996; Lopresti, Di Bella, & Degioanni, 2015; Reeves, Zhao, & Running, 2005; Sibley, Grassini, Thomas, Cassman, & Lobell, 2014), mostly because, in the past there was limited source of data with a sufficient temporal and spatial resolution for accurate within-field crop yield estimates. Nowadays, satellite data have become more accessible (Azzari, Jain, & Lobell, 2016) with more options of high resolution imagery such as Skysat, RapidEye, and Sentinel-2 satellites, and more studies have portrayed the benefits of using high-resolution satellite imagery for identifying within-field yield variation (Azzari et al., 2016; Jin, Azzari, Burke, Aston, & Lobell, 2017; Peralta, Assefa, Du, Barden, & Ciampitti, 2016). Among the high-resolution satellites, the publically accessible Sentinel-2, a joint initiative of the European Commission (EC) and the European Space Agency (ESA), represents a great opportunity towards fine-resolution yield forecast models, since it was designed to provide systematic global acquisitions of high-resolution (10- to 20-m) multi-spectral imagery with a high revisit frequency (5 days at equator) (Drusch et al., 2012).

The potential to forecast yield using satellite information is already known and a wide set of statistical approaches have been explored. Some approaches rely on the statement that total biomass production is closely related to the fraction of photosynthetically active radiation (fAPAR) absorbed by vegetation over the course of the growing season (Monteith, 1977). Estimations of fAPAR are most often derived from VIs (Lobell, 2013), since the linear relationships between those two variables are well-known (Myneni & Williams, 1994). However, considering that most remote sensing data are not available on a daily basis, some interpolation is needed to estimate daily fAPAR.

Empirical relationships between ground-based yield measures and remote sensing data have been considered as the simplest approach to forecast yield with low computational power demanding (Hatfield, Gitelson, Schepers, & Walthall, 2008; Lobell, 2013), and have been successfully implemented in several studies with maize (Bognár et al., 2011; Bu, Sharma, Denton, & Franzen, 2017; Lobell, Thau, Seifert, Engle, & Little, 2015; Peralta et al., 2016; Shanahan et al., 2001; Sibley et al., 2014). The success of this approach is directly related to the selection of ground-truth data to build models. During the model building process the separation of data into training and validation datasets is a common practice allowing self-test model replicability irrespective of the difference between the two datasets in space or time. The selection of training data is known to have a direct impact on the model quality (Hatfield et al., 2008; Schwalbert et al., 2018) but, despite that, the majority of the published scientific literature randomly selected a subset of the data for comprising training or validation data (Assefa et al., 2016; Gholap, Ingole, Gohil, Gargade, & Attar, 2012; Gonzalez-Sanchez, 2014; Peralta et al., 2016; Sheridan, 2013) without following any guideline or statistical procedure.

Moreover, the choice of the statistical model employed to forecast yield has a large impact on the final result (Anselin, Bongiovanni, & Lowenberg-DeBoer, 2004; Peralta et al., 2016). Mostly empirical yield forecasting models based on VIs utilise classical ordinary least squares (OLS)-based on simple or multiple regression techniques (Noureldin, Aboelghar, Saudy, & Ali, 2013; Rembold, Atzberger, Savin, & Rojas, 2013; Shanahan et al., 2001), without properly accounting for the spatial autocorrelation structure amongst these variables (Imran, Zurita-Milla, & Stein, 2013; Peralta et al., 2016). The latter situation can lead to problems with inflated variance and likely resulting in wrong conclusions (Anselin et al., 2004; Bongiovanni, Robledo, & Lambert, 2007).

Models derived from simple empirical relationships usually tend to be time- and space-limited, valid only under similar conditions as when the correlation was established (Hatfield et al., 2008; Lobell, 2013; Tucker, 1979). Currently, the potential to forecast yield using satellite information through empirical models is already known, but the challenge is to extend these tools beyond the environment where the study was done (Hatfield et al., 2008). Lastly, the selection of adequate VIs is also an important step for model development (Peralta et al., 2016). The normalised difference vegetation index (NDVI) (Rouse, Haas, & Schell, 1973) is one the most widely used VIs to assess crop growth and yield (Peralta et al., 2016; Raun, Solie, & Johnson, 2002; Rembold et al., 2013; Solie, Dean Monroe, Raun, & Stone, 2012), and it becomes as a benchmark for researchers developing new VIs (Hatfield et al., 2008). However, there are some constraints related to saturation in medium to high leaf area index (LAI) values with NDVI (Haboudane, Miller, Pattey, Zarco-Tejada, & Strachan, 2004; Nguy-Robertson et al., 2012; Tucker, 1979). Thus, the incorporation of other VIs that still have sensitivity in high LAI values such as green NDVI (NDVIG) (Gitelson, Kaufman, & Merzlyak, 1996) and red-edge NDVI (NDVIre) (Gitelson & Merzlyak, 1994) have been reported improving empirical models (Hatfield et al., 2008; Peralta et al., 2016).

Following this rationale, guidelines for implementing yield forecasting models derived from empirical relationships and for validating their spatio-temporal relevancy still remain unknown. Thus, the objectives of this study were to: i) identify parameters to guide data selection aiming at building yield forecasting models using Sentinel-2 satellite imagery; ii) compare different approaches (OLS vs. spatial correlation) and different VIs during the model building process; iii) perform spatial and temporal model validation using independent datasets to identify potential limitations in up-scaling yield forecasting models. The main hypothesis is that model predictability power increases as the yield frequency distribution of the training data becomes more similar to the validation data even when considering diverse spatio-temporal scales (geography, time, or years).

2. Materials and methods

The analysis was performed on end-season yield monitor data and mid-season Sentinel-2 images, collected during the critical period for maize yield determination (approximately 20 days before and 20 after flowering) (Johnson & Mueller, 2010; Peralta et al., 2016; Sakamoto, Gitelson, & Arkebauer, 2014). Sentinel satellite imagery of selected maize fields located in Brazil (BR) (Fig. 1A and B) and US (Fig. 1C) was utilised for the analysis. Six fields from Rio Grande do Sul (RS) state (2016/2017 season) and seven fields from Mato Grosso (MT) state (five from 2016 season and two from 2017 season) were selected for comprising the BR database. The field size ranged from 20 to 130 ha. It is important to mention that for MT, fields were selected from the second season (mainly cultivated after the soybean) since the first season is harvested around February. Usually during the second season in MT maize yield is lower compare to RS due less favourable weather conditions. In RS, average temperature (considering the last 20 years) during the growing season was 20.4 °C with a cumulative precipitation of 1080 mm, while in MT, average temperature during the second season was 23.7 °C with a cumulative precipitation of 700 mm.

The US database comprised six fields (2016 season), all located in the state of Kansas (KS). This database was only considered as validation data in the last step (spatial validation) where models previously built were used to forecast maize yield in Kansas fields, in order to test our main hypothesis. Information related to harvest date, satellite imagery collection data, and specific coordinates (latitude, longitude) for each field were recorded (Table 1). Most of the BR fields were utilised for training purposes, comprising the training database. Fertilizer application rates, crop management, and tillage practices varied between fields.

This study was divided into four major steps to achieve the proposed objectives (Fig. 2). The four steps were related to: 1) data acquisition and preparation, 2) selection of training data and verification, 3) building yield forecasting models, and 4) spatial and temporal validation (including fields from different growing seasons and geographies).

2.1. *Data acquisition and preparation (Step 1)*

The primary objective of this step was to establish criteria for selecting adequate quality of yield monitor (calibrated) and satellite imagery data. Yield data was submitted to a filter process in order to remove outliers and inliers. In this research, outliers were considered as values out of the mean ± 3 standard deviations (SD) range. According to Chebyshev's theorem (Amidan, Ferryman, & Cooley, 2005), it is

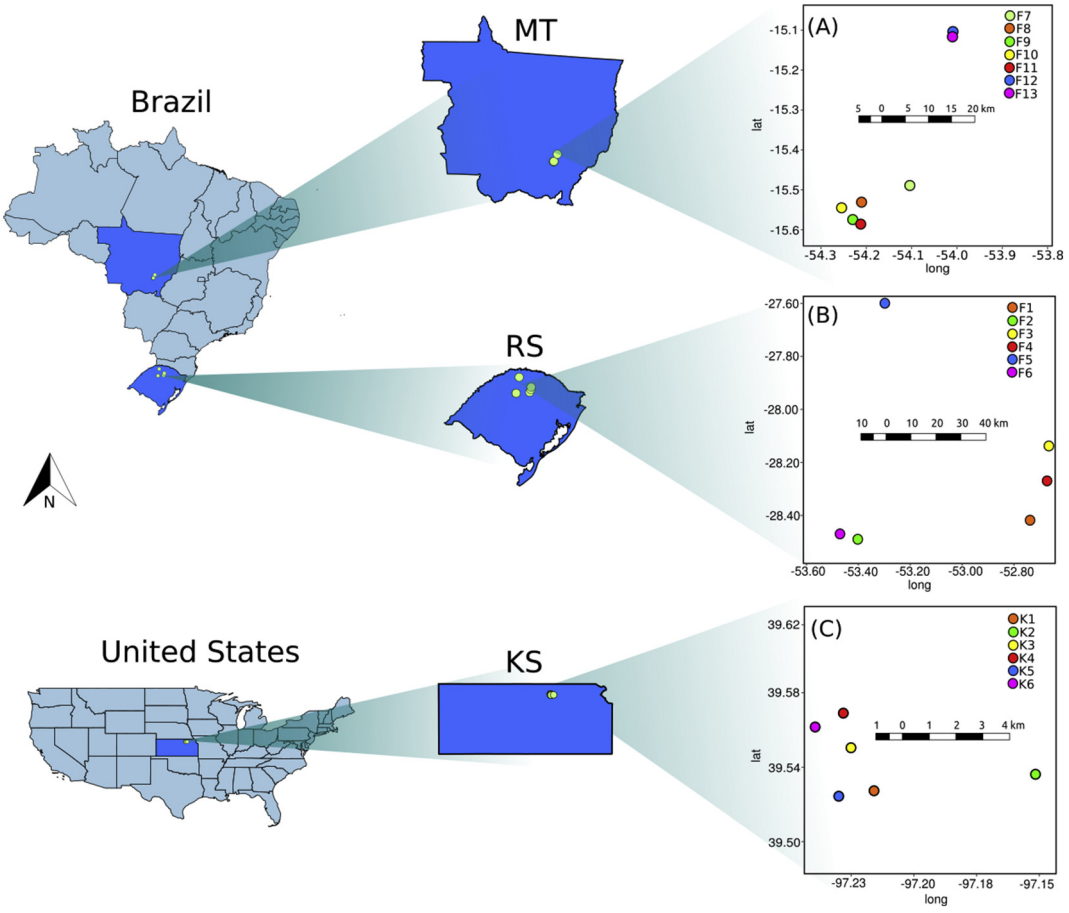

Fig. 1 – Field research studies located in Mato Grosso (MT) state (A), Rio Grande do Sul state (RS) (B) both in Brazil, and the state of Kansas (KS) (C) from US. Circles represent the precise geo-position of the fields within each region and country. Scales bars are in different scales for panels A, B, and C.

Table 1 – Descriptive information of maize yield and satellite data: state, season, geographical position, harvest date and imagery acquisition date.

Field	State	Season	Data	Latitude[a]	Longitude[a]	Harvest date	Imagery date
F1	RS	2016–2017	V	−28.48	−52.78	02/16/2017	11/29/2016
F2	RS	2016–2017	V	−28.53	−53.54	02/21/2017	11/29/2016
F3	RS	2016–2017	V	−28.18	−52.69	02/14/2017	11/29/2016
F4	RS	2016–2017	T	−28.32	−52.71	02/27/2017	11/29/2016
F5	RS	2016–2017	V	−27.62	−53.36	02/18/2017	11/29/2016
F6	RS	2016–2017	T	−28.53	−53.56	02/17/2017	11/29/2016
F7	MT	2016	V	−15.47	−54.01	07/02/2016	04/29/2016
F8	MT	2016	T	−15.57	−54.15	07/06/2016	04/29/2016
F9	MT	2016	V	−15.57	−54.16	07/05/2016	04/29/2016
F10	MT	2016	V	−15.56	−54.17	07/05/2016	04/29/2016
F11	MT	2016	T	−15.58	−54.15	06/30/2016	04/29/2016
F12	MT	2017	V	−15.15	−53.94	06/30/2017	04/24/2017
F13	MT	2017	V	−15.15	−53.94	06/29/2017	04/24/2017
K1	KS	2016	V	39.53	−97.21	09/27/2016	06/20/2016
K2	KS	2016	V	39.54	−97.15	10/01/2016	06/20/2016
K3	KS	2016	V	39.55	−97.22	10/03/2016	06/20/2016
K4	KS	2016	V	39.57	−97.23	09/30/2016	06/20/2016
K5	KS	2016	V	39.53	−97.23	09/22/2016	06/20/2016
K6	KS	2016	V	39.56	−97.24	09/29/2016	06/20/2016

T = Training Database; V = Validation Database.
[a] Decimal coordinates – WGS 84. RS = Rio Grande do Sul. MT = Mato Grosso. KS = Kansas.

inferred that a minimum of 89% of the data is within the mean ± 3 SD, regardless the data distribution. Inliers are data that differ significantly from their neighbourhood but lie within the general range of variation of the data set (Córdoba et al., 2016). Spatial autocorrelation Moran's local index (Ii) (Anselin, 1995) was used to identifying inliers. The Ii is basically applied individually to each neighbourhood and shows the degree of similarity between an observation and its neighbours. In summary, *localmoran* function of the "spdep" R package (Bivand & Piras, 2015) was used to identify inliers. Moreover, the *moran.plot* function was implemented to calculate Ii and perform the Moran scatter plot to identify additional inliers. Further details on this procedure can be found in Córdoba et al. (2016). Lastly, spatial interpolation was performed to estimate maize yield values for areas where yield was not sampled. This procedure was required, even considering that yield monitor data was recorded in a high density (5 × 10 m), because after filtering yield density data was significantly decreased. Aiming at getting similar arrangement for all datasets, equivalent satellite imagery grid structure was used (10 × 10 m). Geostatistical interpolations involving semivariogram adjustment and ordinary kriging were performed, done individually for each field, using R packages "geoR" (Ribeiro & Diggle, 2016) and "gstat" (Pebesma, 2004).

Sentinel-2 images are composed by 10 bands with resolution between 10 × 10 m and 20 × 20 m, in the visible, near infrared, and short-wave infrared part of the spectrum. For this study only 4 bands: 3 (green), 4 (red), 8 (near-infrared), and 8a (red-edge 4) were used, in agreement with the scientific literature in the topic of forecasting crop yields using satellite data, highlighting the importance of those bands (wavelengths) (Bu et al., 2017; DiRienzo et al., 2000; Doraiswamy et al., 2003; Hamar et al., 1996; Lobell et al., 2015; Lopresti et al., 2015; Peralta et al., 2016; Reeves et al., 2005; J. Shanahan et al., 2001; Sibley et al., 2014). The selected bands were employed used to calculate 3 diverse VIs: NDVI, NDVIG, and NDVIre. The selection of these VIs was based on previous researches investigations showing the efficiency of these VIs to forecast final maize yield (Bognár et al., 2011; Bu et al., 2017; Peralta et al., 2016; Shanahan et al., 2001). Sentinel-2 images were collected in an interval between 20 days before flowering and 20 days after flowering, depending on the availability of the image and the cloud interference (Table 1). The red-edge band was resized to 10 m pixel size. Atmospheric correction was performed using the semi-automatic classification plugin in QGIS 2.18 (Congedo, 2016) in order to obtain surface reflectance without the interference of atmospheric gases. The VIs, including NDVI, NDVIre, and NDVIG, were generated using a combination of visible, near-infrared and red-edge bands.

2.2. *Selection of training data (Step 2)*

As previously detailed, only selected BR fields (RS and MT) were used as training data with KS fields left out for the validation step. All fields were randomly sampled (bootstrap with replacement) to generate equal size of data points per field, 800 per field. Since one of the objectives of the paper was to provide guidelines for training data selection, three different methods were evaluated in this study. Steps 2 and 3 for the theoretical framework (Fig. 2) were repeated in a retroactive process for each one the strategies, individually for each BR (MT and RS) dataset. The three data selection strategies tested were: i) selection of two fields with broad yield range and with the data concentrated (more than 50% of the values) between the first and third quartiles of the overall yield frequency distribution, herein termed as the "similar yield distribution"; ii) selection of fields with the smallest, medium, and greatest

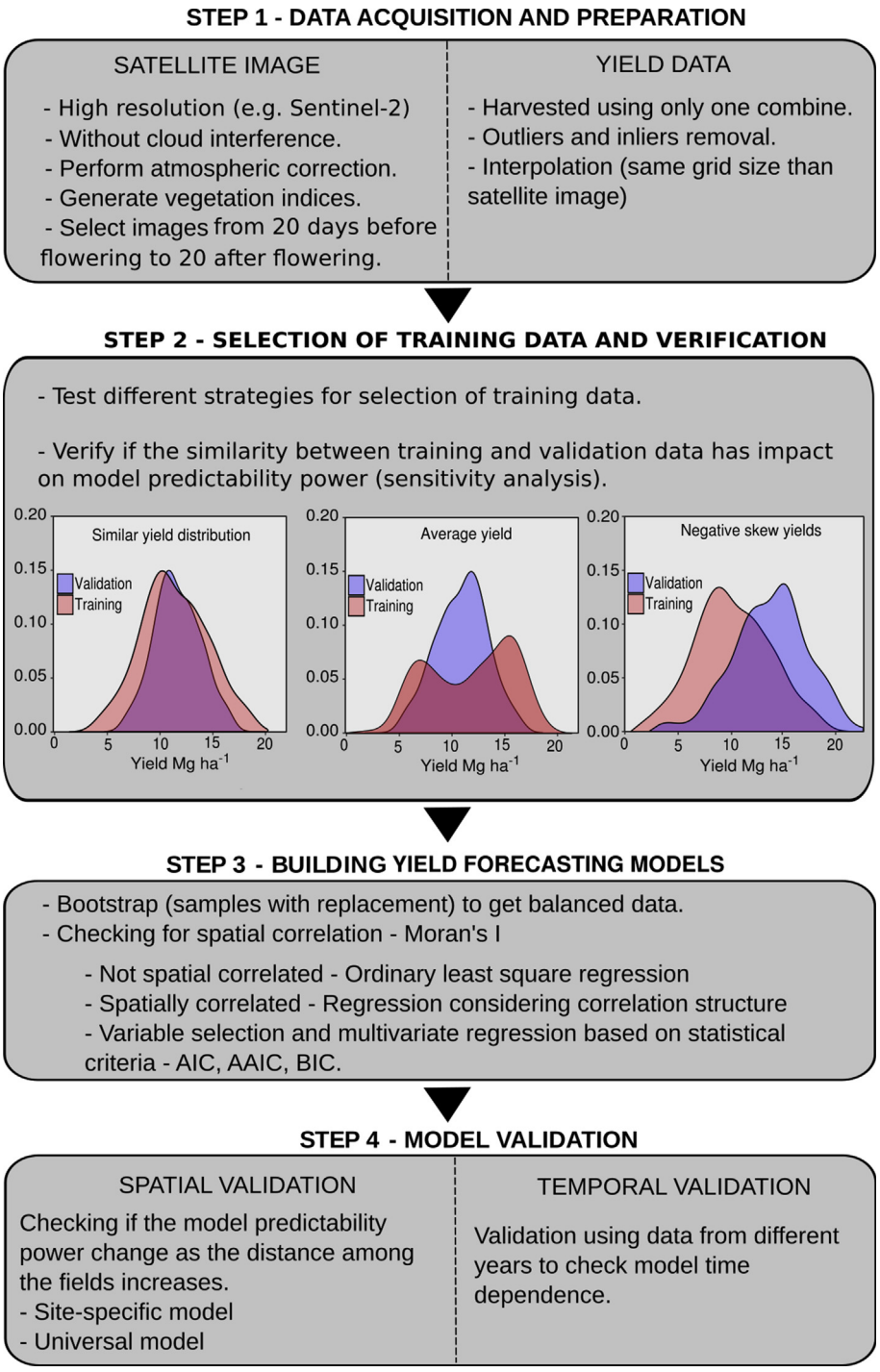

Fig. 2 – Theoretical framework indicating all steps of the analysis: step 1- data acquisition and preparation, step 2- selection of training data and verification, step 3- building yield forecasting models, and step 4- model validation.

average yield among all fields, three field total per region, herein termed as the "average yield", and iii) selection of two fields with the lowest average yields among all fields (left shifted fields in relation to the overall distribution), herein termed as the "negative skew yields" (Supplementary table 1). For each one of the strategies, the remained fields were considered as the verification data. For example, for MT in the strategy "similar yield distribution", two fields were selected and the remaining five fields were utilised for the verification of the developed model.

The similarity of training and verification data distribution was compared using four statistic parameters, mode, inter-quartile Range (IQR) position (range between the first and the third quartiles), skewness, and kurtosis. To compare the statistic parameters a 95% bootstrap percentile confidence interval (CI) (Efron & Tibshirani, 1994) was calculated using the

"boot" package in R (Canty & Ripley, 2017), obtaining a total of 1000 bootstrap replicates to estimate the variability.

2.3. Building yield forecasting models (Step 3)

As aforementioned, this step occurred in parallel to the data selection (Step 2), yield forecasting models were built utilizing the three data selection strategies for training data. As an initial phase, spatial autocorrelation analysis was conducted for yield and VIs (NDVI, NDVIG, and NDVIre) in each field using Moran's test. Moran's I statistic measures the strength of spatial autocorrelation in a response among nearby locations in space as a function of cross-products of the neighbouring weighted deviations from the mean. Moran's I coefficient values near 1 indicate positive and −1 negative autocorrelation. Coefficient near 0 refers to lack of spatial autocorrelation. In this step two approaches were followed for model evaluation between yield and VI relationship: i) implementing linear regression model using the ordinary least squares method, herein termed as "OLS" model, assuming that the errors were independent and identically distributed (i.i.d.), and ii) implementing linear regression considering the spatial structure of the errors (with Gaussian, spherical and exponential spatial correlation of plotted errors) to account for possible violations of the i.i.d. assumption. The last were adjusted using the *gls* function of the "nlme" R package (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2017).

For model selection, stepwise-regression procedure was implemented to determine the variables (VIs) that significantly contributed to yield prediction models. Stepwise forward was implemented using the function *stepAIC* of the "MASS" package (Venables & Ripley, 2002) from the R software. Statistical model comparison was performed using statistical criteria proposed by Akaike (AIC) (Johnson & Omland, 2004) and the coefficient of determination (R^2). Multicollinearity of the remaining bands was also evaluated by computing the variance inflation factor (VIF). A threshold VIF value was established (Zuur, Ieno, & Elphick, 2010) and a VI with VIF greater than 2 was removed from the model. The standardised coefficient was determined using the "lm.beta" R package to check the weight of each VI into the model.

After running all the rounds for Steps 2 and 3 (Fig. 2), statistical evaluation on similarities between training and validation data distribution according to the parameters tested (mode, quartiles, skewness and kurtosis) and model accuracy assessing using Root-mean square error (RMSE) (observed vs predicted yield) were implemented.

As a last phase for this step, two categories of yield forecasting models were built: i) universal model, including both RS and MT training data and ii) site-specific models, obtaining one specific-model per state/region (one for RS and one for MT regions) evaluated.

2.4. Spatial and temporal validation data (Step 4)

After the selection of training data, model development and verification, a validation was performed aiming at verifying spatial and temporal dependency on the models. For testing the first one ("spatial validation"), Kansas database was utilised as validation data. All the six sets of training data (three from RS and three from MT) were tested. The same approach discussed in the previous sections was applied. Yield frequency distribution of all the training data from RS and MT were compared with KS yield frequency distribution. After studying all yield frequency distributions, the most proper model was selected to forecast yield of the KS database (US), comprising six fields. The last validation ("temporal validation") was performed using new MT fields (2017) since only in MT there was data available from two different seasons (2016 and 2017). Basically, yield forecasting model built using data from 2016 was used to estimate 2017 yields. The accuracy of estimation and model fitting was evaluated using the RMSE. In addition, spatial predictions from each model were visually compared with geostatistical interpolation of yield (yield maps).

3. Results

3.1. Selection of training data

Different yield frequency distribution was documented for RS and MT. For RS, average maize yield was 12.7 Mg ha^{-1}, with 50% of the data (IQR) ranging from 10.6 to 14.8 Mg ha^{-1} and with a mode of 12.9 Mg ha^{-1} (Fig. 3A). For MT, average maize yield was 5.5 Mg ha^{-1}, with IQR ranging from 4.5 to 6.4 Mg ha^{-1} and with a mode of 5.7 Mg ha^{-1} (Fig. 3F). In both states, yield frequency distribution was not considered normal according to Shapiro–Wilk test ($P < 0.05$). Furthermore, large within- and between-field variability was documented (Fig. 3B and G). For RS, field 1 was the most productive with a yield average of 14.9 Mg ha^{-1} and with a variation range from 9.2 to 20 Mg ha^{-1}, while field 5 was the least productive with a yield average of 10.3 Mg ha^{-1} and with a variation from 5.0 to 15.2 Mg ha^{-1}. For MT, field 10 was the most productive field with a yield average of 7 Mg ha^{-1} with a variation range from 4.1 to 8.6 Mg ha^{-1}, while field 9 was the least productive with a yield average of 4.2 Mg ha^{-1} and with values ranging from 2.9 to 5.9 Mg ha^{-1}.

The field selection to comprise the training data affected the model quality and, consequently, the predictability power of the model. For RS, three different sets of fields were tested as training data: fields 4 and 6 – similar yield distribution – (Fig. 3C), fields 1, 3 and 4 – average yield – (Fig. 3D) and fields 2 and 4 – negative skew – (Fig. 3E) (Supplementary table 2). The RMSE decreased as the yield frequency distribution of the training becomes more alike to the validation data. When fields 2 and 4 comprised the training data, "negative skew" strategy, the greatest RMSE (1.97 Mg ha^{-1}) was documented (Supplementary table 2). For the "similar yield distribution" strategy, training and validation data were more alike sharing comparable IQR ($P > 0.05$) with a slightly different mode (12.4 vs. 13.0 Mg ha^{-1}) ($P < 0.05$) and RMSE of 1.50 Mg ha^{-1}. The lowest RMSE, 1.48 Mg ha^{-1}, was reported when the number of fields increased from 2 to 3, "average yield" strategy. After this process, training and verification yield frequency distribution resulted in comparable mode ($P > 0.05$) and IQR ($P > 0.05$). Since the selection of one additional field just increased slightly the RMSE (from 1.48 to 1.50 Mg ha^{-1}), only fields 4 and 6 were chosen for posterior analysis – "similar yield distribution" strategy, leaving one more field available for model validation.

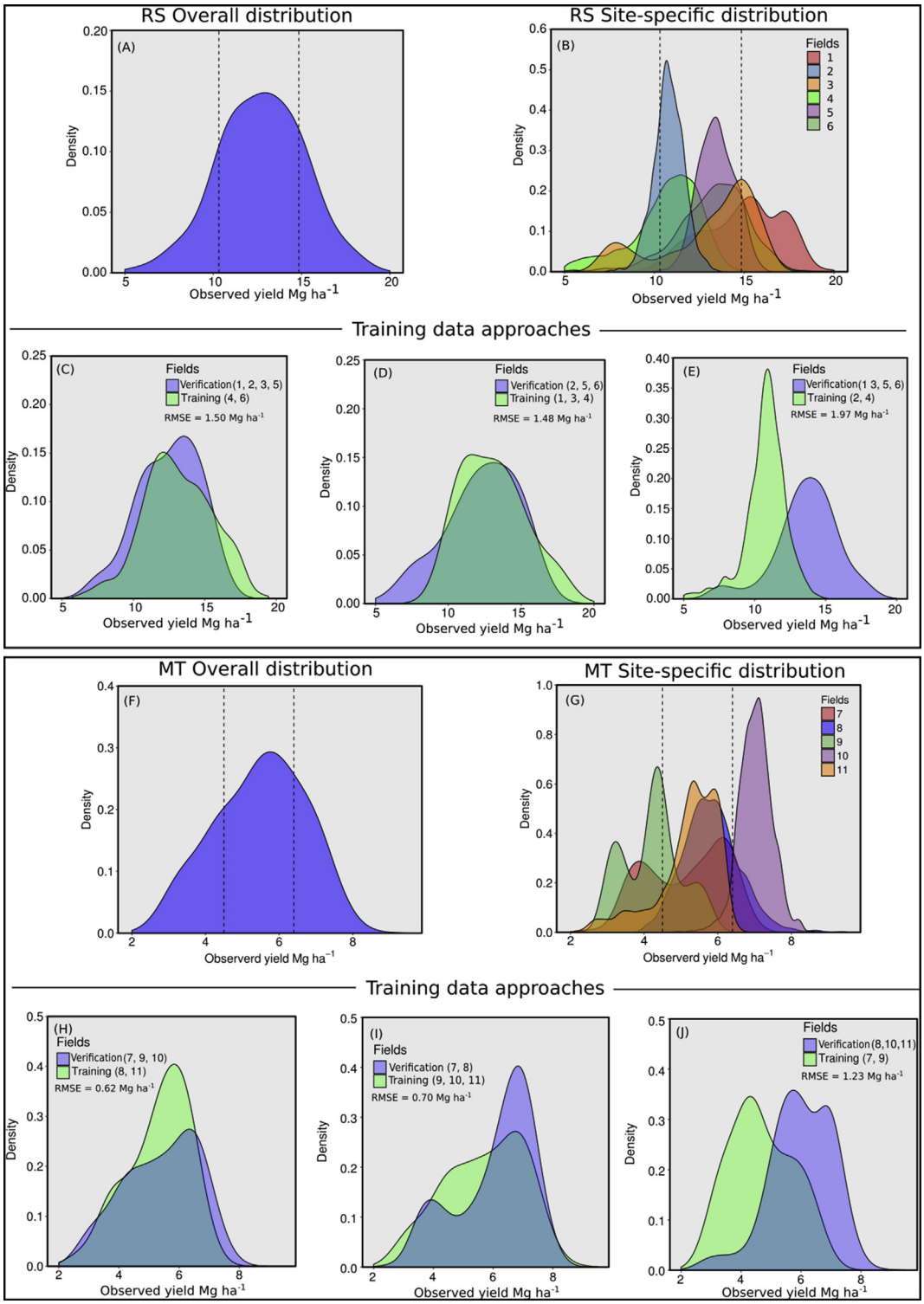

Fig. 3 – Maize yield frequency distributions for RS (A–E) and MT (F–G) fields. (A and F) Overall yield frequency distribution in RS and MT respectively. (B and G) Field level yield frequency distribution in RS and MT respectively. (C, D, E) Training and verification yield frequency distribution for different training data selection strategies in RS. (H, I, J) Training and verification yield frequency distribution for different training data selection strategies in MT. Root mean square error (RMSE) reported in each panel was obtained from the observed and predicted yields using each set of training and verification data.

Likewise, the same criteria aforementioned was applied to MT fields, the "negative skew" strategy of selecting training data shifted fields resulted in significantly different modes, first and third quartiles ($P < 0.05$) and the greatest RMSE (1.23 Mg ha^{-1}) (Fig. 3J). For MT, the "average yield" strategy did not result in the smallest RMSE (0.70 Mg ha^{-1}). This strategy led to statistically equal modes ($P > 0.05$), but different first and third quartile positions ($P < 0.05$) (Fig. 3I). The selection of fields 8 and 11, "similar yield distribution" data training strategy, resulted in non-differences between training and verification modes and IQR ($P > 0.05$) obtaining a RMSE of 0.62 Mg ha^{-1} (Fig. 3H). Following the rationale for field selection for RS, fields 8 and 11 were chosen for posterior analysis. No pattern was observed for average, skewness and kurtosis linking yield data distribution similarities and RMSE for the models from RS and MT (Supplementary table 2).

3.2. Building yield forecasting models

Spatial autocorrelation analysis conducted using Moran's I test (MI) on VIs and yield data are presented in Supplementary Table 3. In general, autocorrelation (Moran's I test) for all variables was positive and statistically significant (exception for F8) indicating that when yield or VI values are geographically in shorter distances are more alike, diminishing the spatial correlation as distance increases. The absence of spatial correlation in F8 was probably due to higher within-field yield homogeneity compared to the other fields.

Following the same rationale, yield forecasting models increase predictability power when a spatial correlation structure was considered. The spatial regression models outperformed the OLS once the AIC values were smaller for the spatial models compared to the OLS ones (Table 2). It indicated that there was a good trade-off between the goodness of fit and the complexity of the model. For the RS model, residuals were assumed following a Gaussian spatial correlation structure, while for the MT and the universal (both RS + MT) models the exponential correlation structure presented the best fit to describe the data (Table 2).

All VIs were kept in the MT and universal models after the stepwise selection, while for RS, only the NDVIre was retained (Table 2). The NDVIre presented the greatest weight for all models. Even some degree of multicollinearity among the indices was expected since near-infrared (NIR) band was a component of all of them, the VIFs were less than 2 for the VIs that remained in the model.

3.3. Spatial and temporal validation of models

In the first step of the spatial validation, the universal model was compared to the site-specific models (state-scale models). The predictability power of the universal model was drastically reduced both for within- (data not shown) and between-field variability (Fig. 4A) compared to the site-specific models (Fig. 4B). The universal model slightly overestimated yield for the MT fields (low productivity) and underestimated yield for the RS fields (high productivity). Site-specific models resulted in RMSE of 1.50 Mg ha^{-1} for RS and 0.62 Mg ha^{-1} for MT.

In the second step of the spatial validation, the RS model was used to forecast yield of one additional dataset comprised of six fields located in KS (US). The RS model was chosen for this purpose since the yield frequency distribution of the RS training data was the closest to the one for KS fields (Fig. 5A). Despite the similarity in yield frequency distribution for KS and RS, differences in mode and IQR ($P > 0.05$) were documented. The RS model presented a good predictability in low productive areas and tended to overestimate yield in high productive zones, resulting in a RMSE of 2.22 Mg ha^{-1} (Fig. 5B).

For the temporal validation, the MT model built with the 2016 data was used to forecast yield for independent fields harvested in 2017. Yield distribution frequency between MT training data (2016) and MT yield data from 2017 was similar with statistically equal mode and IQR ($P > 0.05$) (Fig. 6A). The MT model presented a good predictability power predicting within-field variability of 2017 fields, with a RMSE of 0.95 Mg ha^{-1} (Fig. 6B). Historical weather data showed that the 2016 and 2017 growing seasons were similar, with temperatures slightly above and total precipitation slightly below the

Table 2 – Multiple linear regression models for the ordinary least-square (OLS) and regression considering spatial correlation including the vegetation indices (VIs) obtained from mid-season satellite imagery as predictors of the end-season yield monitor data. Equations are related to model with the lowest AIC. SRE = spatial regression considering exponential correlation of the plotted errors. SRG = spatial regression considering Gaussian correlation of the plotted errors. SRS = spatial regression considering spherical correlation of the plotted errors.

Data	Model	AIC	Equation
RS	OLS	3771	Yield (Mg ha^{-1}) = 2.7*** + 69.88*** (NDVIre) (R^2=0.68)
	SRE	3762	
	SRG	3759	
	SRS	3770	
MT	OLS	3087	Yield (Mg ha^{-1}) = 15.3*** + 81.6*** (NDVIre) − 8.8*** (NDVIG) − 20.3 (NDVI)*** (R^2=0.59)
	SRE	891	
	SRG	1986	
	SRS	894	
Universal	OLS	9985	Yield (Mg ha^{-1}) = −25.6*** − 46.5 (NDVIre)*** + 145.1 (NDVIG)*** − 67.5 (NDVI)*** (R^2=0.32)
	SRE	6750	
	SRG	8959	
	SRS	6836	

Notes: The statistically significant coefficients are indicated by asterisks, where * indicates $P < 0.05$; ** indicates $P < 0.01$; and *** indicates $P < 0.001$. Parameters with no asterisks are therefore not significant at the 0.05 level.

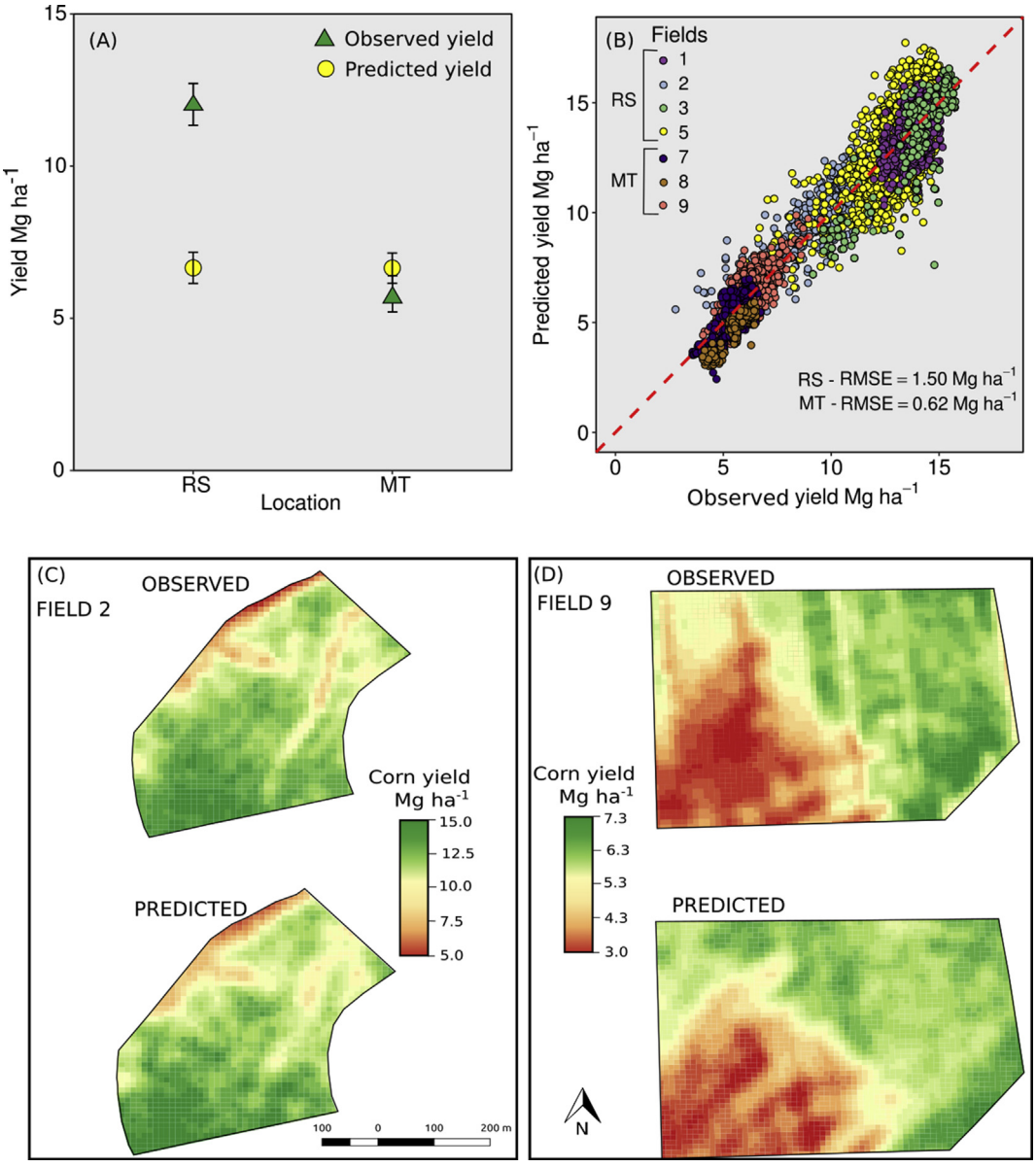

Fig. 4 – Estimated versus observed maize yield. (A) State-level yield prediction using the Universal yield forecasting model. (B) Within-field yield variability prediction using site-specific maize yield forecasting models. A red dashed line is presented in panel B portraying the 1:1 line for the estimated–observed relationship. (C) Observed yield map versus predicted yield map generated based on a site-specific model for RS. (D) Observed yield map versus predicted yield map generated based on a site-specific model for MT. RMSE = Root-mean square error. RS = Rio Grande do Sul. MT = Mato Grosso.

average of the last 17 years (period from 1st January to 31st June) (Fig. 6C).

4. Discussion

4.1. *Building yield forecasting models*

Processes to build empirical models usually involve two steps; construction (training) and validation (Becker-Reshef, Vermote, Lindeman, & Justice, 2010; Hatfield et al., 2008; Peralta et al., 2016). The selection of training and validation data is usually done randomly (Assefa et al., 2016; Lopresti et al., 2015; Peralta et al., 2016), but the selection of training data can directly affect model predictability power (Schwalbert et al., 2018; Sheridan, 2013). The first outcome of this study was related to the similarity between training and validation data and predictability power of the model. Statistical parameters such as mode, first and third quartiles, were implemented to test the similarity between datasets. The selection of fields with a broad level of variability and reduced skewness increased the likelihood of obtaining more representative models. Fields with a high degree of uniformity, narrow variation, in yield are not expected to add useful

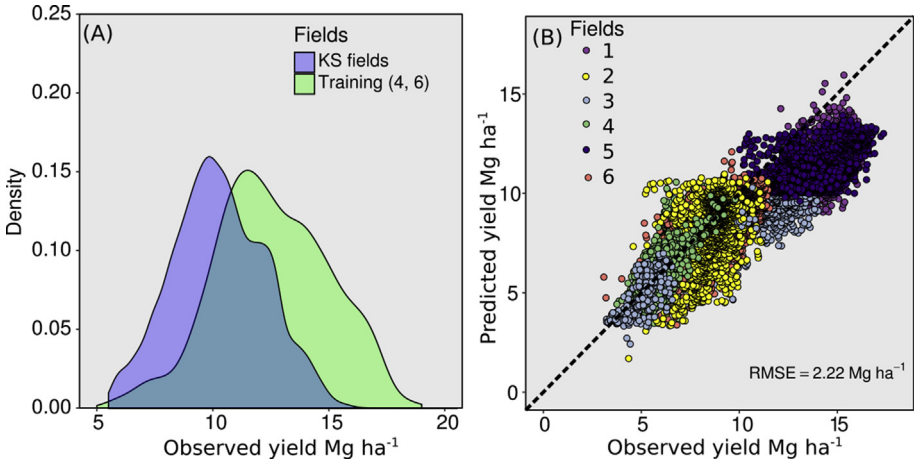

Fig. 5 – (A) Yield frequency distribution for RS (training data – Fields 4 and 6) and for KS fields and (B) Predicted (estimated via RS yield forecasting model) versus KS observed maize yield (end-season yield monitor data). A dashed black line portrays the 1:1 line for the predicted–observed yield relationship. RMSE = Root-mean square error. RS = Rio Grande do Sul. KS = Kansas.

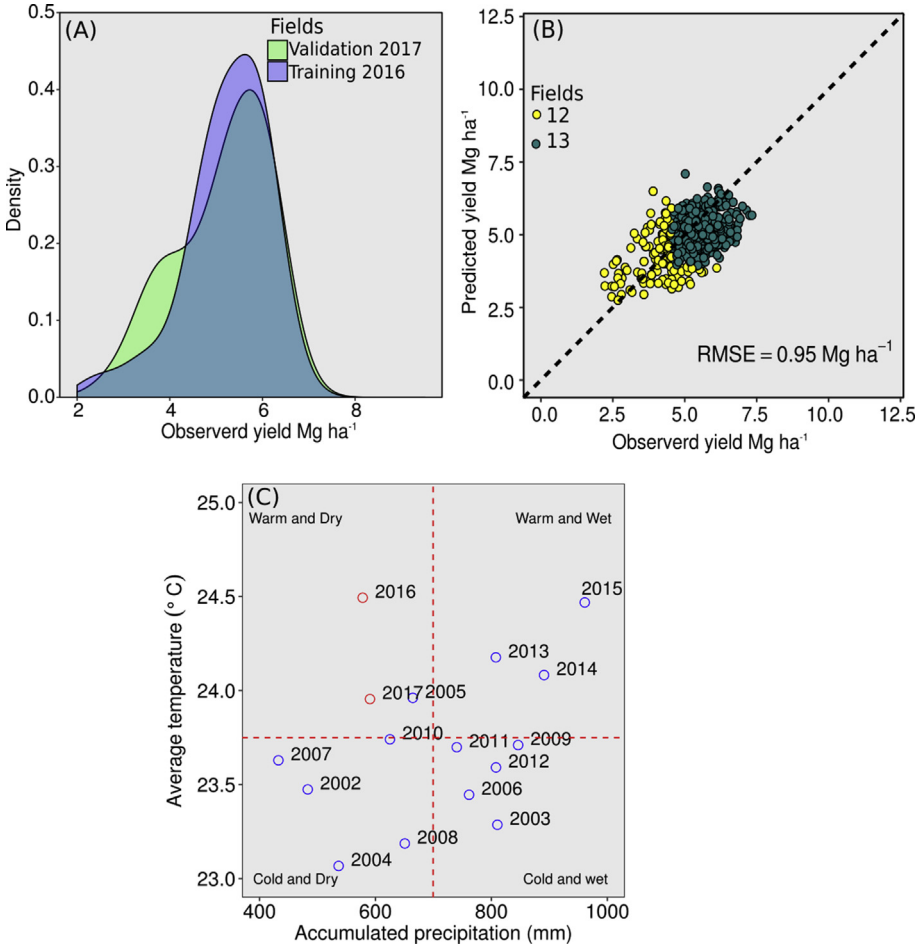

Fig. 6 – (A) Yield frequency distribution for RS (training data – Fields 4 and 6) and for KS (B) Estimated (predicted via RS yield forecasting model) versus KS observed maize yield (end-season yield monitor data). A dashed black line is presented in panel portraying the 1:1 line for the estimated–observed relationship. (C) Average temperature and accumulated precipitation from last 17 years (period from 1st January to 31st June). A dashed red line represents the average from the entire period. RMSE = Root-mean square error. RS = Rio Grande do Sul. MT = Mato Grosso. KS = Kansas.

information to the yield-VI model (Peralta et al., 2016). Fields negatively or positively skewed for yield could bias the model. Negative skewed fields could have a yield-VI relationship affected by biotic or abiotic stress condition after image acquisition (Sadras & Calviño, 2001), while positive skewed fields (with yields towards high values) could face problems related to saturation of VIs, such as NDVI (Hatfield et al., 2008). This study tested statistical parameters related to data distribution (mean, mode, first and third quantile positions, skewness and kurtosis) as potential indicators of similarities in yield frequency, providing guidelines for data selection in building forecast models. Mode and quartile positions were the most suitable parameters driving the selection for training and validation datasets that minimise the RMSE (Supplementary table 2). Similar results were reported by Schwalbert et al. (2018) in a study involving maize yield response to plant density and fertilizer N rates. In summary, this study also presents a novel approach for the data selection process for the training data based on studying yield data distribution.

Additionally, the approach used to build the yield forecasting models as well as the selection of the VIs influenced model predictability. The approach considering spatial correlation of the regression residuals outperformed the method considering the i.i.d assumption. This result, even when expected, since the positive spatial correlation for yield and for VI is already well-known (Bakhsh, Jaynes, Colvin, & Kanwar, 2000; Bresler, Dasberg, Russo, & Dagan, 1981; Jaynes & Colvin, 1997; Morkoc, Blggar, Mlllar, & Nlelsen, 1982; Peralta et al., 2016; Timlin, Pachepsky, Snyder, & Bryant, 1998), suggests that spatial correlation of regression residuals should be accounted for when building yield forecast models (Anselin et al., 2004; DiRienzo et al., 2000; Leiser, Rattunde, Piepho, & Parzies, 2012; Peralta et al., 2016). Until the present time, there are only a few studies showing the benefits of spatial adjustment to models predicting yield from imagery data (Imran et al., 2013; Peralta et al., 2016). Regarding the performance of the VIs as explanatory variables, NDVIre presented the largest weight in the regression, being the most retained index. Recently, Peralta et al. (2016) also reported that NDVIre was more effective to predict yields relative to NDVI and NDVIG. The explanation for this is that the NDVIre is less influenced by changes in leaf area, avoiding saturation issues at medium to high LAI and yield. It is imperative also to mention that for the MT and universal models, NDVIG and NDVI were also retained, reflecting the potential of these indices for predicting yield variation and for fine-tuning the proposed yield forecasting model.

4.2. *Spatial and temporal validation of models*

Empirical models are frequently reported as an efficient tool to forecast cereal yield, and variations in VI can account for more than 80% of the observed within-field yield variation (Shanahan et al., 2001; Wiegand & Richardson, 1990). Despite the high capacity to explain yield variability, even within-field, empirical models are known to be regionally specific (Becker-Reshef et al., 2010; Doraiswamy et al., 2003; Hatfield et al., 2008; Moriondo, Maselli, & Bindi, 2007). Similar constraint was documented in this study since the universal model was not even suitable to forecast yield variations in a state-scale. When yield forecasting models were applied individually for MT or RS, the predictability power increased substantially. The overall yields were lower in MT (second season) than in RS, with maize in MT more adversely affected by abiotic stresses (Minuzzi & Lopes, 2015). When satellite imagery was obtained prior to flowering, abiotic stress in those fields could severely affect final yield (Sadras & Calviño, 2001) and consequently modify the yield-VI relationship. Truly, the model is forecasting the potential yield at the flowering time; thus, under- or over-estimation are probably depending on the conditions during the reproductive growth. Furthermore, as with any purely empirical approach, extrapolation of equations to new locations or years can be problematic (Hatfield et al., 2008; Lobell, 2013; Lopresti et al., 2015; Moriondo et al., 2007). For this study, the yield frequency distribution of RS and KS fields were quite similar resulting in reasonable yield predictability despite a loss in sensitivity to explain within-field yield variability, highlighted by the increase in RMSE in relation to the forecast for the RS fields. Another example when empirical models could overcome the spatial constraint is the study developed by Becker-Reshef et al. (2010), where models developed in KS were successfully applied to forecasting wheat yields in Ukraine. In the same way as the distance in space (geographic distance), distance in time (years) is also expected to decrease model predictability (Bognár et al., 2011). However, in our study, weather conditions led to similar growth environments resulting in comparable yield data frequency distributions for 2016 and 2017 seasons (Fig. 6C); thus, resulting in comparable model predictability. Despite that, the temporal analysis should be cautiously evaluated since it comprises one year and a specific region around the globe. Further testing including more years and other regions presenting comparable weather conditions should be pursued to validate this point.

This study showed that the selection of the fields for comprising the training data directly affected the model structure. Historical yield information is available in platforms such as National Agricultural Statistics Service (NASS), and once knowing the overall yield frequency distribution from a specific region, fields representative to that region can be selected to scale-up the yield forecasting models to county-, district-, and state-scales. One of the main drawbacks of remote sensing based empirical models for estimating yields has been that their application is valid only for the areas those have been calibrated for (Doraiswamy et al., 2003; Hatfield et al., 2008; Lobell, 2013). By means of the current outcomes, it can be implied that independent datasets could portray in similar yield-VI relationship if the following criteria are fulfilled for the yield data distribution: i) IQR and ii) mode statistically similar, and for the satellite imagery: iii) collected at a similar growth stage, even with fields separated by space or time. The latter could provide a foundational knowledge to establish conditions (regions in space and year characteristics) where empirical models could be suitable, and when a new model should be developed. Furthermore, this study presents guidelines for

applicability of yield forecast models where ground-truth data is limited or scarce, providing fundamental information for supporting policy formulation and helping farmers, consumers, and researchers for making informed decisions based on the crop yield forecast report.

5. Conclusions

The likelihood that two independent datasets portray similar yield-VI relationship increases as their yield data distribution becomes more alike, mainly related to the mode and the IQR. In this current study, model performance was more affected by differences in the yield frequency distribution rather than by distance in space (BR and KS) or time within a region (2016 and 2017 seasons). Since RS and MT presented a large difference in yield frequency distribution, the universal model to estimate maize yield in both states presented small predictability power compared to the site-specific models (individual model per state).

The regression model using the NDVI, NDVIG, and NDVIre showed high performance for predicting within-field yield variability. Approaches that adequately account for spatial correlation outperformed the OLS models since yield and VIs were spatially correlated.

This current analysis is among one of the few studies demonstrating the utilization of mid-season high-resolution satellite imagery to forecasting within-field maize yield variation. Future research should be focused on improving the understanding of historical yield distribution at larger scales (county-, district- or state-level) aiming at mapping the potential and limitations of scaling-up yield forecasting models.

Acknowledgements

This study was supported by CAPES Foundation, Ministry of Education of Brazil, Brasilia – DF, Zip Code 70.040-020, Aquarius project (http://w3.ufsm.br/projetoaquarius/index.php/pt/), and Kansas Corn Commission. This is contribution no. 18-073-J from the Kansas Agricultural Experiment Station and process 88887.130848/2016-00 from CAPES. The authors want to thank Drakkar (http://www.drakkar.com.br/), Fabiano Tabaldi and Fabio Gebert for providing organized corn yield monitor information.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.biosystemseng.2018.04.020.

REFERENCES

Amidan, B. G., Ferryman, T. A., & Cooley, S. K. (2005, March). Data outlier detection using the Chebyshev theorem. In *Aerospace Conference, 2005 IEEE* (pp. 3814–3819). IEEE.

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis, 27*, 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x.

Anselin, L., Bongiovanni, R., & Lowenberg-DeBoer, J. (2004). A spatial econometric approach to the economics of site specific nitrogen management in corn production. *American Journal of Agricultural Economics, 86*, 675–687.

Assefa, Y., Prasad, P. V. V., Carter, P., Hinds, M., Bhalla, G., & Ciampitti, I. A. (2016). Yield responses to planting density for US modern corn hybrids: A synthesis-analysis. *Crop Science, 56*, 1–38. https://doi.org/10.2135/cropsci2016.04.0215.

Azzari, G., Jain, M., & Lobell, D. B. (2016). Towards fine resolution global maps of crop yields: Testing multiple methods and satellites in three countries. *Remote Sensing of Environment*, 129–141. https://doi.org/10.1016/j.rse.2017.04.014.

Bakhsh, A., Jaynes, D. B., Colvin, T. S., & Kanwar, R. S. (2000). Spatio-temporal analysis of yield variability for a corn-soybean field in Iowa. *Transactions of the American Society of Agricultural Engineers, 43*, 31–38.

Becker-Reshef, I., Vermote, E., Lindeman, M., & Justice, C. (2010). A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sensing of Environment, 114*, 1312–1323. https://doi.org/10.1016/j.rse.2010.01.010.

Bivand, R., & Piras, G. (2015). Comparing implementations of estimation methods for spatial econometrics. *Journal of Statistical Software, 63*, 1–36.

Bognár, P., Ferencz, C., Pásztor, S., Molnár, G., Timár, G., Hamar, D., et al. (2011). Yield forecasting for wheat and corn in Hungary by satellite remote sensing. *International Journal of Remote Sensing, 32*, 4759–4767. https://doi.org/10.1080/01431161.2010.493566.

Bongiovanni, R. G., Robledo, C. W., & Lambert, D. M. (2007). Economics of site-specific nitrogen management for protein content in wheat. *Computers and Electronics in Agriculture, 58*, 13–24. https://doi.org/10.1016/j.compag.2007.01.018.

Bresler, E., Dasberg, S., Russo, D., & Dagan, G. (1981). Spatial variability of crop yield as a stochastic soil process. *Soil Science Society of America Journal, 45*(3), 600.

Bu, H., Sharma, L. K., Denton, A., & Franzen, D. W. (2017). Comparison of satellite imagery and ground-based active optical sensors as yield predictors in sugar beet, spring wheat, corn, and sunflower. *Agronomy Journal, 109*, 299–308. https://doi.org/10.2134/agronj2016.03.0150.

Canty, A., & Ripley, B. (2017). *boot: Bootstrap R (S-Plus) Functions. R package version 1.3-19*.

Congedo, L. (2016). *Semi-automatic classification plugin – User manual*. https://doi.org/10.13140/RG.2.1.1219.3524.

Córdoba, M. A., Bruno, C. I., Costa, J. L., Peralta, N. R., & Balzarini, M. G. (2016). Protocol for multivariate homogeneous zone delineation in precision agriculture. *Biosystems Engineering, 143*, 95–107. https://doi.org/10.1016/j.biosystemseng.2015.12.008.

DiRienzo, C., Fackler, P., & Goodwin, B. K. (2000, August). Modeling spatial dependence and spatial heterogeneity in county yield forecasting models. In *Proceedings of the American Agricultural Economics Association Annual Meeting, Tampa, FL, USA* (Vol. 1).

Doraiswamy, P. C., Moulin, S., Cook, P. W., & Stern, A. (2003). Crop yield assessment from remote sensing. *Photogrammetric Engineering & Remote Sensing, 69*, 665–674. https://doi.org/10.14358/PERS.69.6.665.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., et al. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment, 120*, 25–36. https://doi.org/10.1016/j.rse.2011.11.026.

Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC Press.

Gholap, J., Ingole, A., Gohil, J., Gargade, S., & Attar, V. (2012). Soil data analysis using classification techniques and soil attribute prediction. *International Journal of Computer Science Issues, 9*(3), 415–418.

Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment, 58*, 289–298. https://doi.org/10.1016/S0034-4257(96)00072-7.

Gitelson, A. A., & Merzlyak, M. N. (1994). Spectral reflectance changes associated with autumn senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. leaves – Spectral features and relation to chlorophyll estimation. *Journal of Plant Physiology, 143*, 286–292. https://doi.org/10.1016/S0176-1617(11)81633-0.

Gonzalez-Sanchez, A. (2014). Predictive ability of machine learning methods for massive crop yield prediction. *Spanish Journal of Agricultural Research, 12*, 313–328. https://doi.org/10.5424/sjar/2014122-4439.

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment, 90*, 337–352. https://doi.org/10.1016/j.rse.2003.12.013.

Hamar, D., Ferencz, C., Lichtenberger, J., Tarcsai, G., & Ferencz-Arkos, I. (1996). Yield estimation for corn and wheat in the Hungarian Great Plain using Landsat MSS data. *International Journal of Remote Sensing, 17*, 1689–1699. https://doi.org/10.1080/01431169608948732.

Hatfield, J. L., Gitelson, A. A., Schepers, J. S., & Walthall, C. L. (2008). Application of spectral remote sensing for agronomic decisions. *Agronomy Journal*. https://doi.org/10.2134/agronj2006.0370c.

Imran, M., Zurita-Milla, R., & Stein, A. (2013). Modeling crop yield in West – African rainfed agriculture using global and local spatial regression. *Agronomy Journal, 105*, 1177–1188.

Jaynes, D. B., & Colvin, T. S. (1997). Spatiotemporal variability of corn and soybean yield. *Agronomy Journal, 89*, 30–37.

Jin, Z., Azzari, G., Burke, M., Aston, S., & Lobell, D. (2017). Mapping smallholder yield heterogeneity at multiple scales in Eastern Africa. *Remote Sensing, 9*(9), 931. https://doi.org/10.3390/rs9090931.

Johnson, D. M., & Mueller, R. (2010). The 2009 cropland data layer. *Photogrammetric Engineering and Remote Sensing, 76*, 1201–1205.

Johnson, J. B., & Omland, K. S. (2004). Model selection in ecology and evolution. *Trends in Ecology & Evolution, 19*, 101–108. https://doi.org/10.1016/j.tree.2003.10.013.

Kantanantha, N., Serban, N., & Griffin, P. (2010). Yield and price for stochastic crop decision planning. *Journal of Agricultural, Biological, and Environmental Statistics, 15*, 362–380. https://doi.org/10.1007/s13253-010-0025-7.

Leiser, W. L., Rattunde, H. F., Piepho, H. P., & Parzies, H. K. (2012). Getting the most out of sorghum low-input field trials in West Africa using spatial adjustment. *Journal of Agronomy and Crop Science, 198*, 349–359. https://doi.org/10.1111/j.1439-037X.2012.00529.x.

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

Lobell, D. B., Thau, D., Seifert, C., Engle, E., & Little, B. (2015). A scalable satellite-based crop yield mapper. *Remote Sensing of Environment, 164*, 324–333. https://doi.org/10.1016/j.rse.2015.04.021.

Lopresti, M. F., Di Bella, C. M., & Degioanni, A. J. (2015). Relationship between MODIS-NDVI data and wheat yield: A case study in Northern Buenos Aires province, Argentina. *Information Processing in Agriculture, 2*(2), 73–84.

Minuzzi, R. B., & Lopes, F. Z. (2015). Desempenho agronômico do milho em diferentes cenários climáticos no Centro-Oeste do Brasil. *Revista Brasileira de Engenharia Agrícola e Ambiental, 19*, 734–740. https://doi.org/10.1590/1807-1929/agriambi.v19n8p734-740.

Monteith, J. L. (1977). Climate and the efficiency of crop production in Britain. *Philosophical Transactions of the Royal Society B, 281*, 277–294.

Moriondo, M., Maselli, F., & Bindi, M. (2007). A simple model of regional wheat yield based on NDVI data. *European Journal of Agronomy, 26*, 266–274. https://doi.org/10.1016/j.eja.2006.10.007.

Morkoc, F., Blggar, J. W., Mlllar, R. J., & Nlelsen, D. R. (1982). Statistical analysis of sorghum yield: A stochastic approach. *Soil Science Society of America Journal, 49*(6), 1342–1349.

Myneni, R. B., & Williams, D. L. (1994). On the relationship between FAPAR and NDVI. *Remote Sensing of Environment, 49*(3), 200–211.

Nguy-Robertson, A., Gitelson, A., Peng, Y., Viña, A., Arkebauer, T., & Rundquist, D. (2012). Green leaf area index estimation in maize and soybean: Combining vegetation indices to achieve maximal sensitivity. *Agronomy Journal, 104*, 1336–1347. https://doi.org/10.2134/agronj2012.0065.

Noureldin, N. A., Aboelghar, M. A., Saudy, H. S., & Ali, A. M. (2013). Rice yield forecasting models using satellite imagery in Egypt. *The Egyptian Journal of Remote Sensing and Space Science, 16*, 125–131. https://doi.org/10.1016/j.ejrs.2013.04.005.

Pebesma, E. J. (2004). Multivariable geostatistics in S: The gstat package. *Computers & Geosciences, 30*(7), 683–691.

Peralta, N., Assefa, Y., Du, J., Barden, C., & Ciampitti, I. (2016). Mid-season high-resolution satellite imagery for forecasting site-specific corn yield. *Remote Sensing, 8*, 848. https://doi.org/10.3390/rs8100848.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2017). *nlme: Linear and nonlinear mixed effects models*.

Raun, W. R., Solie, J. B., & Johnson, G. V. (2002). Improving nitrogen use efficiency in cereal grain production with optical sensing and variable rate application. *Agronomy Journal, 94*, 815–820.

Reeves, M. C., Zhao, M., & Running, S. W. (2005). Usefulness and limits on MODIS GPP for estimating wheat yield. *International Journal of Remote Sensing, 26*, 1403–1421. https://doi.org/10.1080/01431160512331326567.

Rembold, F., Atzberger, C., Savin, I., & Rojas, O. (2013). Using low resolution satellite imagery for yield prediction and yield anomaly detection. *Remote Sensing*. https://doi.org/10.3390/rs5041704.

Ribeiro, P. J., Jr., & Diggle, P. J. (2016). *geoR: Analysis of geostatistical data*.

Rouse, J., Haas, R., & Schell, J. (1973). *Monitoring the vernal advancement and retrogradation (greenwave effect) of natural vegetation* (pp. 1–8). Texas A M Univ.

Sadras, V. O., & Calviño, P. A. (2001). Quantification of grain yield response to soil depth in soybean, maize, sunflower, and wheat. *Agronomy Journal, 93*, 577–583. https://doi.org/10.2134/agronj2001.933577x.

Sakamoto, T., Gitelson, A. A., & Arkebauer, T. J. (2014). Near real-time prediction of U.S. corn yields based on time-series MODIS data. *Remote Sensing of Environment, 147*, 219–231. https://doi.org/10.1016/j.rse.2014.03.008.

Schwalbert, R., Amado, T., Tiago, H., Lincon, S., Yared, A., Prasad, V., et al. (2018). Corn yield response to plant density and nitrogen: Spatial model and yield distribution. *Agronomy Journal, 110*, 1–13.

Shanahan, J. F., Schepers, J. S., Francis, D. D., Varvel, G. E., Wilhelm, W. W., Tringe, J. M., et al. (2001). Use of remote-sensing imagery to estimate corn grain yield. *Agronomy*

Journal, 93, 583–589. https://doi.org/10.2134/agronj2001.933583x.

Sheridan, R. P. (2013). Time-split cross-validation as a method for estimating the goodness of prospective prediction. *Journal of Chemical Information and Modeling, 53*, 783–790. https://doi.org/10.1021/ci400084k.

Sibley, A. M., Grassini, P., Thomas, N. E., Cassman, K. G., & Lobell, D. B. (2014). Testing remote sensing approaches for assessing yield variability among maize fields. *Agronomy Journal, 106*, 24–32. https://doi.org/10.2134/agronj2013.0314.

Solie, J. B., Dean Monroe, A., Raun, W. R., & Stone, M. L. (2012). Generalized algorithm for variable-rate nitrogen application in cereal grains. *Agronomy Journal, 104*, 378–387. https://doi.org/10.2134/agronj2011.0249.

Stone, R. C., & Meinke, H. (2005). Operational seasonal forecasting of crop performance. *Philosophical Transactions of the Royal Society B, 360*, 2109–2124. https://doi.org/10.1098/rstb.2005.1753.

Timlin, D., Pachepsky, Y., Snyder, V., & Bryant, R. B. (1998). Spatial and temporal variability of corn grain yield on a hillslope. *Soil Science Society of America Journal, 62*, 764–773.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment, 8*, 127–150.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (Vol. 200, pp. 183–206). New York: Springer Science & Business Media.

Wiegand, C., & Richardson, A. (1990). Use of spectral vegetation indices to infer leaf area, evapotranspiration and yield: I. Rationale. *Agronomy Journal, 82*, 623–629. https://doi.org/10.2134/agronj1990.00021962008200030038x.

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution, 1*, 3–14. https://doi.org/10.1111/j.2041-210X.2009.00001.x.