Investigating the effects of drought and lockdowns on smallholder and commercial agricultural production in KwaZulu-Natal using remotely sensed data

N.C. Fouotsa Manfouo *, D. Von Fintel

Department of Economics, Stellenbosch University, Private Bag X1 Matieland, Stellenbosch, 7602, South Africa

A R T I C L E   I N F O

Keywords:
Agriculture
Cloud computing
Drought
Google Earth engine
LandSat8
Lockdown
Remote sensing
Sentinel2

A B S T R A C T

Not many efforts have been made so far to understand the effects of both the 2015–2016 drought and the 2020 lockdown measures on the agricultural production of smallholder vis-a-vis commercial farmers in KwaZulu-Natal. Google Earth Engine, and random forest algorithm, are used to generate a dataset that help to investigate this question. A regression is performed on double differenced data to investigate the effects of interest. A k-mean cluster analysis, is also used to determine whether the distribution patterns of crop production changed with drought and disruption of agricultural production input. Results show that: (1) droughts affected the agricultural production of both areas similarly. Crop cover declined in both areas for one season after droughts were broken. Then recovery was driven by greener, more productive crops rather than the expansion of crop area. (2) The response of both areas to the COVID-19 lockdown was also similar. Both smallholder and commercial areas’ Normalised Difference Vegetation Index – a proxy for crop vitality – improved in response to regulations favourable to the sector and improved rainfall. No significant adjustments in crop cover were observed. Production therefore changed primarily at the intensive margin (improved productivity of existing croplands) rather than the extensive (changing the extent of land under cultivation). (3) Cluster analysis allows for a more granular view, showing that the positive impact of lockdowns on agriculture were concentrated in areas with high rainfall and close proximity to metropolitan markets. Both smallholder and commercial farmers therefore are reliant on market access together with favourable environmental conditions for improved production.

1. Introduction

There exists a pronounced dichotomy in the farming system of KwaZulu-Natal: on the one hand, a very dynamic commercial component, on the other hand, a weaker smallholding farming system (Kirsten and Van Zyl, 1998; Pienaar, 2013).

The commercial sector is large scale and capital intensive (Nwafor and van der Westhuizen, 2020), and the leader of the agricultural economy, with exportations, sustainable development strategies; and hence perceived as the thriving farming sector (Tshuma, 2014). This success is attributed to the historical long-term support provided by the government during the apartheid era (Tshuma, 2014).

The improvement of the rural and smallholding agricultural sector only started after the advent of democracy in 1994 (Pienaar, 2013; Thamaga-Chitja and Morojele, 2014). Both the private and governmental efforts now focus on capacity building of smallholders, and streamline support services through crop land extension, training and mentoring of rural farmers, as well as the provision of credit facilities, services, to achieve job creation and food security (Fanazodo and Ncube, 2018). The programs aim to extend mechanised agriculture on idle land (Terblanche, 2013; Mvelase, 2016). After the start of the smallholding development policies, many studies, including the 2017 food security report for South Africa (StatsSA, 2022), support the argument that small-scale farming could be as viable, profitable and as efficient as large-scale farming (Beinart and Delius, 2018). However, smallholders remain vulnerable to climate hazards (drought) and other shocks (Careslsen et al., 2021). This is why the literature on farmers vulnerability, tends to focus on smallholders. There is no recent evidence, in the form of a comparative study, that may investigate whether this historical fact has evolved. The 2015-2016 drought, as well as the potential

* Corresponding author.
E-mail address: ncfouotsamanfouo@sun.ac.za (N.C. Fouotsa Manfouo).

https://doi.org/10.1016/j.heliyon.2022.e11637
Received 25 March 2022; Received in revised form 1 July 2022; Accepted 10 November 2022

2405-8440/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
disruption of the agricultural input supply chain in 2020 due to lockdown measures may help to compare the vulnerability of commercial and smallholder farms, in terms of their agricultural production, namely crop area size and crop productivity. 

Farming output is closely related with temperature and the amount of rainfall received during growth stages (Ray et al., 2018). They also affect decisions on crop area size, because less water availability implies less soil moisture for healthy crop growth (Howden et al., 2007; Kang et al., 2009; Pais et al., 2014; Liang et al., 2017; Ray et al., 2018).

Agricultural production is also affected by COVID-19 pandemic restrictions on movement (from 21st March 2020 till end of year and beyond), through the disruption of the supply chain of the agricultural input, and the distribution channels of agricultural production (Barrett, 2020; Meyer et al., 2022). This inability of farmers to receive agricultural inputs in time or in sufficient quantities affects healthy crop growth, as well as crop production output and crop area (Mthembu et al., 2022). The inability to distribute the agricultural produce leads to the increase of production cost through storage, or complete destruction of perishable goods. This may lead some producers to decrease their production or exit the agricultural market (Tripathi et al., 2021; Meyer et al., 2022). This in turn leads to the reduction of total crop area. So far, no studies have been performed, to compare the effect of drought on agricultural production between commercial and smallholding farms. Existing studies are skewed towards either analysing agricultural production in a specific farm type (Dube and Jury, 2000; Mpandelile et al., 2015; Bahtá et al., 2016; Lottering et al., 2021), a specific crop type (Araujo et al., 2016; Adisa et al., 2019), livestock production (Matlou et al., 2018; Vetter et al., 2020) or are general in the analysis (Phokele and Sylvester, 2012; Baudoin et al., 2017; Abdel-Hamid et al., 2020). This is often due to a lack of reliable data that describe both the smallholdings and commercial farming agricultural production systems separately. A similar gap is observed about the effect of lockdown on agricultural production. Existing studies (Kganyago, 2021; Mthembu et al., 2022) respectively discuss the effect of lockdown, in smallholder farms in KwaZulu-Natal; and the effect of COVID-19 on winter crop in two districts of the Eastern Cape. This literature does not consider a separate analysis between crop farming systems. Therefore, this research provides an analysis of agricultural production, which considers the effects of drought and lockdown, in both commercial and smallholding farms, separately. This is done using the former homelands and non-homelands demarcations, as well as generating a remotely sensed land cover dataset, and considering the case study of KwaZulu-Natal. This study also investigates the effect of proximity to metropolitan areas on agricultural production.

Schoeman et al. (2013) and Ngcofe et al. (2019) produced land cover maps for South Africa in 2013 and 2018, respectively. They contributed to establishing the South African National Land Classification (SANLC 2013 & 2018) with 72 classes including crop land types. The number of studies using remotely sensed images for land cover classification in South Africa has increased in the last decade (see Fig. 1) (Schoeman et al., 2013; Odindi et al., 2016; Ngcofe et al., 2019; Gynamfi-Ampadu et al., 2020; Dlamini et al., 2021); those studies have the merit of showing how new methods are useful for diverse land cover classification tasks and NDVI extraction. While these studies have produced a number of land cover datasets in South Africa for specific years, there is a gap in producing crop land cover data, for the 2015-2016 drought years, as well as the 2019-2020 years, suitable to analyse the effect of the agricultural input supply chain disruption, caused by lockdown.

While a national land cover map for 2020 has recently been released, nobody has produced an annual time series applying consistent methodology that allows a high frequency, longer run perspective on crop land cover dynamics and the productivity of crop lands. Such a dataset is important for assessing the impacts of a range of events on crop production, especially when other factors that occur in years with missing data also affect cultivation decisions. Not accounting for these events could misrepresent the more immediate impacts of specific crises.

Rapid assessment with telephonically administered household surveys is not excluded. For example, the National Income Dynamics Study Corona Rapid Assessment Survey (NIDS-CRAM) was effective at monitoring trends in hunger in South Africa at regular intervals during COVID-19 lockdowns (van der Berg et al., 2021). However, these data do not provide early insights into changes in production, which precede the impacts experienced by households.

This study illustrates these advantages during times of droughts and the period of the initial COVID-19 lockdowns in South Africa. Because this study uses the same methodologies at different periods of time, it can compare the relative impacts of climate events (droughts) to market shocks (lockdowns) in a timely manner.

The contribution of this study to the literature is twofold, namely: (1) generate a consistent time series of crop land classification in KwaZulu-Natal, from 2015 to 2020; (2) present a separate description of the effects of both the 2015-2016 drought and the 2020 lockdown measures, on smallholder and commercial farms in KwaZulu-Natal. The remaining part of this research is presented in four (4) sections, namely: (2) Study area and Dataset; (3) Methodology; (4) Results; (5) Conclusion.

2. Study area and dataset

2.1. Study area

The region of interest is the province of KwaZulu-Natal in South Africa, as shown in Fig. 2. It is one of the largest provinces in the country with an area of 94 361 km² (StatsSA, 2021) and employing 141 000 of South Africa’s 810 000 agricultural workers (StatsSA, 2021). A large proportion of KwaZulu-Natal’s rural population depends on smallholder farming for food security, though access to foods purchased on the market is also a core component of households’ nutritional intake (Adley et al., 2004). On the one hand, many smallholder farmers produce rain-fed crops for own consumption. This type of cultivation is therefore susceptible to drought events, but may not be as responsive to market shocks. This type of farming is mostly located in South Africa’s so-called former “homelands”. These demarcations are a historical remnant of the

---

**Fig. 1.** Survey of previous studies performed on South Africa, using remote sensing data: Figures were obtained by a keyword search on Google Scholar using “remote sensing data for land classification in South Africa”.

**Fig. 2.** KwaZulu-Natal.
apartheid policy of separate development, and reinforced the limits on land ownership put in place by the 1913 Land Act. Households continue to have collective land tenure today. These areas have been historically marginalised by non-democratic governments, and continue to suffer high levels of poverty and - at times - high rates of hunger (Noble and Wright, 2013). On the other hand, areas beyond the homelands are dominated by commercial, mechanised, irrigated farming and private land tenure. These farmers cultivate at scale for the market. While irrigation may mitigate against some of the impacts of drought, these large farmers are arguably more susceptible to changes in market conditions (such as those caused by lockdown measures) than smallholders in the homelands. Separating analysis by homelands and non-homelands regions therefore enables a better characterisation of the market and climate conditions under which food security is supported by commercial integration or by consumption from own production in the region.

2.2. Field data

Georeferenced land cover points for the year 2018 were obtained from GEOTERRAIMAGE (Ngcofe et al., 2019). They identified the land class of 6415 geolocations throughout the country, from which national land cover maps were developed. 33 of the 72 classes in the national land cover data are represented in the field data, and 9 of these classes correspond to various crop types in South Africa. Their study performed a classification of 20 m × 20 m grid cells using Sentinel2 data. The field data was used to validate the overall reliability of their classification task. This study also uses that field data to train models and verify estimates. Fig. 3 is a graphical representation of field data across the country.

2.3. Satellite data

This study combined satellite images (Sentinel2 and Landsat8) from 2015 to 2020. They are remotely sensed and capable of detecting different features of crops (Ban, 2003; Veloso et al., 2017; Van Tricht et al., 2018). Davidson et al. (2017) reported that combining Sentinel2 and Landsat8 data improved crop classification accuracy by up to 8% compared to only using Landsat8. Rozinat et al. (2008) support the use of time series satellite images to perform cropland classifications of high accuracy.

This study uses Landsat8 as the input for image classification (Table 1 presents its spectral bands). A standard Top Of Atmosphere (TOA) calibration is applied and relies on the United States Geological Survey (USGS) remotely sensed data. NDVI is also computed from Landsat8.

Sentinel2 is a Satellite launched by the European Space agency. It is equipped with cameras that record images in thirteen spectral bands, including the visible bands (band 2–4), red edge (RE, band 5–7), near infrared (NIR, band 8), and shortwave infrared (SWIR, band 11–12) with 10 m and 20 m spatial resolutions, respectively.

Rainfall measurements are sourced from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) and contain more than thirty years of data. CHIRPS is a gridded monthly rainfall dataset, and is useful for drought monitoring.

| Table 1. Landsat 8 bands. |
|---------------------------|
| Spectral Bands | Wavelength (μm) | Resolution (m) |
| B2 | band2 - Blue | 0.45 - 0.51 | 30 |
| B3 | band3 - Green | 0.53 - 0.59 | 30 |
| B4 | band4 - Red | 0.64 - 0.67 | 30 |
| B5 | band5 - near infrared (NIR) | 0.85 - 0.88 | 30 |
| B6 | band6 - SWIR 1 | 1.57 - 1.67 | 30 |
| B7 | band7 - SWIR 2 | 2.11 - 2.29 | 30 |
| B8 | band8 - panchromatic | 0.50 - 0.68 | 30 |
| B10 | band10 - thermal IR (TIR) | 10.60 - 11.19 | 30 |
| NDVI | (B4-B3)/(B4 + B3) | -1 - 1 | 30 |

3. Methodology

Fig. 4 gives an overview of the process followed to estimate crop surface area, NDVI and rainfall datasets for 828 electoral wards in KwaZulu-Natal. A detailed explanation of each step is given in §3.1 to §3.4.

3.1. Satellite data preprocessing

Level-1C Sentinel2 images within Google Earth Engine (GEE) (Image Collection ID: COPERNICUS/S2) were used to supplement the field data that was used to train classification models. These images were processed to top of atmosphere (TOA) reflectance values scaled by 10000 through radiometric calibration (Sentinel, 2015), using the cirrus bitmask of QA60 bands. All sentinel photographs where the cloud percentage is greater than 20% cloud cover were filtered out because they reduced classification accuracy. The mosaic optical image was computed from the median bands of the remaining images in each year. The mosaic was used to identify additional training points for crops and non-crops in wards that contained few or no field data. The training dataset was supplemented for three reasons. Firstly, crop types vary significantly by agro-ecological region. Secondly, climatic conditions have differential impacts on the various wards – especially in areas that are drought-prone. Thirdly, the field data were only available for 2018, and it was imperative to include training data from seasons with varying weather conditions. Because this study aims to differentiate between homelands and non-homelands regions, it is essential that predictions are accurate in all parts of the province and in various years. We there-
Flowchart of the steps followed to build a dataset of crop land surface area, NDVI and rainfall datasets at ward level, in KwaZulu-Natal, using GEE, Sentinel2 and Landsat8 remotely sensed data.

Table 2. Visualisation parameters.

| Min  | Max  | Gamma | Bands       | Cloud cover |
|------|------|-------|-------------|-------------|
| 0    | 0.3  | 1.4   | B4_median   |             |
|      |      |       | B3_median   | \leq 20\%   |
|      |      |       | B2_median   |             |

The Landsat8 images were accessed in GEE (Image Collection ID: LANDSAT/LC08/C01/T1_TOA). These products already applied the Top of Atmosphere (TOA) correction, as indicated by the last part of their collection ID (T1_TOA).

Finally, the resulting preprocessed Sentinel2 and Landsat8 scenes were clipped to match the shapefiles of the wards in the province. The parameters used to visualise the resulting preprocessed data are summarised in Table 2. A selected example of the processed satellite image data is given in Fig. 5.

3.2. Features extraction

Appropriate feature extraction is as important as preprocessing for achieving accurate classifications (Ghorbanian and Mohammadzadeh, 2018). A binary feature extraction was performed on the median pixels of 36 images from a pre-processed Sentinel2 mosaic. Separate mosaics were created for each year between 2015 and 2020. Images were only classified into two land use types: crops and non-crops. The reference points for the year 2018, collected from the South African National Land Cover (SANLC 2018) were used and labelled as crop and non-crop features. As noted before, additional training data was collected, and is shown in Fig. 6. Red pins indicate crop pixels and yellow pins represent...
all other non-crop classes. Because Sentinel2 has a higher resolution (10 meters) than Landsat8 (30 meters), this product is used to perform additional feature extraction in wards with either no field data, or few field data points. The additional pixels that were collected were identified by the distinct shapes of fields, being either circular or rectangular, and that have a green colour that is lighter than natural vegetation (see Fig. 7).

Average NDVI was calculated within each ward to proxy for environmental conditions conducive to crop productivity. The B4 and B3 median bands were extracted from the preprocessed annual Landsat8 images to compute NDVI as in equation (1):

$$\text{NDVI} = \frac{B_4 \text{ median} - B_3 \text{ median}}{B_4 \text{ median} + B_3 \text{ median}}$$ (1)

### 3.3. Classification

A single Landsat8 mosaic image of 36 layers was used to perform a random forest classification and to predict crop area for each of the 828 wards in KwaZulu-Natal. The mosaic was preprocessed within GEE, as described in §3.1. The 2018 field data and additional data collected from wards with few or no ground data were used for the training and validation of a classifier for the year of reference (2018). The dataset obtained was randomly split into training (70%) and validation (30%) samples. Random sampling avoids spatial biases (Davidson et al., 2017). Training and validation was repeated to fine tune the random forest algorithm, resulting in 50 to 100 trees, until accuracy was satisfactory. A separate classifier was trained in each ward to perform a binary grid-level classification. The resulting map was visually compared to the preprocessed ward image. If the visual comparison was found satisfactory, then the binary pixels were summed within each ward to obtain the proportion of the ward that was cultivated.

### 3.4. Accuracy assessment

Accuracy assessments of classified crop land use maps for each ward was done both visually and statistically. A statistical check of the classification accuracy, was performed, using the confusion matrix of the classification. This was computed using the validation samples within GEE. Subsequently, various metrics such as proportion accurately predicted and Kappa (κ) coefficients were derived from the confusion matrix. If $\kappa \geq 0.69$, the classification was considered reliable (De Mast, 2007). Visual assessment only followed in cases where the statistical assessment was satisfactory. The high-resolution Sentinel satellite imagery within GEE was used to compare the classification outputs. If the classification was considered unsatisfactory, the training dataset was expanded to improve the classification. This process was performed until the visual accuracy assessment corroborated the statistical comparisons.

### 3.5. Trend analysis

Once the classifications were validated, the analysis moved to understanding differences in the trajectories of crop cover and Normalized Difference Vegetation Index (NDVI) over time. We first distinguished between homeland and non-homeland areas to establish whether drought conditions and market shocks, such as the COVID-19 lockdown, have different impact on areas dominated by commercial and smallholder farmers. A difference in difference analysis (Donald and Lang, 2007) is used to investigate whether there is any statistically significant difference in heterogeneity between homelands and non-homelands before and after the shocks. The significance of such difference helps to assess whether the changes are attributable to the shocks. This method is widely used in the literature of spatial econometrics for impact evaluation (St. Clair and Cook, 2015; Sunak and Madlener, 2016; Wang and Watanabe, 2019; Rousseeuw, 2019; Lee et al., 1995). This case study, is one where differences in both crop area and NDVI are computed for the two groups of interest, in two time periods. Both of the groups receive a treatment (drought and lockdown measures) in the second period. The average crop area and NDVI differences between the homelands and non-homelands after the drought and COVID-19 lockdown measures is subtracted from their average difference before. This should eliminate any biases that could be the result of common patterns over time (Chagas et al., 2016). The difference in differences are given by equations (2), (3) and (4):

$$\Delta \text{Rain}_t = (\text{Rain}_{t, \text{homeland}} - \text{Rain}_{t-1, \text{homeland}}) - (\text{Rain}_{t, \text{non-homeland}} - \text{Rain}_{t-1, \text{non-homeland}})$$ (2)

$$\Delta \text{NDVI}_t = (\text{NDVI}_{t, \text{homeland}} - \text{NDVI}_{t-1, \text{homeland}}) - (\text{NDVI}_{t, \text{non-homeland}} - \text{NDVI}_{t-1, \text{non-homeland}})$$ (3)

$$\Delta \text{Crop}_t = (\text{Crop}_{t, \text{homeland}} - \text{Crop}_{t-1, \text{homeland}}) - (\text{Crop}_{t, \text{non-homeland}} - \text{Crop}_{t-1, \text{non-homeland}})$$ (4)
and the regressions used to investigate their significance are given by equations (5), (6) and (7):

\[
\Delta \text{Rain}_i = \beta_{\text{Rain}i} + \beta_{\text{Rain}i} \cdot \text{homeland} \\
\Delta \text{NDVI}_i = \beta_{\text{NDVI}i} + \beta_{\text{NDVI}i} \cdot \text{homeland} \\
\Delta \text{Crop}_i = \beta_{\text{Crop}i} + \beta_{\text{Crop}i} \cdot \text{homeland}
\]

were, Rain, \text{homeland} and Rain, \text{non-homeland}, stands for the rainfall at time \( t \) in homelands and in non-homelands respectively, at time \( t \); NDVI, \text{homeland} and NDVI, \text{non-homeland}, stands for the NDVI at time \( t \) in homelands and in non-homelands respectively, at time \( t \); and Crop, \text{homeland} and Crop, \text{non-homeland}, stands for the crop proportion at time \( t \) in homelands and in non-homelands respectively, at time \( t \); \beta_{\text{Rain}i}, \beta_{\text{NDVI}i}, \beta_{\text{Crop}i} \) are respectively, the constants and the coefficients related to each variable of interest; \( t \in \{2016, 2020\} \), which correspond to the years of drought and lockdown. The variable “homeland” \( \in \{0, 1\} \), in Equations (5), (6), (7), is a binary variable, such that 1 is homelands, and 0 is non-homelands.

3.6. Clustering

The analysis also considers a data-driven classification of wards into discrete clusters to assess whether the homelands dichotomy is as important as other factors in distinguishing between regions, and to investigate, whether drought and lockdown contributed to any change in the wards characterization. A k-means clustering is used to separate the wards into non-overlapping groups based on crop area, NDVI and rainfall (Likas et al., 2003). Given the set of observations \( X_j \) on the three variables, where \( j \) indexes the 828 wards, and \( t \), the 6 years of interest, such that \( t \in \{2015, 2016, \ldots, 2020\} \), the algorithm partitions the \( \text{card}(i) \times \text{card}(t) \) observations into \( k \) sub-observations, so that the total variance within each sub-group is minimised. The silhouette coefficient, which is the average difference between each point in one cluster from the points in other clusters, provides an idea of how distinct or far apart the clusters are from each other.

4. Results

This section is organised in three parts, namely: (1) validation; (2) the comparison of homelands and non-homelands; (3) cluster characterisation.

4.1. Validation

Fig. 8 presents crop cover classifications of the KwaZulu-Natal region in 2018. The first classification in Fig. 8 is performed using the 1229 validation points obtained from the SANLC 2018 field data. The second classification in Fig. 8 is obtained from the 3057 points which also include the data augmented with Sentinel2 training points. Visual observations using the validation set indicate that the model trained with the augmented data is more accurate at classifying crops. The confusion matrix (see Table 3) shows that only 9% of pixels (64 of 674) are incorrectly classified when only the field data are used for training the classifier. Misclassification reduces to 8% (105 of 1310 pixels) when the augmented data are introduced.

Figs. 9, 10, 11, 12, 13 and 14 display sample predictions from the trained model for one ward (ID:522010004); predictions for the other years (2015-2020) using the 2018 model are also shown. Visual inspection corroborates the statistics provided in the confusion matrix (see Table 3). When the augmented training data are used, average accuracy is high at 89.09%. The Kappa coefficient is 69.96%, which is above the acceptable threshold of 0.69 (see Table 4), as explained in the theory (see §3.4). The Kappa coefficient falls to 0.647 when the classifier is only trained on field data. This emphasises the utility of augmenting the data with additional training points. The visual observations and Kappa statistics give similar conclusions across the 828 wards.

The heatmaps of rainfall and NDVI distribution across KwaZulu-Natal (see Fig. 15), indicate that the spatial distribution of both indicators is positively correlated, which is theoretically intuitive (Wang et al., 2003). However, the correlation with crop cover appears to be weaker, so that many other factors also determine cultivation decisions.

4.2. Comparison of homelands and non-homelands

The classified data allow for an annual longitudinal analysis of 828 wards in KwaZulu-Natal between 2015 and 2020. The study period starts at the peak of a severe drought in the province, as confirmed by data from the South African Weather Service (SAWS) (SAWS, 2021). The period ends during the lockdowns enforced by the South African government to stop the expansion of COVID-19. Fig. 16 and Table 5 present summary statistics from the dataset developed in this study, including rainfall, NDVI and crop production levels. The table also shows cross-sectional mean differences between homeland and non-homeland regions at each point in time, as well as year-on-year differences that
Fig. 9. Observed & Classified Ward ID: 52201004 (2015).

Fig. 10. Observed & Classified Ward ID: 52201004 (2016).

Fig. 11. Observed & Classified Ward ID: 52201004 (2017).

Fig. 12. Observed & Classified Ward ID: 52201004 (2018).
trace growth paths within regions. Finally, differences in the time differences across homeland and other regions are also presented. This quasi-experimental approach traces whether the year-on-year growth trajectory differs across regions, and is regularly used to approach causal interpretation of the impacts of intervening events between two
Fig. 15. Rainfall, NDVI and crop heatmaps.

Fig. 16. Summary of data collected & differentiation between homelands and non-homelands.

4.2.1. The effect of drought

The impacts of drought clearly manifest in the data: in 2015 rainfall and NDVI were at their minimum for the period under study in both homeland and other regions. Homelands and other areas were equally affected: there were no significant differences in rainfall and NDVI across region types during the 2015 drought. However, the impacts of the drought still lingered, as 2016 planting decisions were based on prior years’ rainfall. In both region types, area under crop production fell significantly (by around 5 percentage points) between 2015 and 2016. The difference in this difference is insignificant, emphasising that the drought had equal effect on the cultivation decisions of farmers operating in traditional smallholder areas and commercial farming regions. This decline in area under cultivation occurred despite significant increases in rainfall in the 2016 season ranging from 84 mm more rain in homeland areas to 113 mm in non-homeland areas. NDVI also increased during this time, though the insignificant difference in difference shows that homeland and non-homeland areas were on similar post-drought growth trajectories. These observations emphasise that planting decisions are partially determined by conditions in prior seasons, but that good rains nevertheless correlate with improved yields, as proxied by NDVI. In immediate post-drought seasons, total food production is therefore more dependent on the recovery of productivity of existing fields, and not on growing the extent of land under cultivation. These results corroborate with Hawkins et al. (2022), where drought is perceived as the main driver behind crop abandonment and low crop production. The decrease and recovery phases observed in this study correlates with Vetter et al. (2020) where about 29% of goats herds were lost in 2015 – 2016, followed by a recovery phase, from 2017. This study also corroborates with Ruwanza et al. (2022), where a survey of eighteen studies on the effect of drought in smallholder farms in South Africa, shows that grazing land and crop vegetation loss were among the most experienced consequences of the 2015 – 2016 drought.

The results also show that both commercial and smallholder farmers are sensitive to drought conditions and that the capacity to irrigate crops did not give commercial areas an advantage in coping with the drought. While irrigation allowed commercial farmers to cultivate proportionally more land on a permanent basis (the significant difference across homeland and other regions ranges between 5.6 and 10.5 percentage points across all years), irrigation did not put commercial farming areas on a faster growth trajectory in post-drought years. Similar observations hold for NDVI. While NDVI recovered rapidly in 2016 and 2017 in both region types, the differences in differences were modest, and only significant at 10% in 2017. Furthermore, there are no systematic regional differences in NDVI across time. Our results therefore highlight that the commercial farming advantage is always present, but
that it is driven by larger expanses of land rather than greener fields; however, these advantages do not extend to how the different farming regions adapt to environmental stressors.

4.2.2. The effect of lockdown

Our attention shifts to the effect of lockdown on the proportion of land under cultivation. South Africa’s economy contracted by 7% in the wake of COVID-19 lockdowns in 2020. Despite this economic shock, the agricultural sector expanded by 13% (BFAP, 2021). Agricultural growth was firstly a response to improving rainfall (as seen in Fig. 16) and a recovery from stock disease. Additionally, the sector was allowed to operate as an “essential service” during South Africa’s lockdown, and was shielded against the worst of the economic fallout. Agriculture had a comparative advantage during the lockdown because of the government’s emphasis on continued food provision during the times when households and other firms faced restrictions that were stricter than in most countries (Gustafsson, 2020). Market shocks (either positive or negative) should primarily affect commercial farmers in non-homeland areas.

Fig. 16 indicates that crop area increased in both the homelands and non-homelands between 2019 and 2020. However, Table 5 shows that these changes were not statistically significant. On the other hand, rainfall and NDVI increased significantly over this period. On the other hand, the differences in the differences are not significant in 2020, showing that the growth path in rainfall and NDVI were identical in both areas. Nevertheless, NDVI ended significantly higher in non-homelands areas in 2020.

Our results support the hypothesis that – as with the impact of drought – lockdown-related market shocks did not affect agricultural production differently in homeland and non-homeland regions during the COVID-19 crisis. The extent of crop cultivation remained stable and did not increase in response to better rainfall in both homelands and non-homelands, so that there was no growth at the extensive margin. Instead, higher rainfall was associated with improvements in NDVI. This trend in crop quality presents as one likely explanation for agricultural growth over the period – there was growth at the intensive margin. However, this NDVI growth was not more rapid in commercial farming areas – non-homelands – compared to smallholder farming areas – homelands. Nevertheless, commercial farming areas did end the period with a significant NDVI advantage in 2020. NDVI and growth in agricultural production were therefore not primarily related to more broader market conditions. Instead, agricultural conditions, including better rainfall, present as a better explanation for the improvements in crop productivity and why commercial farmers had a good year during lockdown (BFAP, 2021). More generally, our results show that agricultural output is more closely linked to environmental factors than broader demand factors in the economy. The results of this research are in line with Meyer et al. (2022), which feature how COVID-19 lockdown strategies enforcements have disturbed the distribution of agricultural production inputs and output, but not the agricultural production, due to its classification as an essential service (BFAP, 2021).

4.3. Cluster characterisation

Rather than limiting the study to the homeland dichotomy, k-means cluster analysis was used to identify data-driven geographic differences within KwaZulu-Natal. It’s purpose was to investigate whether there exist, another characterisation of farms, that influences their vulnerability to drought and lockdown restriction measures. The variables used to perform the analysis were rainfall, NDVI, proportion of crop cover and time period. Time differences were included to allow for dynamic changes in the classification of each ward into a specific cluster, depending on prevailing local conditions. Three clusters were identified in the data silhouette analysis (Rousseeuw and Silhouettes, 1987) showing that the clusters are distinct (see the bottom right panel of Fig. 17). The other panels in Fig. 17 characterise the clusters by their median rainfall, NDVI and proportion of area under cropland. The first two clusters have comparatively high rainfall and NDVI which remain stable over time. Only the first cluster expanded its crop cultivation during the 2020 lockdown, differentiating it from the second cluster. By contrast, the third cluster has low rainfall and NDVI, but with similar area under crop production to the second cluster.

Fig. 18 shows the evolution of each ward’s cluster membership over the period of analysis. In 2015, at the peak of the drought, most of the province was classified in cluster 3, with low rainfall and NDVI. As the drought eased, coastal and mountainous areas at the edges of the province move into cluster 2 with improvements in rainfall and NDVI, but not in land area under crop production. Fig. 15 confirms these observations. Importantly, the areas that moved into this cluster were located in both smallholder homeland regions and commercial farming areas in other parts of the province. It is therefore clear that environmental and market factors – rather than farming type – distinguish the regions from each other. As time progressed, and the drought eased, low rainfall cluster 3 started to retreat even more, giving way to medium rainfall cluster 2. As the country headed into lockdown in
2020, cluster 2 emerged strongly around the epicentres of Durban and Pietermaritzburg. This was the only cluster to increase production on the extensive margin in 2020, even if rainfall remained in the mid ranges. These results do point to a local response to lockdown measures that favoured the agricultural sector. Some smallholder homeland areas were also included in this group. This explains why homelands are not very different from non-homelands. However, ward’s proximity to metropolitan markets and favourable rainfall are more important for the improvements in crop production at the intensive margin than if cultivation happened in commercial farming areas.

5. Discussion

This study contributes to the literature on using remotely sensed land classification to estimate the impacts of events on agricultural production. Using a common machine learning methodology, this study shows how generating a consistent time series of cropland allows for high time frequency and geographically disaggregated analysis of weather and market events. It shows that random forest classifiers can be used reliably on one reference dataset – generated by GEOTERRAIM-AGE for 2018 – and be augmented to extrapolate the analysis from 2015 to 2020.

This application confirms the benefits of combining Landsat8 and Sentinel2 data to classify cropland (Davidson et al., 2017; Rozinat and Van der Aalst, 2008). The validation and testing of the random forest classifier required a combination of statistical criteria on reference data, as well as visual observation of Sentinel8 tiles. A robust framework was used to ensure that the classifications match visual observations and can be replicated across other years of interest (see Fig. 4). Using augmented data achieved higher classification accuracy (> 89%) and Kappa index (> 69%) illustrating that including more reference data, which provides improved coverage of heterogeneous agroecological zones in the study area, contributes to higher quality crop classifications. Though the data collected from visual observation of tiles in Google Earth Engine are in general not as credible as those collected in the field, studies performed previously support the proposal that easy, and fast methods to land cover mapping are useful (Xiong et al., 2017).

This work also identified a number of factors that could improve future studies of this nature. More feature engineering, such as adding NDVI, SAVI, BSI, as well as elevation data to the Landsat8 tile may contribute to increasing the performance of crop classification. Investigating the differences in using mean and median bands for classification will help to find out the most suitable bands to use for crop classification in KwaZulu-Natal. In particular, median bands are likely to provide a more stable series when outlier years introduce anomalies to the data. Outliers are likely to become more commonplace as climate change introduces new volatility in rainfall and crop cover, and should be accommodated to produce smoother series for producing long-run trends. However, because the mean is more sensitive to outliers, it could better capture smaller movements that respond to climate shocks, and capture that kind of volatility better. Using additional features will also allow the characterization of multiple crops, so that multi-class classification is possible in the year of reference and can be generalised to other years of interest.

Finally, the descriptive analysis showed how using remotely-sensed data can contribute to understanding differences in the effects of policies across different farming areas – in particular, commercial and smallholder farming regions.

The analysis can also be used to assess how different types of events – natural shocks and market shocks – can be monitored in lower administrative units, especially during times of sudden economic crisis such as COVID-19 lockdowns. More advanced quasi-experimental econometric analysis can be used to extend this descriptive analysis, and to measure the impact of a range of events on the agricultural sector.

The negative relationship between drought and crop production was confirmed in this study. This is intuitive and expected from a conceptual perspective. The non-significant difference of its effect between homelands and non-homelands, is however unexpected, given that non-homelands farms are supposedly commercial farms. As such, they have better irrigation capabilities (Njoko, 2014), which would have contributed to softening the negative effect of drought in non-homelands, as compared to its effects in homelands. This result may suggest a change in the typology of farms in both homelands and non-homelands. This may be indicating that the former characterization of these two entities, is becoming less relevant, especially in a context where more and more so-called small-scale commercial farms are being developed in homelands, in KwaZulu-Natal (Agergaard and Birch-Thomsen, 2006).
The link between the disruption in the agricultural input supply chain, due to the COVID-19 lockdown enforcement measures is non-significant, both in homelands and non-homelands. This is intuitive with the fact that though the lockdown measure were very stringent in South Africa, the consideration of the agricultural sector as an essential service was effective (BFAP, 2021). Thus, the agricultural sector only suffered from the international distribution bottleneck, due to restrictions like flight bans to certain destinations (Meyer et al., 2022). This affected crop output distribution, through the exportation market. More investigations on the link between the distribution bottleneck and the agricultural crop production and productivity can be performed with the extrapolation of land classification to 2021; considering a more precise characterisation of commercial and smallholding farms, other than homelands and non-homelands demarcations.

6. Conclusion

This study expanded our understanding of how smallholder and commercial farmers responded to environmental and market shocks (namely drought and lockdown). Analysis of the data showed that the drought continued to influence farming decisions, even when rainfall and NDVI started to recover. Area under production fell to a minimum in 2016, while other indicators signalled relief from the drought. These patterns were common to both commercial and smallholder farming regions. Immediate recoveries from drought therefore depend on the intensive margin (increases in the quality of crops) rather than expanded area under cultivation, while hesitant farmers decided whether it was worth planting new fields. While irrigation, private land rights and access to other inputs could potentially explain the permanent regional differences, these factors do not mitigate the impacts of drought more prominently in commercial farming areas than in areas dominated by smallholders.

These patterns were also studied in the context of a market shock instead of an environmental shock. South Africa’s 2020 lockdown regulations positioned the agricultural sector as an essential sector to maintain food security. Coupled with good rains, the regulations (that were relatively favourable to agriculture) allowed the sector to defy the pronounced declines recorded in the rest of the economy. Our analysis showed that this increase was dominated by the intensive margin – with greener, high yielding crops – rather than the extensive margin – the area under cultivation remained stable. Importantly, these changes were not restricted to commercial farming regions which are more likely to sell their produce on the market.

Cluster analysis provided alternative regional classifications to unpack this result. This exercise revealed that both homeland and non-homeland regions close to metropolitan areas contributed to the increase in NDVI in response to higher rainfall, but conditional on being closely located to metropolitan areas. There is also some evidence that these areas expanded their area under cultivation in response to these favourable circumstances. Market integration – coupled with favourable environmental conditions – may therefore nevertheless play an essential role in allowing both smallholder and commercial farmers to capitalise on positive economic shocks. These results support previous analyses which illustrate the importance of market integration for allowing smallholders to achieve similar production patterns to commercial farmers (Markussen, 2008).

This study also serves as an example of how remote sensing technologies can improve the study of local economic activity during crises and recoveries. This paper is also a call for closer collaboration between remote-sensing specialists and social scientists.

This paper could be improved in many ways, namely using the 2021 satellite data in the crop classification task, and using more features engineering techniques, to select the best predictors, that would improve the classification accuracy. Discontinuity analysis could also be combined with cluster analysis, in order to investigate further, the effect of closeness to metropolitan areas.

Declarations

Author contribution statement

Noe Careme Fouotsa Manfouo: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Dieter Von Fintel: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Funding statement

Noe Careme Fouotsa Manfouo received funding from Stellenbosch University’s Research Committee, Faculty of Economic and Management Sciences and the School for Data Science and Computational Thinking.

Data availability statement

Data will be made available on request.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

Abdel-Hamid, Ayman, Dubovský, Olena, Graw, Valerie. Greve, Klaus, 2020. Assessing the impact of drought stress on grasslands using multi-temporal SAR data of Sentinel-1: a case study in Eastern Cape, South Africa. Eur. J. Remote Sens. 53 (sup2), 3–16. Taylor & Francis.

Adey, S., Kotze, D.C., Bijrkenoff, F.H.J., 2004. Small scale farming in KwaZulu Natal: experiences from some promising pockets. In: Wiskerke, J.S.C., van der Ploeg, J.D. (Eds.), Seeds of Transition: Essays on Newtly Production, Niches and Regimes in Agriculture. Van Gorcum, Assen, pp. 203–225.

Adiza, Omolola M., Botai, Joel O., Adeola, Abiodun M., Botai, Christina M., Hassen, Abubeker, Darkey, Daniel, Tesfamariam, Eyob, Adisa, Abidemi T., Adisa, Alex F., 2019. Analysis of drought conditions over major maize producing provinces of South Africa. J. Agric. Meteorol. 75 (4), 173–182. The Society of Agricultural Meteorology of Japan.

Agergaard, Jytte, Birch-Thomsen, Torben, 2006. Transitional rural landscapes: the role of small-scale commercial farming in former homelands of Post-Apartheid KwaZulu-Natal. Geogr. Tidsskrift Danish J. Geogr. 106 (2), 87–102. Taylor & Francis.

Araujo, Julio A., Abiodun, Babatunde J., Crespo, Olivier, 2016. Impacts of drought on grape yields in Western Cape, South Africa. Theor. Appl. Climatol. 125 (1), 117–130. Springer.

Ban, Yifang, 2003. Synergy of multitemporal ERS-1 SAR and Landsat TM data for classification of agricultural crops. Can. J. Remote Sens. 29 (4), 518–526. Taylor & Francis.

Baudoin, Marie-Ange, Vogel, Coleen, Nortje, Kirstyn, Naik, Myra, 2017. Living with drought in South Africa: lessons learnt from the recent El Niño drought period. Int. J. Disaster Risk Reduct. 23, 128–137. Elsevier.

Barrett, Christopher B., 2020. Actions now can curb food systems fallout from COVID-19. Nature Food 1 (6), 319–320. Nature Publishing Group.

Bahta, Y.T., Jordaan, A., Muyambo, F., 2016. Communal farmers’ perception of drought in South Africa: policy implication for drought risk reduction. Int. J. Disaster Risk Reduct. 20, 39–50. Elsevier.

Beinart, William, Delius, Peter, 2018. Smallholders and Land Reform: A Realistic Perspective. Centre for Development and Enterprise (CDE).

Bureau for Food and Agricultural Policy, 2021. Perspectives on agriculture’s GDP performance in 2020. Available [online]. [https://www.bfap.co.za/perspectives-on-agricultures-gdp-performance-in-2020/]. (Accessed 3 November 2021).

Card, David, Kruger, Alan B., 1994. Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania. Am. Econ. Rev. 84 (4), 772–793.

Carelse, C.P.R., Ncube, Bongani, Fanadzo, Morris, 2021. Classification and characterisation of smallholder farmers in South Africa: a brief review. South Afr. J. Agric. Ext. 49 (2), 97–106.

Chagas, André L.S., Azzoni, Carlos R., Almeida, Alexandre N., 2016. A spatial difference-in-differences analysis of the impact of sugarcane production on respiratory diseases. Reg. Sci. Urban Econ. 59, 24–36. Elsevier.
