Estimating the impact of crop diversity on agricultural productivity in South Africa

Cecilia Bellora[1], Élodie Blanc[2], Jean-Marc Bourgeon[3,4], and Eric Strobl[4]

[1]CEPII, 75007 Paris, France
[2]MIT Joint Program on the Science and Policy of Global Change
[3]INRA, UMR 210 Économie Publique, 75005 Paris, France
[4]École Polytechnique, Department of Economics, 91128 Palaiseau Cedex, France

1 Introduction

Diversity plays a key role in the resilience to external stresses of farm plants and ani-mals. In particular, crop species diversity increases productivity and production stability (Tilman et al., 2005; Tilman & Downing, 1994; Tilman et al., 1996) in the sense that the probability to found at least one individual that resists to an adverse meteorological phenomenon (for example a drought or a heatwave), or pests and diseases, increases with the diversity within a population. Furthermore, the larger is an homogeneous population, the larger is the number of parasites that use this population as a host and therefore the larger is the probability of a lethal infection (Pianka, 2011). Finally, diversity also allows for species complementarities and therefore a more efficient use of natural resources (Loreau & Hector, 2001). For these reasons, the expansion of monocultures observed in many modern agricultural systems is presented as one of the causes of their increased vulnerability and, consequently, of the increased use of agrochemicals (Landis et al., 2008). In short, crop biodiversity has the potential to enhance resistance to strains due to biotic and abiotic factors and to improve crop production and farm revenues.

Our paper empirically investigates the role of crop biodiversity on agricultural production and on farmer's exposure to risk. We build a probabilistic model to describe crop

survival and productivities according to diversity. From this analytic model, we derive reduced forms that are estimated using data on South African agriculture.

Previous empirical investigations on the role of biodiversity on production has produced sometimes contrasted results. Di Falco & Chavas (2006) report a positive impact of biodiversity on average production mean and on the reduction of its variability, using data on wheat production in Sicily (Italy), Di Falco & Chavas (2009) find positive impacts of crop biodiversity in barley production in Ethiopia. Carew et al. (2009) also find a positive impact of crop diversity on average production of wheat in Canada. Smale et al. (1998), examining the impact of wheat genetic diversity on wheat yield in Pakistan, find a positive impact of diversity in rain fed regions while in irrigated regions higher area concentration on few varieties is associated with higher yields. These results were obtained using either aggregate data (Smale et al., 1998; Carew et al., 2009) which do not allow to control for farm characteristics, or farm-level data from household or farm survey (Di Falco & Chavas, 2006, 2009) with relatively fewer observations. Furthermore, all these studies used pure econometric methods, mainly moment based approaches, Di Falco & Chavas (2006, 2009) stressing the crucial role played by skewness in addition to mean and variance. The functional forms are totally disconnected from the ecology literature and therefore do not allow to go into deep details on the way biodiversity impacts productivity. In economic literature, models of endogenous interaction between biodiversity and crop production have been developed in theoretical papers that analyse the role and the value of biodiversity against specialisation on the most productive crops (Weitzman, 2000; Brock & Xepapadeas, 2003; Bellora & Bourgeon, 2015) but have never been coupled with empirical investigations.

Our results contribute to the existing literature in three main ways. First, we confirm that diversity has a positive and significant impact on produced quantities.[1] An increase in biodiversity is equivalent to a third of the benefits of a comparable increase in irrigation, where irrigation is known to be an important impediment to crop productivity in South Africa, due to unreliable precipitation. Biodiversity also contributes to the reduction of production variance and, thus, to the stabilisation of agricultural revenues. Second, we use a very large dataset of highly disaggregated data at the field level derived from satellite imagery. Thanks to our large dataset, we show that biodiversity has mainly a local impact: biodiversity is a significant predictor of crop production on perimeters having a radius smaller than 1.75 kilometre.[2] Third, our approach is built on the results of ecology literature. Our probabilistic model makes explicit the relationship between

[1]Note that previous studies mainly looked at genetic diversity i.e. intraspecific diversity, while we rather analyse interspecific diversity, i.e. crop diversity.

[2]We quantify biodiversity using an index taken from the ecological literature, based on species richness (i.e. the total number of species) and their relative abundance, the Shannon index. This index captures the fact that biodiversity is high when the total number of species is large and the distribution of their relative abundances is homogeneous.

biodiversity, biotic and abiotic factors that affect agricultural production. It represents how biodiversity impact agricultural production. Stochastic shocks affecting agricultural production are endogenous, in accordance with ecology findings.

In the remainder of the paper, the theoretical model that motivates our empirical investigations is developed in section 2 and section 3 details its empirical implementation. Then, the database on South African agriculture is presented in section 4. In section 5 , we use these data to investigate the impact of crop biodiversity on production levels and variability.

2 The model

A very robust stylised fact in ecology describes the impact of biotic factors on agricultural production: the more area is dedicated to the same crop, the more pests specialise on this crop and the higher is the frequency of their attacks (Pianka, 2011). Relying on this stylised fact, we build a general probabilistic model of crop production where crops are affected by both abiotic (i.e. weather, water availability, soil properties. . .) and biotic (i.e. pests) factors causing pre-harvest losses.[3] More precisely, we consider that the total agricultural production depends on the survival probability of each crop, which is directly linked to the probability of a pest attack. The frequency at which pest attacks occur is linked to the way crops are produced: the more diverse are crops, the lower is the probability of a pest attack, the higher is the survival probability and therefore the higher is the expected agricultural production.

To describe these impacts, we adopt a beta-binomial distribution, which is often used in ecology and plant physiology literature to describe spatial distributions that are not random but clustered, patchy or heterogeneous (Hughes & Madden, 1993; Shiyomi et al., 2000; Chen et al., 2008; Bastin et al., 2012; Irvine & Rodhouse, 2010). We assume that a region (or a country) produces Z different crops on I fields of the same size, each field being sowed with one crop only. Characteristics of field i are gathered in vector $X_i = (x_{i1}, \ldots, x_{iK})$ and are related both to abiotic factors (for example, water availability through irrigation, soil quality, field location. . .) and biotic factors (in particular the state of crop biodiversity around the field we consider). In particular, X_i contains information on the way crops are cultivated and on biodiversity conditions. Depending on the crop cultivated, each field is divided in $n(z)$ patches that are subject to potential lethal strains due, for example, to adverse meteorological conditions or pathogens. We suppose that

[3]Losses due to biotic factors can be significant. Oerke (2006) finds that, during the 2001-2003 period, ~~without crop protection, losses in major c~~rops due to pests were comprised between 50% and 80%, at the world level. Thanks to crop protection, they fall between 29% and 37%. Similar results are found for the US by Fernandez-Cornejo et al. (1998).

a patch on field i is destroyed with probability $1-\lambda_i$ from one (or several) adverse condition, and that otherwise it produces the potential yield $a(z)$ independently of the fate of the other patches on field i or elsewhere.[4] With $n(z)$ patches, the probability of having t patches within field i unaffected (and thus $n(z)-t$ destroyed) follows a binomial distribution

$$\Pr\{\widetilde{T}_i = t|z,\lambda_i\} = \binom{n(z)}{t}\lambda_i^t(1-\lambda_i)^{n(z)-t}$$

where $\widetilde{T}_i$ is the random variable that corresponds to the number of patches that are indeed harvested among the $n(z)$ patches of field i sowed with crop z. We consider that the survival probability of the patches of a field, λ_i, is identically and independently distributed across patches. However, this probability may vary across fields of the same crop (we generally have $\lambda_i \neq \lambda_j$ for any couple of fields (i,j) sowed with the same crop): it depends on natural conditions but also on the characteristics X_i of the field. More precisely, denoting by $S_{ui}(z) = e^{\gamma(z)+\boldsymbol{\theta}_u.X_i}$ and $S_{di}(z) = e^{\beta(z)+\boldsymbol{\theta}_d.X_i}$, the survival probability of patches on a given field is a draw from a Beta distribution given by

$$\Pr\left(\lambda_i = \lambda|X_i,z\right) = \frac{\Gamma\left[S_{ui}(z)+S_{di}(z)\right]}{\Gamma\left[S_{ui}(z)\right]\Gamma\left[S_{di}(z)\right]}\lambda^{S_{ui}(z)-1}(1-\lambda)^{S_{di}(z)-1}$$

where $\Gamma(\cdot)$ is the the gamma function, $\gamma(z)$ and $\beta(z)$ are positive parameters that determine the randomness of the survival probability of a patch of crop z absent any field specific effect, and where the vectors $\boldsymbol{\theta}_u$ and $\boldsymbol{\theta}_d$ capture the influence of each field specific effect on the survival probability of crop z. The expected number of patches among $n(z)$ that are harvested on field i is given by $\mathrm{E}[\widetilde{T}_i|X_i,z] = n(z)\psi(z,X_i)$ where

$$\psi(z,X_i) = \mathrm{E}[\lambda_i|X_i,z] = \frac{S_{ui}(z)}{S_{ui}(z)+S_{di}(z)} \tag{1}$$

is the expected probability that a particular patch of field i is harvested given its characteristics X_i. Absent field specific effects ($\boldsymbol{\theta}_u = \boldsymbol{\theta}_d = 0$), the expected resilience of a particular stand of crop is given by $\gamma(z)/[\gamma(z)+\beta(z)]$. An increase in coefficient $\boldsymbol{\theta}_{uk}$ increases this resilience, while an increase in $\boldsymbol{\theta}_{dk}$ diminishes it, the extent of this effect depending on the corresponding field characteristics. The variance of the number of harvested patches on field i is given by $\sigma_i^2 = n(z)V(z,X_i)$ where

$$V(z,X_i) = \psi(z,X_i)[1-\psi(z,X_i)]\{1+[n(z)-1]\rho(z,X_i)\} \tag{2}$$

[4]Obviously, this is a strong assumption. Pests and/or weather do not necessarily totally destroy a patch, but rather affect the quantity of biomass produced. But, in order to maintain tractability, we consider that a patch is either unaffected either totally destroyed, rather than partially affected, by adverse conditions. Thus, our random variable is the number of harvested patches rather than the share of biomass that is lost on each patch.

with

$$\rho(z, X_i) = [1 + S_{ui}(z) + S_{di}(z)]^{-1}.$$

Equation (2) corresponds to the variance of the survival probability of one patch on a field with characteristics X_i. Compared to the Bernoulli distribution, (2) contains an additional term that accounts for the correlation between patches induced by the common distribution of the survival probability, the correlation coefficient being given by $\rho(z, X_i)$.

The production on field i is given by $\widetilde{Y}_i = a(z)\widetilde{T}_i$. It can be equivalently written as

$$\widetilde{Y}_i = \mathrm{E}[\widetilde{Y}_i](1 + \tilde{\varepsilon}_i) \tag{3}$$

where $\mathrm{E}[\widetilde{Y}_i] = a(z)n(z)\psi(z, X_i)$ and $\tilde{\varepsilon}_i = (\widetilde{T}_i - \mathrm{E}[\widetilde{T}_i])/\mathrm{E}[\widetilde{T}_i]$ has a mean equal to 0 and a variance given by

$$\sigma^2_{\varepsilon_i} = \frac{1 - \psi(z, X_i)}{\psi(z, X_i)} \left(\frac{1}{n(z)} + \frac{n(z) - 1}{n(z)} \rho(z, X_i) \right) \approx \frac{1 - \psi(z, X_i)}{\psi(z, X_i)} \rho(z, X_i) \tag{4}$$

when $n(z)$ is large. This variance is mainly due to the correlation between patches on a field that share the same survival probability, captured by $\rho(z, X_i)$. Indeed, λ_i follows a beta distribution but the parameters of the distribution depend on the field characteristics $\mathbf{X_i}$ and are thus different across fields. In other words, with a sufficiently large number of patches on each field, the difference in the quantities produced is mainly driven by field characteristics.

To sum up, the number of patches that are harvested on field i, $\widetilde{T}_i$, follows a beta-binomial distribution determined by the parameters $\gamma(z), \beta(z), \theta_{uk}, \theta_{dk}$. Parameters θ_{uk} and θ_{dk} determine the impact of the field characteristic x_{ik} on $\widetilde{T}_i$, in addition to the parameters $\gamma(z)$ and $\beta(z)$ that are shared by all fields that grow crop z. Depending on the values of θ_{uk} and θ_{dk}, each characteristic x_{ik} can increase or decrease the expected number of harvested patches on field i and skew the distribution of $\widetilde{T}_i$ to the right or to the left, modifying the probability of extreme events like the loss of all the patches in a field.

In the following, we build an empirical strategy to estimate all the parameters that determine the distribution of $\widetilde{T}_i$, the number of patches harvested in a field, and therefore of $\widetilde{Y}_i$, the total production of the field.

3 Empirical strategy

Starting from the probabilistic model, our aim is to estimate the vector of field specific contributions $\boldsymbol{\theta}_u$ and $\boldsymbol{\theta}_d$ and the crop specific parameters $\gamma(z)$ and $\beta(z)$ on the distribution of the survival probability $\widetilde{T}$ using an (extended) logit estimation procedure.

We first have to derive from each field production the corresponding survival probability $\hat{\lambda}_i$. They are obtained by dividing the field production by the potential maximum production $a(z)n(z)$ derived from $Y_M(z) \equiv \max Y_i(z)$ using $a(z)n(z) = (1+\alpha)Y_M(z)$ where $\alpha \geq 0$. We indeed consider that the maximum production observed is lower than the potential production since it is impossible to actually observe the production of a field absent any field specific effect (i.e. with $\boldsymbol{\theta_u} = \boldsymbol{\theta_d} = 0$). Robustness checks are performed by varying α from 1% to 25%. Using a logit regression of the equation

$$\ln[\hat{\lambda}_i/(1-\hat{\lambda}_i)] = \delta + \boldsymbol{\Delta}.X_i$$

for each type of crop, we obtain the estimate $\hat{\delta}(z)$ of $\beta(z) - \gamma(z)$ and $\hat{\boldsymbol{\Delta}}$ of $\boldsymbol{\theta}_d - \boldsymbol{\theta}_u$. From theses result, we derive an expected patch survival rate per field given by

$$\hat{\psi}_i = 1/(1 + e^{\hat{\delta}+\hat{\boldsymbol{\Delta}}.X_i})$$

and a serie of dispersion values

$$\hat{\varepsilon}_i = (\hat{\lambda}_i - \hat{\psi}_i)/\hat{\psi}_i.$$

Using (4) which can be written as

$$\sigma^2_{\varepsilon_i} = \frac{1-\psi(z,X_i)}{\psi(z,X_i)+S_{ui}}$$

we get, solving for S_{ui}

$$S_{ui} = \frac{1-\psi(z,X_i)(1+\sigma^2_{\varepsilon_i})}{\sigma^2_{\varepsilon_i}}.$$

Recalling that $S_{ui} = \exp(\gamma + \boldsymbol{\theta_u} X_i)$, we construct the variable

$$\hat{Z}_i = \frac{1-\hat{\psi}_i(1+\hat{\varepsilon}_i^2)}{\hat{\varepsilon}_i^2}$$

and we perform an OLS estimation of the equation

$$\ln(\hat{Z}_i) = \gamma(z) + \boldsymbol{\theta}_u.X_i.$$

to obtain $\hat{\gamma}(z)$ and $\hat{\boldsymbol{\theta}}_u$. We then get $\hat{\beta}(z) = \hat{\delta}(z) + \hat{\gamma}(z)$ and $\hat{\boldsymbol{\theta}}_d = \hat{\boldsymbol{\Delta}} + \hat{\boldsymbol{\theta}}_u$.

4 Data

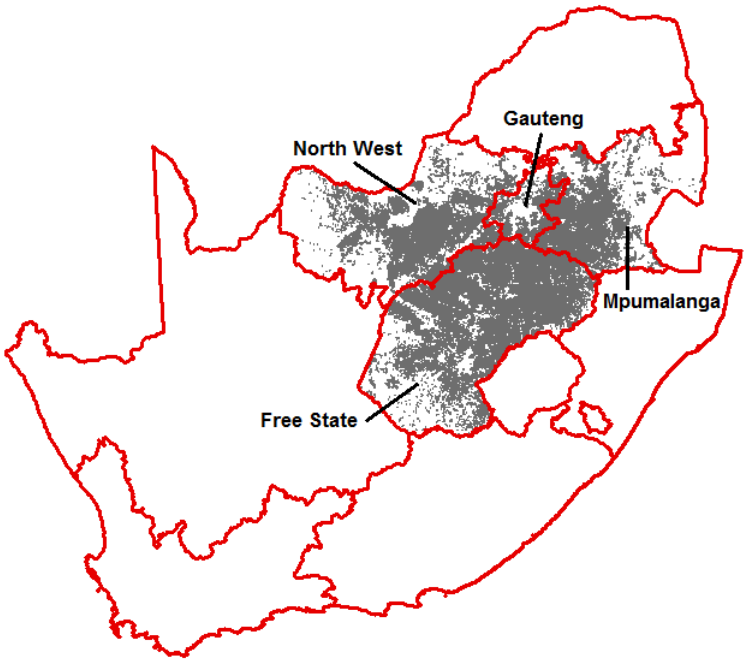
We combine different data sources to build a very detailed database on South African agriculture. First, field boundaries are identified, then agricultural production is characterised on each field by identifying the crops that are grown, measuring the biomass produced, determining the production system (rain fed or irrigated) and quantifying crop interspecific biodiversity.

4.1 Crop fields

Field boundaries, available for the provinces of Free State, Gauteng, North West and Mpumalanga, are determined using the Producer Independent Crop Estimate System (PICES) which combines satellite imagery, Geographic Information System (GIS), point frame statistical platforms, and aerial observations (Ferreira et al., 2006). Satellite imagery of cultivated fields is obtained from the SPOT 5 satellite at a 2.5-m resolution. Plot boundaries are then digitised using GIS and field cloud covered polygons are removed before processing. Over the four regions of interest, PICES distinguishes circa 280 000 fields covering an area of around 6.5 million hectares. To approximately match the resolution of the crop production indicator described below (see 4.2), which is only available at the 250-m resolution, the analysis is limited to fields larger than 6.25 ha. This restricts the sample to 213 110 fields. Figure 1 presents the location of the considered crop fields in South Africa. While the summary statistics in Table 1 show that fields are on average about 28.4 hectares, the large standard deviation (24.8) indicates that they vary substantially in size. For instance, the largest field is 720 hectares large, while the smallest is only 6.25.

Using the digitised satellite images described above, the Agricultural Geo-referenced Information System (AGIS) developed by the South African Department of Agriculture provides information on the crop cultivated on each field. To do so, sample points were selected randomly and surveyed by trained observers from a very light aircraft in order to determine crop type (Ferreira et al., 2006). Crop information collected during the aerial surveys on the sample points was subsequently used as a training set for crop type classification for each field and for accuracy assessment. These estimated crop classifications were then checked against a producer based survey for the Gauteng region. The Gauteng census survey showed that less than 1.8% of crop types had been misclassified. All in all, seven summer crops were distinguished for the provinces of Free State, Gauteng, North West and Mpumalanga for the summer season 2006/2007: cotton, dry beans, groundnuts, maize, sorghum, soybean and sunflower. An example of the distribution of crop types is provided in figure 2. The summary statistics for the entire sample in Table 2 show that maize was the dominant crop cultivated in the three provinces: maize fields represent

Figure 1: Localisation of the considered fields in South Africa



nearly 70 per cent of the total number of fields we consider. Other important crops were sunflower and soybean, standing at 15 and 11 per cent each. In contrast, all other crop types constituted less than 2 per cent individually. One should note that even if one were to adjust the crop type shares by their areas, a similar ranking remains, with a slight redistribution of shares towards the smaller crop types. For instance, the share of maize dropped to 62 per cent of the total crop area.

The AGIS crop boundaries dataset also provides information regarding whether each field is irrigated. As can be seen from Table 1, only about 5% of all fields were part of an irrigation system.

Finally, all fields can be linked to their respective farms with a unique farm identifier. In total the fields were owned by 12 462 different farms, where on average each farm was proprietor of 5 fields. However, ownership differed substantially, with the largest ownership gathering 193 fields, and 3 704 single field farms.

4.2 Crop production measure

We estimate crop biomass production using the satellite derived normalised difference vegetation index (NDVI). Vegetation indices provide consistent spatial and temporal representations of vegetation conditions, when locally derived information is not available. As a matter of fact, numerous studies have demonstrated that NDVI values are significantly correlated with biomass production, and therefore yields, of various crops, including wheat (Das et al., 1993; Gupta et al., 1993; Doraiswamy & Cook, 1995; Hochheim & Bar-

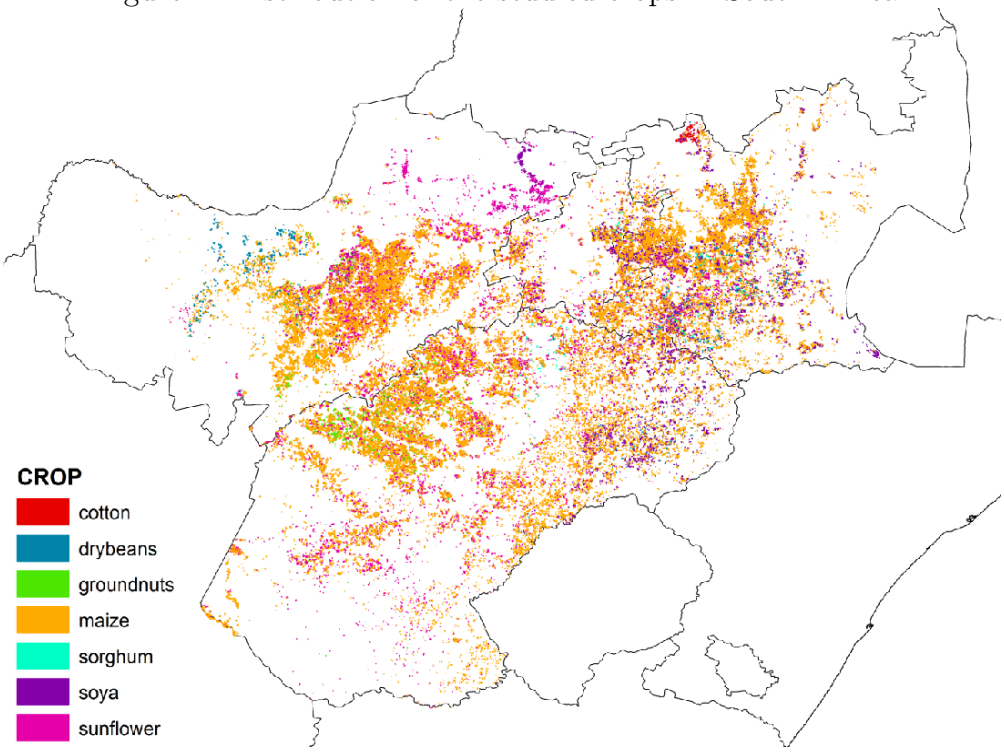
Table 1: Plot summary statistics

Variable	Mean	Standard deviation	Variable	Mean	Standard deviation
All crops			Maize		
NDVI	0.61	0.13	NDVI	0.61	0.13
Water balance	−41.07	11.73	Water balance	−42.60	11.89
Season length (days)	129.64	35.48	Season length (days)	126.98	35.33
Plot area(ha)	28.36	23.47	Plot area(ha)	22.72	2.04
Farm area (ha)	315.2	500.00	Farm area (ha)	193.63	3.53
Irrigation (%)	5.05	–	Irrigation (%)	4.94	21.67
Longitude	27.7	1.48	Longitude	27.69	1.49
Latitude	−27.1	0.99	Latitude	−27.17	0.99
Cotton			Sorghum		
NDVI	0.73	0.05	NDVI	0.69	0.10
Water balance	−28.61	14.54	Water balance	−33.54	7.70
Season length (days)	130.94	52.00	Season length (days)	125.20	26.17
Plot area(ha)	15.63	1.66	Plot area(ha)	19.27	1.91
Farm area (ha)	184.41	2.32	Farm area (ha)	45.63	2.79
Irrigation (%)	56.60	49.80	Irrigation (%)	1.96	13.87
Longitude	29.23	0.06	Longitude	28.83	0.86
Latitude	−25.03	0.06	Latitude	−26.75	0.58
Dry bean			Soybean		
NDVI	0.62	0.14	NDVI	0.71	0.09
Water balance	−38.96	12.39	Water balance	−32.18	7.59
Season length (days)	145.58	38.66	Season length (days)	129.19	30.14
Plot area(ha)	21.43	2.03	Plot area(ha)	19.09	1.91
Farm area (ha)	68.50	3.10	Farm area (ha)	95.98	2.95
Irrigation (%)	3.75	19.00	Irrigation (%)	5.89	23.55
Longitude	27.83	2.15	Longitude	29.06	0.81
Latitude	−26.79	0.71	Latitude	−26.78	0.87
Ground nut			Sunflower		
NDVI	0.52	0.11	NDVI	0.59	0.13
Water balance	−55.56	6.34	Water balance	−44.20	12.10
Season length (days)	115.40	34.67	Season length (days)	131.27	38.01
Plot area(ha)	28.61	2.03	Plot area(ha)	20.54	2.01
Farm area (ha)	73.91	2.96	Farm area (ha)	80.59	3.14
Irrigation (%)	4.10	19.84	Irrigation (%)	7.80	26.81
Longitude	26.15	0.59	Longitude	27.16	1.14
Latitude	−27.67	0.59	Latitude	−27.19	1.11

Table 2: Distribution of the considered crops

Crop	Nb. of fields	Share of total fields (%)	Share of total area (%)
Dry beans	1 227	1.88	1.88
Ground nuts	1 292	2.5	2.50
Maize	45 256	69.77	72.27
Sorghum	715	1.10	0.93
Soybean	6 825	10.52	8.80
Sunflower	9 441	14.56	13.51
Cotton	106	0.16	0.10
Total	64 862	100.00	100.00

Figure 2: Distribution of the studied crops in South Africa



ber, 1998; Labus et al., 2002), sorghum (Potdar, 1993), maize (Hayes & Decker, 1996; Prasad et al., 2006), rice (Nuarsa et al., 2011; Quarmby et al., 1993), soybean (Prasad et al., 2006), barley (Weissteiner & Kuhbauch, 2005), millet (Groten, 1993) and tomato (Koller & Upadhaya, 2005). Moreover, NDVI has also been shown to provide a very good indicator of crop phenological development (Benedetti & Rossini, 1993).

The NDVI index is calculated using ratios of vegetation spectral reflectance over incoming radiation in each spectral band. The NDVI data are extracted from the MOD13Q1 dataset,[5] which gathers together reflectance information collected by the MODerate-resolution Imaging Spectroradiometer (MODIS) instrument operating on NASA's Terra satellite (Huete et al., 2002). From these data, NDVI can be formulated as:

$$NDVI = \frac{NIR - VIS}{NIR + VIS}$$

where the difference between near-infrared reflectance (NIR) and visible reflectance (VIS) values is normalised by the total reflectance and varies between - 1 and 1 (Eidenshink, 1992).[6] Negative and very low values corresponding to water and barren areas were excluded from the analysis by design. Due to the nature of the source data NDVI derived estimates were available as 16-day composite indices at a resolution of 250 m. Area-

[5]Available online from https://lpdaac.usgs.gov/lpdaac/content/view/full/6652

[6]Plant leaves strongly absorb visible light (VIS), while they reflect near-infrared light (NIR). Then, the more leaves a plant has, and therefore the more biomass it produces, the more near-infrared and visible reflectance will be affected. The more biomass is produced, the more the NDVI is close to 1.

weighted averages of the 16-day NDVI values for each crop field are calculated using the "zonal statistics" tool in ArcGIS.

The NDVI can be related to produced quantities by the following relationship, extrapolated by Ma et al. (2001) for soybean:

$$Y = d + bNDVI^c \tag{5}$$

where Y represents the quantities produced (or the yield), d, b and c are three parameters, with c being in [2;35] and b in $[3.10^3; 72.10^3]$ (Ma et al., 2001). Denoting by

$$Z_i = NDVI_i - NDVI_0 \tag{6}$$

with $NDVI_0 = |d/b|^{1/c}$, bZ_i^c gives an estimation of the quantities produced on field i, $Y(i)$.[7]

Crop growing seasons are characterised by the planting date and the phenology cycle, which determines the length of the season. In South Africa, planting generally occur somewhere between October and December in order to reduce the vulnerability to erratic precipitation (Ferreira et al., 2006). However, phenology cycles, and hence growing seasons, can differ substantially among crop types and even for fields of the same crop type. In order to take account of this, we used the TIMESAT program[8] (Jönsson & Eklundh, 2002, 2004) to determine crop and field specific growing seasons. We are then able to approximate the start and end of growing seasons based on distribution properties of the NDVI. Summary statistics in Table 1 show that growing seasons are on average 130 days, with a standard deviation of 35 days.

Finally, as is standard in the literature of satellite derived plant growth measures, we use the maximum NDVI over the growing season as an indicator of crop production (Zhang et al., 2006). It takes on an average value of 0.61 with a standard deviation of 0.13 (see Table 1).

4.3 Crop water balance

An important determinant of crop growth is water availability. A common simple proxy for it is the difference between rainfall and the evaporative demand of the air, i.e, evapotranspiration. To calculate this, we use gridded daily precipitation and reference evapotranspiration data taken from the USGS Early Warning Famine climatic database.[9] More specifically, daily rainfall data, given at the 0.1 degree resolution (ca. 10 km), are gener-

[7] For value smaller than $NDVI_0$, the produced quantities are equal to 0, the NDVI capturing the light reflected by the bare soil.

[8] The algorithm within the TIMESAT software is commonly used to extract seasonality information from satellite time-series data.

[9] http://earlywarning.usgs.gov/fews

ated with the rainfall estimation algorithm RFE (version 2.0) dataset implemented by the National Oceanic and Atmospheric Administration (NOAA) - Climate Prediction Center (CPC) using a combination of rain gauges and satellite observations. Daily reference evapotranspiration data, available at a 1 degree resolution, were calculated using a 6-hourly assimilation of conventional and satellite observational data of air temperature, atmospheric pressure, wind speed, relative humidity and solar radiation extracted from the National Oceanic and Atmospheric Administrationâ€s Global Data Assimilation System. Using these gridded data each field was then assigned a daily precipitation and potential evapotranspiration value over its growing season to then calculate out its average daily water balance. The mean and standard deviation of this measure are given in Table 1.

4.4 Biodiversity index

Among field characteristics, we are particularly interested in crop biodiversity. Diversity measures, extensively used in biology and ecology literature, take into account specie richness (i.e. the number of species present) and evenness (i.e. the distribution of species). In the following, we quantify biodiversity at the field level adopting one of the most widely used indicators, the Shannon index (Shannon, 1948):

$$S_{i\ell} = -\sum_{z} B_l(z) \ln B_l(z) \tag{7}$$

where the subscript i refers to a specific field, ℓ defines the size of the perimeter considered as relevant, and $B_\ell(z)$ is the proportion of area within perimeter ℓ that is of crop z type. S_{il} is then calculated for a given perimeter ℓ, defined by its radius, applied to the centroid of the field i considered. The more diverse crops are and the more equal their abundances, the larger is the Shannon index. When all crops are equally common, all $B(z)$ values will equal $1/Z$ (Z being the total number of crops) and S will be equal to lnZ. On the contrary, the more unequal the abundances of the crops are, the smaller is the index, approaching 0 (and being equal to 0 if $Z = 1$). With respect to other common indicators, like the Simpson's index,[10] the Shannon index is known to put less weight on the more abundant species and to be more sensitive to differences in total species richness and in changes in populations showing small relative abundances (Baumgärtner, 2006). In our specification, the distance threshold for the radius ℓ is 0.75 kilometre; the distance is then increased 250 metres by 250 metres to reach 3 kilometres, the maximum distance considered. We provide summary statistics for the Shannon index in Table 3. Widening the perimeter under consideration increases the value of the Shannon index substantially. For example, the 3 km index is nearly 5 times larger than that of 0.75 km. This strongly suggests that crop types are strongly spatially agglomerated, and thus locally less diverse.

[10]With our notations, the Simpson's index is given by $1 - \sum_z B_\ell^2(z)$.

Table 3: Summary statistics for the Shannon index

	All crops		Dry bean		Ground nut		Maize		Sorghum		Soybean		Sunflower	
ℓ	$\overline{S}$	σ_S	$\overline{S}$	σ_S	$\overline{S}$	σ_S	$\overline{S}$	σ_S	$\overline{S}$	σ_S	$\overline{S}$	σ_S	$\overline{S}$	σ_S
750 m	0.03	0.14	0.17	0.31	0.21	0.31	0.08	0.21	0.11	0.25	0.17	0.29	0.14	0.26
1000 m	0.06	0.18	0.27	0.37	0.35	0.35	0.14	0.26	0.20	0.31	0.28	0.33	0.23	0.31
1250 m	0.07	0.21	0.36	0.39	0.43	0.35	0.19	0.29	0.28	0.35	0.37	0.35	0.30	0.33
1500 m	0.09	0.23	0.45	0.41	0.48	0.34	0.23	0.30	0.35	0.36	0.44	0.36	0.35	0.34
1750 m	0.10	0.24	0.50	0.42	0.51	0.34	0.27	0.31	0.40	0.37	0.49	0.35	0.40	0.34
2000 m	0.12	0.25	0.55	0.43	0.53	0.32	0.31	0.32	0.45	0.37	0.54	0.35	0.43	0.33
2250 m	0.13	0.26	0.60	0.43	0.55	0.31	0.33	0.32	0.49	0.37	0.58	0.34	0.46	0.33
2500 m	0.13	0.27	0.63	0.43	0.56	0.30	0.36	0.32	0.52	0.37	0.61	0.33	0.48	0.32
2750 m	0.14	0.28	0.66	0.43	0.57	0.30	0.38	0.31	0.56	0.36	0.63	0.32	0.50	0.32
3000 m	0.15	0.28	0.68	0.43	0.58	0.29	0.40	0.31	0.59	0.36	0.65	0.32	0.52	0.31

Note: The table reports the mean ($\overline{S}$) and the standard deviation (σ_S) of the distribution of the Shannon index, measured for the different crops we considered, on different perimeters, characterised by their radius, ℓ.

5 Empirical analysis

5.1 Impact on the production level

5.1.1 Regression specification

Our first empirical task is to investigate whether biodiversity affects crop field health. To this end we use the following regression specification:

$$\ln(NDVI_i) = \theta_{0\ell} + \theta_{S\ell} S_{i\ell} + \sum_k \theta_{k\ell} X_{ik\ell} + \mu_m + \epsilon_{i\ell} \tag{8}$$

where $NDVI_i$ is the maximum NDVI over the crop's growing season on field i, $S_{i\ell}$ is the Shannon index for perimeter ℓ as defined above, $X_{i\ell}$ is a vector of k control variables, and μ_m is a farm fixed effect. These effects depend on natural conditions (weather, season length...), field attributes (irrigation, area...) and farm management attributes (pesticides, mechanisation, economies of scale...). Therefore, the vector of control variables X includes crop fixed effects, crop water balance (WB) and its squared value (WB^2), an irrigation dummy indicator, the season length, the logarithm of the field area in hectares ($\ln(AREA)$), the latitude ($LATITUDE$) and longitude ($LONGITUDE$) of the centroid of the field, the percentage of cropland within a defined perimeter that is irrigated (PC_AREA_IRR), and the percentage of land devoted to the same crop that belongs to the same farm, within a defined perimeter (PC_AREA_FARM). One should note that by including farm fixed effects, we are able to capture crop management techniques that are common within farms, such as, for example, pesticide use, as well as farm wide economies of scale. Crop specific dummies allow us to control for the fact that different crops will have different vegetation growth intensity as captured by satellite reflectance data.

One concern in estimating (8) with standard linear estimators is that our dependent variable is theoretically bounded between 0 and 1. However, in our data the maximum

NDVI of fields never reached either of the boundaries â€“ the minimum and maximum values were 0.07 and 0.93, respectively. We can thus proceed with estimating (8) with a standard linear model. In order to make our coefficients more readable, we multiply NDVI by 10 000. Standard errors are clustered at the farm level.

5.1.2 Results

Our estimation results are presented in Table 4. In the first column, we simply include our field specific control variables (vector X). As can be seen, crop water balance has a significant positive and exponentially increasing impact on crop growth. However, while having an irrigation system also acts to increase the productivity of fields, it also makes crops less reliant on water balance (in a linear fashion) as would be expected. The coefficient on season length suggests that the longer the season lasts, the lower crop health is. In other words, the longer the season is, the higher are the probabilities that an adverse event affects crops. Larger fields are less productive per hectare than smaller ones. Finally, being located more in the east results in greater crop health, while being further south or north is inconsequential for field productivity within our sample.

In the second column, we add the Shannon index using a perimeter of 0.75 kilometre to define the neighbourhood ℓ. A greater value of S (i.e., a greater surrounding biodiversity) increases field productivity. Arguably, however, our diversity index may just be capturing the fact that neighbouring areas are different in ways that are correlated with the diversity of crops. To take account of these factors, we thus next include the percentage of the surrounding area that is irrigated and the percentage of the surrounding area of fields of the same crop type that belongs to the same farm. Accordingly, controlling for these factors does indeed reduce the coefficient of S by about 17 per cent, and thus suggests that it is important to control for surrounding area specific characteristics.

We subsequently increased the defined perimeter to calculate the Shannon index to 1 km, as shown in the forth column, adjusting the variables PC_AREA_IRR and PC_AREA_FARM in an analogous fashion. The S coefficient remain statistically significant, but decreases by 30%. As far as control variables are concerned, we find that the share of area irrigated unequivocally increases the biomass production. The share of area belonging to the same farm within the perimeter we consider seems to have no significant impact on the biomass production. Further increasing the perimeter similarly continues to produce a significant positive impact of biodiversity, the coefficient increasing by 48% per cent. However, when further expanding the threshold of our definition of the relevant neighbourhood to more than 1.5 kilometres, biodiversity no longer acts as a significant predictor of crop production, as shown in the last columns of the table.[11] This suggests

[11] We also experimented with increasing the perimeter up to 10 kilometres, but the coefficient on S remains insignificant in all cases.

that biodiversity is relatively locally defined, i.e., within less than 1.75 kilometres, but likely close to 1.25 kilometres. One may also want to note that among our neighbourhood specific controls, only the percentage of area irrigated consistently has a significant (positive) effect on crop field health.

We then apply the same regression specification by crop, i.e. only to the fields dedicated to one of the six crops for which data are available (cotton is not considered in the regressions by crop since the available sample – 106 fields, 0.16% of the total available fields and 0.1 of the total cropland considered – is too small). The results show that, on the one hand, biodiversity has a significant impact on the biomass production of maize, soybean and sunflower, and that the relevant perimeter size of the biodiversity index depends on the crop. Table 5 also reveals that biodiversity has no significant impact on dry bean, ground nuts and sorghum. This can be explained by the fact that each of the latter crops represents less than 2% of the total number of fields. In other words, the area dedicated to these crops is small, the fields are probably sufficiently scattered to not suffer from the proliferation of their pests. Therefore, the biodiversity variation on the perimeter that we consider has a negligible marginal effect on the biomass production. When looking at the crops which biomass production is affected by crop biodiversity, we see that the relevant perimeter for biodiversity varies: biodiversity has a positive and significant impact on the production of maize only for perimeters smaller than 1.5 kilometres whereas the relevant perimeter for soybean is between 1.25 and 2.75 kilometres. Surprisingly, biodiversity has a negative and significant impact on sunflower biomass production on perimeters with a radius larger than 2.25 kilometres. A positive significant impact on sunflower is found only for ℓ equal to one kilometre. This result is probably linked to the fact that the pests that affect the biomass production are not the same for the three crops we consider.[12]

The impact of irrigation also varies and depends on crop characteristics, as shown by tables 6, 7 and 8. Maize is one of the most efficient cultivated plants in South Africa as far as water use is concerned (DAFF, 2014a), hence a positive and significant impact of irrigation. On the contrary, sunflower is highly inefficient in water-use and, as well as soybean, is mostly rain-fed grown.[13] This could explain the absence of a significant impact of irrigation on biomass production for these crops. Finally, unlike maize and sunflower, soybean biomass production is positively affected by the size of the field. This effect could be related to the physiology of the plant or to the higher mechanisation that is allowed by larger fields and that can have a positive impact on the final yield.

[12]The main potential crop losses are caused by weeds, Nevertheless, thanks to the improvement in weed control techniques, the main actual losses come from animals (mainly insects) and pathogens (Oerke, 2007). Maize is mainly attacked by insects (DAFF, 2014a), while sunflower and soybean are mainly attacked by diseases caused by fungi and viruses (DAFF, 2009, 2014b).

[13]Soybean is mostly rain-fed grown because of low profitability and difficult water management. Indeed, water shortage is critical during the pod set stage while excessive water supply prior or after the flowering may jeopardise the final yield.

Table 4: Regression results, all crops

Variables	(1)	(2) $\ell = 0.75$km	(3) $\ell = 1$km	(4) $\ell = 1.25$km	(5) $\ell = 1.5$km	(6) $\ell = 1.75$km	(7) $\ell = 2$km	(8) $\ell = 2.25$ km	(9) $\ell = 2.5$ km	(10) $\ell = 2.75$km	(11) $\ell = 3$km
Shannon index (S_l)		62.53***	44.02**	64.95***	28.87**	25.33	20.83	−1.78	1.87	−9.95	−21.97
		(22.81)	(19.94)	(20.79)	(14.05)	(16.03)	(15.01)	(17.60)	(15.96)	(18.47)	(17.17)
Water balance (WB)	11.66***	11.20***	10.98***	10.94***	10.09***	11.09***	11.27***	11.46***	11.44***	11.47***	11.58***
	(2.52)	(2.47)	(2.50)	(2.47)	(2.46)	(2.48)	(2.50)	(2.49)	(2.51)	(2.49)	(2.50)
WB^2	0.17***	0.17***	0.17***	0.17***	0.16***	0.17***	0.17***	0.17***	0.17***	0.17***	0.17***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Irrigation (IR)	539.13***	507.62***	495.86***	486.96***	495.91***	510.37***	521.80***	528.30***	530.05***	530.39***	533.25***
	(89.97)	(88.46)	(89.43)	(89.81)	(90.10)	(90.10)	(90.10)	(89.66)	(89.12)	(89.62)	(89.66)
WB × IR	−18.28***	−17.84***	−18.32***	−18.47***	−18.59***	−18.46***	−18.26***	−18.27***	−18.28***	−18.37***	−18.37***
	(5.27)	(5.27)	(5.27)	(5.18)	(5.11)	(5.15)	(5.20)	(5.15)	(5.17)	(5.21)	(5.24)
WB^2 × IR	−0.002	0.008	0.003	0.004	0.001	0.00	0.001	0.00	0.00	−0.00	−0.001
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Season length	−5.25***	−5.23***	−5.24***	−5.25***	−5.27***	−5.27***	−5.27***	−5.27***	−5.26***	−5.26***	−5.26***
	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)
ln($AREA$)	−73.0***	−68.51***	−70.62***	−71.09***	−72.29***	−72.30***	−73.06***	−72.88***	−73.30***	−73.51***	−73.71***
	(8.07)	(8.18)	(8.18)	(8.11)	(7.99)	(8.04)	(7.99)	(8.09)	(8.04)	(8.06)	(8.14)
Longitude	1229.69***	1254.00***	1258.71***	1279.73***	1273.08***	1268.87***	1253.39***	1241.49***	1246.15***	1245.12***	1239.04***
	(268.90)	(269.06)	(268.37)	(267.75)	(269.40)	(269.82)	(269.73)	(269.87)	(268.38)	(268.04)	(269.11)
Latitude	−357.88	−344.10	−349.28	−353.55	−347.52	−345.18	−347.73	−350.00	−341.21	−334.49	−337.23
	(521.67)	(485.94)	(488.78)	(489.05)	(510.83)	(515.18)	(522.57)	(522.06)	(534.80)	(540.31)	(539.05)
PC_NEAR_IRR		225.33***	265.33***	329.28***	345.78***	322.21***	305.35***	256.87***	254.75***	249.47***	194.05***
		(32.16)	(33.36)	(27.98)	(34.66)	(40.74)	(38.31)	(44.09)	(41.74)	(49.48)	(48.81)
PC_NEAR_FARM		19.43	11.57	5.29	−9.99	−5.38	−13.33	−6.31	−24.06	−32.40	−36.14
		(11.92)	(11.82)	(14.53)	(17.0)	(16.3)	(17.72)	(21.90)	(25.09)	(27.75)	(31.95)
Number of fields	64 862	64 862	64 862	64 862	64 862	64 862	64 862	64 862	64 862	64 862	64 862
Number of farms	12 462	12 462	12 462	12 462	12 462	12 462	12 462	12 462	12 462	12 462	12 462

Note: ***, ** and * indicate 1, 5 and 10 per cent significance levels, respectively. Robust standard errors are in parentheses, standard errors are clustered at the field level. Farm and crop fixed effects are included but not reported. Treatment (1) includes field specific control variables only, while treatments from (2) to (10) include field specific variables and the Shannon index, calculated for ℓ between 0.75 and 3 km.

Table 5: Impact of biodiversity on crop production

	Crops					
Distance	Dry bean	Ground nut	Maize	Sorghum	Soybean	Sunflower
750 m	−43.31	91.37	85.52**	−2.70	−0.70	36.63
	(94.80)	(160.8)	(33.44)	(142.3)	(31.89)	(61.43)
1000 m	4.94	−86.55	47.01**	202.44	16.03	80.66*
	(90.90)	(182.6)	(23.4)	(128.6)	(29.81)	(47.78)
1250 m	82.51	−37.26	54.20**	143.75	62.71**	43.97
	(74.94)	(203.2)	(23.18)	(145.2)	(31.19)	(39.21)
1500 m	12.70	−305.00	10.78	−26.17	79.49**	−24.11
	(89.57)	(214.0)	(18.30)	(133.4)	(32.88)	(55.66)
1750 m	105.38	−95.21	19.52	6.34	125.99***	−16.10
	(91.12)	(228.5)	(19.82)	(115.5)	(31.35)	(55.89)
2000 m	−30.91	118.9	15.58	42.48	134.98***	−47.67
	(87.49)	(228.4)	(17.17)	(113.8)	(38.30)	(60.05)
2250 m	13.50	−29.43	−1.15	55.41	94.60**	−96.54
	(96.94)	(271.2)	(20.42)	(124.6)	(37.36)	(52.49)
2500 m	70.90	−50.72	16.01	129.71	85.19**	−159.23***
	(117.1)	(345.4)	(16.21)	(184.8)	(32.95)	(54.44)
2750 m	89.75	−42.84	17.17	10.37	66.00*	−199.04**
	(119.2)	(361.1)	(19.06)	(161.8)	(36.32)	(78.24)
3000 m	−25.43	−35.19	11.07	69.33	61.93	−268.42***
	(115.2)	(366.1)	(22.09)	(201.8)	(39.05)	(77.38)

Note: Robust standard errors in parentheses, $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. The table shows the coefficients for the variable S_l for each of the six crops considered, multiplied by 10 000 to facilitate the reading. Complete regression results are reported in tables 6 to 8.

These results confirm the positive impact of crop biodiversity on agricultural production, with the exception of sunflower. But what about farmer's exposure to risks? In other words, does biodiversity also participate in the reduction of production variability, as ecology studies suggest? Previous papers (Di Falco & Chavas, 2006, 2009) find that crop diversity reduces the risk of crop failure. In particular, in Di Falco & Chavas (2009), genetic diversity increases the variance of barley production in Ethiopia as well as its skewness. The skewness effect dominates the variance effect and finally results in the reduction of the exposure of downside risk. Like Di Falco & Chavas (2009), we examine the contribution of diversity to production variance using a second regression, on the error terms of the first regression, in accordance with the theoretical model that supports our investigations.

Table 6: Regression results, maize

Variables	$\ell = 0.75$ km	$\ell = 1$ km	$\ell = 1.25$ km	$\ell = 1.5$ km	$\ell = 1.75$ km	$\ell = 2$ km	$\ell = 2.25$ km	$\ell = 2.5$ km	$\ell = 2.75$ km	$\ell = 3$ km
Shannon index (S_ℓ)	85.52**	47.01**	54.19**	10.78	19.52	15.58	−1.15	16.01	17.17	11.07
	(33.44)	(23.40)	(23.18)	(18.30)	(19.82)	(17.17)	(20.43)	(16.21)	(19.06)	(22.09)
Water balance (WB)	5.34	5.00	4.94	5.00	5.42	5.55	5.64	5.56	5.64	5.76
	(3.56)	(3.68)	(3.70)	(3.68)	(3.68)	(3.66)	(3.65)	(3.68)	(3.67)	(3.65)
WB^2	0.12**	0.11**	0.11**	0.11**	0.12**	0.12**	0.12**	0.12**	0.12**	0.12**
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Irrigation (IR)	731.07***	731.27***	723.09***	727.73***	743.11***	752.58***	756.74***	758.90***	759.19***	759.81***
	(114.79)	(116.41)	(117.38)	(116.10)	(114.64)	(113.94)	(113.97)	(115.11)	(114.89)	(114.64)
WB × IR	−9.01	−9.04	−9.59	−9.86	−9.88	−9.66	−9.61	−9.51	−9.63	−9.65
	(6.53)	(6.54)	(6.59)	(6.55)	(6.51)	(6.50)	(6.49)	(6.55)	(6.53)	(6.53)
WB^2 × IR	0.06	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	(0.09)	(0.08)	(0.09)	(0.08)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Season length	−5.70***	−5.70***	−5.69***	−5.72***	−5.72***	−5.72***	−5.72***	−5.71***	−5.71***	−5.70***
	(0.40)	(0.39)	(0.40)	(0.40)	(0.40)	(0.40)	(0.40)	(0.40)	(0.40)	(0.40)
ln($AREA$)	−63.89***	−66.32***	−66.99***	−67.88***	−68.23***	−69.17***	−68.93***	−69.53***	−69.95***	−70.36***
	(8.88)	(8.95)	(8.91)	(8.86)	(8.81)	(8.73)	(8.77)	(8.82)	(8.89)	(8.98)
Longitude	962.56***	972.54***	981.79***	980.43***	968.68***	955.94***	948.98***	956.40***	953.72***	951.89***
	(292.06)	(292.32)	(294.53)	(296.82)	(297.18)	(297.05)	(296.18)	(295.65)	(294.77)	(294.44)
Latitude	175.12	170.91	175.41	178.10	179.65	178.65	178.88	186.16	178.53	181.66
	(540.15)	(541.23)	(543.88)	(559.37)	(567.85)	(580.03)	(581.60)	(590.99)	(595.01)	(599.89)
PC_NEAR_IRR	283.56***	346.00***	405.00***	417.23***	309.50***	285.39***	294.35***	287.82***	226.33***	192.80***
	(37.99)	(44.14)	(48.21)	(58.94)	(48.03)	(55.80)	(49.70)	(54.84)	(61.47)	(70.50)
PC_NEAR_FARM	20.26*	10.77	4.25	−6.86	−5.45	−21.88	−19.42	−39.83	−50.13**	−66.95**
	(11.09)	(12.85)	(16.10)	(15.96)	(15.50)	(17.95)	(21.95)	(25.67)	(24.99)	(27.44)
Number of farms	10 551	10 551	10 551	10 551	10 551	10 551	10 551	10 551	10 551	10 551
Number of fields	45 256	45 256	45 256	45 256	45 256	45 256	45 256	45 256	45 256	45 256

Note: ***, ** and * indicate 1, 5 and 10 per cent significance levels, respectively. Robust standard errors are in parentheses, standard errors are clustered at the field level. Farm and crop fixed effects are included but not reported. Treatments include field specific variables and the Shannon index, calculated for ℓ between 0.75 and 3 km.

Table 7: Regression results, soybean

Variables	$\ell = 0.75$ km	$\ell = 1$ km	$\ell = 1.25$ km	$\ell = 1.5$ km	$\ell = 1.75$ km	$\ell = 2$ km	$\ell = 2.25$ km	$\ell = 2.5$ km	$\ell = 2.75$ km	$\ell = 3$ km
Shannon index (S_l)	−0.70	16.03	62.71**	79.49**	125.99**	134.98***	94.60**	85.19**	66.00*	61.93
	(31.89)	(29.81)	(31.19)	(32.88)	(31.35)	(38.30)	(37.36)	(32.95)	(36.32)	(39.05)
Water balance (WB)	10.00	9.92	9.99	10.46*	10.45*	10.48*	10.42*	10.32	10.25	10.14
	(6.24)	(6.15)	(6.21)	(6.27)	(6.25)	(6.28)	(6.27)	(6.25)	(6.18)	(6.23)
WB^2	.29	0.29**	0.29**	0.30**	0.30***	0.30**	0.30**	0.30**	0.30**	0.30**
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Irrigation (IR)	151.49	150.57	136.72	158.20	141.36	152.55	150.47	148.87	145.32	150.21
	(140.09)	(137.92)	(139.14)	(140.83)	(142.59)	(142.08)	(142.21)	(142.27)	(141.43)	(141.63)
WB × IR	−12.47**	−12.57	−12.87	−12.28	−13.01	−12.45	−12.85	−13.25	−13.43	−13.12
	(8.92)	(9.01)	(8.84)	(8.91)	(8.93)	(8.85)	(8.87)	(8.99)	(8.96)	(8.88)
WB^2 × IR	−0.11	−0.11	−0.12	−0.11	−0.12	−0.11	−0.12	−0.13	−0.13	−0.13
	(0.20)	(0.20)	(0.19)	(0.19)	(0.20)	(0.20)	(0.19)	(0.20)	(0.20)	(0.20)
Season length	−4.65***	−4.68***	−4.68***	−4.68***	−4.69***	−4.68***	−4.69***	−4.71***	−4.70***	−4.72***
	(0.45)	(0.46)	(0.47)	(0.47)	(0.47)	(0.46)	(0.46)	(0.46)	(0.46)	(0.46)
ln($AREA$)	45.15***	42.68***	42.28***	40.76***	41.91***	40.87***	40.24**	40.32**	39.96**	39.69**
	(16.18)	(15.99)	(15.94)	(15.71)	(15.55)	(15.59)	(15.61)	(15.79)	(15.84)	(16.01)
Longitude	2138.97***	2156.27***	2171.44***	2161.12***	2151.44***	2195.33***	2173.92***	2168.77***	2156.57***	2162.75***
	(681.75)	(677.05)	(677.36)	(677.17)	(673.02)	(672.67)	(685.60)	(682.83)	(679.82)	(681.58)
Latitude	−1304.35	−1361.76*	−1389.11*	−1392.56*	−1420.42*	−1411.74*	−1392.86*	−1364.11*	−1319.19	−1318.97
	(797.04)	(799.14)	(790.734)	(787.69)	(777.57)	(790.14)	(810.05)	(817.98)	(841.64)	(834.58)
PC_NEAR_IRR	46.76	36.31	82.94	−9.08	140.24*	125.91	69.44	79.74**	229.69**	175.23
	(74.00)	(65.64)	(67.14)	(79.89)	(82.74)	(86.57)	(92.96)	(89.39)	(102.08)	(116.12)
PC_NEAR_FARM	66.73***	66.17***	48.85**	42.61	53.38	29.22	18.62	27.83	−2.71	−8.67
	(22.42)	(18.25)	(24.46)	(32.70)	(34.27)	(40.97)	(47.60)	(45.45)	(46.83)	(52.83)
Number of farms	2 248	2 248	2 248	2 248	2 248	2 248	2 248	2 248	2 248	2 248
Number of fields	6 825	6 825	6 825	6 825	6 825	6 825	6 825	6 825	6 825	6 825

Note: ***, ** and * indicate 1, 5 and 10 per cent significance levels, respectively. Robust standard errors are in parentheses, standard errors are clustered at the field level. Farm and crop fixed effects are included but not reported. Treatments include field specific variables and the Shannon index, calculated for ℓ between 0.75 and 3 km.

Table 8: Regression results, sunflower

Variables	$\ell = 0.75$ km	$\ell = 1$ km	$\ell = 1.25$ km	$\ell = 1.5$ km	$\ell = 1.75$ km	$\ell = 2$ km	$\ell = 2.25$ km	$\ell = 2.5$ km	$\ell = 2.75$ km	$\ell = 3$ km
Shannon index (S_l)	36.63	80.66*	43.97*	−24.11	−16.10	−47.67	−96.54*	−159.23***	−199.04**	−268.42***
	(61.43)	(47.77)	(39.21)	(55.66)	(55.89)	(60.05)	(52.49)	(54.45)	(78.24)	(77.38)
Water balance (WB)	17.49***	17.18***	17.06***	17.10***	16.97***	16.94***	17.14**	17.28***	17.38**	17.38***
	(5.15)	(5.12)	(5.14)	(5.16)	(5.18)	(5.19)	(5.20)	(5.17)	(5.15)	(5.15)
WB^2	0.18**	0.17**	0.17**	0.17**	0.17***	0.17**	0.17**	0.17**	0.18**	0.18**
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Irrigation (IR)	418.04	387.46	345.82	357.07	351.61	351.84	356.13	346.43	339.11	351.20
	(283.66)	(286.04)	(288.54)	(294.03)	(292.16)	(293.16)	(288.87)	(281.65)	(281.27)	(277.03)
WB × IR	−44.13***	−44.43***	−44.43***	−44.72***	−44.45***	−44.43***	−44.66***	−45.33***	−45.39***	−45.59***
	(15.11)	(15.48)	(15.51)	(15.66)	(15.66)	(15.86)	(15.83)	(15.58)	(15.51)	(15.32)
WB^2 × IR	−0.47**	−0.47**	−0.46**	−0.47**	−0.46**	−0.46**	−0.47**	−0.48**	−0.47**	−0.48**
	(0.20)	(0.21)	(0.21)	(0.21)	(0.21)	(0.21)	(0.21)	(0.21)	(0.21)	(0.21)
Season length	−2.27**	−2.29**	−2.34**	−2.35**	−2.36***	−2.37***	−2.37***	−2.36***	−2.35**	−2.36***
	(0.89)	(0.90)	(0.89)	(0.90)	(0.90)	(0.90)	(0.90)	(0.90)	(0.90)	(0.89)
ln($AREA$)	−176.03***	−178.38***	−177.74***	−179.14***	−177.89***	−176.27***	−175.74***	−176.58***	−174.99***	−176.80***
	(22.77)	(23.59)	(22.81)	(22.91)	(22.90)	(22.60)	(22.70)	(22.58)	(22.74)	(22.65)
Longitude	2652.50***	2642.40***	2639.27***	2580.98***	2584.23***	2528.84***	2485.28***	2460.29***	2437.61***	2447.18***
	(711.86)	(710.87)	(707.95)	(718.08)	(704.95)	(717.52)	(717.78)	(712.56)	(691.20)	(684.86)
Latitude	−3750.76***	−3740.78***	−3745.43***	−3704.66***	−3737.15***	−3769.75***	−3699.76***	−3628.35***	−3600.40***	−3503.20***
	(1127.16)	(1129.86)	(1131.27)	(1145.88)	(1144.54)	(1129.58)	(1128.09)	(1125.61)	(1123.87)	(1132.23)
PC_NEAR_IRR	−263.41***	−78.04	117.72	77.98	155.50	213.33	157.13	180.39	314.26*	121.37
	(93.61)	(89.82)	(119.24)	(119.55)	(145.13)	(140.64)	(154.78)	(158.57)	(168.79)	(171.09)
PC_NEAR_FARM	50.74	2.79	−4.08	−13.08	4.65	54.37	76.09	78.60	103.05	126.00
	(36.94)	(32.15)	(49.64)	(61.62)	(64.84)	(76.64)	(82.19)	(90.14)	(94.93)	(99.30)
Number of farms	3 994	3 994	3 994	3 994	3 994	3 994	3 994	3 994	3 994	3 994
Number of fields	9 441	9 441	9 441	9 441	9 441	9 441	9 441	9 441	9 441	9 441

Note: ***, ** and * indicate 1, 5 and 10 per cent significance levels, respectively. Robust standard errors are in parentheses, standard errors are clustered at the field level. Farm and crop fixed effects are included but not reported. Treatments include field specific variables and the Shannon index, calculated for ℓ between 0.75 and 3 km.

5.2 Impact on the variability of production

5.2.1 Regression specification

Following the empirical strategy developed in section 3, we now use the error terms of the regression presented in (8) to estimate the contribution of crop biodiversity to production variance. Our theoretical model is built to illustrate the way crop biodiversity impacts agricultural production level and variance. It therefore assumes that the production variance is mainly driven by biodiversity: it shows how to capture the effects of explanatory variables other than S_i on the average production but not on its variance. Nevertheless, in the following regression, we also include the vector X to investigate the main drivers of agricultural production variance and their relative importance. Then, the regression specification is given by

$$\epsilon_i^2 = \delta_{0\ell} + \sum_k \delta_{k\ell} X_{ik\ell} + \delta_{S\ell} S_{i\ell} + \epsilon'_{i\ell} \tag{9}$$

where the vector $X_{i\ell}$ is the same vector of k control variables as the one described in section 5.1.1.

5.2.2 Results

Regression results show that crop biodiversity has a significant and negative impact on agricultural production variance both when we consider all crops together (table 9) and maize (table 10) alone. This impact is verified only for perimeters having a radius equal or smaller to 2 kilometres.

More generally, variance of agricultural production (both for maize and for all crops) appears to be driven by two main effects: cultivation intensity and water management. A low cultivation intensity has a negative impact on variance (leading to its decrease), as it is shown by the contribution of the Shannon index and of the area of the field, measured by $\ln(AREA)$. An increase in the surrounding biodiversity, i.e. a decrease in cultivation intensity, measured by the Shannon index, leads to a decrease in production variance, all else being equal. In parallel, when the area of the considered field increases, variance increases. Indeed, a larger field forms a bigger island for parasitic species. Their attacks being more frequent, the variance of the production increases. Water management also has a significant impact. Water balance has a positive impact on variance, as expected. Indeed, water balance is measured by evapotranspiration. The larger is evapotranspiration, i.e. the higher is the water stress of plants, the greater is the production impediment and the higher is the variance in the production. Irrigation makes crops less reliant on water balance, in a non linear fashion this time, but does not have a significant impact on variance by itself. Furthermore, irrigated surrounding fields have a negative significant

impact on variance. They probably allow some indirect increase in water availability. This effect confirms the role that water management can have in the reduction of production variance.

Table 9: Contribution to production variability – Regression results, all crops

Variables	$\ell = 0.75$km	$\ell = 1$km	$\ell = 1.25$km	$\ell = 1.5$km	$\ell = 1.75$km	$\ell = 2$km
Shannon index (S_ℓ)	-20.19^{**}	-12.04^{***}	-24.04^{***}	-13.31^{**}	-19.06^{***}	-17.72^{***}
	(7.80)	(4.18)	(4.27)	(5.88)	(5.53)	(5.43)
Water balance (WB)	6.05^{***}	6.07^{***}	6.12^{***}	6.02^{***}	5.99^{***}	5.93^{***}
	(1.79)	(1.80)	(1.78)	(1.78)	(1.77)	(1.78)
WB^2	0.06^{***}	0.06^{***}	0.06^{***}	0.06^{***}	0.06^{***}	0.07^{***}
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Irrigation (IR)	-41.78	-41.13	-36.87	-43.84	-44.71	-48.74
	(35.63)	(35.93)	(36.75)	(37.47)	(37.15)	(36.62)
WB × IR	-8.30^{***}	-8.15^{***}	-8.19^{***}	-8.26^{***}	-8.21^{***}	-8.05^{***}
	(1.72)	(1.74)	(1.77)	(1.77)	(1.76)	(1.72)
WB^2 × IR	-0.15^{***}	-0.15^{***}	-0.15^{***}	0.15^{***}	0.15^{***}	-0.14^{***}
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Season length	-0.16	-0.16	-0.17	-0.17	-0.17	-0.17
	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)
ln($AREA$)	21.02^{***}	20.81^{***}	19.92^{***}	20.27^{***}	20.26^{***}	20.79^{***}
	(3.71)	(3.74)	(3.72)	(3.70)	(3.69)	(3.68)
Longitude	95.89	91.88	84.08	86.08	87.28	94.57
	(80.03)	(81.26)	(82.57)	(83.31)	(83.13)	(81.93)
Latitude	27.88	28.54	34.80	39.56	32.19	29.75
	(129.13)	(129.19)	(135.95)	(133.71)	(137.06)	(137.47)
PC_NEAR_IRR	-62.70^{***}	-55.32^{***}	-61.44^{***}	-48.31^{***}	-46.57^{**}	-32.64
	(13.62)	(17.14)	(17.13)	(17.95)	(22.05)	(21.40)
PC_NEAR_FARM	11.81^{***}	9.20^{***}	5.42	4.98	3.90	3.72
	(3.26)	(3.06)	(3.42)	(11.09)	(4.32)	(3.78)

Note: ***, ** and * indicate 1, 5 and 10 per cent significance levels, respectively. Robust standard errors are in parentheses, standard errors are clustered at the field level. Farm and crop fixed effects are included but not reported. Treatments include field specific variables and the Shannon index, calculated for ℓ between 0.75 and 3 km. Treatments in which the Shannon index has not a significant contribution are not reported.

Table 10: Contribution to production variability – Regression results, maize

Variables	$\ell = 0.75$km	$\ell = 1$km	$\ell = 1.25$km	$\ell = 1.75$km	$\ell = 2$km
Shannon index (S_ℓ)	-30.75^{***}	-16.10^{***}	-26.96^{***}	-18.54^{**}	-14.73^{**}
	(11.25)	(7.71)	(4.94)	(7.25)	(7.23)
Water balance (WB)	8.67^{***}	8.62^{***}	8.58^{***}	8.63^{***}	0.08^{***}
	(3.32)	(3.30)	(3.28)	(3.33)	(3.34)
WB^2	0.08^{***}	0.08^{***}	0.08^{***}	0.08^{***}	0.08^{***}
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Irrigation (IR)	-113.87	-116.72	-110.52	-117.98	-119.53
	(76.28)	(75.18)	(75.37)	(77.50)	(77.29)
WB × IR	-14.01^{***}	-14.10^{***}	-13.77^{***}	-13.86^{***}	-13.64^{***}
	(3.23)	(3.19)	(3.20)	(3.27)	(3.27)
WB^2 × IR	-0.25^{***}	-0.25^{***}	-0.24^{***}	-0.24^{***}	-0.24^{***}
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Season length	-0.24	-0.25	-0.26	-0.25	-0.25
	(0.17)	(0.17)	(0.17)	(0.17)	(0.18)
$\ln(AREA)$	29.12^{***}	28.75^{***}	28.00^{***}	28.39^{***}	28.78^{***}
	(3.69)	(3.74)	(3.71)	(3.72)	(3.73)
Longitude	93.10	84.20	83.79	91.38	97.90
	(81.91)	(82.91)	(84.46)	(85.28)	(84.83)
Latitude	24.26	36.82	45.38	36.88	33.14
	(107.98)	(106.66)	(110.04)	(111.12)	(110.44)
PC_NEAR_IRR	-42.91^{***}	-32.33^{***}	-25.37^{**}	-1.59	18.44
	(15.35)	(11.81)	(11.82)	(15.78)	(18.78)
PC_NEAR_FARM	13.45	10.52^{***}	1.01	-2.12	-1.73
	(4.63)	(3.74)	(4.01)	(4.76)	(5.56)

Note: ***, ** and * indicate 1, 5 and 10 per cent significance levels, respectively. Robust standard errors are in parentheses, standard errors are clustered at the field level. Farm and crop fixed effects are included but not reported. Treatments include field specific variables and the Shannon index, calculated for ℓ between 0.75 and 3 km. Treatments in which the Shannon index has not a significant contribution are not reported.

6 Conclusion

Using a new large database built from satellite imagery, we confirm that crop biodiversity has a positive impact on agricultural production, with the exception of sunflower. Maintaining a large diversity of crops in the landscape increases agricultural production level while decreasing its variance. These impacts, that were previously described at regional scales, are robust when we consider a whole country. Furthermore, we show the consistency of these results with the underlying ecologic and agricultural mechanisms. For this purpose, we build a probabilistic model in which stochastic factors linked to biodiversity, namely pests, are endogenous, as it is shown in the ecology literature.

Nevertheless, the impacts of biodiversity on the skewness of production, i.e. on the probability of extreme events, is not tested at this stage, while Di Falco & Chavas (2009) found that they are significant. Our reduced form approach could be developed with a third step to test for these effects. Alternatively, a structural approach could be built on

the theoretical model we use to describe the relationships between crop production and biodiversity. The parameters of the beta-binomial distribution could be estimated to fully describe the impacts of biodiversity on the production distribution.

However, from the results we already have, crop diversification can be seen as a possible strategy to increase agricultural productivity or to maintain its level while decreasing the use of pesticides.

References

Bastin, G., Scarth, P., Chewings, V., Sparrow, A., Denham, R., Schmidt, M., O'Reagain, P., Shepherd, R., & Abbott, B. (2012). Separating grazing and rainfall effects at regional scale using remote sensing imagery: a dynamic reference-cover method. *Remote Sensing of Environment*, 121(0), 443–457.

Baumgärtner, S. (2006). *Measuring the diversity of what? And for what purpose? A conceptual comparison of ecological and economic biodiversity indices.* Technical report, Department of Economics, University of Heidelberg.

Bellora, C. & Bourgeon, J.-M. (2015). *Agricultural trade, biodiversity effects and food price volatility.* working papier 5417, CESifo.

Benedetti, R. & Rossini, P. (1993). On the use of NDVI profiles as a tool for agricultural statistics: the case study of wheat yield estimate and forecast in Emilia Romagna. *Remote Sensing of Environment*, 45(3), 311–326.

Brock, W. A. & Xepapadeas, A. (2003). Valuing biodiversity from an economic perspective: A unified economic, ecological, and genetic approach. *American Economic Review*, 93(5), 1597–1614.

Carew, R., Smith, E. G., & Grant, C. (2009). Factors influencing wheat yield and variability : evidence from Manitoba, Canada. *Journal of Agricultural and Applied Economics*, 41(3), 625–639.

Chen, J., Shiyomi, M., Bonham, C., Yasuda, T., Hori, Y., & Yamamura, Y. (2008). Plant cover estimation based on the beta distribution in grassland vegetation. *Ecological Research*, 23(5), 813–819.

DAFF (2009). *Sunflower.* Technical report, Department of Agriculture, Forestry and Fisheries, Republic of South Africa.

DAFF (2014a). *Maize.* Technical report, Department of Agriculture, Forestry and Fisheries, Republic of South Africa.

DAFF (2014b). *Soybeans.* Technical report, Department of Agriculture, Forestry and Fisheries, Republic of South Africa.

Das, D. K., Mishra, K. K., & Kalra, N. (1993). Assessing growth and yield of wheat using remotely-sensed canopy temperature and spectral indices. *International Journal of Remote Sensing*, 14(17), 3081–3092.

Di Falco, S. & Chavas, J. (2006). Crop genetic diversity, farm productivity and the management of environmental risk in rainfed agriculture. *European Review of Agricultural Economics*, 33(3), 289–314.

Di Falco, S. & Chavas, J.-P. (2009). On crop biodiversity, risk exposure, and food security in the highlands of Ethiopia. *American Journal of Agricultural Economics*, 91(3), 599–611.

Doraiswamy, P. C. & Cook, P. W. (1995). Spring wheat yield assessment using NOAA AVHRR data. *Canadian Journal of Remote Sensing*, 21, 43–51.

Eidenshink, J. C. (1992). The 1990 conterminous U.S. AVHRR dataset. *Photogrammetric Engineering and Remote Sensing*, 58(6), 809–813.

Fernandez-Cornejo, J., Jans, S., & Smith, M. (1998). Issues in the economics of pesticide use in agriculture: a review of the empirical evidence. *Review of Agricultural Economics*, 20(2), 462–488.

Ferreira, S. L., Newby, T., & du Preez, E. (2006). Use of remote sensing in support of crop area estimates in South Africa. In *ISPRS WG VIII/10 Workshop 2006, Remote Sensing Support to Crop Yield Forecast and Area Estimates*.

Groten, S. M. E. (1993). NDVI – crop monitoring and early yield assessment of Burkina Faso. *International Journal of Remote Sensing*, 14(8), 1495–1515.

Gupta, R., Prasad, S., Rao, G., & Nadham, T. (1993). District level wheat yield estimation using NOAA/AVHRR NDVI temporal profile. *Advances in Space Research*, 13(5), 253–256.

Hayes, M. J. & Decker, W. L. (1996). Using NOAA AVHRR data to estimate maize production in the United States Corn Belt. *International Journal of Remote Sensing*, 17(16), 3189–3200.

Hochheim, K. & Barber, D. G. (1998). Spring wheat yield estimation for Western Canada using NOAA NDVI data. *Canadian journal of remote sensing*, 24, 17–27.

Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., & Ferreira, L. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1-2), 195–213.

Hughes, G. & Madden, L. (1993). Using the beta-binomial distribution to describe aggegated patterns of disease incidence. *Phytopathology*, 83(7), 759–763.

Irvine, K. M. & Rodhouse, T. J. (2010). Power analysis for trend in ordinal cover classes: implications for long-term vegetation monitoring. *Journal of Vegetation Science*, 21(6), 1152–1161.

Jönsson, P. & Eklundh, L. (2002). Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*, 40(8), 1824–1832.

Jönsson, P. & Eklundh, L. (2004). TIMESAT – a program for analyzing time-series of satellite sensor data. *Computers & Geosciences*, 30(8), 833–845.

Koller, M. & Upadhaya, S. K. (2005). Prediction of processing tomato yield using a crop growth model and remotely sensed aerial images. *Transactions of the ASAE*, 48(6), 2335–2341.

Labus, M. P., Nielsen, G. A., Lawrence, R. L., Engel, R., & Long, D. S. (2002). Wheat yield estimates using multi-temporal NDVI satellite imagery. *International Journal of Remote Sensing*, 23(20), 4169?4180.

Landis, D. A., Gardiner, M. M., van der Werf, W., & Swinton, S. M. (2008). Increasing corn for biofuel production reduces biocontrol services in agricultural landscapes. *Proceedings of the National Academy of Sciences*, 105(51), 20552–20557.

Loreau, M. & Hector, A. (2001). Partitioning selection and complementarity in biodiversity experiments. *Nature*, 412(6842), 72–76.

Ma, B. L., Dwyer, L. M., Costa, C., Cober, E. R., & Morrison, M. J. (2001). Early prediction of soybean yield from canopy reflectance measurements. *Agronomy Journal*, 93(6), 1227.

Nuarsa, I. W., Nishio, F., Nishio, F., Hongo, C., & Hongo, C. (2011). Relationship between rice spectral and rice yield using MODIS data. *Journal of Agricultural Science*, 3(2).

Oerke, E. (2006). Crop losses to pests. *Journal of Agricultural Science*, 144(1), 31–43.

Oerke, E. (2007). *Encyclopedia of Pest Management, Volume 2*, chapter Crop losses to animal pests, plant pathogens and weeds, (pp. 116 – 120). CRC Press.

Pianka, E. R. (2011). *Evolutionary Ecology*. E-book.

Potdar, M. B. (1993). Sorghum yield modelling based on crop growth parameters determined from visible and near-IR channel NOAA AVHRR data. *International Journal of Remote Sensing*, 14(5), 895–905.

Prasad, A. K., Chai, L., Singh, R. P., & Kafatos, M. (2006). Crop yield estimation model for Iowa using remote sensing and surface parameters. *International Journal of Applied Earth Observation and Geoinformation*, 8(1), 26–33.

Quarmby, N. A., Milnes, M., Hindle, T. L., & Silleos, N. (1993). The use of multi-temporal NDVI measurements from AVHRR data for crop yield estimation and prediction. *International Journal of Remote Sensing*, 14(2), 199–210.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27, 379–423, 623–656.

Shiyomi, M., Takahashi, S., & Yoshimura, J. (2000). A measure for spatial heterogeneity of a grassland vegetation based on the beta-binomial distribution. *Journal of Vegetation Science*, 11(5), 627–632.

Smale, M., Hartell, J., Heisey, P. W., & Senauer, B. (1998). The contribution of genetic resources and diversity to wheat production in the Punjab of Pakistan. *American Journal of Agricultural Economics*, 80(3), 482–493.

Tilman, D. & Downing, J. A. (1994). Biodiversity and stability in grasslands. *Nature*, 367(6461), 363–365.

Tilman, D., Polasky, S., & Lehman, C. (2005). Diversity, productivity and temporal stability in the economies of humans and nature. *Journal of Environmental Economics and Management*, 49(3), 405–426.

Tilman, D., Wedin, D., & Knops, J. (1996). Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature*, 379(6567), 718–720.

Weissteiner, C. J. & Kuhbauch, W. (2005). Regional yield forecasts of malting barley (Hordeum vulgare L.) by NOAA-AVHRR remote sensing data and ancillary data. *Journal of Agronomy and Crop Science*, 191(4), 308–320.

Weitzman, M. L. (2000). Economic profitability versus ecological entropy. *The Quarterly Journal of Economics*, 115(1), pp. 237–263.

Zhang, X., Friedl, M. A., & Schaaf, C. B. (2006). Global vegetation phenology from moderate resolution imaging spectroradiometer (MODIS): evaluation of global patterns and comparison with in situ measurements. *Journal of Geophysical Research*, 111(G4).