



The use of remote sensing to derive maize sowing dates for large-scale crop yield simulations

Ehsan Eyshi Rezaei^{1,2} · Gohar Ghazaryan³ · Javier González³ · Natalie Cornish⁴ · Olena Dubovyk³ · Stefan Siebert¹

Received: 4 July 2020 / Revised: 21 September 2020 / Accepted: 6 November 2020 / Published online: 30 November 2020
© The Author(s) 2020

Abstract

One of the major sources of uncertainty in large-scale crop modeling is the lack of information capturing the spatiotemporal variability of crop sowing dates. Remote sensing can contribute to reducing such uncertainties by providing essential spatial and temporal information to crop models and improving the accuracy of yield predictions. However, little is known about the impacts of the differences in crop sowing dates estimated by using remote sensing (RS) and other established methods, the uncertainties introduced by the thresholds used in these methods, and the sensitivity of simulated crop yields to these uncertainties in crop sowing dates. In the present study, we performed a systematic sensitivity analysis using various scenarios. The LINTUL-5 crop model implemented in the SIMPLACE modeling platform was applied during the period 2001–2016 to simulate maize yields across four provinces in South Africa using previously defined scenarios of sowing dates. As expected, the selected methodology and the selected threshold considerably influenced the estimated sowing dates (up to 51 days) and resulted in differences in the long-term mean maize yield reaching up to 1.7 t ha^{-1} (48% of the mean yield) at the province level. Using RS-derived sowing date estimations resulted in a better representation of the yield variability in space and time since the use of RS information not only relies on precipitation but also captures the impacts of socioeconomic factors on the sowing decision, particularly for smallholder farmers. The model was not able to reproduce the observed yield anomalies in Free State (Pearson correlation coefficient: 0.16 to 0.23) and Mpumalanga (Pearson correlation coefficient: 0.11 to 0.18) in South Africa when using fixed and precipitation rule-based sowing date estimations. Further research with high-resolution climate and soil data and ground-based observations is required to better understand the sources of the uncertainties in RS information and to test whether the results presented herein can be generalized among crop models with different levels of complexity and across distinct field crops.

Keywords Crop modeling · Remote sensing · Drought · Maize · Sowing date · South Africa · MODIS

Introduction

According to household surveys, more than 13 million people have limited access to food in South Africa (Bvenura

and Afolayan 2015). Drought, caused by large anomalies in annual and interannual precipitation, is the major yield-limiting factor in the rainfed cropping systems of South Africa (Vogel and van Zyl 2016). Most of the maize-growing areas in South Africa are rainfed (Schultz 2006) with a water demand between 450 and 600 mm depending on various climatic and soil variables (du Plessis 2003).

Management decisions, such as the selection of sowing dates, can considerably affect a summer crop's yield in semi-arid regions (Aguirrezábal et al. 2009). The sowing date can directly influence crop yield and yield quality, mainly by changing the temporal patterns of temperature and soil moisture experienced by crops during the growing period (Marteau et al. 2011; Srivastava et al. 2016). Shifting the maize sowing date in the dry regions of South Africa by only 15 days could increase crop yield by 10% (Abraha and Savage 2006). Early sowing dates allow maize to avoid high temperatures and

✉ Ehsan Eyshi Rezaei
ehsaneyshi.rezaei@zalf.de

¹ Department of Crop Sciences, University of Göttingen, 37075 Göttingen, Germany

² Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg, Germany

³ Center for Remote Sensing of Land Surfaces (ZFL), University of Bonn, 53113 Bonn, Germany

⁴ RSS Remote Sensing Solutions GmbH, Dingolfinger Strasse 9, 81673 Munich, Germany

terminal drought during the grain-filling period. On the other hand, late and medium sowing dates can increase the crop growth rate during the early growth period, leading to faster development in leaf area and increasing the plant's ability to intercept solar radiation (Cirilo and Andrade 1994a, b). The effects of management practices, such as the effect of the optimized sowing date on crop yield, are generally assessed by conducting field experiments (Chen et al. 2011). However, carrying out field experiments is expensive and time-consuming, and the results are limited by the environmental conditions at the study location (Rinaldi 2001; Heng et al. 2007). Hence, testing the suitability of management practices such as optimized sowing dates and cultivar selection is a challenging issue at larger scales (Therond et al. 2011).

Crop growth models have been developed to improve our understanding of these processes and to upscale the effects of management strategies on crop yields and agricultural production from the field level to larger scales (Gaiser et al. 2010). Therefore, employing tools such as crop models supports enhancing the decision-making process for sustainable maize production under variable climatic conditions in South Africa (Zinyengere et al. 2015). In addition, crop models can be used for projections using climate scenarios or to explore the responses of crop yields to altered farm management practices. Most of the crop models used for large-scale modeling experiments are developed at the field level and require a large amount of input data for parameterization and calibration (Van Bussel et al. 2011). However, high-resolution model input data, including sowing dates, are usually not available at a large scale (Rezaei et al. 2015; Zhao et al. 2015). It is therefore common to use a fixed sowing date in large-scale simulation experiments, but failure to consider the spatiotemporal variability of crop model inputs could introduce large uncertainty into the model outputs (Waha et al. 2015).

Various precipitation-based approaches have been introduced to estimate the sowing dates of crops and to consider the spatiotemporal variability of sowing dates in Africa (Dodd and Jolliffe 2001; Laux et al. 2008). The sowing date is determined when the cumulative precipitation at the beginning of the rainy season reaches a certain threshold in a specific period of time (Tachie-Obeng et al. 2013; Waongo et al. 2013). The selection of an appropriate cumulative precipitation threshold, which is related to the temperature pattern and soil characteristics, can be a challenging task. In general, cumulative precipitation-based approaches are relatively reasonable; however, they cannot completely explain African farmers' decisions since other independent variables, including risk management practices, workload, and variability in soil characteristics, can largely influence their decisions (Waha et al. 2012). In addition, high-quality precipitation data based on a suitable density of climate stations are often limited in Africa (Paeth and Diederich 2011), and this sparsity of data can influence the reliability of rule-based methods for large-scale simulation experiments.

Using remote sensing (RS) to identify the phenology of crops such as maize (Viña et al. 2004) constitutes a promising approach, especially when the aim is to detect the start of the sowing date in semiarid regions with limited ground-based observations (Clinton et al. 2010; Atkinson et al. 2012; Brown et al. 2012). RS information can deliver crop canopy variables such as the leaf area index (LAI), phenology, and evapotranspiration over large spatial areas and thus inform large-scale crop models (Jin et al. 2018). However, the full exploitation of RS information can be hampered by factors such as the presence of clouds and/or dust and topographical effects and by a limited number of observations during the growing season of crops (Zhang et al. 2003; Tucker et al. 2005). In this sense, various RS sensors (MODIS, RapidEye, MERIS, etc.) and indices (LAI, NDVI, EVI, WDRVI, etc.) have been used to detect phenological dynamics to partly overcome these limitations (Jeganathan et al. 2010; Pan et al. 2012; Siachalou et al. 2015). Unlike the flowering stage, which can be related to the maximum LAI (Tagliapietra et al. 2018), estimating the sowing date is more challenging because of the difference between the sowing date and the time when the reflections from green leaves are detectable by the sensors. This time gap is generally addressed by assuming a fixed number of days between sowing and emergence (Dimou et al. 2018), which can introduce considerable uncertainty to the estimated sowing date. The scale mismatch between the low-resolution climate and soil data (particularly in Africa) used by crop models and the finer resolution of RS observations is another limitation of using RS information for large-scale crop modeling experiments (Doraiswamy et al. 2005). In summary, RS information and precipitation data are widely used to independently estimate the sowing date. However, to date, the differences in the sowing dates estimated by different methods have not been quantified.

At present, little is known about the sensitivity of crop model outputs to (i) changes in the fixed sowing date, (ii) variations in the cumulative precipitation threshold, and (iii) the selected time gap between sowing and emergence in RS-derived sowing estimations. In this context, the objectives of the current study are to perform a systematic sensitivity analysis to compare the long-term sowing date estimates based on the MODIS phenology product at the native resolution with those derived with a precipitation rule and to assess the impact of distinct sowing date estimates on the rainfed maize yield simulated for South Africa.

Materials and methods

Study area

This analysis was performed for the provinces of Free State, Gauteng, Mpumalanga, and North West (Fig. 1), collectively covering 329,318 km² and producing approximately 85% of the maize harvest in South Africa (DAFF 2019). Rainfed

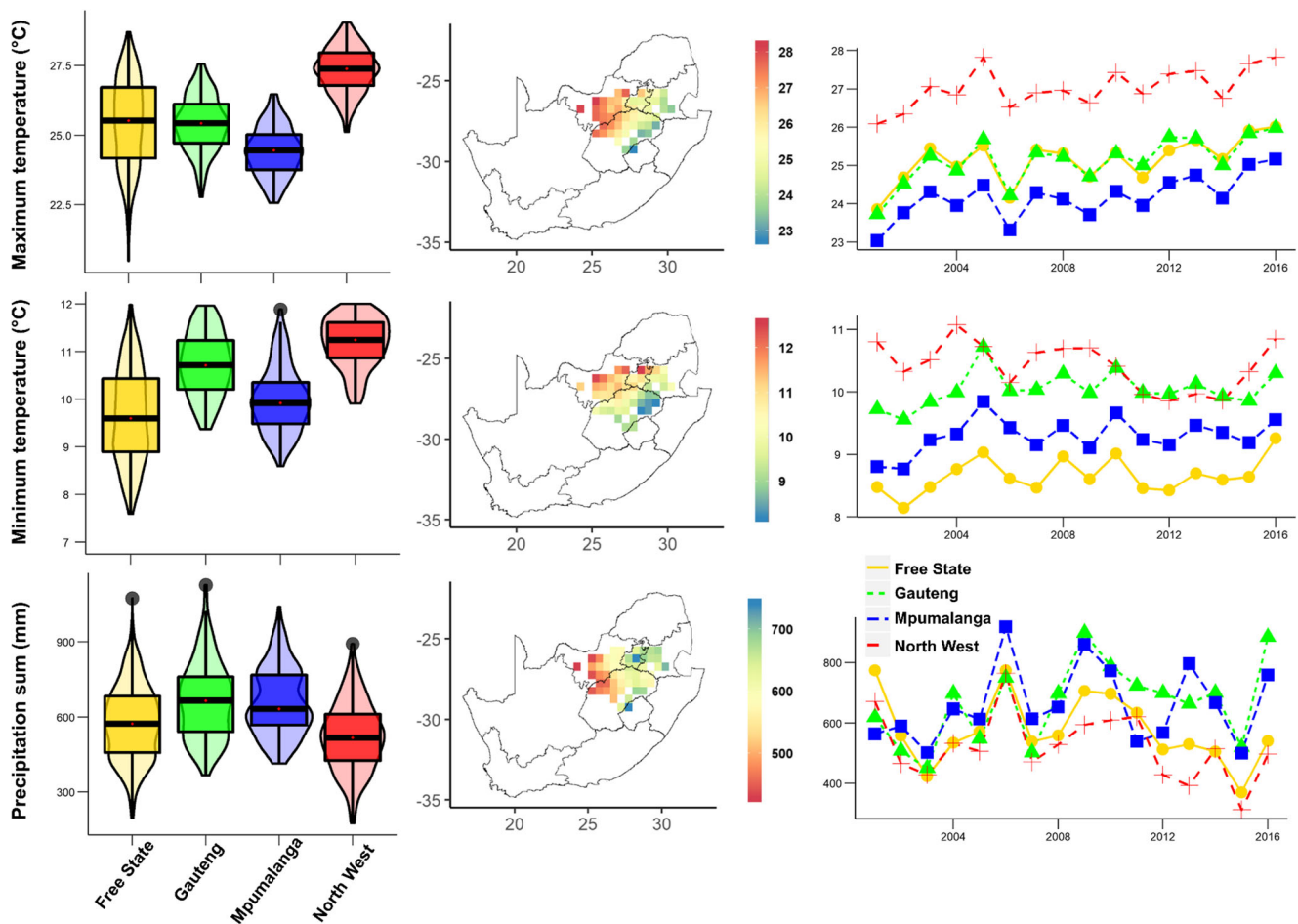


Fig. 1 Violin boxplots (a), spatial pattern (b), and interannual variability (c) of the annual precipitation sum and the mean daily maximum and mean daily minimum temperatures in the period 2001–2016 across maize growing grids in Free State, Gauteng, Mpumalanga, and North West Provinces, South Africa. The boxes are related to the 25th and 75th

percentiles; the whiskers extend to 1.5 times the interquartile range, and the horizontal line indicates the median. No overlapping of the notches among boxes offers evidence of statistically significant differences among their medians. Black and red dots reflect outliers and means, respectively

farming is the most common cultivation pattern of maize with-in these provinces. An analysis of climate data for the period 2001–2016 showed that maize-growing areas received, on average, 520 mm (North West) to 665 mm (Gauteng) of annual precipitation (Weedon et al. 2014). North West Province has experienced a remarkably higher mean daily maximum temperature (27.1 °C) than the other studied provinces (24.2 °C to 25.2 °C). The mean daily minimum temperature recorded in the provinces varied by 1.8 °C, with Free State experiencing the lowest minimum temperature (8.7 °C) during the study period (Fig. 1).

General data analysis workflow

Analyzing the sensitivity of model outputs to sowing date estimation methods involved various steps, including data preparation, RS-based data extraction, model setup, and model output analysis (Fig. 2). In the first step, non-maize grids

were filtered out across the study areas in South Africa using a maize crop mask. Then, curves of vegetation indices were extracted for each of the remaining grid cells from the time series of the MODIS global vegetation phenology (MCD12Q2) product (Xiao et al. 2013) for the period 2001–2016. Due to the lack of high-resolution climate and soil data as model inputs, the green-up dates extracted from MODIS pixels (500-m resolution) were aggregated to 0.5° × 0.5° grids, consistent with the spatial resolution of available climate and soil data. Five sensitivity scenarios were defined for estimates of the fixed sowing date, precipitation rule-based sowing date, and RS-based sowing date (Fig. 2). The thresholds of cumulative precipitation were defined as 10, 15, 20, 25, and 30 mm. Fixed periods of 10, 20, 30, 40, and 50 days were subtracted from the detected green-up dates as sensitivity scenarios to account for the time gap between sowing and green-up (emergence of leaves), which is when plants become visible to satellites. The fixed sowing date also varied

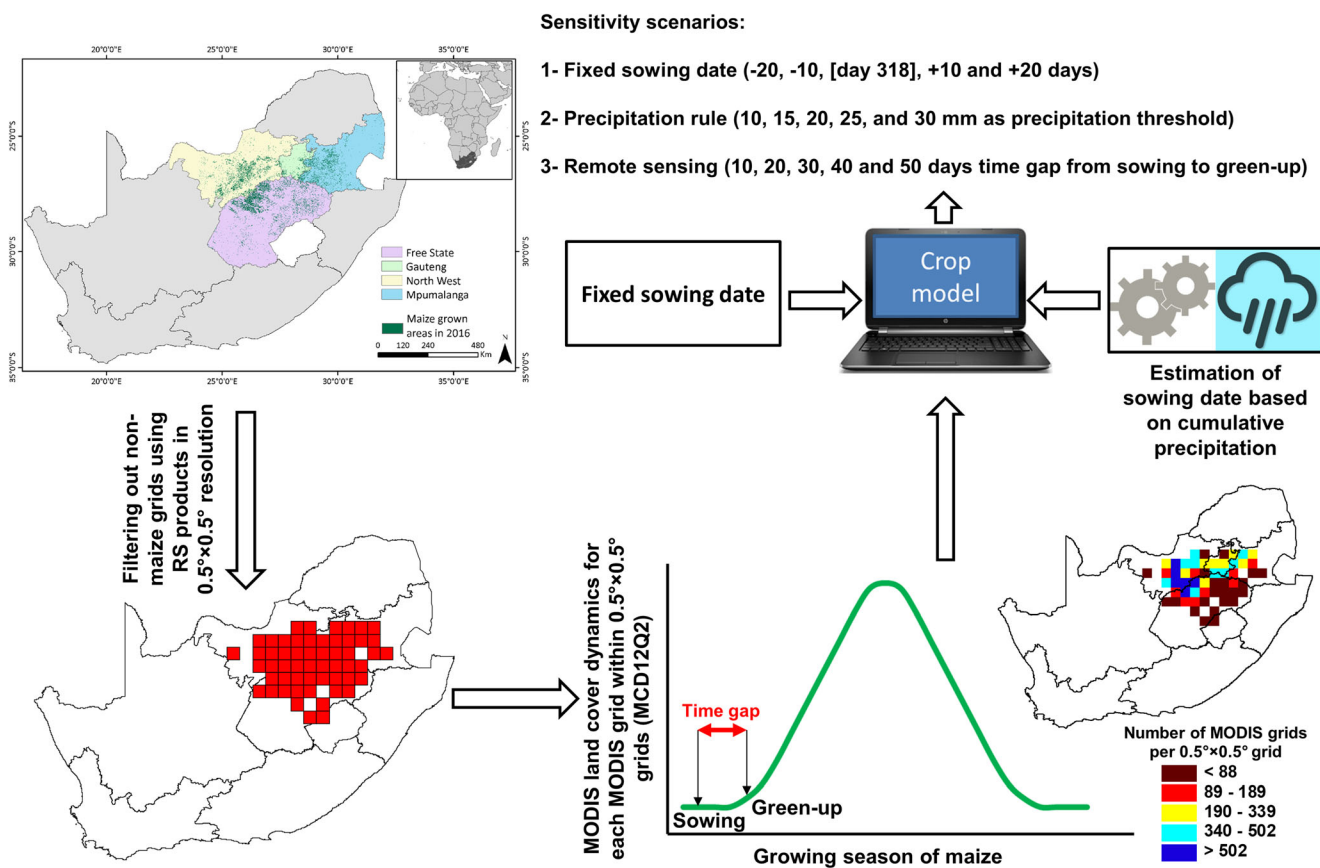


Fig. 2 Study region and the workflow for processing the remote sensing information used to inform the crop model to analyze the effects of different sowing date estimation methods on crop model results

in the range between -20 and $+20$ days relative to the standard sowing date.

The crop model was executed using the sowing dates estimated by using the different methods and thresholds. Next, the spatial and temporal patterns of the estimated sowing dates and simulated yields using different methods were analyzed. Significant differences in the sowing dates and simulated yields obtained using different sowing date estimation methods were tested using ANOVA followed by a Tukey test (HSD) at $P = 0.05$. Anomalies of the simulated yield with different sowing estimation methods were compared with the anomalies of the observed yield in the period from 2003 to 2015.

Processing of climate, soil, and management data

The WATCH Forcing climate dataset, which was developed using ERA-Interim reanalysis data (Weedon et al. 2014), was used as a climate input for the crop model. This dataset includes daily and 3-hourly values of temperature, precipitation, radiation, and wind speed at a global scale ($0.5^\circ \times 0.5^\circ$) for the period 1979–2016 (Weedon et al. 2014; Müller Schmied et al. 2016). The maximum and minimum values of the 3-hourly

temperature data were calculated and assumed to represent the daily maximum and minimum temperatures required by the crop model.

Both physical (saturation, field capacity, wilting point) and chemical (total nitrogen density) soil characteristics with a spatial resolution of $0.5^\circ \times 0.5^\circ$ were provided by the ISRIC-WISE dataset and the Global Gridded Surfaces of Selected Soil Characteristics, respectively (Batjes 2012). Soil depth information was obtained from the FAO soil depth dataset and restricted to 1 m to be compatible with the other soil information (Batjes 1997).

The long-term mean reference sowing dates of rainfed maize in South Africa was obtained from the MIRCA2000 dataset (Portmann et al. 2010). These sowing dates were used for simulations with fixed sowing dates and to define a time window of potential sowing dates for the precipitation rule-based sowing date estimations. Since there is no large-scale phenological database of rainfed maize for South Africa, the long-term mean harvest date was derived from MIRCA2000, and the temperature sum from sowing to harvest was calculated under the assumption that 60% of the temperature sum contributed to the vegetative phase while the remaining 40% contributed to the reproductive phase (van Bussel et al. 2015).

The nitrogen fertilizer application rate of maize in South Africa was obtained from a global dataset of fertilizer application rates (Potter et al. 2010).

Processing of RS information

As a primary data source of RS-based phenology estimates, the MODIS MCD12Q2 Version 6 product was used, as this product offers a validated dataset providing high-quality and consistent outputs, such as green-up, peak, and maturity stage, which are applicable for studies at regional to global scales. In this product, the seasonal dynamics of vegetation are derived using time series of the 2-band enhanced vegetation index (EVI2), which in turn is based on the MODIS nadir bidirectional reflectance distribution function (BRDF)-adjusted reflectance (NBAR) at a spatial resolution of 500 m (Friedl et al. 2019). The dataset was preprocessed using quality masks to exclude low-quality pixels (Stanimirova et al. 2019).

The estimated start of the season, “Greenup_1” (the date when EVI2 exceeded 15% of the EVI2 segment amplitude during the first cycle), was converted to the DOY and extracted for only maize-grown areas using crop distribution information to reduce contamination from neighboring land use classes and to decrease the impact of mixed pixels. The “Greenup_2” product, which accounts for the second growing period, was excluded after the initial test, as it did not provide data for maize-growing areas. The crop layers for the provinces of Free State, North West, and Mpumalanga were obtained from the South African Department of Agriculture, Forestry and Fisheries (DAFF 2017); these maps were generated based on the classification of SPOT imagery. For Gauteng Province, crop information was provided by the Sen2-Agri project (Defourny et al. 2019) based on Sentinel-2 information for 2016. When available, crop information from several years was considered for the analysis.

Sowing date estimation based on daily precipitation

The sum of daily precipitation was accumulated (in scenarios ranging from 10 to 30 mm) in moving 7-day intervals during the potential sowing period to determine the sowing date. The start of the potential sowing period was defined as the period between the first day of the month provided by MIRCA2000 as the sowing month and the following 50 continuous days. The first day of the first 7-day period with a precipitation sum ranging from 10 to 30 mm (with an interval of 5 mm) was defined as the day of sowing (Srivastava et al. 2016). The last day of the potential sowing period was selected as the sowing date if each 7-day precipitation sum in the potential sowing period was smaller than the predefined threshold.

Crop model description

SIMPLACE (Scientific Impact assessment and Modeling Platform for Advanced Crop and Ecosystem management) is a modeling framework based on the concept of encapsulating the solution of a modeling problem in discrete, replaceable, and interchangeable software units called Sim-components or sub-models (Rezaei et al. 2015). A specific combination of submodels within the modeling framework is called a model solution. The SIMPLACE < LINTUL-5, Heat > solution employed in this study combines the LINTUL-5 crop model (Wolf 2012) with a module implemented in the SIMPLACE platform to quantify the effect of heat stress on grain yield (Rezaei et al. 2015). The yield-limiting factors considered in the crop model were drought, heat, and nitrogen stress. Due to the lack of experimental data for the model calibration, we used the SIMPLACE model parameters used in a preceding study to simulate the maize yield across Africa (Rezaei and Gaiser 2017) (Table 1). The drought stress reduction factor is based on the ratio between actual and potential transpiration and influences leaf area expansion, root growth, biomass accumulation, and the partitioning of photosynthetic assimilates (Wolf 2012).

Comparison of estimated sowing dates and simulated yields

We first pooled the sowing date estimates for grids and years assigned to the studied provinces to compute the summary statistics (mean, median, and quantiles) obtained from different sowing date estimation methods and sensitivity scenarios. Outliers were defined based on the distance from the interquartile range (Krzywinski and Altman 2014). Then, we calculated the annual anomalies of the estimated sowing dates and simulated yields across the years to analyze the differences caused by the various sowing estimation methods and scenarios on the interannual yield variability. Simulated yields were compared with the yields of rainfed yellow maize available at the field level for the period from 2003 to 2015 (DAFF 2019). The field-scale observations were aggregated to a resolution of $0.5^\circ \times 0.5^\circ$, and the yield anomaly was calculated for the grids (29 grids) that had observations for the entire period.

Statistical analysis

Differences in estimated sowing dates and simulated yields were evaluated using ANOVA (aov function in R) followed by a Tukey test (HSD) at $P=0.05$ employing the agricolae package (De Mendiburu and Simon 2015) in R. The familywise 95% confidence intervals indicate significant multiple (pairwise) comparisons of means among the sowing estimation methods.

Table 1 Crop parameters of SIMPLACE < LINTUL-5, Heat > used in the current study

Parameter	Description	Unit	Value
TSUM1	Temperature sum from emergence to anthesis	°C day ⁻¹	1085–1712
TSUM2	Temperature sum from anthesis to maturity	°C day ⁻¹	723–1141
RUE-0.0	Radiation use efficiency at development stage 0*	g MJ ⁻¹ m ⁻²	3.0
RUE-1.50	Radiation use efficiency at development stage 1.50	g MJ ⁻¹ m ⁻²	3.0
RUE-2.0	Radiation use efficiency at development stage 2.0	g MJ ⁻¹ m ⁻²	2.4
ROOTDM	Maximum rooting depth	m	1
ROOTDI	Initial rooting depth	m	0.1
RRDMAX	Maximum rate of increase in rooting depth	m	0.012
TDWI	Initial total crop dry weight	kg ha ⁻¹	5
HSTCritical	Critical temperature threshold which grain yield start to damage	°C	35
RTNMIN	Fraction of soil mineral N coming available per day	-	0.005

Results and discussion

Differences in sowing dates estimated using RS and the precipitation rule

The sowing dates were remarkably different among the estimation methods, thresholds, and provinces (Fig. 3). Low precipitation thresholds resulted in relatively similar sowing dates for all studied provinces at day 320 (Fig. 3a), while the estimated sowing date differed among the studied provinces when using high precipitation thresholds. The distribution of sowing dates estimated by using the precipitation rule changed from unimodal (one peak) to multimodal (multiple peaks) from day 320 to day 340 in North West and Free State (Fig. 3a). There is a remarkable difference in the estimated sowing dates between the hot and dry provinces of North West and Free State and the other provinces (Fig. 3b). The earliest RS-derived

sowing dates were obtained for Mpumalanga (290–326) (Fig. 3b), and the RS-based sowing dates for North West and Free State were 11 days earlier than the sowing dates estimated for the other provinces (Fig. 3b).

The use of RS information remarkably increased the variability of the estimated sowing dates compared to those estimated with the precipitation rule. Using the precipitation rule resulted in sowing windows spanning 21–46 days depending on the specific precipitation threshold (Fig. 3). This indicates relatively small inter-annual variability at the beginning of the rainy season in South Africa. However, the actual sowing date window can be much wider because, in addition to climatic factors, other management factors can affect the choice of sowing date (Srivastava et al. 2016). For instance, the optimum temperature range, land preparation, and availability of labor can impact the timing of sowing, particularly for smallholder farmers in South Africa

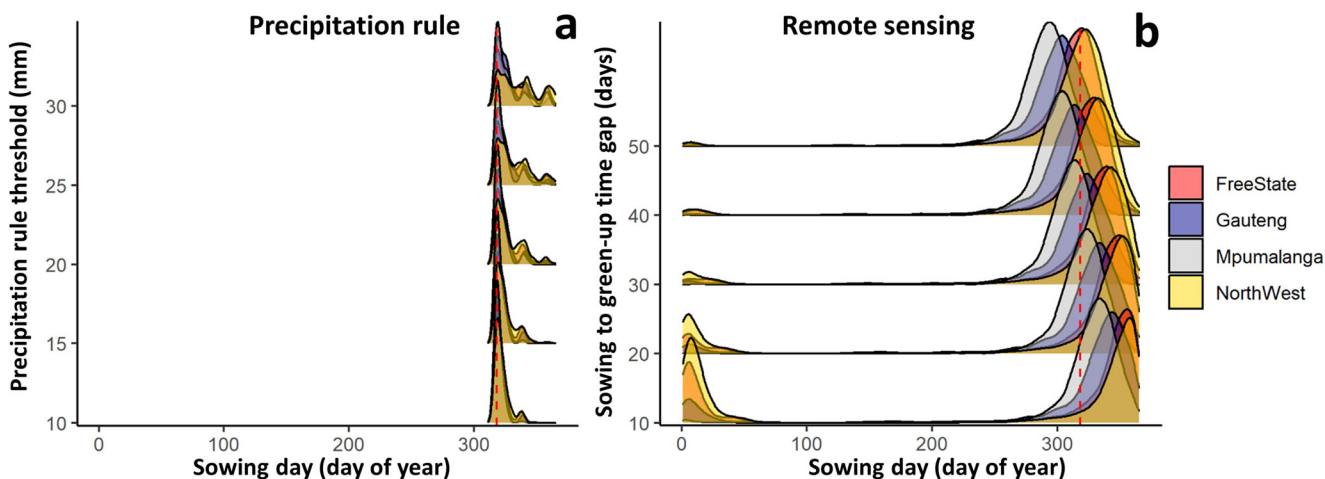


Fig. 3 Density ridgeline of the estimated sowing dates of maize using the RS and precipitation rule sensitivity scenarios in the period 2001–2016 across maize growing grids in Free State, Gauteng, Mpumalanga, and North West Provinces, South Africa. The red line indicates a fixed sowing date

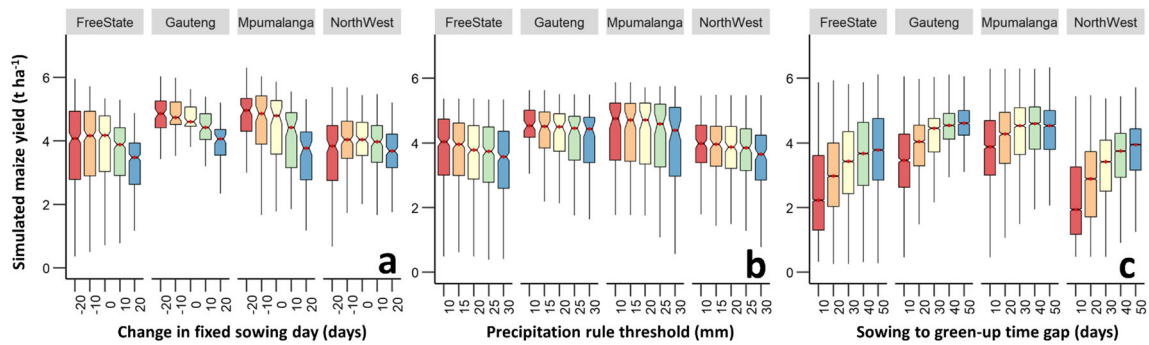


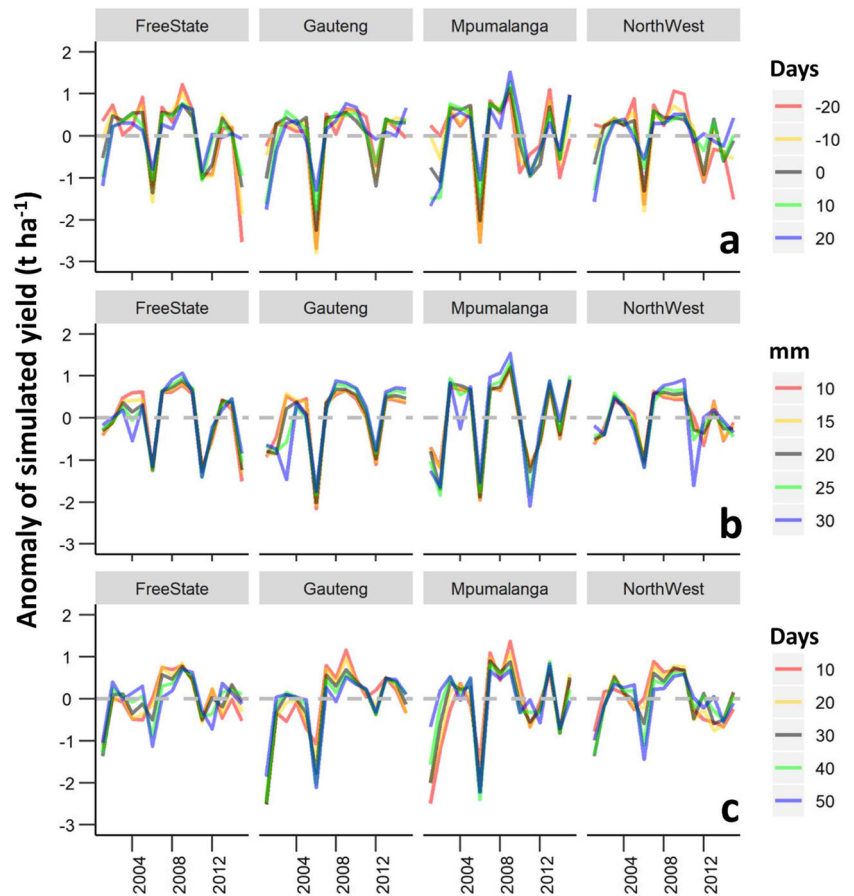
Fig. 4 Boxplots of simulated yields of maize using sensitivity scenarios based on fixed sowing dates (a), sowing dates estimated by the precipitation rule (b), and RS-based sowing dates (c) in the period 2001–2016 across maize growing grids in Free State, Gauteng, Mpumalanga, and North West Provinces, South Africa. The boxes refer

to the 25th and 75th percentiles; the whiskers extend to 1.5 times the interquartile range, and the horizontal line indicates the median, where no overlapping of the notches among boxes offers evidence of statistically significant differences among their medians. The red dots represent the mean values

(Waddington et al. 1991). A previous study showed that the sowing date of maize can range from October to January in Free State Province (Moeletsi 2017). Potentially, RS observations can better capture the variability in sowing dates caused by these differences in the onsets of rainfall in cropping systems and socioeconomic limitations (Kasampalis et al. 2018). Using fixed sowing dates or using sowing dates based on the

precipitation rule did not account for the wide range of cultivars (short-, medium-, and long-season cultivars) grown in South Africa, which may have different optimum sowing dates (Moeletsi et al. 2013). The effectiveness of using RS information in tracking the development stage was confirmed in previous studies for different crops grown in various environments (Zhang et al. 2003; Viña et al. 2004).

Fig. 5 Anomaly of the simulated yield of maize using sensitivity scenarios based on fixed sowing dates (a), sowing dates estimated by the precipitation rule (b), and RS-based sowing dates (c) in the period 2001–2016 across maize growing grids in Free State, Gauteng, Mpumalanga, and North West Provinces, South Africa



Effects of differences in the estimated sowing dates on the simulated yield

The simulated maize yield calculated as the mean across grids and years was in the range between 2.51 and 4.04 t ha⁻¹ (Fig. 4). The mean simulated yield in Mpumalanga and Gauteng was 0.6 t ha⁻¹ higher than that in North West and Free State. Early sowing dates resulted increases in the mean simulated yields in all studied provinces except North West (Fig. 4a). Increasing the cumulative precipitation rule threshold from 10 to 30 mm reduced the mean simulated yield between 7 and 12% depending on the studied province (Fig. 4b). Expanding the time gap between sowing and green-up to 50 days resulted in a 0.56 to 1.48 t ha⁻¹ increase in the simulated yield, particularly in the hot-dry provinces (Fig. 4c). The standard deviation of the simulated yield anomaly was 40% larger for provinces with higher simulated yields, such as Gauteng and Mpumalanga, than for Free State and North West (Fig. 5). In general, early sowing dates, large precipitation thresholds, and large differences between the sowing and green-up dates resulted in larger anomalies of simulated yields (Fig. 5). An analysis of the spatial patterns of the simulated yields showed a decreasing trend of the simulated yield from west (3.0 to 6.0 t ha⁻¹) to east (0.3 to 3.0 t ha⁻¹), particularly when the RS-based sowing date estimations were employed across the maize-growing areas (Fig. 6). This pattern was less

pronounced when using fixed sowing dates and precipitation rule-based sensitivity scenarios (Fig. 6).

The results of the Tukey test showed a significant difference among the methodologies for estimating sowing dates (Fig. 7). There was also a significant difference between the simulated yields using the RS-based sowing estimations and the other sowing estimation methods (Fig. 7). However, there was no significant difference in the simulated yield between the fixed date and precipitation rule-based sowing estimations (Fig. 7).

The simulated yield was significantly affected by the selected method and the sensitivity scenario used to estimate the sowing date (Figs. 3 and 4). These differences highlight the importance of selecting a robust and suitable sowing date estimation method and threshold in large-scale simulation experiments, particularly when ground-based observations are not available (Ogutu et al. 2018). The maize sowing dates in the semiarid regions of Africa are generally tied to the start of the rainy season (Sacks et al. 2010). However, the assumption that all farmers (particularly smallholder farmers) in certain provinces decide to sow at the optimum time (e.g., at the fixed sowing date given in large-scale modeling experiments) is an oversimplification of real conditions (Srivastava et al. 2016), since the farmers in large growing areas of South Africa have different levels of agronomic knowledge (Muzangwa et al. 2017) and different capacities to deal with labor peaks. In

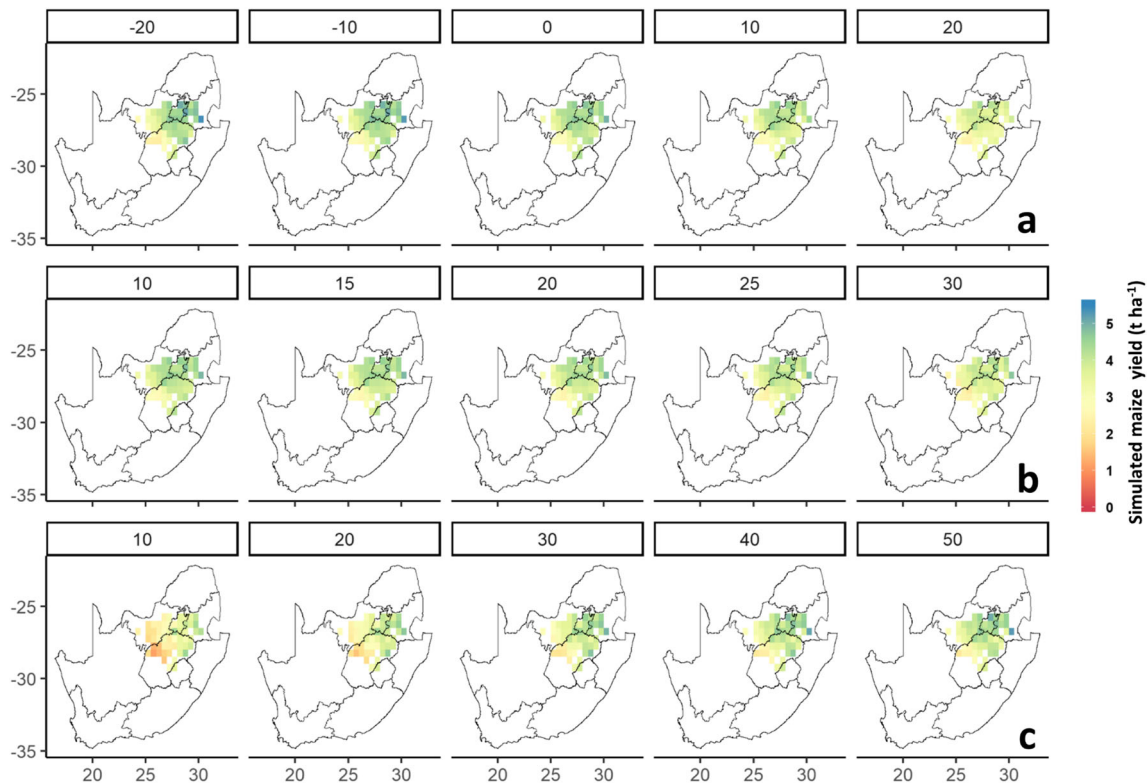


Fig. 6 Spatial pattern of the mean yield of maize simulated using sensitivity scenarios based on fixed sowing dates (a), sowing dates estimated by the precipitation rule (b), and RS-based sowing dates (c)

in the period 2001–2016 across maize growing grids in Free State, Gauteng, Mpumalanga, and North West Provinces, South Africa

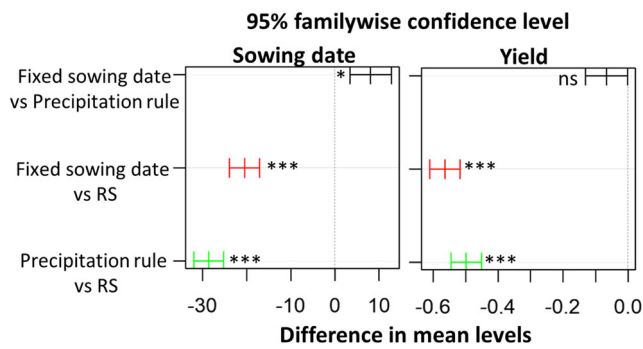


Fig. 7 Multiple (pairwise) comparisons among the means of the estimated sowing dates (a) and simulated yield (b) of maize using RS dates, fixed sowing dates and the dates estimated with the precipitation rule using the Tukey test in the period 2001–2016 across maize growing grids in Free State, Gauteng, Mpumalanga, and North West Provinces, South Africa. ns, nonsignificant; * and *** $p \leq 0.001$ and 0.05 significance levels, respectively

addition, farmers cannot know at the beginning of the growing season how the climatic conditions will subsequently develop. Therefore, it is, in principle, impossible to determine the optimal sowing date in real time.

The anomaly of the simulated yield using the RS estimation method indicated a stronger ($r = 0.30$ to 0.79) correlation with the anomaly of the observed yield, particularly in Gauteng and North West, than the correlations between the simulated and observed yield anomalies when using a fixed sowing date ($r = 0.12$ to 0.66) or a precipitation rule-based sowing date ($r = 0.10$ to 0.54) (Fig. 8).

The RS-based sowing date estimations resulted in a better representation of the spatiotemporal variability of yield, at least for Gauteng and Mpumalanga Provinces (Fig. 6). Using the fixed sowing or precipitation rule-based sowing date estimation methods resulted in a loss of information regarding the spatial yield pattern (Fig. 6). The failure to detect the variability of the observed yield using a fixed sowing date

can be related to the use of low-resolution climate and soil inputs for our large-scale simulations. Using the model inputs at a low spatial resolution ($0.5^\circ \times 0.5^\circ$ in our case) ignored the heterogeneity of precipitation, temperature, soil characteristics, and agronomic management (such as the sowing date) within the grids. Such variability can be partly counterbalanced by using RS-derived information. The use of RS products for estimating the sowing date could be a better approach than using a fixed sowing date. Using a fixed sowing date does not account for the interannual precipitation pattern and ignores the spatial pattern of sowing dates, assuming that all farmers sow crops at the same time (Wolf et al. 2015). Using the precipitation rule-based sowing estimation approach always depends on the selected start date of the potential sowing period, which in turn determines how the period with continuous precipitation events is found.

Based on the results of our sensitivity analysis, the thresholds leading to earlier sowing date estimations resulted in higher simulated yields in most of the provinces. This result may be caused by the fact that the time between sowing and emergence was not impacted by soil moisture in the topsoil in the crop model used in this study. In reality, emergence would be delayed if the topsoil is very dry (Blackshaw 1991). The results of this sensitivity analysis may therefore change when using other crop models accounting for the relationship between soil moisture and the duration between sowing and emergence.

Limitations of the data and methods used in this study

Through this study, we improved our understanding of the impact of distinct sowing date estimation methods and the selection of appropriate thresholds (for each method) on simulated maize yields. However, we also need to acknowledge

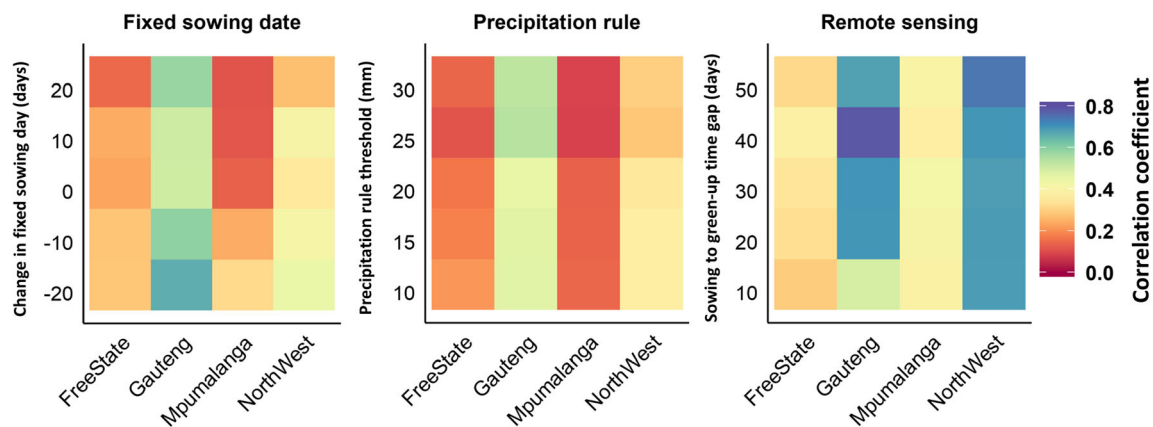


Fig. 8 Correlation coefficients between the anomaly of simulated yield using sensitivity scenarios based on RS-based sowing dates, fixed sowing dates and the dates estimated with the precipitation rule and the observed

yield in the period 2003–2015 across maize growing grids in Free State, Gauteng, Mpumalanga, and North West Provinces, South Africa

some shortcomings of our input data and of the methods presented here. We were not able to conduct an extensive model calibration due to a lack of time series of experimental data for crop phenology. We employed yields observed at the field scale in the period 2003–2015 for a comparison to yield anomalies simulated using distinct methods and thresholds for detecting the sowing date. Therefore, we can quantify the sensitivity of yields to the choice of the sowing date estimation method, but we cannot determine which method works most precisely (in the yield simulation) for the study period. However, this may not change the merit of the present study, as an extensive calibration would improve only the fit of the simulated yields to the observed yields but not the difference among the sowing estimation methodologies. In addition, improving the model itself was not the objective of this study.

We also did not have access to information about the phenological characteristics (differences in thermal requirements) and the proportions of local and commercial maize cultivars in the specifically studied provinces of South Africa. Additionally, maize maps were only available for a few study years, which implies the assumption for dataless years that maize was sown at the same place with no rotation or fallow periods in between. The crop model also did not account for biotic stress factors such as pests or diseases of maize in the study region, which can influence the selected sowing dates by farmers and thus affect crop yield (Martin and Shepherd 2009; Bennett et al. 2012).

Finally, the inconsistency between the spatial resolutions of remote sensing information (high resolution) and those of model inputs such as climate and soil (low resolution) can reduce the improvement achieved by using RS information due to the required aggregation of RS information to a coarser resolution. We also did not have information about the spatial extent of irrigated maize, which may have a completely different sowing window from that of rainfed maize.

Conclusions

We conclude that the choice of method for estimating the sowing date and the selected threshold causes considerable differences in estimated sowing dates and in crop yields simulated by using these sowing dates as inputs. To improve the accuracy of yield simulations, it is therefore essential to reduce the uncertainty in the spatiotemporal pattern of sowing dates. Using RS-derived sowing date estimations can result in a better representation of the variability of crop yield compared to using a fixed sowing date. Considerable spatiotemporal variability was observed in the estimated sowing date using RS-derived information, particularly in the dry provinces. Improving the reliability of RS sources (e.g., through higher

temporal and spatial resolutions and extensive RS product validation) and of techniques for extracting phenological metrics could help to reduce the number of outliers. The sensitivity of the model results to the sowing date estimations was substantially larger in drier and warmer provinces owing to the higher variability of the exposure of crops to heat and drought caused by varying sowing dates.

Acknowledgements Open Access funding enabled and organized by Projekt DEAL.

Authors' contributions E.E.R. initiated the study, set up the crop model, and analyzed the crop model results. G.G., J.G., and O.D. processed the remote sensing data used in the study. N.C. provided the cropland masks. S.S. coordinated the study. E.E.R. and S.S. wrote the paper with contributions from all coauthors. All authors have read and agreed to the published version of the manuscript.

Funding This research was funded by the German Federal Ministry of Education and Research (Project: GlobeDrought, grant nos. 02WGR1457A (G.G., J.G., O.D.), 02WGR1457D (N.C.), 02WGR1457F (E.E.R., S.S.)).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abraha MG, Savage MJ (2006) Potential impacts of climate change on the grain yield of maize for the midlands of KwaZulu-Natal, South Africa. *Agric Ecosyst Environ* 115:150–160. <https://doi.org/10.1016/j.agee.2005.12.020>
- Aguirrezabal L, Martre P, Pereyra-Irujo G et al (2009) Management and breeding strategies for the improvement of grain and oil quality. In: *Crop physiology*. Academic Press, pp 387–421. <https://doi.org/10.1016/B978-0-12-374431-9.00016-5>
- Atkinson PM, Jeganathan C, Dash J, Atzberger C (2012) Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. *Remote Sens Environ* 123:400–417. <https://doi.org/10.1016/j.rse.2012.04.001>
- Batjes NH (1997) A world dataset of derived soil properties by FAO-UNESCO soil unit for global modelling. *Soil Use Manag* 13:9–16. <https://doi.org/10.1111/j.1475-2743.1997.tb00550.x>

- Batjes NH (2012) ISRIC-WISE derived soil properties on a 5 by 5 arc-minutes global grid (ver. 1.2). Report 2012/01. ISRIC — World Soil Information, Wageningen, p 57. <http://research.wur.nl/en/datasets/isric-wise-derived-soil-properties-on-a-5-by-5-arc-minutes-global>
- Bennett AJ, Bending GD, Chandler D, Hilton S, Mills P (2012) Meeting the demand for crop production: the challenge of yield decline in crops grown in short rotations. *Biol Rev* 87:52–71. <https://doi.org/10.1111/j.1469-185X.2011.00184.x>
- Blackshaw RE (1991) Soil temperature and moisture effects on downy brome vs. winter canola, wheat and rye emergence. *Crop Sci* 31: 1034–1040. <https://doi.org/10.2135/cropsci1991.0011183X003100040038x>
- Brown ME, de Beurs KM, Marshall M (2012) Global phenological response to climate change in crop areas using satellite remote sensing of vegetation, humidity and temperature over 26years. *Remote Sens Environ* 126:174–183. <https://doi.org/10.1016/j.rse.2012.08.009>
- Bvenura C, Afolayan AJ (2015) The role of wild vegetables in household food security in South Africa: a review. *Food Res Int* 76:1001–1011. <https://doi.org/10.1016/j.foodres.2015.06.013>
- Chen X-P, Cui Z-L, Vitousek PM, Cassman KG, Matson PA, Bai JS, Meng QF, Hou P, Yue SC, Romheld V, Zhang FS (2011) Integrated soil–crop system management for food security. *Proc Natl Acad Sci* 108:6399–6404. <https://doi.org/10.1073/pnas.1101419108>
- Cirilo AG, Andrade FH (1994a) Sowing date and maize productivity: I. crop growth and dry matter partitioning. *Crop Sci* 34:1039–1043. <https://doi.org/10.2135/cropsci1994.0011183X003400040037x>
- Cirilo AG, Andrade FH (1994b) Sowing date and maize productivity: II. Kernel number determination. *Crop Sci* 34:1044–1046. <https://doi.org/10.2135/cropsci1994.0011183X003400040038x>
- Clinton NE, Potter C, Crabtree B, Genovese V, Gross P, Gong P (2010) Remote sensing–based time-series analysis of cheatgrass (*Bromus tectorum* L.) phenology. *J Environ Qual* 39:955–963. <https://doi.org/10.2134/jeq2009.0158>
- DAFF (2017) South Africa field crop boundary data layers (provincial). Department of Agriculture, Forestry and Fisheries, Government of South Africa, Pretoria
- DAFF (2019) Objective yield data for maize 2004–2017. Department of Agriculture, Forestry and Fisheries, Government of South Africa, Pretoria
- De Mendiburu F, Simon R (2015) Agricolae - ten years of an open source statistical tool for experiments in breeding, agriculture and biology. *PeerJ Prepr* 3:e1404v1. <https://doi.org/10.7287/peerj.preprints.1404v1>
- Defoumy P, Bontemps S, Bellemans N, Cara C, Dedieu G, Guzzonato E, Hagolle O, Inglada J, Nicola L, Rabaute T, Savinaud M, Udrou C, Valero S, Bégué A, Dejoux JF, el Harti A, Ezzahar J, Kussul N, Labbassi K, Lebourgeois V, Miao Z, Newby T, Nyamugama A, Salh N, Shelestov A, Simonneaux V, Traore PS, Traore SS, Koetz B (2019) Near real-time agriculture monitoring at national scale at parcel resolution: performance assessment of the Sen2-Agri automated system in various cropping systems around the world. *Remote Sens Environ* 221:551–568. <https://doi.org/10.1016/j.rse.2018.11.007>
- Dimou M, Meroni M, Rembold F (2018) Development of a national and sub-national crop calendars data set compatible with remote sensing derived land surface phenology. EUR 29327 EN, Publications Office of the European Union, Luxembourg. <https://doi.org/10.2760/25859>
- Dodd DES, Jolliffe IT (2001) Early detection of the start of the wet season in semiarid tropical climates of western Africa. *Int J Climatol* 21: 1251–1262. <https://doi.org/10.1002/joc.640>
- Doraiswamy PC, Sinclair TR, Hollinger S, Akhmedov B, Stern A, Prueger J (2005) Application of MODIS derived parameters for regional crop yield assessment. *Remote Sens Environ* 97:192–202. <https://doi.org/10.1016/j.rse.2005.03.015>
- du Plessis J (2003) Maize production. Directorate of Agricultural Information Services, Department of Agriculture South Africa report, pp 1–34. <https://www.arc.agric.za/arc-gci/Fact%20Sheets%20Library/Maize%20Production.pdf>
- Friedl M, Gray J, Sulla-Menashe D (2019) MCD12Q2 MODIS/Terra+Aqua Land Cover Dynamics Yearly L3 Global 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. <https://doi.org/10.5067/MODIS/MCD12Q2.006>
- Gaiser T, Judex M, Hiepe C, Kuhn A (2010) Regional simulation of maize production in tropical savanna fallow systems as affected by fallow availability. *Agric Syst* 103:656–665. <https://doi.org/10.1016/j.agsy.2010.08.004>
- Heng LK, Asseng S, Mejahed K, Rusan M (2007) Optimizing wheat productivity in two rain-fed environments of the West Asia–North Africa region using a simulation model. *Eur J Agron* 26:121–129. <https://doi.org/10.1016/j.eja.2006.09.001>
- Jeganathan C, Dash J, Atkinson PM (2010) Mapping the phenology of natural vegetation in India using a remote sensing-derived chlorophyll index. *Int J Remote Sens* 31:5777–5796. <https://doi.org/10.1080/01431161.2010.512303>
- Jin X, Kumar L, Li Z, Feng H, Xu X, Yang G, Wang J (2018) A review of data assimilation of remote sensing and crop models. *Eur J Agron* 92:141–152. <https://doi.org/10.1016/j.eja.2017.11.002>
- Kasampalis AD, Alexandridis KT, Deva C et al (2018) Contribution of remote sensing on crop models: a review. *J Imaging* 4:52. <https://doi.org/10.3390/jimaging4040052>
- Krzywinski M, Altman N (2014) Visualizing samples with box plots. *Nat Methods* 11:119–120. <https://doi.org/10.1038/nmeth.2813>
- Laux P, Kunstmann H, Bárdossy A (2008) Predicting the regional onset of the rainy season in West Africa. *Int J Climatol* 28:329–342. <https://doi.org/10.1002/joc.1542>
- Marteau R, Sultan B, Moron V, Alhassane A, Baron C, Traoré SB (2011) The onset of the rainy season and farmers' sowing strategy for pearl millet cultivation in Southwest Niger. *Agric For Meteorol* 151: 1356–1369. <https://doi.org/10.1016/j.agrformet.2011.05.018>
- Martin DP, Shepherd DN (2009) The epidemiology, economic impact and control of maize streak disease. *Food Secur* 1:305–315. <https://doi.org/10.1007/s12571-009-0023-1>
- Moeletsi ME (2017) Mapping of maize growing period over the Free State Province of South Africa: heat units approach. *Adv Meteorol* 2017:7164068–7164011. <https://doi.org/10.1155/2017/7164068>
- Moeletsi ME, Moopisa SG, Walker S, Tsubo M (2013) Development of an agroclimatological risk tool for dryland maize production in the Free State Province of South Africa. *Comput Electron Agric* 95: 108–121. <https://doi.org/10.1016/j.compag.2013.04.006>
- Müller Schmied H, Adam L, Eisner S, Fink G, Flörke M, Kim H, Oki T, Portmann FT, Reinecke R, Riedel C, Song Q, Zhang J, Döll P (2016) Variations of global and continental water balance components as impacted by climate forcing uncertainty and human water use. *Hydrol Earth Syst Sci* 20:2877–2898. <https://doi.org/10.5194/hess-20-2877-2016>
- Muzangwa L, Mkeni PNS, Chiduzo C (2017) Assessment of conservation agriculture practices by smallholder farmers in the eastern Cape province of South Africa. *Agronomy* 7:7030046
- Ogutu GEO, Franssen WHP, Supit I, Omondi P, Hutjes RWA (2018) Probabilistic maize yield prediction over East Africa using dynamic ensemble seasonal climate forecasts. *Agric For Meteorol* 250–251: 243–261. <https://doi.org/10.1016/j.agrformet.2017.12.256>
- Paeth H, Diederich M (2011) Postprocessing of simulated precipitation for impact research in West Africa. Part II: a weather generator for daily data. *Clim Dyn* 36:1337–1348. <https://doi.org/10.1007/s00382-010-0840-0>
- Pan Y, Li L, Zhang J, Liang S, Zhu X, Sulla-Menashe D (2012) Winter wheat area estimation from MODIS-EVI time series data using the Crop Proportion Phenology Index. *Remote Sens Environ* 119:232–242. <https://doi.org/10.1016/j.rse.2011.10.011>

- Portmann FT, Siebert S, Döll P (2010) MIRCA2000—global monthly irrigated and rainfed crop areas around the year 2000: a new high-resolution data set for agricultural and hydrological modeling. *Glob Biogeochem Cycles* 24:1–24. <https://doi.org/10.1029/2008GB003435>
- Potter P, Ramankutty N, Bennett EM, Donner SD (2010) Characterizing the spatial patterns of global fertilizer application and manure production. *Earth Interact* 14:1–22. <https://doi.org/10.1175/2009EI288.1>
- Rezaei EE, Gaiser T (2017) Change in crop management strategies. ZEF - Discussion Papers on Development Policy No. 239. <https://doi.org/10.2139/ssrn.3005416>
- Rezaei EE, Siebert S, Ewert F (2015) Impact of data resolution on heat and drought stress simulated for winter wheat in Germany. *Eur J Agron* 65:69–82. <https://doi.org/10.1016/j.eja.2015.02.003>
- Rinaldi M (2001) Application of EPIC model for irrigation scheduling of sunflower in Southern Italy. *Agric Water Manag* 49:185–196. [https://doi.org/10.1016/S0378-3774\(00\)00148-7](https://doi.org/10.1016/S0378-3774(00)00148-7)
- Sacks WJ, Deryng D, Foley JA, Ramankutty N (2010) Crop planting dates: an analysis of global patterns. *Glob Ecol Biogeogr* 19:607–620. <https://doi.org/10.1111/j.1466-8238.2010.00551.x>
- Schultz RE (2006) Agricultural crop production in South Africa: back ground. WRC report 1489/1/06, Section 16.1, Pretoria. <http://sarva2.dirisa.org/resources/documents/beeH/Section%2016.1%20Agric%20Crops%20Backgr%20Info.pdf>
- Siachalou S, Mallinis G, Tsakiri-Strati M (2015) A hidden Markov models approach for crop classification: linking crop phenology to time series of multi-sensor remote sensing data. *Remote Sens* 7:3633–3650
- Srivastava AK, Mboh CM, Gaiser T, Webber H, Ewert F (2016) Effect of sowing date distributions on simulation of maize yields at regional scale - a case study in Central Ghana, West Africa. *Agric Syst* 147:10–23. <https://doi.org/10.1016/j.agsy.2016.05.012>
- Stanimirova R, Cai Z, Melaas EK et al (2019) An empirical assessment of the MODIS land cover dynamics and TIMESAT land surface phenology algorithms. *Remote Sens* 11:2201. <https://doi.org/10.3390/rs11192201>
- Tachie-Obeng E, Akponikpè PBI, Adiku S (2013) Considering effective adaptation options to impacts of climate change for maize production in Ghana. *Environ Dev* 5:131–145. <https://doi.org/10.1016/j.envdev.2012.11.008>
- Tagliapietra EL, Streck NA, da Rocha TSM, Richter GL, da Silva MR, Cera JC, Guedes JVC, Zanon AJ (2018) Optimum leaf area index to reach soybean yield potential in subtropical environment. *Agron J* 110:932–938. <https://doi.org/10.2134/agronj2017.09.0523>
- Therond O, Hengsdijk H, Casellas E, Wallach D, Adam M, Belhouchette H, Oomen R, Russell G, Ewert F, Bergez JE, Janssen S, Wery J, van Ittersum MK (2011) Using a cropping system model at regional scale: low-data approaches for crop management information and model calibration. *Agric Ecosyst Environ* 142:85–94. <https://doi.org/10.1016/j.agee.2010.05.007>
- Tucker CJ, Pinzon JE, Brown ME, Slayback DA, Pak EW, Mahoney R, Vermote EF, el Saleous N (2005) An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. *Int J Remote Sens* 26:4485–4498. <https://doi.org/10.1080/01431160500168686>
- Van Bussel LGJ, Ewert F, Leffelaar PA (2011) Effects of data aggregation on simulations of crop phenology. *Agric Ecosyst Environ* 142:75–84. <https://doi.org/10.1016/j.agee.2010.03.019>
- van Bussel LGJ, Stehfest E, Siebert S, Müller C, Ewert F (2015) Simulation of the phenological development of wheat and maize at the global scale. *Glob Ecol Biogeogr* 24:1018–1029. <https://doi.org/10.1111/geb.12351>
- Viña A, Gitelson AA, Rundquist DC, Keydan G, Leavitt B, Schepers J (2004) Monitoring maize (*Zea mays* L.) phenology with remote sensing. *Agron J* 96:1139–1147. <https://doi.org/10.2134/agronj2004.1139>
- Vogel C, van Zyl K (2016) Drought: in search of sustainable solutions to a persistent, ‘wicked’ problem in South Africa. In: *Climate change adaptation strategies – an upstream-downstream perspective*. Springer International Publishing, Cham, pp. 195–211
- Waddington S, Mudhara M, Hlatshwayo M, Kunjeku P (1991) Extent and causes of low yield in maize planted late by smallholder farmers in subhumid areas of Zimbabwe. *Farming Syst Bull East South Africa, CIMMYT* 9:15–31
- Waha K, van Bussel LGJ, Müller C, Bondeau A (2012) Climate-driven simulation of global crop sowing dates. *Glob Ecol Biogeogr* 21:247–259. <https://doi.org/10.1111/j.1466-8238.2011.00678.x>
- Waha K, Huth N, Carberry P, Wang E (2015) How model and input uncertainty impact maize yield simulations in West Africa. *Environ Res Lett* 10:24017. <https://doi.org/10.1088/1748-9326/10/2/024017>
- Waongo M, Laux P, Traoré SB, Sanon M, Kunstmann H (2013) A crop model and fuzzy rule based approach for optimizing maize planting dates in Burkina Faso, West Africa. *J Appl Meteorol Climatol* 53:598–613. <https://doi.org/10.1175/JAMC-D-13-0116.1>
- Weedon GP, Balsamo G, Bellouin N, Gomes S, Best MJ, Viterbo P (2014) The WFDEI meteorological forcing data set: WATCH forcing data methodology applied to ERA-interim reanalysis data. *Water Resour Res* 50:7505–7514. <https://doi.org/10.1002/2014WR015638>
- Wolf J (2012) User guide for LINTUL5: simple generic model for simulation of crop growth under potential, water limited and nitrogen, phosphorus and potassium limited conditions. Wageningen UR, pp 1–63. <https://library.wur.nl/WebQuery/wurpubs/reports/444719>
- Wolf J, Ouattara K, Supit I (2015) Sowing rules for estimating rainfed yield potential of sorghum and maize in Burkina Faso. *Agric For Meteorol* 214–215:208–218. <https://doi.org/10.1016/j.agrformet.2015.08.262>
- Xiao W, Sun Z, Wang Q, Yang Y (2013) Evaluating MODIS phenology product for rotating croplands through ground observations. *J Appl Remote Sens* 7:073562. <https://doi.org/10.1117/1.jrs.7.073562>
- Zhang X, Friedl MA, Schaaf CB, Strahler AH, Hodges JCF, Gao F, Reed BC, Huete A (2003) Monitoring vegetation phenology using MODIS. *Remote Sens Environ* 84:471–475. [https://doi.org/10.1016/S0034-4257\(02\)00135-9](https://doi.org/10.1016/S0034-4257(02)00135-9)
- Zhao G, Siebert S, Enders A, Rezaei EE, Yan C, Ewert F (2015) Demand for multi-scale weather data for regional crop modeling. *Agric For Meteorol* 200:156–171. <https://doi.org/10.1016/j.agrformet.2014.09.026>
- Zinyengere N, Crespo O, Hachigonta S, Tadross M (2015) Crop model usefulness in drylands of southern Africa: an application of DSSAT. *South African J Plant Soil* 32:95–104. <https://doi.org/10.1080/02571862.2015.1006271>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.